Analyse fonctionnelle, analyse spectrale

Philippe Jaming

Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France. CNRS, IMB, UMR 5251, F-33400 Talence, France.

Email address: Philippe.Jaming@math.u-bordeaux.fr

Contents

Chapter 1. Topologie en dimension infinie 5
1. Compacité 5
2. Au tour du thèorème de Baire 20
3. Densité 26

Chapter 2. Application de l'analyse fonctionnelle à l'analyse de Fourier 31
1. Rappel de base sur les séries de Fourier 31
2. Convergence et divergence des séries de Fourier 37
3. Quelques applications des séries de Fourier 51

Chapter 3. Opérateurs bornés, opérateurs compacts 61
1. Opérateurs bornés, critère de Schur, adjoint 61
2. Adjoint d'un opérateur 66
3. Opérateurs compacts 70
4. Opérateurs de Hilbert-Schmidt 72
5. Théorie spectrale 76

CHAPTER 1

Topologie en dimension infinie

1. Compacité

1.1. Définitions et premières propriétés. Commençons par la définition topologique d'un ensemble compact (au sens de Borel-Lebesgue):

DÉFINITION 1.1. Soit (X,d) un espace métrique *complet*. On dit que X est compact si de tout recouvrement de X par des ouverts de X, $X=\bigcup_{i\in I}U_i$, U_i ouvert, on peut extraire un sous-recouvrement fini: $X=\bigcup_{i\in F}U_i$ avec $F\subset I$, *fini*.

Notons qu'en passant au complémentaire, X est compact si de toute famille de fermés d'intersection vide, $\bigcap_{i\in I}F_i=\emptyset$, F_i fermé, on peut extraire une sous-famille finie qui soit d'intersection vide: il existe $F\subset I$ fini tel que $\bigcap_{i\in F}F_i=\emptyset$.

Avant d'explorer cette notion et son utilisation, rappelons le résultat suivant du cours de topologie:

THÉORÈME 1.2. *Soit* (X,d) *un espace métrique* complet. *On a équivalence entre les proprétés suivantes:*

(1) *X est un espace compact.*
(2) *X est totalement borné: pour tout $\varepsilon>0$, il existe $F\subset X$* fini *tel que* $X=\bigcup_{x\in F}B(x,\varepsilon)$.
(3) *X est séquentiellement compact (compacité de Bolzano-Weierstrass) pour toute suite $(x_n)_{n\geq 1}\subset X$, il existe $(n_k)_{k\geq 1}$ une suite strictement croissante d'entiers tel que $(x_{n_k})_k$ converge dans X.*

REMARQUE 1.3. Un ensemble totalement borné est borné: en effet, il existe E fini tel que $X=\bigcup_{x\in E}B(x,1)$. Soit $x_0\in E$ et $R=\max_{x\in E}d(x_0,x)$ alors, si $y\in X$, il existe $x\in E$ tel que $d(x,y)\leq 1$. Mais alors $d(x_0,y)\leq d(x_0,x)+d(x,y)\leq R+1$ c'est-à-dire $y\in B(x_0,R+1)$.

REMARQUE 1.4. Fréquemment $X\subset E$ où E est un espace métrique complet (par exemple, un espace de Banach). On note d la distance sur E qui est aussi la distance sur X. Rappelons que les boules de X sont alors les boules de E restreintes à X: $B_X(x,r)=B_E(x,r)\cap X$. Par exemple, dans $[0,1]$, $[0,1/2)=(-1/2,1/2)\cap[0,1]$ est une boule ouverte.

On dira alors que X est totalement borné si, pour tout $\varepsilon>0$, il existe F fini tel que $X\subset\bigcup_{x\in F}B_E(x,\varepsilon)$ ce qui implique que $X=\bigcup_{x\in F}B_E(x,\varepsilon)\cap X=\bigcup_{x\in F}B_X(x,\varepsilon)$.

Enfin, rappelons que si $X\subset E$ avec E complet, alors X est complet si et seulement si X est fermé dans E.

COROLLAIRE 1.5. *Soit (X,d)un espace métrique compact. Alors X est séparable.*

Rappelons que X est séparable s'il existe un ensemble dénombrable dense $\mathcal{D}$ dans X. Ainsi, pour tout $x\in X$ et tout $\varepsilon>0$, il existe un élément $x_\varepsilon\in\mathcal{D}$tel que $d(x,x_\varepsilon)<\varepsilon$.

DÉMONSTRATION. Il est facile de construire cet ensemble: pour $n \in \mathbb{N}^*$, il existe un ensemble fini F_n tel que $X = \bigcup_{x\in F_n} B(x, 1/n)$. Alors $\mathcal{D} = \bigcup_{n\in\mathbb{N}^*} F_n$ est dense et dénombrable puisque réunion dénombrable d'ensembles finis. □

EXEMPLE 1.6. Dans $\mathbb{R}$, tout intervalle fermé borné $[a, b]$ est compact.

Il y a deux façons de faire cela: en montrant que $[a, b]$ est séquentiellement compact. On prend une suite $(x_n) \in [a, b]$ dont on va extraire une sous-suite de Cauchy qui est donc convergente. Quitte à remplacer x_n par $a + \frac{x_n-a}{b-a}$, on peut supposer que $[a, b] = [0, 1]$.

On pose $I_0 = [0, 1]$ et on découpe I_0 en deux intervalles de longeur $1/2$, $[0, 1/2]$ et $[1/2, 1]$. L'un au moins des deux, qu'on notera I_1 contient une infinité de termes de la suite $(x_n)_{n\in\mathbb{N}}$ *i.e.* $N_1 = \{n \in \mathbb{N}, x_n \in I_1\}$ est infini. On pose $n_1 = \min N_1$. Puis on recommence en découpant I_1 en deux intervalles de longueur $1/2^2$. L'un au moins des deux, qu'on notera I_2 contient une infinité de termes de la suite $(x_n)_{n\in N_1}$ *i.e.* $N_2 = \{n \in N_1 \setminus \{n_1\}, x_n \in I_1\}$ est infini. On pose $n_2 = \min N_2$.

Par récurrence, on construit ainsi une suite N_k de parties emboitées de $\mathbb{N}$ et une suite strictement croissante $n_k = \min N_k$, $N_{k+1} \subset N_k \setminus \{n_k\}$ et une suite d'intervalles I_k de longueur 2^{-k} telle que $(x_n)_{n\in N_k} \subset I_k$. On vérifie immédiatement que $(x_{n_k})_k$ est une suite de Cauchy donc converge.

Alternativement, il est bien plus rapide de remarquer que $[0, 1]$ est totalement borné (et complet puisque fermé dans $\mathbb{R}$ qui est complet). En effet $[0, 1] = \bigcup_{j=0}^{2^n-1} [2^{-n}j, 2^{-n}(j + 1)]$ est un recouvrement par des boules fermées de diamètre 2^{-n}. Si on veut un recouvrement par des boules ouvertes, il suffit de légèrement augmenter leur diamètre, par exemple en remplaçant chaque $[2^{-n}j, 2^{-n}(j + 1)] = \big[2^{-n}(j + 1/2) - 2^{-n-1}, 2^{-n}(j + 1/2) + 2^{-n-1}\big]$ par $(2^{-n}(j+1/2)-2^{-n-1/2}, 2^{-n}(j+1/2)+2^{-n-1/2})\cap[0, 1]$.

D'autres exemples s'obtiennent par produit:

LEMME 1.7. *Soient X, Y deux espaces métriques compacts. Alors $X \times Y$ est compact.*

REMARQUE 1.8. Il faut munir $X \times Y$ d'une distance telle que (x_k, y_k) converge si et seulement si (x_k) et (y_k) convergent toutes deux. Le plus simple est de prendre $d\big((x, y), (x', y')\big) = \max\big(d_X(x, y), d_Y(x', y')\big)$ Dans ce cas $B_{X\times Y}\big((x, y), r\big) = B_X(x, r) \times B_Y(y, r)$. Il est alors facile de voir que $X \times Y$ est complet si X et Y le sont.

DÉMONSTRATION. Soit $\big((x_k, y_k)\big)_k$ une suite d'éléments de $X \times Y$, comme $(x_k)_k$ est une suite de X, elle admet une sous-suite $(x_{n_k^1})$ qui converge. Comme $(y_{n_k^1})_k$ est une suite de Y, elle admet une sous-suite $(y_{n_k^2})$ qui converge et, évidemment, $(x_{n_k^2})$ converge encore. Donc $\big((x_{n_k^2}, y_{n_k^2})\big)_k$ converge et $X \times Y$ est bien compact. □

EXEMPLE 1.9. Dans $(\mathbb{R}^d, \|\cdot\|_\infty)$, les boules fermées $[-R, R]^d$ sont compactes.

Il se trouve que ce lemme se prolonge aux produits finis, mais aussi aux produits infinis, $X = \prod_{j\in J} X_j$ où chaque X_j est un espace topologique. Nous reviendrons au paragraphe suivant sur le cas des produits dénombrables d'espaces métriques compacts, mais avant cela, mentionnons à titre culturel le cas général.

Un élément de X est donc une famille $x = (x_j)_{j\in J}$ avec $x_j \in J$ et on note $x_j = p_j(x)$ (la projection de x sur X_j). On eut munir X d'une topologie, appelée topologie produit. C'est la topologie la moins fine (*i.e.* celle qui a le moins d'ouverts possibles) pour que toutes ces applications soient continues. On peut montrer que les ouverts sont précisément les ensembles de la forme suivante: il existe $F \subset J$ fini tel que $0 = \prod_{j\in J} A_j$ avec $A_j = X_j$ si $j \notin F$ et A_j est un ouvert de X_j si $j \in F$.

THÉORÈME 1.10 (Tikhonov). *Si chaque X_j est compact, alors $X = \prod_{j\in J} X_j$ est compact.*

Ce théorème sort (largement) du cadre de ce cours est nécessite l'axiome du choix, et lui est même équivalent. Toutefois, une démonstration plus simple est possible lorsque J est *dénombrable* et que chaque X_j est un espace métrique.

On suppose donc que, pour $j \in \mathbb{N}$, (X_j, d_j) est un espace métrique compact et $X = \prod_{j=0}^{\infty} X_j$. Les éléments de X sont donc des suites $x = (x_j)_{j\in\mathbb{N}}$ dont le j-ième terme appartient à X_j. Commençons par définir une distance sur X par

$$d(x,y)=\sum_{j=0}^{\infty}\frac{\min\bigl(d_j(x_j,y_j),1\bigr)}{2^{j+1}}.$$

On vérifie sans peine que d est bien défini et est une distance sur X.

On peut montrer comme suit que cette distance définit la topologie produit. Il est facile de voir que, pour tout j, $x\to x_j$ est continu. En effet, si $\varepsilon>0$, $\varepsilon<2^{-j-1}$ alors si $d(x,y)<\varepsilon$ donc

$$\frac{\min\bigl(d_j(x_j,y_j),1\bigr)}{2^{j+1}}\le d(x,y)<\frac{1}{2^{j+1}}.$$

Ainsi, tout d'abord $d_j(x_j,y_j)<1$ et ensuite $d_j(x_j,y_j)<2^n d(x,y)\to 0$ quand $x\to y$. Cela signifie que cette topologie est plus fine que la topologie grossière

Par ailleurs, $B(x,\varepsilon)$, $x\in X$ et $\varepsilon>0$ une boule de X. Soit N un entier tel que $\sum_{j>N}2^{-j-1}<\varepsilon/2$. Si $y\in\prod X_j$ est tel que pour $j=0,\dots,N$ $d_j(y_j,x_j)<\varepsilon/2$ *i.e.* $y_j\in B_j(x_j,\varepsilon/2)$ (une boule de X_j et $y_j\in X_j$ arbitraire pour $j>N$ alors

$$\begin{aligned}d(x,y)&=\sum_{j\le N}\frac{\min(d_j(y_j,x_j),1)}{2^{-j-1}}+\sum_{j>N}\frac{\min(d_j(y_j,x_j),1)}{2^{-j-1}}\\&\le\sum_{j\le N}\frac{\varepsilon/2}{2^{-j-1}}+\sum_{j>N}\frac{1}{2^{-j-1}}\\&<\frac{\varepsilon}{2}\sum_{j=0}^{\infty}\frac{1}{2^{-j-1}}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

Ainsi $B(x,\varepsilon)$ contient la boule $\left(\prod_{j\le N}B_j(x_j,\varepsilon/2)\right)\times\prod_{j>N}X_j$ de la topologie produit. Cela signifie que cette topologie est moins fine que la topologie grossière.

Théorème 1.11 (Théoréme de Tikhonov dénombrable). *Pour $j\in\mathbb{N}$ soit (X_j,d_j) un espace métrique compact. Soit $X=\prod_{j\in J}X_j$ muni de la distance que nous venons de définir. Alors X est compact.*

Proof. Maintenant que $X=\prod X_j$ est un espace métrique, on peut utiliser le fait que les compacts sont les compacts séquentiels. Soit donc $(x^{(k)})_{k\in\mathbb{N}}$ une suite d'éléments de X. Chaque $x^{(k)}$ est lui-même une suite $x^{(k)}=(x_0^{(k)},x_1^{(k)},\dots)$ avec $x_j^{(k)}=p_j(x^{(k)})\in X_j$. Comme X_j est compact, il sera possible d'en extraire une sous-suite $\bigl(x_j^{(k)}\bigr)_{k\in I_j}$. Nous allons maintenant construire les I_j de la façon suivante:

– chaque $I_j\subset\mathbb{N}$ est infini;
– $I_{j+1}\subset I_j$ et, en posant $n_j=minI_j$, $n_{j+1}>n_j$;
– pour chaque j, il existe $x_j\in X_j$ tel que $x_j=\lim_{\substack{k\to+\infty\\k\in I_j}}x_j^{(k)}=\lim_{\substack{k\to+\infty\\k\in I_j}}p_j(x^{(k)})$.

On commence par $\bigl(p_0(x^{(k)})\bigr)_{k\in\mathbb{N}}\subset X_0$. Comme X_0 est compact, on peut extraite une sous-suite convergente de cette suite. En d'autres termes, il existe $I_0\subset\mathbb{N}$ infini tel que $x_0=\lim_{\substack{k\to+\infty\\k\in I_0}}p_0(x^{(k)})$ existe.

Supposons maintenant $I_0,I_1,\dots,I_j$ construits. On considère alors $\tilde I_j=I_j\backslash\{n_j\}$ et la suite $\bigl(p_{j+1}(x^k)\bigr)_{k\in\tilde I_j}$ de X_{j+1}. Comme X_{j+1} est compact, on peut en extraire une sous-suite convergente: il existe donc $I_{j+1}\subset\tilde I_j\subset I_j$ (ce qui garanti $\min I_{j+1}>\min I_j$) et $x_{j+1}=\lim_{\substack{k\to+\infty\\k\in I_{j+1}}}p_{j+1}(x^{(k)})$.

Enfin, on considère la suite $(x^{(n_k)})_{k\ge0}$ qui est une suite extraite de $(x^{(k)})_{k\ge0}$. Montrons qu'elle converge vers $x:=(x_j)_{j\in\mathbb{N}}$. Tout d'abord, comme $(x^{(n_k)})_{k\ge j}$ est une suite extraite de $(x^{(k)})_{k\in I_j}$ on a

$p_j\big(x^{(n_k)}\big) \to x_j$ pour tout j. Fixons alors $\varepsilon > 0$ et prenons N assez grand pour que $\sum_{j>N} 2^{-j-1} < \varepsilon/2$. Il existe ensuite K tel que, si $k \geq K$, pour tout $j \leq N$ on a $d_j\big(p_j\big(x^{(n_k)}\big), x_j\big) < \varepsilon/2$. Mais alors

$$\begin{aligned} d\big(x^{(n_k)}, x\big) &\leq \sum_{j\leq N} \frac{d_j\big(p_j\big(x^{(n_k)}\big), x_j\big)}{2^{j+1}} + \sum_{j>N} \frac{1}{2^{j+1}} \\ &\leq \frac{\varepsilon}{2} \sum_{j\leq N} \frac{1}{2^{j+1}} + \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

On a donc bien $x^{(n_k)} \to x$. □

La démonstration précédente porte le nom de *procédé diagonal de Cantor*. Nous utiliserons ce résultat dans la (deuxième) démonstration du théorème d'Ascoli. Pour comprendre l'idée, imaginons qu'on représente le procédé d'extraction dans un tableau. On met la première suite extraite dans le première colonne, la deuxième suite dans la deuxième colonne et ainsi de suite... La suite extraite choisie est la suite des éléments de la diagonale.

Le lemme suivant permet d'obtenir beaucoup d'autres exemples:

LEMME 1.12. *Soit (X, d) un espace métrique complet. Si X est compact alors toute partie fermée E de X est encore compacte.*

DÉMONSTRATION. Soit (x_n) une suite de E alors elle admet une sous-suite $(x_{n_k})_k$ qui converge dans X *i.e.* il existe $x \in X$ tel que $\lim x_{n_k} = x$. Comme E est fermé, $x \in E$ donc la sous-suite $(x_{n_k})_k$ converge dans E □

EXEMPLE 1.13. Dans $(\mathbb{R}^d, \|\cdot\|_\infty)$, X est compact si et seulement si X est fermé borné.

En effet, un compact est complet donc fermé dans $\mathbb{R}^d$ et totallement borné donc borné. Réciproquement, si X est borné, il est inclus dans un $[-R, R]^d$ qui est compact. Si X est de plus fermé, il sera alors compact.

COROLLAIRE 1.14. *Soient (X, d), (Y, d) deux espaces métriques complets et $f : X \to Y$ une fonction continue. Si X est compact alors $f(X)$ aussi. En particulier f est bornée*

DÉMONSTRATION. Soit (y_n) une suite de $f(X)$. Il existe donc $x_n \in X$ tel que $y_n = f(x_n)$. Comme X est compact, on peut extraire une sous-suite convergente (x_{n_k}). Par continuité, $y_{n_k} = f(x_{n_k})$ converge également.

Alternativement, si $(V_i)_{i\in I}$ est un recouvrement ouvert de $f(X)$, alors $U_i = f^{-1}(V_i)$ est ouvert et $(U_i)_{i\in I}$ recouvre X. Comme X est compact, on peut extraire un sous-recouvrement fini $(U_i)_{i\in F}$ recouvre X et alors $(V_i)_{i\in F}$ recouvre encore $f(X)$.

Pour la deuxième assertion, rappelons qu'un compact est totalement borné donc borné, en particulier, $f(X)$ est borné, c'est-à-dire que f est borné. □

COROLLAIRE 1.15. *Soit (X, d), un espace métrique complet et $f : X \to \mathbb{R}$ une fonction continue. Si X est compact alors f est bornée et atteint ses bornes.*

DÉMONSTRATION. On a déja montré que f est bornée, donc $M = \sup f$ et $m = \inf f$ existent (et sont finis). Par définition, pour tout n, il existe $x_n \in X$ tel que $M - 1/n \leq f(x_n) \leq M$. Par compacité de X, on peut extraire une sous-suite convergente (x_{n_k}) de (x_n). En notant x sa limite et en passant à la limite dans l'inégalité $M - 1/n_k \leq f(x_{n_k}) \leq M$ on trouve $f(x) = M$. Le raisonnement pour l'inf est similaire. □

Donnons un deuxième exemple d'application de la compacité aux fonctions continues:

PROPOSITION 1.16. *Soient (X, d_X), (Y, d_Y) deux espaces métriques complets et $f : X \to Y$ une fonction continue. Si X est compact alors f est uniformément continue.*

DÉMONSTRATION. Soit $\varepsilon > 0$. Si f est continue, pour tout $x \in X$, il existe $\eta_x > 0$ tel que, si $d_X(x, y) \leq 2\eta_x$ alors $d_Y\big(f(y), f(x)\big) \leq \varepsilon$.

Remarquons que $\bigcup_{x\in X} B(x, \eta_x)$ est un recouvrement ouvert de X, il existe donc $E \subset X$ fini tel que $\bigcup_{x\in E} B(x, \eta_x)$ recouvre encore X. Soit $\eta = \inf_{x\in E} \eta_x$, comme E est fini, $\eta > 0$.

Soient maintenant $y, z \in X$ avec $d_X(y,z) \leq \eta$. Comme $\bigcup_{x\in E} B(x,\eta_x)$ recouvre encore X, il existe $x \in E$ tel que $y \in B(x,\eta_x)$ *i.e.* $d_X(x,y) \leq \eta_x$. Par ailleurs, $d_X(x,z) \leq d_X(x,y)+d_X(y,z) \leq \eta_x+\eta \leq 2\eta_x$. Par définition de η_x, on a donc $d_Y\big(f(y),f(x)\big) \leq \varepsilon$ et $d_Y\big(f(z),f(x)\big) \leq \varepsilon$. Avec l'inégalité triangulaire, on en déduit que $d_Y\big(f(y),f(z)\big) \leq 2\varepsilon$ d'où l'uniforme continuité. □

Continuons avec les applications. Rappelons que deux distances d, d' sur un espace métrique X sont équivalents s'il existe $C \geq 1$ tel que, pour tous $x, y \in X$

$$\frac{1}{C}d'(x,y) \leq d(x,y) \leq Cd'(x,y).$$

Il est facile de voir que deux distances équivalentes définissent les mêmes suites de Cauchy et les mêmes suites convergentes. En particulier, s'il est complet pour une distance, il l'est aussi pour l'autre. De plus, si X est totalement borné pour la distance d, *i.e.* pour $\varepsilon > 0$, il existe un recouvrement fini de X par des boules (pour d) de rayon ε, $X = \bigcup_{x\in E} B(x,\varepsilon)$. Mais

$$B(x,\varepsilon) = \{y\,:\, d(x,y) \leq \varepsilon\} \subset \{y\,:\, C^{-1}d'(x,y) \leq \varepsilon\} = B'(x, C\varepsilon)$$

donc $X = \bigcup_{x\in E} B'(x,C\varepsilon)$. Ainsi X admet aussi des recouvrements fini par des boules pour d' de rayon arbitrairement petit et X est encore totalement borné pour d'. Donc (X,d) est compact si et seulement si (X,d') est compact.

Par exemple, on peut munir $\mathbb{C}^d$ de deux normes naturelles

$$\|z\|_\infty = \max_{j=1}^{d} |z_j| \quad \text{and} \quad \|z\|_{r,\infty} \max_{j=1}^{d} \max(|\Re z_j|, |\mathrm{Im}\, z_j|)$$

la seconde revenant à identifier $\mathbb{C}^d$ avec $\mathbb{R}^{2d}$. Pour la seconde, on sait donc déjà que les compacts sont les fermés bornés. Mais ces deux normes sont équivalents puisque $\|z\|_{r,\infty} \leq \|z\|_\infty \leq \sqrt{2}\|z\|_{r,\infty}$. Ainsi, dans $(\mathbb{C}^d, \|\cdot\|_\infty)$ aussi, les compacts sont les fermés bornés.

Théorème 1.17. *Soit E un espace vectoriel de dimension finie. Sur E toutes les normes sont équivalentes. Les parties compactes (quelle que soit la norme sur E) sont les parties fermées bornées.*

Démonstration. Commençons par fixer une base $(e_1, \ldots, e_d)$ et munissons E de la norme $\|x\| = \left\|\sum_{j=1}^d x_j e_j\right\| = \max_{j=1,\ldots,d} |x_j|$ quand $x = \sum x_i e_i$. Ceci permet d'identifier E avec $\mathbb{R}^d$ muni de la norme $\|\cdot\|_\infty$ si E est réel et avec $\mathbb{R}^d$ muni de la norme $\|\cdot\|_\infty$ si E est complexe. Dans la suite, nous allons supposer que E es réel, le cas complexe étant simimaire. On a déjà vu que dans $(\mathbb{R}^d, \|\cdot\|_\infty)$ les compacts sont les fermés bornés.

Montrons maintenant que toutes les normes sont équivalentes, cette propriété ne dépendra donc plus de la norme choisie.

Soit $\|\cdot\|$ une norme sur $\mathbb{C}^d$. Alors, en écrivant $x = \sum x_i e_i$

$$\|x\| = \left\|\sum_{i=1}^d x_i e_i\right\| \leq \sum_{i=1}^d |x_i|\|e_i\| \leq \left(\sum_{i=1}^d \|e_i\|\right)\|x\|_\infty.$$

Cela montre de plus que $x \to \|x\|$ est continue $(\mathbb{R}^d, \|\cdot\|_\infty) \to \mathbb{R}$ puisque $\|x-y\| \leq \left(\sum_{i=1}^d \|e_i\|\right)\|x-y\|_\infty \to 0$ quand $x \to y$. Par ailleurs, $S = \{x \in \mathbb{R}^d\,:\, \|x\|_\infty = 1\}$ est fermé dans la boule unité et est donc compact. Ainsi, $\|x\|$ atteint son minimum m sur S en un point $x_0 \in S$: $m = \|x_0\| > 0$ puisque $x_0 \neq 0$. Mais alors, si $x \in \mathbb{R}^d$ soit $x = 0$ et $\|x\| = \|x\|_\infty = 0$ soit $x \neq 0$ et alors $\left\|\dfrac{x}{\|x\|_\infty}\right\|_\infty = 1$ donc $\left\|\dfrac{x}{\|x\|_\infty}\right\| \geq m$ d'où $\|x\| \geq m\|x\|_\infty$. □

Corollaire 1.18. *Soit E un espace vectoriel normé de dimension finie et $X \subset E$. Alors X est totallement borné si et seulement si X est borné.*

Démonstration. On a déjà vu que si X est totallement borné alors X est borné. Inversement, si X est borné alors son adhérence $\bar{X}$ est fermée bornée. Ainsi $\bar{X}$ est compact donc totalement borné. Ainsi, pour tout $\varepsilon > 0$, il existe une partie finie F de $\bar{X}$ telle que $\bar{X} \subset \bigcup_{x\in F} B(x,\varepsilon)$. Mais pour chaque $x \in \bar{X}$, il existe $y_x \in X$ tel que $x \in B(y_x, \varepsilon)$ et alors $B(x,\varepsilon) \subset B(y_x, 2\varepsilon)$. On écrit alors $\tilde{F} = \{y_x : x \in F\}$ et on a $X \subset \bar{X} \subset \bigcup_{x\in\tilde{F}} B(x, 2\varepsilon)$. Ainsi X est totalement borné. $\square$

Notons qu'en modifiant légèrement l'argument ci-dessus, on peut montrer que X est totalement borné si, pour tout $\varepsilon > 0$ il existe $F \subset E$ fini (pas forcément $F \subset X$) tel que $X \subset \displaystyle\bigcup_{x\in F} B(x,\varepsilon)$.

Corollaire 1.19. *Soit $T : \mathbb{R}^d \to \mathbb{R}^D$ une application linéaire, alors T est continue.*

Démonstration. Il suffit que ce soit vrai pour $\mathbb{R}^d$ et $\mathbb{R}^D$ muni chacun de sa norme $\|\cdot\|_\infty$. On note $[a_{i,j}]_{1\le i\le D, 1\le j\le d}$ la matrice de T dans la base canonique et $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. Alors

$$Tx = \left(\sum_{j=1}^d a_{1,j}x_j, \ldots, \sum_{j=1}^d a_{D,j}x_j\right)$$

donc

$$\begin{aligned}\|Tx\|_\infty &= \sup_{i=1,\ldots,D}\left|\sum_{j=1}^d a_{i,j}x_j\right| \le \sup_{i=1,\ldots,D}\sum_{j=1}^d |a_{i,j}||x_j| \\ &\le \left(\sup_{i=1,\ldots,D}\sum_{j=1}^d |a_{i,j}|\right)\|x\|_\infty\end{aligned}$$

comme annoncé. $\square$

Nous avons vu au premier semestre des exemples de normes non-équivalentes et des exemples d'applications linéaires non continues en dimension infinie. Les choses se compliquent donc en dimension infinie et cela provient du manque de compacité de la boule unité.

Exemple 1.20. Soit H un espace de Hilbert de dimension infinie, alors la boule unité n'est pas compacte:

Quitte à passer à un sous-espace séparable, on peut supposer que H lui-même est séparable. Alors H admet une base orthonormée $(e_n)_{n\in\mathbb{N}}$. En particulier $\|e_n\| = 1$ c'est donc une suite dans la boule unité. Mais, pour $n \neq m$

$$\|e_n - e_m\|^2 = \|e_n\|^2 + \|e_m\|^2 - 2\Re\langle e_n, e_m\rangle = 1 + 1 - 0 = 2$$

donc on ne peut pas extraire de sous-suite de Cauchy de $(e_n)_{n\in\mathbb{N}}$, a forciori de sous-suite convergente.

Ce cas est typique de la dimension infinie:

Théorème 1.21 (Riesz). *Soit $(E, \|\cdot\|)$ un espace vectoriel normé. On a équivalence entre*

(1) *La boule unité fermée de E est compacte;*
(2) *E est de dimension finie.*

Démonstration. On a déjà vu que si E est de dimension finie, alors la boule unité fermée de E est compacte. Dans le cas où E est un espace de Hilbert, cela provient de l'existence de bases orthonormées qui elle-même provient de l'existence de projections sur des sous-espaces fermés. La démonstration va consister en une construction d'une suite infinie équidistante. Celle-ci va résulter du lemme suivant:

Lemme 1.22. *Soit E un espace vectoriel normé et $F \subset E$ un sous-espace fermé strict, $F \neq E$. Alors il existe $x_0 \in E \setminus F$ avec $\|x_0\| = 1$ et, pour tout $y \in F$, $\|x - y\| \ge \dfrac{1}{2}$ i.e. $dist(x_0, F) \ge \dfrac{1}{2}$.*

Si F est de dimension finie (donc nécessairement fermé), on a même $dist(x_0, F) \ge 1$.

Montrons d'abord que ce lemme permet de conclure.

Soit $x_1 \in E$ avec $\|x_1\| = 1$ et appliquons le lemme à $F_1 = Vect(x_1)$. Il existe alors $x_2 \in E$ avec $\|x_2\| = 1$ et $\|x_2 - x_1\| \ge 1$. Appliquons le lemme à $F_2 = Vect(x_1, x_2)$. Il existe alors $x_3 \in E$ avec $\|x_3\| = 1$ et $dist(x_3, F_2) \ge 1$, en particulier $\|x_2 - x_1\| \ge 1$ et $\|x_3 - x_1\| \ge 1$.

Une fois $x_1, \dots, x_n$ construits, on applique le lemme à F_n l'espace engendré par les vecteurs déjà construits, $F_n = \mathrm{Vect}(x_1, \dots, x_n)$. Comme E est de dimension infini, $F_n \neq E$ et F_n est évidemment fermé puisque de dimension finie. On obtient donc un vecteur $x_{n+1} \in E$ avec $\|x_{n+1}\| = 1$ et $dist(x_{n+1}, F_2) \geq 1$ donc $\|x_{n+1} - x_j\| \geq 1$ pour $j = 1, \dots, n$.

On a ainsi construit une suite $(x_n)_{n\geq 1}$ de la boule unité de E qui vérifie $\|x_n - x_m\| \geq 1$ si $n \neq m$. Une telle suite ne peut contenir de sous-suite convergente donc la boule unité de E n'est pas compacte. $\square$

Démonstration du lemme. Soit $x \in E \setminus F$ et

$$d = dist(x, F) = \inf_{y\in F} \|x - y\|.$$

Comme $0 \in F$, $d \leq \|x\|$ et, si $y \in F$ est tel que $\|y\| \geq 3\|x\|$,

$$\|x - y\| \geq \|y\| - \|x\| \geq 3\|x\| - \|x\| = 2\|x\| > d.$$

Ainsi

$$d = \inf_{y\in F, \|y\|\leq 3\|x\|} \|x - y\|.$$

Premier cas: si F est de dimension finie, $\{y \in F, \|y\| \leq 3\|x\|\}$ est un compact, il existe donc $y_0 \in F$ avec $\|y_0\| \leq 3\|x\|$ tel que $\|x - y_0\| = d$. Si on avait $d = 0$ alors $x = y_0 \in F$ ce qui contredit l'hypothèse $x \in E \setminus F$. On peut donc poser $x_0 = \dfrac{x - y_0}{\|x - y_0\|} = \dfrac{x - y_0}{d}$ et on remarque que $\|x_0\| = 1$. De plus, si $y \in F$

$$\|x_0 - y\| = \left\| \frac{x - y_0}{d} - y \right\| = \frac{\|x - (y_0 + dy)\|}{d} \geq 1$$

puisque $y_0 + dy \in F$ (qui est un espace vectoriel).

Cas général: si on avait $d = 0$, il existerait $y_n \in F$ avec $\|x - y_n\| \leq d + \dfrac{1}{n} = \dfrac{1}{n}$. Mais alors, si $m \geq n$,

$$\|y_n - y_m\| \leq \|x - y_n\| + \|x - y_m\| \leq \frac{1}{n} + \frac{1}{m}.$$

Ainsi (y_n) est de Cauchy dans E complet, elle converge donc vers un $y \in E$. Comme $y_n \in F$ et que F est fermé, on a de plus que $y \in F$. Mais, la norme étant continue $0 = \lim \|x - y_n\| = \|x - y\|$ donc $x - y = 0$ ce qui contredit à nouveau $x \notin F$. Ainsi $d > 0$.

On prend alors $y_0 \in F$ tel que $d \leq \|x - y_0\| \leq 2d$ (n'importe quel nombre $> d$ pourrait remplacer $2d$ ici) et on pose à nouveau $x_0 = \dfrac{x - y_0}{\|x - y_0\|}$. On remarque $\|x_0\| = 1$. De plus, si $y \in F$

$$\|x_0 - y\| = \left\| \frac{x - y_0}{\|x - y_0\|} - y \right\| = \frac{\|x - (y_0 + \|x - y_0\|y)\|}{\|x - y_0\|} \geq \frac{d}{2d} = \frac{1}{2}$$

puisque $y_0 + \|x - y_0\|y \in F$. $\square$

1.2. Trois critères de compacité en dimension infinie. *

Nous avons vu que la compacité est utile mais pas simple à établir en dimension infinie. Nous allons dans cette partie donner des critères de compacité dans les principaux espaces que nous rencontrerons: $\mathcal{C}(K)$, ℓ^p et L^p. Ces critères vont tous résulter d'un lemme qui permettra de se ramener à la dimension finie:

Lemme 1.23. *Soit X un espace métrique. On suppose que, pour tout $\varepsilon > 0$, il existe $\delta > 0$, un espace métrique W et une application $\Phi : X \to W$ telles que*

(i) *$\Phi(X)$ est totalement borné;*
(ii) *si $d\big(\Phi(x), \Phi(y)\big) < \delta$ alors $d(x, y) < \varepsilon$.*

Alors X est totalement borné.

*La présentation ici est inspirée de Hanche-Olsen H., Holden H. *The Kolmogorov–Riesz compactness theorem.* Expo. Math., **28** (2010), pp. 385-394.

Pour utiliser ce lemme, on va en général prendre $W = \mathbb{R}^d$ muni d'une norme. En effet, dans ce cas, $E \subset \mathbb{R}^d$ est totalement borné si et seulement si E est borné. (Car si E est borné, $\bar{E}$ aussi et est donc compact, donc peut être recouvert par des ouverts arbitrairement petits).

Démonstration. Soit $\varepsilon > 0$ et W, δ, ϕ comme dans l'énoncé. Comme $\Phi(X)$ est totalement borné, il existe $V_1, \ldots, V_n$ des boules de rayon $\delta/2$ qui recouvrent $\Phi(X)$.

Il en suit que $\Phi^{-1}(V_1), \ldots, \Phi^{-1}(V_n)$ recouvrent X. De plus, la propriété *ii*) implique que chaque $\Phi^{-1}(V_j)$ a un diamètre au plus ε: si $x, y \in \Phi^{-1}(V_j)$ alors $\Phi(x), \Phi(y) \in V_i$ donc $d\big(\Phi(x), \Phi(y)\big) \leq \delta$ donc $d(x,y) \leq \varepsilon$. Ainsi, chaque $\Phi^{-1}(V_j)$ est inclus dans une boule de rayon ε (en prenant n'importe quel point comme centre). Ces boules fournissent le recouvrement de X cherché. $\square$

Nous pouvons maintenant établir le premier critère de compacité:

Théorème 1.24 (Arzéla-Ascoli). *Soit K un espace métrique compact et $X \subset \mathcal{C}(K)$ (l'espace des fonctions continues sur K). Alors on a équivalence entre*

(1) *X est compact*
(2) *X vérifie les deux propriétés suivantes*
   (a) *X est ponctuellement borné: pour tout $x \in K$, il existe $M_x > 0$ tel que, pour tout $f \in X$, $|f(x)| \leq M_x$.*
   (b) *X est équi-continue: pour tout $\varepsilon > 0$ et tout $x \in K$, il existe η_x tel que, si $d(x,y) < \eta_x$ alors, pour tout $f \in X$, $|f(x) - f(y)| < \varepsilon$.*
   (c) *X est fermé.*

Démonstration. $1 \Rightarrow 2$: Si X est compact, alors il est fermé et borné: il existe M tel que, pour tout $f \in X$, $\|f\|_\infty = \sup_{x\in K} |f(x)| \leq M$ donc pour tout $x \in K$ et tout $f \in X$, $|f(x)| \leq M$ en particulier, X est ponctuellement borné (c'est plus fort puisque M ne dépend pas de x). Reste donc à voir que X est équi-continue.

On fixe $\varepsilon > 0$, comme X est compact, il est totalement borné, il existe donc $f_1, \ldots, f_N$ tel que $X \subset \bigcup_{k=1}^N B(f_k, \varepsilon)$. En d'autres termes, pour tout $f \in X$, il existe $k \in \{1, \ldots, N\}$ tel que $\|f - f_k\|_\infty < \varepsilon$.

Mais chaque f_k est continu sur le compact K, elle est donc uniformément continue. Il existe donc η_k tel que, si $d(x,y) < \eta_k$ alors $|f_k(x) - f_k(y)| < \varepsilon$. On pose $\eta = \min \eta_k > 0$.

Supposons maintenant que $d(x,y) < \eta$ et que $f \in X$. Soit k tel que $\|f - f_k\|_\infty < \varepsilon$. Alors

$$|f(x) - f(y)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_j(y) - f(y)| \leq 3\varepsilon$$

d'où l'équicontinuité des $f \in X$ (c'est même plus fort puisque le η trouvé ici ne dépend pas de x).

$2 \Rightarrow 1$: C'est évidemment le sens le plus important. *iii*) nous dit que X est fermé dans $\mathcal{C}(K)$ qui est complet, donc X est un espace métrique complet. Reste à voir que *i*) et *ii*) impliquent que X est totalement bornée. Soit donc $\varepsilon > 0$.

Pour tout $x \in K$, il existe $\eta_x > 0$ tel que, si $d(x,y) < \eta_x$ alors,

$$\text{pour tout } f \in X,\ |f(x) - f(y)| < \varepsilon. \tag{1.1}$$

On recouvre alors le compact K par les boules $K = \bigcup_{x\in K} B(x, \eta_x)$ et on en extrait un sous-recouvrement ouvert fini: $K = \bigcup_{j=1}^N B(x_j, \eta_{x_j})$. On notera $\eta_j = \eta_{x_j}$. Enfin, on définit l'application

$$\Phi : \begin{array}{ccc} X & \rightarrow & \mathbb{R}^N \\ f & \mapsto & \Phi(f) = \big(f(x_1), \ldots, f(x_N)\big) \end{array}.$$

D'abord $\Phi(X)$ est totalement borné. Comme $\Phi(X) \subset \mathbb{R}^N$, il suffit de voir que $\Phi(X)$ est bornée. Mais X étant ponctuellement borné, pour chaque $j \in \{1, \ldots, N\}$ il existe M_j tel que, pour tout $f \in X$, $|f(x_j)| \leq M_j$. Ainsi, en posant $M = \max M_j < +\infty$, pour tout $f \in X$, $\|\Phi(f)\|_\infty \leq M$.

Ensuite supposons que $f, g \in X$ sont telles que $\|\Phi(f) - \Phi(g)\|_\infty \leq \varepsilon$, c'est-à-dire que pour $j = 1, \ldots, N$, $|f(x_j) - g(x_j)| \leq \varepsilon$. Alors, si $y \in K$, il existe $j \in \{1, \ldots, N\}$ tel que $d(y, x_j) < \eta_j$ et alors

$$\begin{array}{rcl} |f(y) - g(y)| & \leq & |f(y) - f(x_j)| + |f(x_j - g(x_j)| + |g(x_j) - g(y)| \\ & < & \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{array}$$

où les premier et troisième termes sont majorés avec (1.1). Ainsi $\|f-g\|_\infty \leq 3\varepsilon$. Il suffit donc d'appliquer le lemme 1.23. □

Une démonstration un peu plus classique de la deuxième partie.

Étape 1. *L'équi-continuité est uniforme:*†

pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pout tout $x \in K$, si $d(x,y) < \eta_x$ alors, pour tout $f \in X$, $|f(x) - f(y)| < \varepsilon$

Si ce n'était pas le cas, il existerait $\varepsilon_0 > 0$ tel que, pour tout n, il existe $x_n, y_n \in K$ avec $d(x_n, y_n) < 1/n$ et $f_n \in X$ tel que $|f_n(x_n) - f_n(y_n)| > \varepsilon_0$.

Mais K est compact, donc quitte à passer à une sous-suite, on peut supposer que x_n converge vers un $x \in K$. Quitte à encore passer à une sous-suite, on peut également supposer que y_n converge. Comme $d(x_n, y_n) \to 0$, on a $y_n \to x$.

Mais, d'après l'équicontinuité, il existe η_x tel que, si $d(x,y) < \eta_x$ alors, pour tout $f \in X$, $|f(x) - f(y)| < \varepsilon_0/2$. Prenons donc n assez grand pour que $d(x_n, x) < \eta_x$ et $d(y_n, x) < \eta_x$. On a donc $|f_n(x) - f_n(x_n)| < \varepsilon_0/2$ et $|f_n(x) - f_n(y_n)| < \varepsilon_0/2$ donc, avec l'inégalité triangulaire

$$|f_n(x_n) - f_n(y_n)| \leq |f_n(x) - f_n(x_n)| + |f_n(x) - f_n(y_n)| < \varepsilon_0/2 + \varepsilon_0/2 = \varepsilon_0$$

ce qui contredit l'inégalité $|f_n(x_n) - f_n(y_n)| > \varepsilon_0$.

Étape 2. *Conclusion.*

Soit donc X une partie de $\mathcal{C}(K)$ qui est ponctuellement bornée et uniformément équi-continue. Soit (f_n) une suite d'éléments de X. On cherche à extraire de (f_n) une suite qui converge uniformément sur K. Pour cela, il suffit que cette suite soit uniformément de Cauchy (voir cours du premier semestre pour la complétude de $\mathcal{C}(K)$).

On a déjà vu qu'un compact est séparable. On prend donc une partie $\{x_k\}_{k\in\mathbb{N}}$ dénombrable dense dans K. À chaque x_k on associe $X_k = \overline{\{f(x_k) \,:\, f \in X\}}$ l'adhérence de l'ensemble des images de x_k par les éléments de f. Comme X est ponctuellement borné, X_k est une partie bornée de $\mathbb{C}$. Comme X_k est aussi fermée, X_k est compact. D'après le théorème de Tikhonov dénombrable, il en résulte que $\tilde{X} = \prod_{k\in\mathbb{N}} X_k$ est compact.

Mais maintenant, si on pose $y_n = \{f_n(x_k) \,:\, k \in \mathbb{N}\} \in \tilde{X}$, $(y_n)_{n\geq 0}$ est une suite dans un compact, elle admet donc une sous-suite $(y_{n_k})_k$ qui converge. Quitte à remplacer la suite f_n initiale par sa sous-suite f_{n_k}, nous pouvons supposer que (y_n) converge (ce qui a pour seul but d'alléger la notation). Mais si (y_n) converge alors chacune de ses coordonées converge en particulier, est de Cauchy: pour tout k, $\big(f_n(x_k)\big)_n$ est de Cauchy.

On fixe maintenant $\varepsilon > 0$. Il existe donc $\eta > 0$ tel que, si $d(x,y) < \eta$, pour tout n, $|f_n(x) - f_n(y)| < \varepsilon$. Mais maintenant, si $x \in K$, par densité, il existe k tel que $d(x_k, x) < \eta$. En d'autres termes les boules $\{B(x_k, \eta), k \in \mathbb{N}\}$ recouvrent K, un nombre fini d'entre elles recouvrent donc encore K *i.e.* il existe $F \subset \mathbb{N}$ fini tel que $K = \displaystyle\bigcup_{k\in F} B(x_k, \eta)$. Ensuite, comme chaque $\big(f_n(x_k)\big)_n$ est de Cauchy, $k = 0, \ldots, N$, il existe N tel que, si $m, n \geq N$ et $k \in F$ alors $|f_n(x_k) - f_m(x_k)| \leq \varepsilon$.

Soit maintenant $x \in X$ et $k \in F$ tel que $x \in B(x_k, \eta)$ *i.e.* $d(x, x_k) < \eta$. Alors pour tout $m, n \geq N$

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_n(x_k)| + |f_n(x_k) - f_m(x_k)| + |f_m(x_k) - f_m(x_k)| \leq 3\varepsilon.$$

Ainsi f_n est *uniformément de Cauchy* sur K donc converge dans $\mathcal{C}(K)$. □

Le critère de compacité dans ℓ^p est très similaire:

Théorème 1.25 (Fréchet). *Soit $1 \leq p < +\infty$ et $X \subset \ell^p$ alors X est compact si et seulement si*

(i) *X est fermé;*
(ii) *X est ponctuellement borné: pour tout i, il existe M_i tel que, pour tout $(x_n) \in X$, $|x_i| \leq M$*
(iii) *pour tout $\varepsilon > 0$ il existe $N > 0$ tel que, pour tout $(x_n) \in X$, $\sum_{n\geq N} |x_n|^p \leq \varepsilon^p$.*

†C'est pour cette raison que le théorème est souvent ennoncé avec cette hypothèse un peu plus dificile à vérifier.

Démonstration. Dans la démonstration de ce lemme, nous utiliserons la notation suivante. Soit $N \geq 1$ un entier et $\pi_N : \ell^p \to \ell^p$ l'application qui à une suite $y = (y_n)_{n \geq 0}$ associe la suite $(z_n)_n$ définie par $z_n = y_n$ si $n < N$ et $z_n = 0$ si $n \geq N$. Notons également $\tilde{\pi}_N y = y - \pi_N y$ et remarquons que $\|\pi_N y\|_p, \|\tilde{\pi}_N y\|_p \leq \|y\|_p$.

Comme ℓ^p est complet, la condition $i)$ signifie simplement que X muni de la norme de ℓ^p est un espace métrique complet. Il reste donc à voir que les deux autres conditions sont équivalentes au fait que X est totalement borné.

Supposons d'abord que X est totalement borné. En particulier, X est borné, il existe donc M tel que, pour tout $x = (x_n) \in X$, $\|x_n\|_p \leq M$. Fixons $i \in \mathbb{N}$, alors $|x_i| \leq \|x\|_p \leq M$ donc X est bien ponctuellement borné. Ensuite, comme X est totalement borné, il existe $x^1, \ldots, x^M \in X$ tel que, pour tout $x \in X$, il existe $j \in \{1, \ldots, M\}$ tel que $\left\|x - x^{(j)}\right\|_p \leq \varepsilon/2$. Par ailleurs, pour chaque j, il existe N_j tel que $\sum_{n \geq N_j} |x_n^{(j)}|^p \leq (\varepsilon/2)^p$. On pose alors $N = \max N_j$ et on remarque que N a été choisi pour que $\left\|\tilde{\pi}_N x^{(j)}\right\|_p \leq \varepsilon/2$ si $j \in \{1, \ldots, M\}$. Mais alors

$$\left\|\tilde{\Pi}_N x\right\|_p \leq \left\|\tilde{\Pi}_N x - \tilde{\Pi}_N x^{(j)}\right\|_p + \left\|\Pi x^{(j)}\right\|_p \leq \left\|x - x^{(j)}\right\|_p + \left\|\tilde{\Pi}_N x^{(j)}\right\|_p \leq \varepsilon.$$

Montrons maintenant la réciproque. Soit $\varepsilon > 0$ et soit $N > 0$ donné par la condition $iii)$. Soit $E_N = \Pi_N \ell^p$ et remarquons que E_N est un sous-espace de dimension finie de ℓ^p. Considérons alors π_N comme une application $\Pi_N : X \to E_N$.

D'après $ii)$ pour chaque $n = 1, \ldots, N-1$ il existe M_k tel que, pour tout $(x_k)_k \in X$, $|x_n| \leq M_n$ donc

$$\|\pi_N x\|_p = \left(\sum_{n=0}^{N-1} |x_n|^p\right)^{1/p} \leq M_N := \left(\sum_{n=0}^{N-1} M_n^p\right)^{1/p}.$$

Notons que M_N ne dépend pas de $x \in X$ donc $\pi_N(X)$ est borné dans E_N. Comme E_N est de dimension finie, $\pi_N(X)$ est donc totalement borné.

Par ailleurs, la propriété $iii)$ se lit: pour tout $x \in X$ $\|x - \Pi_N x\| = \left\|\tilde{\Pi}_N x\right\| \leq \varepsilon$. Alors si $\|\pi_N(x) - \pi_N(y)\|_p \leq \varepsilon$,

$$\|x - y\|_p \leq \|x - \Pi_N x\| + \|\Pi_N(x) - \Pi_N(y)\|_p + \|\Pi_N(y) - y\|_p \leq 3\varepsilon.$$

On peut donc appliquer le lemme 1.23 et en déduire que X est compact. □

Le cas des espaces $L^p(\mathbb{R}^d)$ est un peu plus complexe:

Théorème 1.26 (Kolmogorov-Riesz). *Soit $1 \leq p < +\infty$ et $X \subset L^p(\mathbb{R}^d)$ alors X est compact si et seulement si*

(i) *X est fermé;*
(ii) *X est borné: il existe $M > 0$ tel que pour tout $f \in X$, $\|f\|_p \leq M$;*
(iii) *X est équi-intégrable: pour tout $\varepsilon > 0$ il existe $R > 0$ tel que, pour tout $f \in X$, $\int_{\|x\|_\infty \geq R} |f(x)|^p \,\mathrm{d}x \leq \varepsilon^p$.*
(iv) *Les translations sont équi-continues sur X: pour tout $\varepsilon > 0$ il existe $\rho > 0$ tel que, si $\|y\|_\infty < \rho$ alors pour tout $f \in X$, $\|f - \tau_y f\|_p \leq \varepsilon$.*

Exemple 1.27. Il est facile de voir que la boule unité de $L^p(\mathbb{R})$ n'est pas équi-intégrable. En effet, pour tout $a > 0$, $f_a = \mathbf{1}_{[a,a+1]}$ appartient à la boule unité, mais si $a > R$,

$$\int_{|x|>R} |f_a(x)|^p \,\mathrm{d}x = \int_a^{a+1} 1 \,\mathrm{d}x = 1.$$

La boule unité de L^p n'est pas non plus équi-continue. En effet, si on regarde $f_a(x) = \mathbf{1}_{[0,1]}(x) e^{iax}$, alors f_a appartient bien à la boule unité de L^p mais pour $h \leq 1$,

$$|f_a - \tau_h f_a| = \mathbf{1}_{[0,h]} + |e^{iax} - e^{ia(x-h)}| \mathbf{1}_{[h,1]} + \mathbf{1}_{[1,1+h]} = \mathbf{1}_{[0,h]} + |1 - e^{iah}| \mathbf{1}_{[h,1]} + \mathbf{1}_{[1,1+h]}$$

donc

$$\|f_a - \tau_h f_a\|_p^p = 2h + (1-h)|1 - e^{iah}|^p.$$

Il en résulte que $\left\|f_{1/h}-\tau_h f_{1/h}\right\|_p^p \to |1-e^i|^p \neq 0$ quand $h\to 0$.

1.3. Une application: le théorème de convergence monotone de Dini.

THÉORÈME 1.28 (Dini). *Soit (X,d) un espace métrique compact. Soit $(f_n)_{n\geq 0}\in\mathcal{C}(X,\mathbb{R})$ et $f\in\mathcal{C}(X,\mathbb{R})$ des fonctions continues à valeurs rélles. On suppose que, pour $x\in X$, $f_n(x)$ est croissante (resp. décroissante) et que $f_n(x)\to f(x)$. Alors $f_n\to f$ uniformément sur X,*

$$\sup_{x\in X}\big(f(x)-f_n(x)\big)=\sup_{x\in X}|f(x)-f_n(x)|\to 0.$$

REMARQUE 1.29. Les hypothèses sont essentielles. Considérons $f_n(x)=1-x^n$. Alors sur $(0,1)$ (qui n'est pas compact), $f_n(x)\to 1$ mais la convergence n'est pas uniforme puisque $sup_{x\in(0,1)}|f_n(x)-1|=sup_{x\in(0,1)}x^n=1$.

Par contre, sur $[0,1]$, qui est compact, $f_n(x)\to f(x)=\begin{cases}1 & \text{si } 0\leq x<1\\ 0 & \text{si } x=1\end{cases}$ qui n'est pas continu. Comme f_n est continue, et que sa limite ne l'est pas, la convergence ne peut être uniforme. Cela peut évidemment se voir directement puisque $sup_{x\in(0,1)}|f_n(x)-f(x)|=sup_{x\in(0,1)}x^n=1$.

DÉMONSTRATION. En remplaçant f_n par $-f_n$ et f par $-f$, on voit qu'il suffit de considérer le cas où, pour x fixé, $f_n(x)$ est croissante.

On pose $g_n(x)=f(x)-f_n(x)$ de sorte que pour x fixé, $g_n(x)$ est continue, positive, décroissante et converge (simplement) vers 0. On pose $M_n=\{g_n\}_\infty$ et nous voulons montrer que $M_n\to 0$.

On fixe $\varepsilon>0$ et on pose $O_n=g_n^{-1}\big((-\infty,\varepsilon)\big)=\{x\in X\ :g_n(x)<\varepsilon\}$ qui est donc un ouvert de X. Si $x\in X$, comme $g_n(x)\to 0$, il existe n tel que $0\leq g_n(x)<\varepsilon$ donc $x\in O_n$. Ainsi $\displaystyle\bigcup_{n\in\mathbb{N}}O_n$ est un recouvrement de X par des ouverts. Comme X est compact, il existe une partie finie $F\subset\mathbb{N}$ telle que $\displaystyle\bigcup_{n\in F}O_n$ est encore un recouvrement de X. On pose alors $N=\max\{n\in F\}$ et on remarque que la décroissance de $g_n(x)$ implique que, pour tout $n\in F$, $n\leq N$ donc $O_n\subset O_N$. Ainsi $X\subset\displaystyle\bigcup_{n\in F}O_n\subset O_N\subset X$. Enfin si $m\geq N$, $O_N\subset O_m$ donc $X=O_m$. En d'autres termes, pour tout $x\in X$, $0\leq g_m(x)<\varepsilon$, donc $M_m\leq\varepsilon$. On a bien $M_n\to 0$. □

1.4. Compacité faible. Rappelons que dans un espace de Banach X, la convergence faible est définie comme suit: on dit que $(x_n)_{n\geq 0}$ converge faiblement vers x si pour toute forme linéaire continue ℓ sur X, $\ell(x_n)\to\ell(x)$. Dans le cas d'un espace de Hilbert H, ceci équivaut à ce que, pour tout $y\in H$ $\langle x_n,y\rangle\to\langle x,y\rangle$. Nous avons aussi vu l'exemple suivant:

EXEMPLE 1.30. Soit $(e_n)_{n\geq 0}$ une base orthonormée d'un espace de Hilbert (séparable) alors
– on ne peut pas extraire de $(e_n)_{n\geq 0}$ de sous suite convergente dans H
– $(e_n)_{n\geq 0}$ converge faiblement vers 0.
En effet, tout $x\in H$ s'écrit

$$x=\sum_{n\geq 0}\langle x,e_n\rangle e_n\quad\text{avec}\quad\|x\|^2=\sum_{n\geq 0}|\langle x,e_n\rangle|^2.$$

En particulier, cette dernière série étant convergente, son terme général tend vers 0. Ainsi, pour tout $x\in H$, $\langle e_n,x\rangle=\overline{\langle x,e_n\rangle}\to 0=\langle 0,x\rangle$.

Ce cas particulier est assez typique:

THÉORÈME 1.31 (Banach-Alaoglu). *Soit H un espace de hilbert séparable. Alors la boule unité de H est faiblement séquentiellement compact. Plus précisément, de toute suite bornée, on peut extraire une sous-suite faiblement convergente.*

DÉMONSTRATION. Comme H est séparable, H possède une base orthonormée $(e_n)_{n\geq 0}$. Soit donc $(x_k)_{k\geq 0}$ une suite bornée par M. Avec Cauchy-Schwarz, pour tout $n\geq 0$,

$$|\langle x_k,e_n\rangle|\leq\|x_k\|\|e_n\|\leq M$$

i.e. $(\langle x_k, e_n\rangle)_{k\geq 0}$ est une suite (scalaire) bornée. En particulier, $\langle x_k, e_0\rangle$ étant bornée, elle admet une sous-suite convergente $\left\langle x_{k^0_\ell}, e_0\right\rangle \to a_0$ quand $\ell \to +\infty$.

Comme $\left\langle x_{k^0_\ell}, e_1\right\rangle$ est également bornée, elle admet elle aussi une sous-suite convergente $\left\langle x_{k^1_\ell}, e_1\right\rangle \to a_1$ quand $\ell \to +\infty$. Bien évidemment, on a encore $\left\langle x_{k^1_\ell}, e_0\right\rangle \to a_0$ puisque c'est une suite extraite de $\left\langle x_{k^0_\ell}, e_0\right\rangle$.

Supposons maintenant qu'on ait construit des suites d'entiers $\{k^1_\ell, \ell \in \mathbb{N}\} \subset \cdots \subset \{k^\ell_l, \ell \in \mathbb{N}\} \subset \{k^0_\ell, \ell \in \mathbb{N}\} \subset \mathbb{N}$ telle que, pour $j = 0, \ldots, n$, $a_j = \lim\limits_{\ell\to+\infty} \left\langle x_{k^j_\ell}, e_j\right\rangle$ existe. Alors, comme $\left\langle x_{k^n_\ell}, e_{n+1}\right\rangle$ est bornée, on peut en extraire une sous-suite telle que $a_j = \lim\limits_{\ell\to+\infty} \left\langle x_{k^{n+1}_\ell}, e_{n+1}\right\rangle$ existe.

On considère alors $n_\ell = k^\ell_\ell$ (le ℓ-ième terme de la ℓ-ième suite). On remarque que $(\langle x_{n_\ell}, e_j\rangle)_{l\geq j}$ est une suite extraite de $(\left\langle x_{k^j_\ell}, e_j\right\rangle)_\ell$ donc converge: $\langle x_{n_\ell}, e_j\rangle \to a_j$ quand $\ell \to +\infty$.

Remarque 1.32. Ceci s'appelle le procédé diagonal de Cantor.

On fixe maintenant J, on a

$$\sum_{j=0}^{J} |a_j|^2 = \lim_{\ell\to+\infty} \underbrace{\sum_{j=0}^{J} |\langle x_{n_\ell}, e_j\rangle|^2}_{\leq \|x_{n_\ell}\|^2} \leq M^2$$

en particulier, $(a_j)_{j\geq 0} \in \ell^2$ avec $\|(a_j)\|_2 \leq M$. On peut donc définir $x = \sum_{j=0}^{+\infty} a_j e_j \in H$ et on a $a_j = \langle x, e_j\rangle$ et $\|x\|^2 = \sum_{j=0}^{+\infty} |a_j|^2 \leq M^2$.

On va montrer que x est la limite faible de $(x_{n_\ell})_\ell$. Soit maintenant $y \in H$, on veut donc montrer que $\langle x_{n_\ell}, y\rangle \to \langle x, y\rangle$. On sait déjà que c'est le cas si $y = e_j$ donc encore si $y \in Vect\{e_j\}_{j\geq 0}$ (les combinaisons linéaires finies des e_j). On veut montrer que c'est vrai pour tout $y \in H = \overline{Vect\{e_j\}_{j\geq 0}}$. On fixe donc $y \in H$ et $\varepsilon > 0$. Comme $H = \overline{Vect\{e_j\}_{j\geq 0}}$ il existe $y_\varepsilon \in Vect\{e_j\}_{j\geq 0}$ tel que $\|y - y_\varepsilon\|_H < \varepsilon$. On vient de voir que $\langle x_{n_\ell}, y_\varepsilon\rangle \to \langle x, y_\varepsilon\rangle$, il existe donc $L \geq 0$ tel que, si $\ell \geq L$, $|\langle x_{n_\ell}, y_\varepsilon\rangle - \langle x, y_\varepsilon\rangle| < \varepsilon$. Enfin

$$\begin{aligned}
&|\langle x_{n_\ell}, y\rangle - \langle x, y\rangle| \\
&\qquad \leq |\langle x_{n_\ell}, y\rangle - \langle x_{n_\ell}, y_\varepsilon\rangle| + |\langle x_{n_\ell}, y_\varepsilon\rangle - \langle x, y_\varepsilon\rangle| + |\langle x, y_\varepsilon\rangle - \langle x, y\rangle| \\
&\qquad = |\langle x_{n_\ell}, y - y_\varepsilon\rangle| + |\langle x_{n_\ell}, y_\varepsilon\rangle - \langle x, y_\varepsilon\rangle| + |\langle x, y_\varepsilon - y\rangle| \\
&\qquad \leq \|x_{n_k}\|\|y - y_\varepsilon\| + |\langle x_{n_\ell}, y_\varepsilon\rangle - \langle x, y_\varepsilon\rangle| + \|x\|\|y - y_\varepsilon\| < (2M+1)\varepsilon
\end{aligned}$$

comme annoncé. □

Remarque 1.33. Le théorème s'étend de façon presque directe à deux cas plus généraux

– L'espace de Hilbert H n'a pas besoin d'être séparable. On considère d'abord $H_0 = \overline{Vect(x_k)}$ le sous-espace fermé engendré par la suite. Ainsi H_0 est un espace de Hilbert séparable. On extrait alors de (x_k) une sous-suite x_{n_k} qui converge faiblement dans H_0: il existe $x \in H_0$ tel que, pour tout $y_0 \in H_0$, $\langle x_{n_k}, y_0\rangle \to \langle x, y_0\rangle$. Ensuite, si $y \in H$, on écrit $y = y_0 + z$ avec $y_0 \in H_0$ et $z \in H_0^\perp$ un vecteur orthogonal à H_0. Alors

$$\langle x_{n_k}, y\rangle = \langle x_{n_k}, y_0\rangle + \langle x_{n_k}, z\rangle = \langle x_{n_k}, y_0\rangle \to \langle x, y_0\rangle = \langle x, y\rangle.$$

– H peut être remplacé par n'importe quel espace de Banach X dont le dual est séparable. La base orthonormée $(e_k)_{k\geq 0}$ est alors remplacée par une suite $(y_k)_{k\geq 0}$ qui engendre X'. Ainsi, le théorème reste vrai dans $L^p(\mu)$, $1 < p < +\infty$ pour une mesure μ raisonnable (par exemple la mesure de Lebesgue sur un ouvert de $\mathbb{R}^d$).

Remarque 1.34. Si X' est séparable, la convergence faible dans X est une convergence métrique. Plus précisément, on prend $\{\ell_n\}$ une famille dénombrable de formes linéaires continues sur X dense dans

X'. On munit X de la distance

$$d(x,y) = \sum_{n\geq 0} 2^{-n} \frac{|\ell_n(x-y)|}{1+|\ell_n(x-y)|}.$$

On peut alors montrer que $x_n \to y$ faiblement si et seulement si $d(x_n, y) \to 0$.

2. Au tour du thèorème de Baire

Rappelons qu'une partie $\mathcal{D}$ d'un espace métrique X est dense si, pour tout ouvert O de X, $O \cap \mathcal{D}$ est non vide. En particulier, si on se contente de boules ouvertes, pour tout $\varepsilon > 0$ et tout $x \in X$, il existe $y \in \mathcal{D}$ tel que $d(x,y) < \varepsilon$. Notons que $\mathcal{D}$ est dense si et seulement si $\mathcal{A} = X \setminus \mathcal{D}$ est d'intérieur vide *i.e.* pour tout ouvert O, O n'est pas inclus dans $\mathcal{A}$, c'est-à-dire $O \cap (X \setminus \mathcal{A}) \neq \emptyset$ puisque $X \setminus \mathcal{A} = \mathcal{D}$.

THÉORÈME 2.1 (Théorème de Baire). *Soit (X,d) un espace métrique complet non vide.*

(1) *Toute intersection dénombrable d'ouverts denses est encore dense.*
(2) *Toute réunion dénombrable de fermés d'intérieur vide est d'intérieur vide.*
(3) *Si $X = \bigcup_{n\in\mathbb{N}} F_n$, F_n fermé, alors il existe n_0 tel que F_{n_0} n'est pas d'intérieur vide.*

REMARQUE 2.2. Rappelons qu'une intersection dénombrable d'ouverts n'est pas nécessairement ouverte (une telle intersection est appelée un G_δ en topologie). De même, une réunion dénombrable de fermés n'est pas nécessairement fermée (une telle réunion est appelée un F_σ en topologie). Par exemple

$$\bigcap_{n\geq 0}]-1/n, 1/n[= \{0\} \quad \text{et} \quad \bigcup_{n\geq 0} [-1+1/n, 1-1/n] =]-1,1[.$$

DÉMONSTRATION. Si les O_i sont des ouverts alors les $F_i = X \setminus O_i$ sont des fermés. De plus O_i est dense si et seulement si F_i est d'intérieur vide. Enfin $X \setminus \bigcap_i O_i = \bigcup_i F_i$. Ainsi les propositions 1) et 2) sont équivalentes. Enfin 3) est un cas particulier de 2) puisque X n'est pas d'intérieur vide, au moins l'un des F_i n'est pas d'intérieur vide sinon 2) serait contredit. Il nous reste donc à montrer 1).

Soient $(O_n)_{n\in\mathbb{N}}$ une famille d'ouverts denses et soit U un ouvert.

Comme O_0 est dense, il existe $x_0 \in O_0 \cap U$ et comme $O_0 \cap U$ est ouvert, il existe $r_0 > 0$ tel que $B(x_0, r_0) \subset O_0 \cap U$. Ensuite, comme O_1 est dense, il existe $x_1 \in O_1 \cap B(x_0, r_0/2)$ qui est ouvert. Il existe donc $r_1 \leq r_0/2$ tel que $B(x_1, r_1) \subset O_1 \cap B(x_0, r_0/2) \subset O_1 \cap O_2 \cap U$.

On suppose qu'on a construit $r_0, \ldots, r_n$ avec $r_k \leq r_{k-1}/2 \leq r_0 2^{-k}$ et $x_0, \ldots, x_n$ tel que $B(x_n, r_n) \subset O_1 \cap \cdots \cap O_n \cap U$. Comme O_{n+1} est dense, il existe $x_{n+1} \in O_{n+1} \cap B(x_n, r_n/2)$ qui est ouvert. Il existe ensuite $r_{n+1} \leq r_n/2 \leq r_0 2^{-(n+1)}$ tel que $B(x_{n+1}, r_{n+1}) \subset O_1 \cap \cdots \cap O_n \cap O_{n+1} \cap B(x_n, r_n/2) \cap U$.

Comme les boules sont emboitées, si $m \geq n$, $x_m \in B(x_n, r_n)$ donc

$$d(x_m, x_n) < r_n \leq r_0 2^{-n}.$$

Ainsi (x_n) est de Cauchy et donc convergente. On note x sa limite. Mais, si $m \geq n \geq 1$, $x_m \in B(x_n, r_n)$ donc en passant à la limite

$$x \in \overline{B(x_n, r_n)} \subset \overline{B(x_{n-1}, r_{n-1}/2)} \subset B(x_{n-1}, r_{n-1}) \subset O_{n-1} \cap U.$$

Comme $n \geq 1$ est arbitraire, on a bien $x \in U \cap \bigcap_{k\geq 0} O_k$ et $\bigcap_{k\geq 0} O_k$ est bien dense. □

EXEMPLE 2.3. Comme $\mathbb{R} = \bigcup_{x\in\mathbb{R}} \{x\}$, et qu'un singleton est d'intérieur vide, $\mathbb{R}$ n'est pas dénombrable. De même, $\mathbb{R}^2$ n'est pas réunion dénombrable de droites, cercles...

On considére l'espace métrique $X = (\mathbb{Q}, d)$ où d est la distance de $\mathbb{R}$, $d(x,y) = |x-y|$. Ainsi $O \subset \mathbb{Q}$ est ouvert s'il existe $U \subset \mathbb{R}$ ouvert tel que $O = U \cap \mathbb{Q}$. Notons qu'un singleton est encore d'intérieur vide pour cette distance mais que $\mathbb{Q}$ ne l'est évidemment pas. Comme $\mathbb{Q} = \bigcup_{x\in\mathbb{Q}} \{x\}$ il en résulte que $\mathbb{Q}$ n'est pas complet (il y a évidemment des démonstrations plus directe, par exemple en approchant $\sqrt{2}$ par des rationnels et en montrant que $\sqrt{2}$ n'est pas rationnel).

COROLLAIRE 2.4. *Soit X un espace métrique complet, et $(F_n)_{n\in\mathbb{N}}$ des espaces fermés. Si $X = \bigcup_{n\in\mathbb{N}} F_n$ alors la réunion des intérieurs des F_n, $\bigcup_{n\in\mathbb{N}} \mathring{F}_n$ est un ouvert dense.*

Démonstration. Soit $\Omega = \bigcup_{n\in\mathbb{N}} \mathring{F}_n$ alors Ω est un ouvert. Avec la propriété 3) du théorème de Baire, l'un au moins des F_n est d'intérieur non vide donc Ω n'est pas vide.

On note alors $E_n = F_n \cap (X \setminus \Omega)$ de sorte que E_n soit fermé, d'intérieur vide puisque $\mathring{E}_n \subset \mathring{F}_n \cap (X \setminus \Omega) \subset \Omega \cap (X \setminus \Omega) = \emptyset$. Il résulte du théorème de Baire que $\bigcup E_n$ est d'intérieur vide. Mais $\bigcup E_n = \bigcup F_n \cap (X \setminus \Omega) = X \cap (X \setminus \Omega) = X \setminus \Omega$ et on a déjà vu que cela signifie que Ω est dense. $\square$

Donnons un exemple type de fermé d'intérieur vide:

Lemme 2.5. *Soit $(E, \|\cdot\|)$ un espace vectoriel normé et $F \subset E$ un sous espace vectoriel.*
(1) *Si F est de dimension finie, alors F est fermé.*
(2) *Si $F \neq E$, alors F est d'intérieur vide.*

Démonstration. Le premier point a déjà été utilisé plusieurs fois et devrait être acquis à l'issue d'une L3. Rappelons la démonstration. Supposons que F soit de dimension finie et soit $e_1, \ldots, e_d$ une base de F. Tout $u \in F$ s'écrit de façon unique $u = \sum_{j=1}^d x^j e_j$. On peut alors définir deux normes sur F, la norme de E, $\|u\|$ et la norme $\|u\|_\infty = \max |x^j|$. Comme F est de dimension finie, ces deux normes sont équivalentes: pour tout $v \in F$, $A\|v\|_\infty \leq \|v\| \leq B\|v\|_\infty$.

Soit $u_n = \sum_{j=1}^d x_n^j e_j$ une suite d'éléments de F qui converge vers $u \in E$. Comme $\|u_n - u_m\| \geq A\|u_n - u_m\|_\infty$, la suite des coordonnées $X_n = (x_n^1, \ldots, x_n^d) \in \mathbb{C}^d$ est de Cauchy dans $\mathbb{C}^d$ donc converge. On note $X = (x^1, \ldots, x^d) \in \mathbb{C}^d$ sa limite et $\tilde{u} = \sum_{j=1}^d x^j e_j \in F$. Comme $\|u_n - \tilde{u}\| \leq B\|u_n - \tilde{u}\|_\infty = B\|X_n - X\|_\infty$, on en déduit que $u_n \to \tilde{u}$. Par unicité de la limite $u = \tilde{u} \in F$ et F est bien fermé.

Soit maintenant F un sous-espace de E d'intérieur non-vide. Il existe alors $x_0 \in F$ et $r > 0$ tel que $B(x_0, r) \subset F$. En remarquant que $B(-x_0, r) = \{-x, x \in B(x_0, r)\}$ et en utilisant le fait que F soit un espace vectoriel, on en déduit que $B(-x_0, r) \subset F$. Enfin, si $\|y\| \leq r$ alors $y = \frac{1}{2}\big((x_0 + y) + (-x_0 + y)\big)$ et $x_0 + y \in B(x_0, r) \subset F$ alors que $-x_0 + y \in B(-x_0, r) \subset F$ donc $y \in F$ *i.e.* $B(0, r) \subset F$. Soit maintenant $x \in E$, si $x = 0$ on a évidemment $x \in F$ alors que si $x \neq 0$ on écrit $x = \frac{2\|x\|}{r} y$ et on remarque que

$$\|y\| = \left\| \frac{r}{2\|x\|} x \right\| = \frac{r}{2\|x\|} \|x\| = \frac{r}{2} < r$$

donc $y \in B(0, r) \subset F$ et alors $x = \frac{2\|x\|}{r} y \in F$ puisque F est un sous-espace vectoriel. $\square$

Corollaire 2.6. *Soit $(E, \|\cdot\|)$ un espace vectoriel normé avec une base (algébrique) dénombrable alors E n'est pas complet.*

Remarque 2.7. Rappelons qu'une base algébrique d'un espace vectoriel E est une famille $(e_i)_{i\in I}$ tel que, pour tout $x \in E$, il existe une partie *finie* $F \subset I$ et des scalaires $(x_f)_{f\in F}$ tel que $x = \sum_{f\in F} x_f e_f$ *i.e.* tout élément de E est combinaison linéaire finie d'éléments de la famille $(e_i)_{i\in I}$. C'est une conséquence de l'axiome du choix que tout espace vectoriel admet une base algébrique. Celle-ci n'est en général d'aucune utilité en analyse car il n'y a aucun contrôle des coefficients $(x_f)_{f\in F}$ en fonction de x (les applications $x \to x_f$ ne sont pas nécessairement continues).

Il existe beaucoup de notions de bases en analyse, parmi lesquelles la notion de base orthonormée. Une base orthonormée (en dimension infinie) n'est pas une base algébrique dénombrable et un espace de Hilbert est (par définition) complet.

Exemple 2.8. Soit $\mathbb{R}[X]$ l'espace des polynômes à coefficients réels. Cet espace a une base algébrique dénombrable, par exemple, sa base canonique $\{X^j\}_{j\in\mathbb{N}}$. On peut munir $\mathbb{R}[X]$ de nombreuses normes, par exemple $\|P\|_\infty = \sup_{t\in[0,1]} |P(t)|$ ou $\|P\|_1 = \int_0^1 |P(t)|\,\mathrm{d}t$. Alors $(\mathbb{R}[X], \|P\|_\infty)$ et $(\mathbb{R}[X], \|P\|_1)$ ne sont pas complets. Bien évidemment, cela peut se voir directement en approchant uniformément sur $[0, 1]$ une fonction continue qui n'est pas un polynôme par des polynômes (voir plus loin).

Démonstration. Soit $(e_n)_{n\in\mathbb{N}}$ une base algébrique dénombrable de E et notons $E_n = Vect(e_0, \ldots, e_n)$. Alors $E = \bigcup_{n\in\mathbb{N}} E_n$. Mais chaque E_n est fermé (de dimension fini) et d'intérieur vide (car $E_n \neq E$). Ainsi, les conclusions du théorème de Baire ne sont pas satisfaites, donc ses hypothèses non-plus, en l'occurence, E n'est pas complet. □

THÉORÈME 2.9 (Banach-Steinhaus). *Soient E et F des espaces de Banach et $T_n : E \to F$ linéaires bornées telles que, pout tout $x \in E$, $(T_n x)$ est une suite bornée. Alors $(\|T_n\|)_n$ est une suite bornée.*

Avant de démontrer ce théorème, expliquons comment celui-ci peut être utilisé

- Supposons que $(\|T_n\|)_n$ ne soit pas bornée, alors il existe $x \in E$ tel que $(T_n x)$ n'est pas borné, en particulier, $T_n x$ diverge. Une version un peu plus fine du théorème montre qu'il existe même un G_δ-dense U de X tel que, pour tout $x \in U$, $T_n x$ diverge.

  Nous verrons comment utiliser cela pour montrer qu'il existe des fonctions intégrables dont la série de Fourier diverge. Ce théorème a aussi une utilisation positive:
- Soit $\mathcal{D}$ un sous-espace dense de E tel que pour tout $x \in \mathcal{D}$, il existe $Tx \in F$ tel que $T_n x \to Tx$. Notons que cela implique la linéarité de $T : \mathcal{D} \to F$.

  Supposons que, pour tout $x \in E$, $(T_n x)$ est une suite bornée, *i.e.* qu'il existe M_x tel que, pour tout n, $\|T_n x\| \le M_x$. Alors il existe M tel que, pour tout $x \in E$, et pour tout n, $\|T_n x\| \le M\|x\|$. Mais alors, en passant à la limite, pour tout $x \in \mathcal{D}$, $\|Tx\| \le M\|x\|$. Par suite, T se prolonge en une application linéaire continue $E \to F$ (voir cours du premier semestre).

  Dans certains cas, on peut voir que pour tout $x \in \mathcal{D}$, $T_n x \to x$ *i.e.* $T = I$ l'identité. Son prolongement est donnu: c'est l'identité. Ainsi si, pour tout $x \in E$, $(T_n x)$ est borné, alors pour tout $x \in E$, $T_n x \to x$.

Démontrons maintenant ce théorème central de l'analyse.

DÉMONSTRATION. On suppose que pour tout $x \in E$, il existe $M_x > 0$ tel que, pour tout n, $\|T_n x\| \le M_x$.

On note

$$\begin{aligned}\Omega_k &= \{x \in E \ : \ \text{il existe } n \text{ tel que } \|T_n x\|_F > k\} \\ &= \bigcup_{n \in \mathbb{N}} \{x \in E \ : \ \|T_n x\|_F > k\}.\end{aligned}$$

Ainsi Ω_k est ouvert. Par ailleurs

$$\bigcap_{k \in \mathbb{N}} \Omega_k = \{x \in E \ : \ \text{pour tout } k, \text{ il existe } n \text{ tel que } \|T_n x\|_F > k\} = \emptyset$$

puisque $\|T_n x\|$ est borné. Comme E est complet, le théorème de Baire implique que l'un des Ω_k, disons Ω_K, n'est pas dense. Il existe donc $x_0 \in X$ et $r_0 > 0$ tels que $B(x_0, r_0) \cap \Omega_K = \emptyset$.

Écrivons $x \in B(x_0, r_0)$ sous la forme $x = x_0 + r_0 y$ avec $y \in B(0,1)$. Comme $x \notin \Omega_K$, on a, pour tout n, $\|T_n x\| \le K$. Mais

$$\|T_n x\| = \|rT_n y + T_n x_0\| \ge r\|T_n y\| - \|T_n x_0\| \ge r\|T_n y\| - M_{x_0}.$$

Par suite, pour tout $y \in B(0,1)$, $\|T_n y\| \le \dfrac{K + M_{x_0}}{r}$ *i.e.* $\|T_n\| \le \dfrac{K + M_{x_0}}{r}$. □

THÉORÈME 2.10 (Application ouverte). *Soient E et F deux espaces de Banach et $T : E \to F$ linéaire, borné et injective, alors T est ouverte, c'est-à-dire que, pour tout ouvert U, $T(U)$ est ouvert.*

Ainsi, si $T : E \to E$ est linéaire bornée bijective *alors l'application inverse T^{-1} est automatiquement continue (on dit que T est inversible).*

REMARQUE 2.11. Rappelons qu'en dimension infinie, un application $T : E \to E$ injective n'est pas forcément bijective (en dimension finie, cette propriété provient du théorème du rang qui n'a pas de sens en dimension infinie). Par ailleurs, une application linéaire n'est pas forcément continue.

Ce théorème dit que si une application linéaire continue T est bijective, alors son inverse T^{-1} est continue. Rappelons que T^{-1} est définie en algèbre linéaire: c'est l'application qui à $y \in F$ associe l'unique $x \in E$ tel que $y = Tx$, application nécessairement linéaire.

La continuité de T^{-1} est évidemment la même chose que le fait que T soit ouverte: pour tout ouvert U, $(T^{-1})^{-1}(U) = T(U)$ est alors ouvert.

Remarque 2.12. Dire que T est ouvert équivaut à ce qu'il existe $R, r > 0$ tel que $B(0,r) \subset T\big(B(0,R)\big)$.

En effet, si T est ouvert, $T\big(B(0,R)\big)$ est un ouvert contenant $T(0) = 0$ donc une boule ouverte centrée en 0. Réciproquement, supposons qu'il existe $R_0, r_0 > 0$ tel que $B(0,r_0) \subset T\big(B(0,R_0)\big)$. Alors pour $R > 0$

$$T\big(B(0,R)\big) = T\big(\frac{R}{R_0}B(0,R_0)\big) = \frac{R}{R_0}T\big(B(0,R_0)\big) \supset \frac{R}{R_0}B(0,r_0) = B(0,\frac{r_0}{R_0}R).$$

Prenons un ouvert U, et $b \in T(U)$, on cherche une boule ouverte centrée en b dans $T(U)$. Écrivons $b = Ta$ avec $a \in U$ et soit $\rho > 0$ tel que $B(a,\rho) \in U$. Mais $B(a,\rho) = a + B(0,\rho)$ donc

$$\begin{aligned} T(U) \quad &\supset \quad T\big(B(a,\rho)\big) = T\big(a + B(0,\rho)\big) = T(a) + T\big(B(0,\rho)\big) \\ &\supset \quad b + B(0,\frac{r_0}{R_0}\rho) = B(b,\frac{r_0}{R_0}\rho). \end{aligned}$$

Démonstration du théorème. Comme T est surjective,

$$F = \bigcup_{n\in\mathbb{N}} T\big(B(0,n)\big) = \bigcup_{n\in\mathbb{N}} \overline{T\big(B(0,n)\big)}$$

(puisque cette dernière réunion est plus grande que la première et aussi incluse dans F). Comme F est un espace de Banach, le théorème de Baire donne l'existence d'un N tel que $\overline{T\big(B(0,N)\big)}$ est d'intérieur non vide.

Il existe donc une boule $B(y,r) \subset \overline{T\big(B(0,N)\big)}$. Mais alors

$$B(-y,r) = -B(y,r) \subset -\overline{T\big(B(0,N)\big)} = \overline{T\big(-B(0,N)\big)} = \overline{T\big(B(0,N)\big)}.$$

Ainsi,

$$\begin{aligned} B(0,r) \subset B(y,r) + B(-y,r) \subset \overline{T\big(B(0,N)\big)} + \overline{T\big(B(0,N)\big)} \\ \subset \overline{T\big(B(0,N)\big) + T\big(B(0,N)\big)} = \overline{T\big(B(0,2N)\big)}. \end{aligned}$$

Enfin, si $s > 0$, en écrivant $s = \frac{s}{r}r$ et en multipliant par s/r, on trouve que

$$\text{pour tout } s > 0, \quad B(0,s) \subset \overline{T\big(B(0,cs)\big)} \tag{2.2}$$

avec $c = \frac{2N}{r}$.

Soit $z \in B(0,1) \subset \overline{T\big(B(0,c)\big)}$. Il existe donc $x_1 \in B(0,c)$ tel que $\|z - Tx_1\| < 1/2$ c'est-à-dire $z - Tx_1 \in B(0,1/2)$. On applique alors (2.2) avec $s = 1/2$: $z - Tx_1 \in \overline{T\big(B(0,2^{-1}c)\big)}$. Ainsi, il existe $x_2 \in B(0,2^{-1}c)$ tel que $\|z - Tx_1 - Tx_2\| < 2^{-2}$. On construit ainsi par récurrence une suite (x_n) tel que $x_n \in B(0,2^{-n+1}c)$ et $\|z - T(x_1 + \cdots + x_n)\| < 2^{-n}$. En effet, une fois que x_n est construit, $z - T(x_1 + \cdots x_n) \in B(0,2^{-n}) \subset \overline{T\big(B(0,c2^{-n})\big)}$, il existe donc $x_{n+1} \in B(0,c2^{-n})$ tel que $\|z - T(x_1 + \cdots x_n) - Tx_{n+1}\| < 2^{-n-1}$.

Mais maintenant, comme $\|x_n\| \le 2^{-n+1}c$, la série $\sum x_n$ est normalemement convergente, donc convergente (dans le Banach E). On note $x = \displaystyle\sum_{n=1}^{\infty} x_n$ alors $\|x\| \le c\sum_{n=1}^{\infty} 2^{-n+1} = 2c$. Comme T est continue, $T(x_1 + \cdots + x_n) \to Tx$.

Par ailleurs, $\|z - T(x_1 + \cdots + x_n)\| \le 2^{-n} \to 0$ donc $T(x_1 + \cdots + x_n) \to z$. On a donc montré que, pour tout $z \in B(0,1)$, il existe $x \in B(0,2c)$ tel que $z = Tx$. On a déjà vu que cela signifie que T est ouverte. □

Remarque 2.13. Si on considére $E = c_{0,0}$ l'espace des suites à support fini muni de la norme $\|\cdot\|_\infty$ et $T(x_n)_n = (n^{-1}x_n)_n$ alors T est continue bijective avec $T^{-1}(x_n)_n = (nx_n)_n$ mais T^{-1} n'est pas continue. Évidemment E n'est pas complet.

THÉORÈME 2.14 (Théorème du graphe fermé). *Soient E, F deux espaces de Banach et $T : E \to F$ une application linéaire. Alors T est continue si et seulement si son graphe $Gr(T) = \{(x, Tx) \in E \times F : x \in E\}$ est fermé.*

Notons que si X, Y sont des espaces métriques et $f : X \to Y$ est continue alors son graphe est fermé. En effet, si $(x_n, y_n) \in Gr(f)$ converge vers (x, y) alors $x_n \to x$ et $y_n = f(x_n) \to y$. Comme f est continue, $f(x_n) \to f(x)$ donc $y = f(x)$, *i.e.* $(x, y) \in Gr(f)$.

La linéarité est essentielle pour la réciproque. En effet, définissons $f : \mathbb{R} \to \mathbb{R}$ par $f(x) = 1/x$ si $x \neq 0$ et $f(0) = 0$. Alors $Gr(f) = \{(0,0)\} \cup \{(x, 1/x), x > 0\} \cup \{(x, 1/x), x < 0\}$ et chacune de ces parties est fermée dans $\mathbb{R}^2$ (si $(x_n, 1/x_n)$ convergent alors x_n ne tend pas vers 0). Alternativement, si on définit $F : \mathbb{R}^2 \to \mathbb{R}$ par $F(x, y) = xy$ alors $\{(x, 1/x), x > 0\} \cup \{(x, 1/x), x < 0\} = F^{-1}(\{1\})$ est fermé et $Gr(f) = F^{-1}(\{1\}) \cup \{(0, 1)\}$.

DÉMONSTRATION. Munissons $E \times F$ de la norme $\|(x, y)\| = \|x\|_E + \|y\|_F$ qui en fait un espace de Banach et supposons que $Gr(T)$ est fermé. Notons que, comme T est linéaire, $Gr(T)$ est un sous-espace vectoriel de $E \times F$. Ainsi $Gr(T)$ est lui-même un espace de Banach.

On considère maintenant les deux applications $\pi_1 : Gr(T) \to E$ et $\pi_2 : Gr(T) \to F$ données par $\pi_1(x, y) = x$ et $\pi_2(x, y) = y$. Ce sont clairement des applications linéaires continues. De plus, π_1 est une bijection $Gr(T) \to E$, son application réciproque π_1^{-1} est donc continue. Comme $T = \pi_2 \circ \pi_1^{-1}$, on en déduit que T est continue. □

Une autre démonstration est possible à partir du résultat suivant:

PROPOSITION 2.15. *Soit E un espace de Banach et soient $\|\cdot\|_1$ et $\|\cdot\|_2$ deux normes sur E. On suppose qu'il existe $C > 0$ tel que, pour tout $x \in E$ $\|x\|_1 \leq C\|x\|_2$. Si E est complet pour les deux normes alors elles sont équivalentes.*

DÉMONSTRATION. En effet, considérons $T = Id$ (l'identité $Tx = x$) vu comme application linéaire $(E, \|\cdot\|_2) \to (E, \|\cdot\|_1)$. Alors l'inégalité $\|x\|_1 \leq C\|x\|_2$ signifie que T est continue. Comme T est bijective, l'inverse $T^{-1} : (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$ est également continue, c'est-à-dire qu'il existe $C' > 0$ tel que pour tout $x \in E$, $\|T^{-1}x\|_2 \leq C'\|x\|_1$. Il suffit de remarquer que $T^{-1}x = x$. □

Donnons maintenant une deuxième démonstration du théorème du graphe fermé:

DÉMONSTRATION. On munit E de deux normes, sa norme usuelle $\|x\|_E$ et la norme $\|x\|_T = \|x\|_E + \|Tx\|_F$. Clairement $\|x\|_E \leq \|x\|_T$.

Soit maintenant (x_n) une suite de Cauchy dans E pour la norme $\|\cdot\|_T$. Mais comme $\|x_n - x_m\|_E \leq \|x_n - x_m\|_T$, (x_n) est de Cauchy dans E (muni de la norme $\|\cdot\|_E$). Elle converge donc vers un $x \in E$. De même $\|Tx_n - Tx_m\|_F = \|T(x_n - x_m)\|_F \leq \|x_n - x_m\|_T$ donc (Tx_n) est de Cauchy dans F. Ainsi (Tx_n) converge vers un $y \in F$. Mais alors (x_n, Tx_n) est une suite convergente de $Gr(T)$ qui est fermé donc sa limite $(x, y) \in Gr(T)$ c'est-à-dire $y = Tx$ et donc $\|x_n - x\|_T = \|x_n - x\|_E + \|Tx_n - Tx\|_F \to 0$ et x_n converge vers x en norme $\|\cdot\|_T$. Ainsi E est complet pour les deux normes, il existe donc C tel que, pour tout x, $\|x\|_E + \|Tx\|_F = \|x\|_T \leq C\|x\|_E$ soit $\|Tx\|_F \leq (C-1)\|x\|_E$. □

3. Densité

Le but de cette section est de montrer la densité des polynômes dans l'espaces de fonctions continues sur $[a, b]$ et dans $L^2([a, b])$. Il y a de nombreuses façons de faire cela, la méthode que nous prenons ici est celle qui utilise des principes généraux.

3.1. Densité dans $\mathcal{C}(K)$, K compact.

3.1.1. *Première méthode: par approximation de l'unité.*

THÉORÈME 3.1. *Soient $-\infty < a < b < +\infty$ deux réels, alors les polynômes sont denses dans $\mathcal{C}([a, b])$. Ainsi, pour tout fonction f continue sur $[a, b]$, il existe une suite de polynômes P_n telle que $\sup_{t \in [a,b]} |f(t) - P_n(t)| \to 0$.*

La première démonstration est due à Weierstrass‡

DÉMONSTRATION. Nous commençons par rappeler que si $\gamma(t) = e^{-\pi t^2}$ alors $\gamma_\lambda(t) = \dfrac{1}{\lambda}\gamma\left(\dfrac{t}{\lambda}\right)$ est une approximation de l'unité donc que si $f \in \mathcal{C}_0(\mathbb{R})$, alors $f * \gamma_\lambda \to f$ uniformément quand $\lambda \to 0$.

Soit alors $f \in \mathcal{C}([a,b])$. Començons par prolonger f en une fonction g de $\mathcal{C}_0(\mathbb{R})$ avec $\|g\|_\infty \leq \|f\|_\infty$ comme suit:

$$g(x) = \begin{cases} f(x) & \text{si } x \in [a,b] \\ f(a)(x-a+1) & \text{si } x \in [a-1,a] \\ f(b)(b+1-x) & \text{si } x \in [b,b+1] \\ 0 & \text{sinon} \end{cases}.$$

Ainsi $\gamma_\lambda * g \to g$ uniformément sur $\mathbb{R}$. Soit donc $\varepsilon > 0$ et λ tel que $\sup_{t\in\mathbb{R}} |\gamma_\lambda * g(t) - g(t)| \leq \varepsilon$. En particulier, pour $t \in [a,b]$, $|\gamma_\lambda * g(t) - f(t)| \leq \varepsilon$.

Nous allons maintenant approcher $\gamma_\lambda * g(t)$ par des polynômes. Pour cela, remarquons que

$$\gamma_\lambda * g(t) = \int_{a-1}^{b+1} \frac{1}{\lambda} g(s) e^{-\pi(s-t)^2/\lambda^2}\,\mathrm{d}s.$$

Mais, si $t \in [a,b]$ et $s \in [a-1,b+1]$ alors $s-t \in [-(b-a)-1,(b-a)+1]$. Comme le rayon de convergence de la série exponentielle est infinie, il existe N tel que, pour tout $u \in [-(b-a)-1,(b-a)+1]$,

$$\left| e^{-\pi u^2/\lambda^2} - \sum_{k=0}^{N} (-1)^k \frac{\pi^k}{\lambda^{2k} k!} u^{2k} \right| \leq \varepsilon.$$

Mais alors, pour $t \in [a,b]$

$$\begin{aligned} &\left| \int_{a-1}^{b+1} \frac{1}{\lambda} g(s) e^{-\pi(s-t)^2/\lambda^2}\,\mathrm{d}s - \int_{a-1}^{b+1} \frac{1}{\lambda} g(s) \sum_{k=0}^{N} (-1)^k \frac{\pi^k}{\lambda^{2k}k!}(s-t)^{2k}\,\mathrm{d}s \right| \\ &\qquad \leq \int_{a-1}^{b+1} \frac{1}{\lambda} |g(s)| \left| e^{-\pi(s-t)^2/\lambda^2} - \sum_{k=0}^{N} (-1)^k \frac{\pi^k}{\lambda^{2k}k!}(s-t)^{2k} \right| \mathrm{d}s \\ &\qquad\qquad \leq \varepsilon \int_{a-1}^{b+1} \frac{1}{\lambda} |g(s)|\,\mathrm{d}s \leq \frac{(b-a+2)\|f\|_\infty}{\lambda}\varepsilon. \end{aligned}$$

On conclue en remarquant que

$$\int_{a-1}^{b+1} \frac{1}{\lambda} g(s) \sum_{k=0}^{N} (-1)^k \frac{\pi^k}{\lambda^{2k}k!}(s-t)^{2k}\,\mathrm{d}s$$

est un polynôme de degré $2N$ en t. □

Notez que la convergence de $\gamma_\lambda * g$ vers g ne requière pas d'arguments de densité (on est dans $\mathcal{C}_0(\mathbb{R})$ ici, pas dans L^p).

DÉMONSTRATION. Une démonstration un peu plus directe est possible.§

On suppose sans perte de généralité que $a = \frac{1}{4}, b = \frac{3}{4}$ en remplaçant $f(x)$ par $f\big(\frac{1}{4} + \frac{1}{2(b-a)}(x-a)\big)$ puis on prolonge f en une fonction continue sur $\mathbb{R}$, à support $[0,1]$ et qui soit affine sur $[0,1/4]$ et sur $[3/4,1]$. On suppose aussi que $f \neq 0$.

‡Weierstrass, Über die analytische Darstellbarkeit sogenannter willkürlicher Functioneneiner reellen Veränderlichen, Verl. d. Kgl. Akad. d. Wiss. Berlin **2**(1885) 633–639.

§D. Jackson *A Proof of Weierstrass's Theorem* The American Mathematical Monthly, **41** (1934), 309-312 qui simplifie légèrement E. Landau, *Uber die Approximation einer stetigen Funktion durch eine ganze rationale Funktion*, Rendiconti del Circolo Matematico di Palermo, vol. 25 (1908), pp. 337-345

On pose alors $J_n = \int_{-1}^{1} (1-u^2)^n \, \mathrm{d}u$ et

$$P_n(x) = \frac{1}{J_n} \int_0^1 f(t)\big(1-(t-x)^2\big)^n \, \mathrm{d}t.$$

En développant $\big(1-(x-t)^2\big)^n$ on voit immédiatement que P_n est un polynôme de degré $2n$. Par ailleurs on remarque que si $0 \leq x \leq 1$ alors $f = 0$ sur $[-1+x, 0]$ et sur $[1, 1+x]$ donc

$$P_n(x) = \frac{1}{J_n} \int_{-1+x}^{1+x} f(t)\big(1-(t-x)^2\big)^n \, \mathrm{d}t = \frac{1}{J_n} \int_{-1}^{1} f(t+x)\big(1-t^2\big)^n \, \mathrm{d}t$$

avec un changement de variable. Il en résulte que

$$P_n(x) - f(x) = \frac{1}{J_n} \int_{-1}^{1} \big(f(t+x) - f(x)\big)(1-t^2)^n \, \mathrm{d}t.$$

Soit maintenant $\varepsilon > 0$. Comme f est uniformément continue sur $\mathbb{R}$, il existe $\eta > 0$ tel que, si $|t| \leq \eta$, $|f(t+x) - f(x)| \leq \varepsilon$ pour tout $x \in \mathbb{R}$.

Par ailleurs, $|f(t+x) - f(x)| \leq 2\|f\|_\infty$ donc, pour $|t| \geq \eta$, $1 \leq t^2/\eta^2$ et alors $|f(t+x) - f(x)| \leq 2\|f\|_\infty t^2/\eta^2$. En additionnant les deux estimation, on obtient

$$|f(t+x) - f(x)| \leq \varepsilon + \frac{2\|f\|_\infty}{\eta^2} t^2.$$

Il en résulte que

$$\begin{aligned} |P_n(x) - f(x)| &= \frac{1}{J_n} \int_{-1}^{1} \left(\varepsilon + \frac{2\|f\|_\infty}{\eta^2} t^2\right)(1-t^2)^n \, \mathrm{d}t \\ &= \varepsilon + \frac{2\|f\|_\infty}{\eta^2} \frac{1}{J_n} \int_{-1}^{1} t^2(1-t^2)^n \, \mathrm{d}t. \end{aligned}$$

Notons $K_n = \int_{-1}^{1} t^2(1-t^2)^n \, \mathrm{d}t$ et intégrons par parties

$$K_n = \frac{1}{2} \int_{-1}^{1} t(1-t^2)^n 2t \, \mathrm{d}t = \frac{1}{2(n+1)} \int_{-1}^{1} (1-t^2)^{n+1} \, \mathrm{d}t = \frac{J_{n+1}}{2(n+1)}.$$

Enfin, pour $t \in [-1, 1]$, $(1-t^2)^{n+1} \leq (1-t^2)^n$ donc $J_{n+1} \leq J_n$ donc $\dfrac{K_n}{J_n} \leq \dfrac{1}{2(n+1)}$. Il en résulte que

$$|P_n(x) - f(x)| \leq \varepsilon \frac{2\|f\|_\infty}{\eta^2} \frac{1}{2(n+1)} \leq 2\varepsilon$$

pour peu que $n \geq \eta^2/(\varepsilon\|f\|_\infty)$. □

Nous laissons en exercice de montrer qu'on peut démontrer par la même méthode que

$$\frac{1}{H_n} \int_{-\pi}^{\pi} f(t) \cos^{2n} \frac{t-x}{2} \, \mathrm{d}t, \quad H_n = \int_{-\pi}^{\pi} \cos^{2n} \frac{t-x}{2} \, \mathrm{d}t$$

est une suite de polynômes trigonométriques qui converge uniformément vers f.

3.1.2. *Seconde méthode: les polynômes de Bernstein.* Soit f une fonction continue sur $[0, 1]$, le polynôme de Bernstein associé à f est

$$B_n f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}$$

où les $\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$ sont les coefficients binomiaux.

Théorème 3.2 (Bernstein). *Pour $f \in \mathcal{C}([0, 1])$, $B_n f \to f$ uniformément.*

Démonstration. Pour commencer, il convient de calculer $B_n f$ quand $f_0(t) = 1$, $f_1(t) = t$ et $f_2(t) = t^2$:

$$B_n f_0(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = (x+1-x)^n = 1$$

avec la formule du binôme de Newton. On écrit alors

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

et en dérivant par rapport à x

$$n(x+y)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} k x^{k-1} y^{n-k}$$

donc

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} x^k y^{n-k} = x(x+y)^{n-1}.$$

En prenant $y = 1-x$ on trouve $B_n f_1(x) = x$ (pour $n \geq 1$ car $B_0 f_1 = 0$). Pour obtenir $B_n f_2$ on redérive et on obtient, pour $n \geq 2$

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k^2}{n^2} x^k y^{n-k} = \frac{x}{n} \frac{\partial}{\partial x} [x(x+y)^{n-1}] = \frac{x}{n}((x+y)^{n-1} + (n-1)x(x+y)^{n-2}).$$

En prenant $y = 1-x$ on trouve

$$B_n f_2(x) = \frac{x}{n}(1+(n-1)x) = x^2 + \frac{1}{n} x(1-x).$$

On constate que

$$B_n f_2(x) - f_2(x) = \frac{1}{n} x(1-x) \tag{3.3}$$

en particulier $B_n f_2 \to f_2$ uniformément sur $[0,1]$.

Soit maintenant $f \in \mathcal{C}([0,1]$. On utilise le fait que $B_n f_0 = 1$ pour écrire

$$B_n f(x_0) - f(x_0) = B_n f(x) - f(x) B_n f_0(x) = \sum_{k=0}^{n} \left[f\left(\frac{k}{n}\right) - f(x) \right] \binom{n}{k} x^k (1-x)^{n-k}$$

donc pour $x \in [0,1]$,

$$|B_n f(x) - f(x)| \leq \sum_{k=0}^{n} \left| f\left(\frac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k} \tag{3.4}$$

Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que, si $|y-x| \leq \eta$ alors $|f(y) - f(x)| \leq \eta$. (Remarquons pour la suite que η ne dépend pas de x). Soit $I = \{k \in \{0,\ldots,N\} : |\frac{k}{n} - x| \leq \eta$ et $J = \{0,\ldots,N\} \setminus I$. On a

$$\begin{aligned} \sum_{k \in I} \left| f\left(\frac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k} &\leq \varepsilon \sum_{k \in I} \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \varepsilon \sum_{k=0}^{N} \binom{n}{k} x^k (1-x)^{n-k} = \varepsilon. \end{aligned} \tag{3.5}$$

Par ailleurs, si $k \in J$, alors $|\frac{k}{n} - x| \geq \eta$ que nous allons écrire $1 \leq \frac{1}{\eta^2}\left(\frac{k}{n} - x\right)^2$. D'autre part, $\left|f\left(\frac{k}{n}\right) - f(x)\right| \leq 2\|f\|_\infty$. Il en résulte que

$$
\begin{aligned}
(3.6) \quad \sum_{k\in J} \left| f\left(\frac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k(1-x)^{n-k} \\
&\leq \sum_{k\in J} 2\|f\|_\infty \binom{n}{k} \frac{1}{\eta^2}\left(\frac{k}{n} - x\right) x^k(1-x)^{n-k} \\
&\leq \frac{2\|f\|_\infty}{\eta^2} \sum_{k=0}^{N} \binom{n}{k} \left[\left(\frac{k}{n}\right)^2 - 2x\frac{k}{n} + x^2\right] x^k(1-x)^{n-k} \\
&= \frac{2\|f\|_\infty}{\eta^2}[B_nf_2(x) - 2xB_nf_1(x) + x^2B_nf_0(x)] \\
&= \frac{2\|f\|_\infty}{\eta^2 n} x(1-x) \leq \frac{\|f\|_\infty}{2\eta^2 n}
\end{aligned}
$$

avec (3.3). Il suffit alors de prendre N tel que $\frac{\|f\|_\infty}{2\eta^2 N} \leq \varepsilon$ et on obtient $|B_nf(x) - f(x)| \leq 2\varepsilon$ pour $n \geq N$.

Notons que, comme η est indépendant de x, N aussi, la convergence est donc uniforme. □

Remarque 3.3. Il faut remarquer que les polynômes de Bernstein ne sont pas des polnômes d'interpolation de f. On a en général $B_nf(j/n) \neq f(j/n)$ (bien que $B_nf(j/n)$ soit une approximation de $f(j/n)$ si n est assez grand).

On pourrait se demander pour quoi ne pas considérer les polynômes d'interpolation de Lagrange. Il se trouve que la suite ainsi construite ne converge pas uniformément vers f. Cela peut se montrer à l'aide du théorème de Banach-Steinhaus.

CHAPTER 2

Application de l'analyse fonctionnelle à l'analyse de Fourier

1. Rappel de base sur les séries de Fourier

Dans ce chapitre, on notera

- Pour $1 \leq p < +\infty$,

$$L^p(\mathbb{T}) = \left\{ f\ \mathbb{R} \to \mathbb{C},\ f \text{ 1-périodique},\ \|f\|_p = \left(\int_0^1 |f(t)|^p \,\mathrm{d}t \right)^{\frac{1}{p}} < +\infty \right\};$$

- Pour $p = +\infty$, on modifie

$$L^\infty(\mathbb{T}) = \{ f\ \mathbb{R} \to \mathbb{C},\ f \text{ 1-périodique},\ \|f\|_\infty = \operatorname{supess} |f(t)| < +\infty \};$$

  En d'autres termes, $L^p(\mathbb{T}) = \{ f\ \mathbb{R} \to \mathbb{C},\ f$ 1-périodique $f\mathbf{1}_{[0,1]} \in L^p(\mathbb{R})\}$ muni de la norme $\|f\|_{L^p(\mathbb{T})} = \left\| f\mathbf{1}_{[0,1]} \right\|_{L^p(\mathbb{R})}$. Ainsi $L^p(\mathbb{T})$ s'identifie à $L^p(0,1)$.
- $\mathcal{C}(\mathbb{T}) = \{ f\ \mathbb{R} \to \mathbb{C},\ f$ 1-périodique et continue$\}$. Alors, si $f \in \mathcal{C}(\mathbb{T})$, f est borné et on note, $\|f\|_\infty = \sup |f(t)|$.

Pour $f \in L^p$, dire que f est 1-périodique signifie que $f(t+1) - f(t) = 0$ presque partout et évidemment, on identifie deux fonctions qui sont égales presque partout. Par ailleurs, on noter simplement $\|f\|_p$ si aucune ambiguité n'existe. Remarquons aussi qu'une fonction $f \in L^p(\mathbb{R})$, $p < +\infty$ ne peut être 1-périodique, à moins d'être nulle.

Exercice.

(1) Montrer que $L^p(\mathbb{T}) \subset L^p_{loc}(\mathbb{R}) \subset L^1_{loc}(\mathbb{R})$.
(2) Si $f \in L^1_{loc}(\mathbb{R})$ est 1-périodique, alors pour tout $a \in \mathbb{R}$, $\int_a^{a+1} f(t)\,\mathrm{d}t = \int_0^1 f(t)\,\mathrm{d}t$. En d'autres termes, l'intégrale d'une fonction 1-périodique sur une période ne dépend pas de l'endroit où commence la période.
(3) Si $q \leq p$ alors $L^p(\mathbb{T}) \subset L^q(\mathbb{T})$ et, pour tout $f \in L^p(\mathbb{T})$, $\|f\|_q \leq \|f\|_p$.

Définition 1.1. Pour $f \in L^1(\mathbb{T})$, on définit

- les coefficients de Fourier $c_k(f) = \displaystyle\int_0^1 f(t) e^{-2i\pi kt}\,\mathrm{d}t$, $k \in \mathbb{Z}$;
- les sommes partielles de la série de Fourier de f,

$$S_N f(t) = \sum_{k=-N}^{N} c_k(f) e^{2i\pi kt} \qquad N \geq 0;$$

- la somme de la série de Fourier de f,

$$Sf(t) = \lim_{N \to +\infty} \sum_{k=-N}^{N} c_k(f) e^{2i\pi kt}$$

Attention, on note souvent de façon abusive $Sf(t) = \sum_{k\in\mathbb{Z}} c_k(f)e^{2i\pi kt}$, mais au sens strict, cette dernière quantité signifie

$$Sf(t) = \lim_{M,N\to+\infty} \sum_{k=-M}^{N} c_k(f)e^{2i\pi kt}$$

$$\lim_{M\to+\infty} \sum_{k=-M}^{-1} c_k(f)e^{2i\pi kt} + \lim_{N\to+\infty} \sum_{k=0}^{N} c_k(f)e^{2i\pi kt}.$$

Il est important de remarquer que, quand on parle de séries de Fourier, les indices dans les sommes sont symétriques.

Remarquons que cette difficulté n'a pas lieu si $\sum_{k\in\mathbb{Z}} |c_k(f)|$ converge *i.e.* $\lim_{M,N\to+\infty} \sum_{k=-M}^{N} |c_k(f)|$ a une limite. Comme $\sup_{t\in\mathbb{R}} |c_k(f)e^{2i\pi kt}| = |c_k(f)|$, cela signifie que la série de Fourier converge normalement dans $\mathcal{C}(\mathbb{T})$. En particulier, sa limite est continue.

Exemple 1.2. Soient $0 \le a < b \le 1$ et soit f la fonction 1-périodique telle que $f(x) = \mathbf{1}_{[a,b]}(x)$ si $x \in [0,1]$. Alors $c_0(f) = b - a$ (remarquez que $c_0(f)$ est la moyenne de f) et, si $k \neq 0$, alors

$$c_k(f) = \int_a^b e^{-2i\pi kt}\,\mathrm{d}t = \left[\frac{e^{-2i\pi kt}}{-2i\pi k}\right]_a^b$$

$$= \frac{e^{-2i\pi kb} - e^{-2i\pi ka}}{-2i\pi k} = e^{-2i\pi k\frac{a+b}{2}}\frac{\sin \pi k(b-a)}{\pi k}.$$

Remarquons que $c_k(f) \to 0$ quand $k \to \pm\infty$.

Notation 1.3. On notera dans la suite e_k la fonction 1-périodique continue définie par $e_k(t) = e^{2i\pi kt}$.

Notons que $|e_k| = 1$ et que la 1-périodicité vient du fait que $e^{2i\pi m} = 1$ si $m \in \mathbb{Z}$. De plus $e_{-k}(t) = e_k(-t) = \overline{e_k(t)}$ et $e_je_k = e_{j+k}$. Enfin $\int_0^1 e_k(t)\,\mathrm{d}t = \delta_{k,0}$ où $\delta_{k,j}$ est le symbole de Kronecker

$$\delta_{k,j} = \begin{cases} 1 & \text{si } k = j \\ 0 & \text{sinon.} \end{cases}$$

On notera

$$\mathcal{P}_N = \{\sum_{k=-N}^{N} c_k e^{2ik\pi t} \ : c_k \in \mathbb{C}\}$$

l'espace des polynômes trigonométriques de degré N, et $\mathcal{P} = \bigcup_{N\ge0} \mathcal{P}_N$ l'espace des polynômes trigonométriques.

Remarque 1.4. Si on note $z = e^{2i\pi t}$, on peut remarquer que $e^{-2i\pi t} = \bar{z} = \frac{1}{z}$. Ainsi, le polyôme trigonométrique $\sum_{k=-N}^{N} c_k e^{2ik\pi t} = \sum_{k=-N}^{N} c_k z^k$ est un polynôme de Laurent (restreint au cercle unité).

Exemple 1.5. Si $f = \sum_{k=-N}^{N} c_k e_k \in \mathcal{P}_N$ alors $c_k(f) = \begin{cases} c_k & \text{si } |k| \le N \\ 0 & \text{si } |k| > N \end{cases}$. En particulier $S_M(f) = f$ si $M \ge N$ et $c_k(f) \to 0$ quand $k \to +\infty$.

Cela résulte directement de la propriété $c_k(e_j) = \delta_{k,j}$ qui vient du fait que

$$c_k(e_j) = \int_0^1 e_k(t)\bar{e}_j(t)\,\mathrm{d}t = \int_0^1 e_{k-j}(t)\,\mathrm{d}t = \delta_{k-j,0} = \delta_{k,j}.$$

Lemme 1.6 (Lemme de Riemann-Lebesgue). *Si $f \in L^1(\mathbb{T})$, alors pour tout $k \in \mathbb{Z}$, $|c_k(f)| \le \|f\|_1$. De plus $c_k(f) \to 0$ quand $k \to \pm\infty$.*

En d'autres termes, l'application linéaire $f \to \big(c_k(f)\big)_{k\in\mathbb{Z}}$ est continue $L^1(\mathbb{T}) \to c_0(\mathbb{Z})$.

Démonstration. On peut soit adapter la démonstration du premier semestre pour la transformée de Fourier, soit remarquer que $c_k(f) = \mathcal{F}[f\mathbf{1}_{[0,1]}](k)$ et utiliser le lemme de Riemann-Lebesgue pour la transformée de Fourier.

Pour la démonstration directe, il faut utiliser la densité des fonctions étagées dans $L^1(\mathbb{T})$. On peut aussi utiliser la densité des polynômes trigonométriques dans $L^1(\mathbb{T})$. Il faut toutefois prendre garde que cette densité peut se démontrer à l'aide du lemme de Riemann-Lebesgue. □

Définition 1.7. Pour $f, g \in \mathcal{C}(\mathbb{T})$, on définit leur convolution

$$f * g(t) = \int_0^1 f(s)g(t-s)\,\mathrm{d}s.$$

Alors $f * g \in \mathcal{C}(\mathbb{T})$. De plus, $(f, g) \to f * g$ s'étend en un opérateur bilinéaire $L^p(\mathbb{T}) \times L^q(\mathbb{T}) \to L^r(\mathbb{T})$ à condition que $\dfrac{1}{p} + \dfrac{1}{q} \leq 1 + \dfrac{1}{r}$. Enfin, si $\dfrac{1}{p} + \dfrac{1}{q} = 1$ alors $f * g \in \mathcal{C}(\mathbb{T})$.

La première partie est directe. Pour la deuxième, $L^p(\mathbb{T}) \times L^q(\mathbb{T}) \to L^r(\mathbb{T})$ quand $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$ se fait comme au premier semestre (c'est une utilisation de la dualité et de l'inégalité de Hölder) et ensuite $L^r(\mathbb{T}) \subset L^s(\mathbb{T})$ si $s < r$ *i.e.* si $\dfrac{1}{r} \leq \dfrac{1}{s}$. La dernière se fait également comme pour la convolution sur $\mathbb{R}$. Notons aussi que $f * g = g * f$.

Remarque 1.8. Remarquons que, si $f \in L^1(\mathbb{T})$,

$$\begin{aligned} f * e_k(t) &= \int_0^1 f(s)e^{2i\pi k(t-s)}\,\mathrm{d}s = \int_0^1 f(s)e^{2i\pi kt}e^{-2i\pi ks}\,\mathrm{d}s \\ &= \int_0^1 f(s)e^{-2i\pi ks}\,\mathrm{d}s e^{2i\pi kt} = c_k(f)e_k(t). \end{aligned}$$

Ainsi, si on définit l'opérateur $L^p(\mathbb{T}) \to L^p(\mathbb{T})$ donné par $T_f\ : g \to f * g$, alors $Te_k = c_k(f)e_k$. Ainsi e_k est vecteur propre de T_f pour la valeur propre $c_k(f)$.

Lemme 1.9. *Soient* $1 \leq p, q \leq \infty$. *Si* $f \in L^p(\mathbb{T})$ *et* $g \in L^q(\mathbb{T})$ *alors* $f * g \in L^1(\mathbb{T})$ *et* $c_k(f * g) = c_k(f)c_k(g)$.

Démonstration. Comme $L^p(\mathbb{T}), L^q(\mathbb{T}) \subset L^1(\mathbb{T})$, il suffit d'appliquer l'inégalité de Young pour voir que $f * g \in L^1(\mathbb{T})$. De plus

$$\int_{\mathbb{T}\times\mathbb{T}} |f(s)g(t-s)|\,\mathrm{d}s\,\mathrm{d}t < +\infty$$

ce qui justifie l'utilisation du théorème de Fubini dans le calcul suivant

$$\begin{aligned} c_k(f * g) &= \int_0^1 \int_0^1 f(s)g(t-s)\,\mathrm{d}s\, e^{-2i\pi kt}\,\mathrm{d}t \\ &= \int_0^1 \int_0^1 f(s)g(t-s)e^{-2i\pi k(t-s)}e^{-2i\pi ks}\,\mathrm{d}t\,\mathrm{d}s \\ &= \int_0^1 f(s) \int_s^{s+1} g(x)e^{-2i\pi kx}\,\mathrm{d}x\, e^{-2i\pi ks}\,\mathrm{d}s \\ &= \int_0^1 f(s) \int_0^1 g(x)e^{-2i\pi kx}\,\mathrm{d}x\, e^{-2i\pi ks}\,\mathrm{d}s \\ &= \int_0^1 f(s)c_k(g)\, e^{-2i\pi ks}\,\mathrm{d}s = c_k(f)c_k(g) \end{aligned}$$

où on a utilisé le changement de variable $x = t - s$ et le fait que l'intégrale d'une fonction 1-périodique sur une période ne dépend pas de l'endroit où commence la période. □

Revenons aux séries de Fourier. Notons qu'alors

$$S_N f = \sum_{k=-N}^{N} c_k(f) e_k f = \left(\sum_{k=-N}^{N} e_k\right) * f.$$

Mais si $e_k(t) \neq 1$, c'est-à-dire, si $t \notin \mathbb{Z}$,

$$\begin{aligned}\sum_{k=-N}^{N} e_k(t) &= \sum_{k=-N}^{N} e^{2i\pi kt} = \frac{e^{-2i\pi Nt} - e^{2i\pi(N+1)t}}{1 - e^{2i\pi t}} \\ &= \frac{e^{i\pi t}}{e^{i\pi t}} \frac{e^{-i\pi(2N+1)t} - e^{i\pi(2N+1)t}}{e^{-i\pi t} - e^{i\pi t}} = \frac{-2i\sin(2N+1)\pi t}{-2i\sin\pi t}.\end{aligned}$$

Pour $t \in \mathbb{Z}$, on remarque que $\sum_{k=-N}^{N} e_k(t) = \sum_{k=-N}^{N} 1 = 2N+1$ qui s'obtient aussi par prolongement par continuité de la formule précédente.

Définition 1.10. Le *noyau de Dirichlet* (de degré N) est le polynôme trigonométrique de degré N

$$D_N(t) = \sum_{k=-N}^{N} e_k(t) = \begin{cases} 2N+1 & \text{si } t \in \mathbb{Z} \\ \frac{\sin(2N+1)\pi t}{\sin\pi t} & \text{si } t \notin \mathbb{Z}. \end{cases}$$

Alors $S_N(f) = D_N * f$.

Les propriétés clés du noyau de Dirichlet sont les suivantes

Proposition 1.11. *Le noyau de Dirichlet D_N a les propriétés suivantes:*

(1) *D_N est un polynôme trigonométrique de degré N.*
(2) *D_N est une fonction $\mathcal{C}^\infty$, 1-périodique et paire. En particulier, $D_N \in L^1(\mathbb{T})$.*
(3) $\int_{1/2}^{1/2} D_N(t)\,\mathrm{d}t = 1.$
(4) $\int_{1/2}^{1/2} |D_N(t)|\,\mathrm{d}t \geq \frac{2}{\pi}\ln N.$

Démonstration. Par définition $D_N(t) = \sum_{k=-N}^{N} e_k(t)$ est un polynôme trigonométrique de degré N et est donc dans $\mathcal{C}^\infty$ et 1-périodique. La parité est directe soit avec la deuxième expression de D_N soit avec l'observation $e_k(-t) = e_{-k}(t)$.

On observe immédiatement que

$$\int_{-1/2}^{1/2} e_k(t)\,\mathrm{d}t = \int_0^1 e^{2i\pi kt}\,\mathrm{d}t = \begin{cases} 1 & \text{si } k = 0 \\ \frac{e^{2i\pi k}-1}{2i\pi k} = 0 & \text{si } k \neq 0 \end{cases}$$

et par linératitié de l'intégrale, on en déduit que

$$\int_{1/2}^{1/2} D_N(t)\,\mathrm{d}t = \sum_{k=-N}^{N} \int_{-1/2}^{1/2} e_k(t)\,\mathrm{d}t = 1.$$

Pour la dernière propriété, il faut d'abord tracer le graphe de D_N sur $[0,1/2]$. Remarquons que $t \to \sin\pi t$ est croissante et positive sur $[0,1/2]$. Par ailleurs, $t \to \sin(2N+1)\pi t$ s'annule et change de signe en $\frac{j}{2N+1}$, $j = 0,\ldots,N$. De plus, pour $j \in \{1,\ldots,N\}$ et $t \in I_j := [j/(2N+1),(j+1)/(2N+1)]$, en utilisant le fait que $\sin t \leq t$, on obtient

$$|D_N(t)| = \frac{|\sin(2N+1)\pi t|}{\sin\pi t} \geq \frac{|\sin(2N+1)\pi t|}{\pi t} \geq \frac{(2N+1)|\sin(2N+1)\pi t|}{j+1}.$$

Il en résulte que

$$\begin{aligned}\int_0^{1/2}|D_n(t)| &\geq \sum_{j=0}^{N-1}\int_{I_j}|D_n(t)|\,\mathrm{d}t=\sum_{j=0}^{N-1}\int_{j/(2N+1)}^{(j+1)/(2N+1)}\frac{(2N+1)|\sin(2N+1)\pi t|}{j+1}\,\mathrm{d}t\\ &\geq \sum_{j=0}^{N-1}\frac{1}{j+1}\int_j^{j+1}|\sin\pi t|\,\mathrm{d}t=\frac{2}{\pi}\sum_{j=1}^{N}\frac{1}{j}\\ &\geq \frac{2}{\pi}\int_1^N\frac{\mathrm{d}x}{x}=\frac{2}{\pi}\ln N.\end{aligned}$$

□

Rappelons que nous avons vu la notion d'approximation de l'unité dans $L^p(\mathbb{R})$ au premier semestre. La notion s'étend naturellement à $L^p(\mathbb{T})$. Nous laissons en exercice l'adaption de la démonstration pour obtenir le résultat suivant:

THÉORÈME 1.12 (Approximation de l'unité dans $L^p(\mathbb{T})$). *Soit $k_n\in L^1(\mathbb{T})$ une suite de fonctions telles que*

(1) *pour tout n, $\displaystyle\int_{-1/2}^{1/2}k_n(t)\,\mathrm{d}t=1$;*

(2) *Il existe $C>0$ tel que, pour tout n, $\displaystyle\int_{-1/2}^{1/2}|k_n(t)|\,\mathrm{d}t\leq C$;*

(3) *Pour tout $\delta>0$, $\displaystyle\int_{\delta\leq|t|\leq1/2}|k_n(t)|\,\mathrm{d}t\to0$ quand $n\to+\infty$.*

Alors, si $1\leq p<+\infty$ et $f\in L^p(\mathbb{T})$, $\|k_n*f-f\|_p\to0$ et si $f\in\mathcal{C}(\mathbb{T})$, $\|k_n*f-f\|_\infty\to0$.
On dit que $(k_n)_n$ est une approximation de l'unité.

Malheureusement, $(D_N)_{N\geq1}$ *n'est pas une approximation de l'unité*, on n'a donc aucune garantie (pour l'instant) que $S_Nf=D_N*f\to f$. Avant d'étudier cette question plus avant, nous allons étudier la *Césaro convergence* de S_Nf vers f. Rappelons que nous avons introduit cette notion au premier semestre et qu'on dit que a_n tend vers a au sens de Césaro si $\dfrac{a_0+a_1+\cdots+a_{n-1}}{n}\to a$ et que si $a_n\to a$ alors a_n tend vers a au sens de Césaro mais que la réciproque est fausse.

Considérons donc

$$\sigma_Nf=\frac{S_0f+\cdots+S_{N-1}f}{N}=\frac{D_0*f+\cdots+D_{N-1}*f}{N}=\left(\frac{1}{N}\sum_{n=0}^{N-1}D_n\right)*f.$$

On introduit donc

$$\begin{aligned}F_N(t) &= \frac{1}{N}\sum_{n=0}^{N-1}D_n(t)=\frac{1}{N}\sum_{n=0}^{N-1}\frac{\sin(2n+1)\pi t}{\sin\pi t}=\frac{1}{N\sin\pi t}\sum_{n=0}^{N-1}\mathrm{Im}\,e^{i(2n+1)\pi t}\\ &= \frac{1}{N\sin\pi t}\mathrm{Im}\left(\sum_{n=0}^{N-1}e^{i(2n+1)\pi t}\right)=\frac{1}{N\sin\pi t}\mathrm{Im}\,\frac{e^{i\pi t}-e^{i(2N+1)\pi t}}{1-e^{i2\pi t}}\\ &= \frac{1}{N\sin\pi t}\mathrm{Im}\,e^{iN\pi t}\frac{e^{-iN\pi t}-e^{iN\pi t}}{e^{-i\pi t}-e^{i\pi t}}=\frac{1}{N\sin\pi t}\mathrm{Im}\,e^{iN\pi t}\frac{-2i\sin N\pi t}{-2i\sin\pi t}\\ &= \frac{(\sin N\pi t)^2}{N(\sin\pi t)^2}.\end{aligned}$$

Ce calcul n'est valable que pour $t\notin\mathbb{Z}$ mais pour $t\in\mathbb{Z}$, on obtient le résultat par continuité.

DÉFINITION 1.13. Le *noyau de Fejér* est défini par

$$F_N(t)=\frac{1}{N}\sum_{n=0}^{N-1}D_n(t)=\begin{cases}\frac{1}{N}\left(\frac{\sin N\pi t}{\sin\pi t}\right)^2 & \text{si } t\notin\mathbb{Z}\\ N & \text{si } t\in\mathbb{Z}\end{cases}.$$

et les sommes de Fejér sont définies par

$$\sigma_N f = F_N * f = \frac{1}{N}\sum_{n=0}^{N-1} S_n f.$$

Exercice. Montrer que

$$F_N(t) = \sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)e^{2i\pi kt}$$

et en déduire que

$$\sigma_N f(t) = \sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)c_k(f)e^{2i\pi kt}.$$

THÉORÈME 1.14. *Le noyau de Féjer est une approximation de l'unité. Ainsi, si* $1 \leq p < +\infty$ *et* $f \in L^p$, $\sigma_N f \to f$ *dans* L^p *i.e.* $\|\sigma_N f - f\|_p \to 0$ *et si* $f \in \mathcal{C}(\mathbb{T})$, $\sigma_N f \to f$ *uniformément i.e.* $\|\sigma_N f - f\|_\infty \to 0$.

DÉMONSTRATION. Tout d'abord

$$\int_{-1/2}^{1/2} F_N(t)\,\mathrm{d}t = \int_{-1/2}^{1/2}\frac{1}{N}\sum_{n=0}^{N-1} D_n(t)\,\mathrm{d}t = \frac{1}{N}\sum_{n=0}^{N-1}\int_{-1/2}^{1/2} D_n(t)\,\mathrm{d}t = \frac{1}{N}\sum_{n=0}^{N-1} 1 = 1.$$

Ensuite, on remarque avec le calcul ci-dessus que $F_N \geq 0$ donc

$$\int_{-1/2}^{1/2} |F_N|(t)\,\mathrm{d}t = \int_{-1/2}^{1/2} F_N(t)\,\mathrm{d}t = 1.$$

Enfin F_N est paire donc si $0 < \delta < 1/2$,

$$\int_{\delta<|t|<1/2} |F_N|(t)\,\mathrm{d}t = 2\int_{\delta}^{1/2}\frac{1}{N}\left(\frac{\sin N\pi t}{\sin \pi t}\right)^2\mathrm{d}t \leq \frac{2}{N}\int_{\delta}^{1/2}\frac{\mathrm{d}t}{\sin \pi t} \to 0$$

quand $N \to +\infty$. □

Remarquons enfin que $S_N f$ est un polynôme trigonométrique donc $\sigma_N f$ aussi (ou utiliser l'exercice ci-dessus). On en déduit donc:

COROLLAIRE 1.15. *Les polynômes trigonométriques sont denses dans* $L^p(\mathbb{T})$, $1 \leq p < +\infty$ *et dans* $\mathcal{C}(\mathbb{T})$.

Par ailleurs, si $f, g \in L^1(\mathbb{T})$ sont tels que $c_k(f) = c_k(g)$ pour tout k alors $\sigma_N f = \sigma_N g$. En passant à la limite, on en déduit que $f = g$. Ainsi

COROLLAIRE 1.16. *L'application* $L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ *donnée par* $f \to \big(c_k(f)\big)_{k\in\mathbb{Z}}$ *est injective.*

On peut montrer que cette application n'est pas surjective.

2. Convergence et divergence des séries de Fourier

2.1. La théorie L^2. Le cas $L^2(\mathbb{T})$ est particulier puisqu'il s'agît d'un espace de Hilbert avec pour produit scalaire

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t.$$

On remarque qu'alors

$$\langle e_k, e_\ell\rangle = \int_0^1 e^{2i\pi kt}\overline{e^{2i\pi \ell t}}\,\mathrm{d}t = \int_0^1 e^{2i\pi(k-\ell)t}\,\mathrm{d}t = \delta_{k,\ell}.$$

Ainsi $(e_k)_{k\in\mathbb{Z}}$ est un système orthonormé. Comme par ailleurs l'espace engendré par les $(e_k)_{k\in\mathbb{Z}}$ est l'espace de polynômes trigonométriques qui est dense dans $L^2(\mathbb{T})$, on en déduit que c'est une base orthonormée. On remarque enfin que

$$\langle f, e_k\rangle = \int_0^1 f(t)\overline{e^{2i\pi kt}}\,\mathrm{d}t = \int_0^1 f(t)e^{-2i\pi kt}\,\mathrm{d}t = c_k(f).$$

En utilisant la théorie des espaces de Hilbert, on en déduit

Théorème 2.1. *Soit $f \in L^2(\mathbb{T})$ alors*

(1) *La série de Fourier de f converge vers f dans $L^2(\mathbb{T})$, i.e. $\|f - S_N f\|_2 \to 0$ ou encore, au sens L^2*

$$\lim_{N\to+\infty} \sum_{k=-N}^{N} c_k(f)e^{2i\pi kt} = f$$

qu'on écrit

$$f(t) = \sum_{k=-\infty}^{+\infty} c_k(f)e^{2i\pi kt}. \tag{2.7}$$

(2) *On a l'identité de Plancherel*

$$\int_0^1 |f(t)|^2\,\mathrm{d}t = \sum_{k=-\infty}^{+\infty} |c_k(f)|^2.$$

(3) *On a l'identité de Parseval, si $f, g \in L^2(\mathbb{T})$,*

$$\int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t = \sum_{k=-\infty}^{+\infty} c_k(f)\overline{c_k(g)}.$$

Remarque 2.2. En utilisant la théorie hilbertienne, on a légèrement mieux

$$\lim_{M,N\to+\infty} \sum_{k=-M}^{N} c_k(f)e^{2i\pi kt} = f.$$

Il faut remarquer que cette limite L^2 n'est *pas* une limite presque partout. Du cours du premier semestre, on sait qu'il existe $M_k, N_k \to +\infty$ tel que

$$\lim_{k\to+\infty} \sum_{k=-M_k}^{N_k} c_k(f)e^{2i\pi kt} = f(t) \text{ p.p.}$$

C'est un résultat très profond dû à Carleson ($p = 2$, 1966) et Hunt (cas $1 < p < +\infty$, 1968) que si $f \in L^p(\mathbb{T})$

$$\lim_{N\to+\infty} \sum_{k=-N}^{N} c_k(f)e^{2i\pi kt} = f(t) \text{ p.p.}$$

2.2. Convergence ponctuelle - théorème de Dirichlet.

Théorème 2.3 (Dirichlet). *Soit $f \in L^1(\mathbb{T})$ et $t_0 \in \mathbb{R}$. On suppose qu'il existe $\eta > 0$ tel que, sur $]t_0-\eta, t_0[$ et sur $]t_0, t_0+\eta[$, f est de classe $\mathcal{C}^1$ et de plus*

$$f(t_0^-) = \lim_{t\to t_0^-} f(t) \quad \textit{et} \quad f(t_0^+) = \lim_{t\to t_0^+} f(t)$$

existent (si f est discontinue en t_0, c'est un saut) et f a des dérivées à droite et à gauche en t_0:

$$f'(t_0^-) = \lim_{t\to t_0^-} \frac{f(t) - f(t_0^-)}{t - t_0} \quad \textit{et} \quad f'(t_0^+) = \lim_{t\to t_0^+} \frac{f(t) - f(t_0^+)}{t - t_0}$$

existent. Alors la série de Fourier de f converge en t_0 et

$$\sum_{k=-N}^{N} c_k(f)e^{2i\pi k t_0} = \frac{f(t_0^-)+f(t_0^+)}{2}.$$

Démonstration. Notons $\tilde{f}(t)=f(t_0+t)$ qui vérifie les mêmes propriétés que f, mais avec $t_0=0$. Il est facile de voir que $c_k(\tilde{f})=c_k(f)e^{2i\pi k t_0}$ et alors

$$\begin{aligned} S_N(\tilde{f})(t) &= \sum_{k=-N}^{N} c_k(\tilde{f})e^{2i\pi k t} = \sum_{k=-N}^{N} c_k(f)e^{2i\pi k t_0}e^{2i\pi k t} \\ &= \sum_{k=-N}^{N} c_k(f)e^{2i\pi k(t+t_0)} = S_N(f)(t+t_0). \end{aligned}$$

Quitte à remplacer f par $\tilde{f}$, on peut donc supposer que $t_0=0$.

On remarque ensuite que D_N est paire et que $\displaystyle\int_{-1/2}^{1/2} D_N(t)\,\mathrm{d}t=1$, on a donc

$$\int_{-1/2}^{0} D_N(t)\,\mathrm{d}t = \int_{0}^{1/2} D_N(t)\,\mathrm{d}t = \frac{1}{2}.$$

Mais

$$S_N f(0) = D_N * f(0) = \int_{-1/2}^{1/2} D_N(0-t)f(t)\,\mathrm{d}t = \int_{-1/2}^{1/2} D_N(t)f(t)\,\mathrm{d}t$$

d'où

$$\begin{aligned} S_N f(0) - \frac{f(0^-)+f(0^+)}{2} &= \int_{-1/2}^{0} D_N(t)f(t)\,\mathrm{d}t - \int_{-1/2}^{0} D_N(t)\,\mathrm{d}t\, f(0^-) \\ &\quad + \int_{0}^{1/2} D_N(t)f(t)\,\mathrm{d}t - \int_{0}^{1/2} D_N(t)\,\mathrm{d}t\, f(0^+) \\ &= \int_{-1/2}^{0} D_N(t)\big(f(t)-f(0^-)\big)\,\mathrm{d}t \\ &\quad + \int_{0}^{1/2} D_N(t)\big(f(t)-f(0^+)\big)\,\mathrm{d}t. \end{aligned}$$

Il suffit donc de montrer que

$$\int_{0}^{1/2} D_N(t)\big(f(t)-f(0^+)\big)\,\mathrm{d}t \to 0$$

quand $N\to+\infty$. En remplaçant f par $f(-t)$, la première intégrale tendra également vers 0. Pour cela, rappelons l'expression du noyau de Dirichlet:

$$\begin{aligned} \int_{0}^{1/2} D_N(t)\big(f(t)-f(0^+)\big)\,\mathrm{d}t &= \int_{0}^{1/2} \frac{\sin(2N+1)\pi t}{\sin \pi t}\big(f(t)-f(0^+)\big)\,\mathrm{d}t \\ &= \operatorname{Im} \int_{0}^{1/2} \frac{t}{\sin \pi t}\,\frac{f(t)-f(0^+)}{t}\,e^{2i\pi\frac{2N+1}{2}t}\,\mathrm{d}t \\ &= \operatorname{Im} \mathcal{F}[\varphi]\left(\frac{2N+1}{2}\right) \end{aligned}$$

où $\mathcal{F}$ désigne la transformée de Fourier sur $\mathbb{R}$ et

$$\varphi(t) = \mathbf{1}_{[0,1/2]}\frac{t}{\sin \pi t}\,\frac{f(t)-f(0^+)}{t}.$$

Remarquons que $\varphi \in L^1(\mathbb{R})$ car $\mathbf{1}_{[0,1/2]}\dfrac{t}{\sin \pi t}$ est bornée ($\sin \pi t$ ne s'annule qu'en 0 sur $[0,1/2]$ et $\frac{t}{\sin \pi t}$ se prolonge par continuité en 0) et

$$\frac{f(t)-f(0^+)}{t} = \frac{f(t)-f(0^+)}{t}\mathbf{1}_{[0,\eta]} + \frac{f(t)-f(0^+)}{t}\mathbf{1}_{[\eta,1/2]}.$$

La première est continue sur son support $[0,\eta]$ qui est compact donc dans L^1, la seconde est bornée par

$$\frac{1}{\eta}|f(t)|\mathbf{1}_{[\eta,1/2]} + \frac{|f(0^+)|}{\eta}\mathbf{1}_{[\eta,1/2]}$$

qui est encore dans L^1 car $f \in L^1([0,1/2])$ et le deuxième terme est borné à support compact. On a donc écrit que

$$\int_0^{1/2} D_N(t)\big(f(t)-f(0^+)\big)\,\mathrm{d}t = \operatorname{Im}\mathcal{F}[\varphi]\left(\frac{2N+1}{2}\right)$$

où $\varphi \in L^1(\mathbb{R})$. D'après le Lemme de Riemann-Lebesgue (du premier semestre), $\mathcal{F}[\varphi]\left(\frac{2N+1}{2}\right) \to 0$ donc sa partie imaginaire aussi. $\square$

2.3. Régularité, convergence, l'espace $H^1(\mathbb{T})$. Commençons par une remarque, si $g \in \mathcal{C}(\mathbb{T})$ et f une primitive de g, alors f n'est pas forcément 1-périodique. Plus précisément, on a alors $f(t) = f(0) + \displaystyle\int_0^t g(s)\,\mathrm{d}s$ donc

$$\begin{aligned} f(t+1) = f(0) + \int_0^{t+1} g(s)\,\mathrm{d}s &= f(0) + \int_0^1 g(s)\,\mathrm{d}s + \int_1^{t+1} g(s)\,\mathrm{d}s \\ &= f(0) + \int_0^1 g(s)\,\mathrm{d}s + \int_0^t g(s)\,\mathrm{d}s = f(t) + \int_0^1 g(s)\,\mathrm{d}s. \end{aligned}$$

Donc f est 1-périodique si et seulement si $\displaystyle\int_0^1 g(s)\,\mathrm{d}s = 0$ *i.e.* g et de moyenne nulle sur une période. Notons que dans ce raisonnement, g n'a pas besoin d'être continue, il suffit qu'elle soit intégrable.

Nous allons définir

$$H^1(\mathbb{T}) = \Big\{ f\ \mathbb{R} \to \mathbb{C} \text{ telle qu'il esite } c \in \mathbb{C},\ g \in L^2(\mathbb{T}) \\ \text{avec } \int_0^1 g(t)\,\mathrm{d}t = 0,\ f(t) = c + \int_0^t g(s)\,\mathrm{d}s \Big\}.$$

La première observation est que $L^2(\mathbb{T}) \subset L^1_{loc}(\mathbb{R})$ donc $c + \displaystyle\int_0^t g(s)\,\mathrm{d}s$ est bien définie sur $\mathbb{R}$ et est 1-périodique. Ensuite, si $t, t_0 \in \mathbb{R}$, il existe $R > 0$ tel que $t, t_0 \in [-R, R]$ et alors

$$|f(t)-f(t_0)| = \left|\int_{t_0}^t g(s)\,\mathrm{d}s\right| \leq \int_0^1 \mathbf{1}_{[t_0,t]}(s)|g(s)|\,\mathrm{d}s \leq |t-t_0|^{1/2}\big\|g\mathbf{1}_{[-R,R]}\big\|_2.$$

En particulier, f est continue donc dans $L^2(\mathbb{T})$. En évaluant en $t = 0$ on obtient $c = f(0)$ donc $f(t) = f(0) + \displaystyle\int_0^t g(s)\,\mathrm{d}s$. La fonction g est notée f'. En effet, c'est la dérivée au sens des distributions de f: si $\varphi \in \mathcal{C}_c^\infty$

$$\begin{aligned} \langle f', \varphi\rangle &= -\langle f, \varphi'\rangle = -\int_{\mathbb{R}} f(0)\varphi'(t)\,\mathrm{d}t - \int_{\mathbb{R}}\int_0^t g(s)\,\mathrm{d}s\varphi'(t)\,\mathrm{d}t \\ &= -\int_{\mathbb{R}}\int_{\mathbb{R}} \operatorname{sgn}(t)\mathbf{1}_{[0,t]}(s)g(s)\varphi'(t)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

En effet, la première intégrale est nulle car φ est à support dans un intervalle $[-R,R]$ donc

$$\int_{\mathbb{R}} f(0)\varphi'(t)\,\mathrm{d}t = \int_{-2R}^{2R} f(0)\varphi'(t)\,\mathrm{d}t = f(0)\varphi(2R) - f(0)\varphi(-2R) = 0$$

et pour la deuxième, on peut appliquer Fubini puisque $|\varphi'(t)| \le \|\varphi'\|_\infty \mathbf{1}_{[-R,R]}$ donc

$$\begin{aligned}|\operatorname{sgn}(t)\mathbf{1}_{[0,t]}(s)g(s)\varphi'(t)| &\le \|\varphi'\|_\infty \mathbf{1}_{[0,t]}(s)\mathbf{1}_{[-R,R]}(t)||g(s)|\\ &\le \|\varphi'\|_\infty \mathbf{1}_{[-R,R]}(s)\mathbf{1}_{[-R,R]}(t)||g(s)| \in L^1(\mathbb{R})\end{aligned}$$

puisque $L^2(\mathbb{T}) \subset L^1(\mathbb{T})$ donc g est intégrable sur tout intervalle de longueur 1 et donc localement intégrable. Mais alors

$$\langle f', \varphi\rangle = -\int_{\mathbb{R}} g(s) \int_{\mathbb{R}} \mathbf{1}_{[0,t]}(s)\varphi'(t)\,\mathrm{d}t\,\mathrm{d}s.$$

Revenons donc à Fubini. On remarque que si $t \ge 0$ et $s \in [0,t]$ alors $t \ge s$ et si $t \le 0$ et $\int_0^t g(s)\,\mathrm{d}s = -\int_t^0 g(s)\,\mathrm{d}s$ donc $t \le s$

$$\begin{aligned}-\int_{\mathbb{R}}\int_0^t g(s)\,\mathrm{d}s\varphi'(t)\,\mathrm{d}t &= \int_{-\infty}^0\int_t^0 g(s)\,\mathrm{d}s\,\varphi'(t)\,\mathrm{d}t - \int_0^{+\infty}\int_0^t g(s)\,\mathrm{d}s\,\varphi'(t)\,\mathrm{d}t\\ &= \int_{-\infty}^0\int_{-\infty}^s \varphi'(t)\,\mathrm{d}t\,g(s)\,\mathrm{d}s - \int_0^{+\infty}\int_s^{+\infty}\varphi'(t)\,\mathrm{d}t\,g(s)\,\mathrm{d}s\\ &= \int_{-\infty}^0 g(s)\varphi(s)\,\mathrm{d}s - \int_0^{+\infty} -\varphi(s)\,g(s)\,\mathrm{d}s\\ &= \int_{\mathbb{R}} f(s)\varphi(s)\,\mathrm{d}s = \langle g,\varphi\rangle.\end{aligned}$$

Inversement, si $f \in L^2$ est tel que sa dérivée au sens des distribution est dans $L^2(\mathbb{T})$ et de moyenne nulle, on peut définir

$$\tilde{f}(t) = \int_0^t f'(s)\,\mathrm{d}s.$$

Alors le calcul qu'on vient de faire montre que la dérivée au sens des distributions vérifie $\tilde{f}' - f' = 0$ donc $\tilde{f} - f$ est une constante.

Définition 2.4. L'espace de Hardy est définit par

$$H^1(\mathbb{T}) = \{f \in L^2(\mathbb{T}) \,:\, f' \in L^2(\mathbb{T})\}$$

et on munit cet espace de la norme

$$\|f\|_{H^1}^2 = \int_0^1 |f(t)|^2\,\mathrm{d}t + \int_0^1 |f'(t)|^2\,\mathrm{d}t.$$

Exercice. Montrer que $\|f\|_{H^1}$ définit bien une norme.

Cette norme est clairement issue du produit scalaire

$$\langle f,g\rangle_{H^1} = \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t + \int_0^1 f'(t)\overline{g'(t)}\,\mathrm{d}t.$$

Lemme 2.5. *Les fonctions de H^1 sont Hölderiennes d'exposant $1/2$ et s'écrivent $f(x) = f(0) + \int_0^x f'(t)\,dt$. De plus*

$$\|f\|_\infty \le 2\|f\|_{H^1} \quad \text{et} \quad \sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|^{1/2}} \le \|f\|_{H^1}.$$

En particulier $\mathcal{C}^1(\mathbb{T}) \subset H^1(\mathbb{T}) \subset \mathcal{C}(\mathbb{T})$ et si $(f_k) \in H^1$ converge vers $f \in H^1$ alors $f_k \to f$ uniformément.

DÉMONSTRATION. Soient $f \in H^1(\mathbb{T})$ et $x, y \in \mathbb{R}$. Écrivons $f(x) = f(0) + \int_{\mathbb{R}} \operatorname{sgn}(x)\mathbf{1}_{[0,x]}(t)f'(t)\,\mathrm{d}t$ avec $f' \in L^2(\mathbb{T})$.

Nous allons d'abord montrer que

$$|f(x) - f(y)| \leq \|f\|_{H^1}|x-y|^{1/2}.$$

Il suffit évidemment de considérer le cas $x \leq y$. Commençons par regarder $x \leq y < x+1$. Alors

$$\begin{aligned}|f(x) - f(y)| = \left|\int_x^y f'(t)\,\mathrm{d}t\right| &\leq \int_x^y |f'(t)|\,\mathrm{d}t = \int_x^{x+1} \mathbf{1}_{[x,y]}|f'(t)|\,\mathrm{d}t \\ &\leq |x-y|^{1/2}\left(\int_x^{x+1}|f'(t)|^2\,\mathrm{d}t\right)^{1/2} = |x-y|^{1/2}\left(\int_0^1 |f'(t)|^2\,\mathrm{d}t\right)^{1/2} \leq (x-y|^{1/2}\|f\|_{H^1}\end{aligned}$$

avec Cauchy-Schwarz puis la 1-périodicité de f'.

Si $y > x+1$, il existe un entier $k \geq 0$ tel que $x+k \leq y < x+k+1$. Par 1-périodicité

$$|f(x) - f(y)| = |f(x+k) - f(y)| \leq |x+k-y|^{1/2}\|f\|_{H^1}$$

avec le cas précédent. Mais $|x+k-y| = y-x-k \leq y-x = |x-y|$, on a donc encore

$$|f(x) - f(y)| \leq\leq |x-y|^{1/2}\|f\|_{H^1}. \tag{2.8}$$

Cela montre bien que f est Hölderienne 1/2 avec

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^{1/2}} \leq \|f\|_{H^1}.$$

De plus, en prenant $y = 0$ dans (2.8) on trouve, pour $x \in [0,1]$,

$$|f(x)| \leq |f(0)| + |x|^{1/2}\|f\|_{H^1} \leq 2\|f\|_{H^1}.$$

En prenant le supremum sur $[0,1]$ puis en utilisant la 1-périodicité, on trouve bien $\|f\|_\infty \leq 2\|f\|_{H^1}$. □

Il en résulte donc que

$$\begin{aligned}|f(x)| \leq |f(0)| + \left|\int_0^x f'(t)\,\mathrm{d}t\right| &\leq |f(0)| + \int_0^1 |f'(t)|\,\mathrm{d}t \\ &\leq |f(0)| + \left(\int_0^1 |f'(t)|^2\,\mathrm{d}t\right)^{1/2}\end{aligned}$$

donc

$$|f(x)|^2 \leq 2|f(0)|^2 + 2\int_0^1 |f'(t)|^2\,\mathrm{d}t$$

puis en intégrant sur $[0,1]$,

$$\int_0^1 |f(x)|^2\,\mathrm{d}x \leq 2|f(0)|^2 + 2\int_0^1 |f'(t)|^2\,\mathrm{d}t$$

et finalement $\|f\|_{H^1}^2 \leq 3|f(0)|^2 + 3\int_0^1 |f'(t)|^2\,\mathrm{d}t$. Inversement, le même raisonnement donne

$$|f(0)|^2 \leq 2|f(x)|^2 + 2\int_0^1 |f'(t)|^2\,\mathrm{d}t$$

et en intégrant entre 0 et 1

$$|f(0)|^2 \leq 2\int_0^1 |f(x)|^2\,\mathrm{d}x + 2\int_0^1 |f'(t)|^2\,\mathrm{d}t$$

d'où $|f(0)|^2 + \displaystyle\int_0^1 |f'(t)|^2 \,\mathrm{d}t \leq 3\|f\|_{H^1}^2$. Certains auteurs définissent donc la norme H^1 comme étant $|||f|||_{H^1}^2 = |f(0)|^2 + \displaystyle\int_0^1 |f'(t)|^2 \,\mathrm{d}t$. On vient de montrer que cette norme est équivalente à la précédente:

$$\frac{1}{\sqrt{3}}|||f|||_{H^1} \leq \|f\|_{H^1} \leq \sqrt{3}|||f|||_{H^1}.$$

L'intérêt de la deuxième norme est qu'il est un tout petit peu plus facile de montrer:

Théorème 2.6. *L'espace $H^1(\mathbb{T})$ est complet et est donc un espace de Hilbert.*

Démonstration. Soit $(f_k)_k$ une suite de Cauchy dans $H^1(\mathbb{T})$ et écrivons $f_k(x) = f_k(0) + \displaystyle\int_0^x f_k'(t)\,\mathrm{d}t$.

On a $|f_k(0) - f_\ell(0)| \leq |||f_k - f_\ell|||_{H^1}$ donc $\big(f_k(0)\big)_k$ est une suite de Cauchy dans $\mathbb{C}$. Elle est donc convergente et on note c sa limite. De même $\|f_k' - f_\ell'\|_2 \leq |||f_k - f_\ell|||_{H^1}$ donc $\big(f_k'\big)_k$ est une suite de Cauchy dans $L^2(\mathbb{T})$. Elle est donc convergente et on note g sa limite, $g \in L^2(\mathbb{T})$. On définit alors $f(x) = c + \displaystyle\int_0^x g(t)\,\mathrm{d}t$ donc $f \in H^1(\mathbb{T})$ et $f' = g$. Il reste à voir que $f_k \to f$ dans H^1, ce qui est évident puisque

$$|||f_k - f|||_{H^1}^2 = |f_k(0) - c|^2 + \int_0^1 |f_k'(t) - g(t)|^2 \,\mathrm{d}t \to 0$$

puisque chacun des deux termes de la somme tend vers 0. □

Lemme 2.7. *Les polynômes trigonométriques (donc aussi les fonctions $\mathcal{C}^\infty(\mathbb{T})$) sont denses dans $H^1(\mathbb{T})$.*

Démonstration. Soit $\varepsilon > 0$ et $f \in H^1$. Comme $f' \in L^2$, les sommes partielles de sa série de Fourier convergent vers f' dans L^2, il existe donc N tel que $\|f' - S_N(f')\|_2 \leq \varepsilon$. On définit alors $g(x) = f(0) + \int_0^x S_N(f')(t)\,\mathrm{d}t$.

Notons que, comme f' est de moyenne nulle, $c_0(f') = 0$ donc $S_N(f')$ est de moyenne nulle. Ainsi, on a bien $g_N \in H^1(\mathbb{T})$. Notez que $S_N(f')$ n'a pas de terme constant, donc sa primitive g est encore un polynôme trigonométrique. Enfin $|||f - g|||_{H^1} = \|f' - S_N\| \leq \varepsilon$. □

On peut espérer qu'un certain nombre de propriétés des fonctions de $\mathcal{C}^1(\mathbb{T})$ restent valables dans $H^1(\mathbb{T})$. C'est en particulier le cas de la formule d'intégration par parties:

Lemme 2.8 (Formule d'inégration par parties). *Soient $u, v \in H^1(\mathbb{T})$ alors pour $a, b \in \mathbb{R}$*

$$\int_a^b u(t)v'(t)\,dt = u(b)v(b) - u(a)v(a) - \int_a^b u'(t)v(t)\,dt.$$

En particulier

$$\int_0^1 u(t)v'(t)\,dt = -\int_0^1 u'(t)v(t)\,dt.$$

Démonstration. Remarquons que ceci a bien un sens puisque u, v sont continues donc bornées sur $[a, b]$ et que $u', v' \in L^2(\mathbb{T}) \subset L^1_{loc}(\mathbb{R})$. De plus, la continuité de u, v donne un sens à $u(b)v(b) - u(a)v(a)$. En particulier, par périodicité, $u(1) = u(0)$ et $v(1) = v(0)$, donc la deuxième formule résulte directement de la première.

Cette dernière est évidemment vraie si $u, v \in \mathcal{C}^1$. Comme $\mathcal{C}^1$ est dense dans H^1, si $u, v \in H^1$, il existe des suites (u_k), (v_k) de fonctions de classe $\mathcal{C}^1$ telles que $u_k \to u$ et $v_k \to v$ dans H^1. En particulier, $u_k \to u$ uniformément et $v_k' \to v'$ dans $L^2(\mathbb{T})$ donc dans $L^1([a, b])$. Ainsi

$$\int_a^b u_k(x)v_k'(x)\,\mathrm{d}x \to \int_a^b u(x)v'(x)\,\mathrm{d}x.$$

De même, $\displaystyle\int_a^b u_k'(x)v_k(x)\,\mathrm{d}x \to \int_a^b u'(x)v(x)\,\mathrm{d}x$. La convergence uniforme impliquant la convergence ponctuelle, on a aussi $u_k(b)v_k(b) - u_k(a)v_k(a) \to u(b)v(b) - u(a)v(a)$. En passant à la limite dans la formule d'intégration par parties de $\mathcal{C}^1$,

$$\int_a^b u_k(x)v_k'(x)\,\mathrm{d}x = u_k(b)v_k(b) - u_k(a)v_k(a) - \int_a^b u_k'(x)v_k(x)\,\mathrm{d}x$$

on trouve la formule d'intégration par parties de H^1. $\square$

Corollaire 2.9 (Dirichlet). *Soit $f \in H^1(\mathbb{T})$ alors $c_k(f') = 2i\pi k c_k(f)$. De plus la série de Fourier de f converge uniformément vers f et*

$$\|f\|_{H^1}^2 = \sum_{k\in\mathbb{Z}} (1+4\pi^2k^2)|c_k(f)|^2$$

donc

$$H^1(\mathbb{T}) = \left\{ f \in L^1(\mathbb{T}) \ : \sum_{k\in\mathbb{Z}} (1+4\pi^2k^2)|c_k(f)|^2 \right\}.$$

Démonstration. En appliquant la formule d'intégration par parties $u = f$ et $v = e^{-2i\pi kt}$ on trouve

$$c_k(f') = \int_0^1 f'(t)e^{-2i\pi kt}\,\mathrm{d}t = -\int_0^1 f(t)(-2i\pi k)e^{-2i\pi kt}\,\mathrm{d}t = 2i\pi k c_k(f).$$

On peut remarquer qu'on retrouve le fait que $c_0(f') = \displaystyle\int_0^1 f'(t)\,\mathrm{d}t = 0$ par périodicité de f.

Comme f est continue, on sait grâce à Fejér que $S_N(f) \to f$ au sens de Césaro. Mais par ailleurs, pour $k \neq 0$, $c_k(f') = c_k(f)/2i\pi k$ donc

$$\begin{aligned}\sum_{k\neq 0} \left\| c_k(f)e^{2i\pi kt} \right\|_\infty = \sum_{k\neq 0} |c_k(f)| = \sum_{k\neq 0} \frac{|c_k(f)|}{2\pi|k|} \\ \leq \left(\sum_{k\neq 0} |c_k(f)|^2 \right)^{1/2} \left(\sum_{k\neq 0} \frac{1}{4\pi^2|k|^2} \right)^{1/2}\end{aligned}$$

avec Cauchy-Schwarz. Il résulte de Parseval et $f \in L^2(\mathbb{T})$ que la première somme est finie et la deuxième l'est clairement. Ainsi $S_N(f)$ est normalement convergente donc uniformément convergente. On sait déjà que sa limite au sens de Césaro est f donc sa limite usuelle aussi.

Enfin, Parseval nous dit que

$$\|f\|_{H^1}^2 = \|f\|_2^2 + \|f'\|_2^2 = \sum_{k\in\mathbb{Z}} |c_k(f)|^2 + \sum_{k\in\mathbb{Z}} 4\pi^2k^2|c_k(f)|^2.$$

La caractérisation de $H^1(\mathbb{T})$ en terme de séries de Fourier en résulte immédiatement puisqu'on a alors $S_N(f)$ converge dans H^1. $\square$

Remarque 2.10. La démonstration est un peu plus simple si $f \in C^1$. On n'utilise H^1 que pour étendre la formule d'intégration par parties.

On peut aussi considérer les classes de Hölder $\mathcal{C}^\alpha$ *i.e.* les fonctions qui vérifient $|f(x) - f(y)| \leq C|x-y|^\alpha$, $\alpha > 0$. On peut adapter la démonstration du théorème de Dirichlet de convergence ponctuelle des séries de Fourier pour vérifier que si $f \in \mathcal{C}^\alpha$ $S_N(f) \to f$ ponctuellement si $\alpha > 0$. Un résultat un peu plus fin dû à Bernstein montre qu'on a même convergence normale quand $\alpha > 1/2$. Par contre, ce dernier résultat n'est plus vrai pour $\alpha \leq 1/2$.

Le lecteur intéressé pourra consulter le livre de Katznelson "Introduction to harmonic analysis" ou celui de Zygmund "Fourier series".

Corollaire 2.11. *Soit $f \in L^1(\mathbb{T})$ et supposons que f est de classe $\mathcal{C}^1$ sur un intervalle $(\alpha,\beta) \subset \mathbb{R}$. Soient $\alpha < a < b < \beta$. Alors $S_N(f) \to f$ uniformément sur $[a,b]$.*

Démonstration. Si $\beta-\alpha>1$ il n'y a rien à faire puisqu'alors $f\in\mathcal{C}^1(\mathbb{T})$ et le théorème de Dirichlet permet de conclure. On suppose donc que $\beta-\alpha\leq 1$.

Considérons alors $\alpha<c<a<b<d<\beta$ de sorte que f est $\mathcal{C}^1$ sur (c,d) et on peut prolonger f de (c,d) à $\mathbb{R}$ en une fonction g, 1-périodique de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Ainsi $S_N(f)$ converge uniformément vers g sur $\mathbb{R}$. En appliquant le principe de localisation, il en résulte que $S_N(f)$ converge également uniformément vers $g=f$ sur $[a,b]$. □

2.4. Un exemple: Le phénomène de Gibbs. On considère la fonction de Heaviside périodique:

$$H(t)=\begin{cases}0 & \text{si } t\in(-1/2,0)\\ 1 & \text{si } t\in(0,1/2)\end{cases}$$

et en 0 et en 1/2 on peut prendre n'importe quelle valeur, disons 1/2. Cette fonction vérifie les hypothèses du théorème de Dirichlet. On calcule, $c_0(H)=\dfrac{1}{2}$ et

$$c_k(H)=\int_0^{1/2}e^{-2i\pi kt}\,\mathrm{d}t=\frac{e^{-i\pi k}-1}{-2i\pi k}=\begin{cases}0 & \text{si } k \text{ est paire, } k\neq 0\\ \dfrac{1}{ik\pi} & \text{si } k \text{ est impaire,}\end{cases}.$$

La série de Fourier de H est donc

$$\begin{aligned}S(f)&=\frac{1}{2}+\sum_{p\in\mathbb{Z}}\frac{e^{2i\pi(2p+1)t}}{(2p+1)i\pi}=\frac{1}{2}+\sum_{p=0}^{+\infty}\frac{e^{2i\pi(2p+1)t}-e^{-2i\pi(2p+1)t}}{i(2p+1)\pi}\\ &=\frac{1}{2}+\frac{2}{\pi}\sum_{p=0}^{+\infty}\frac{\sin 2\pi(2p+1)t}{2p+1}.\end{aligned}\tag{2.9}$$

Le théorème de Dirichlet nous dit donc que

$$S[H]=\frac{1}{2}+\frac{2}{\pi}\sum_{p=0}^{+\infty}\frac{\sin 2\pi(2p+1)t}{2p+1}=\begin{cases}0 & \text{si } t\in]-1/2,0[\\ 1 & \text{si } t\in]0,1/2[\\ 1/2 & \text{si } t=0 \text{ ou } \pm\frac{1}{2}\end{cases}.$$

et que, grâce à Dirichlet, cette série converge simplement pour tout $t\in\mathbb{R}$ (vers H sur $]k-1/2,k[\cup]k,k+1/2[$ et vers 1/2 en $k/2$, $k\in\mathbb{Z}$).

Le principe de localisation montre que de plus la convergence est uniforme sur $[\delta,1/2-\delta]$ et sur $[-1/2+\delta,-\delta]$ pour $0<\delta<1/4$. En particulier, il existe C_δ tel que $|S_N(f)(t)-H(t)|\leq C_\delta$ pour tout $t\in[\delta,1/2-\delta]$.

La convergence ne peut être uniforme sur $[0,1/2]$ ni sur $[-1/2,0]$ car $S_N(f)$ étant continue, sa limite uniforme serait continue aussi or on sait que $S_N(0),S_N(1/2),S_N(-1/2)\to 1/2$.

Notez aussi que Parseval donne

$$\frac{1}{2}=\int_{-1/2}^{1/2}|H(t)|^2\,\mathrm{d}t=\sum_{k\in\mathbb{Z}}|c_k(H)|^2=\frac{1}{4}+2\sum_{p=0}^{+\infty}\frac{1}{(2p+1)^2\pi^2}$$

soit $\displaystyle\sum_{p=0}^{\infty}\frac{1}{(2p+1)^2}=\frac{\pi^2}{8}$. Ensuite

$$S:=\sum_{n=1}^{+\infty}\frac{1}{n^2}=\sum_{p=1}^{\infty}\frac{1}{(2p)^2}+\sum_{p=0}^{\infty}\frac{1}{(2p+1)^2}=\frac{1}{4}S+\frac{\pi^2}{8}$$

et en résolvant l'équation en S, $\displaystyle\sum_{n=1}^{+\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$.

Étudions l'absence de convergence uniforme de plus près. Pour cela, considérons les sommes partielles

$$S_{2N+1}(H)(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{p=0}^{N}\frac{\sin 2\pi(2p+1)t}{2p+1}.$$

On écrit $S_{2N+1}(H)(t) = \dfrac{1}{2} + \Phi_N(t)$ et on remarque que Φ_N est 1-périodique, impaire. De plus

$$\Phi_N(1/2-t) = \frac{2}{\pi}\sum_{p=0}^{N}\frac{\sin\big(\pi(2p+1) - 2\pi(2p+1)t\big)}{2p+1} = \Phi_N(t).$$

Il suffit donc d'étudier Φ_N sur $]0,1/4]$.

En dérivant, on trouve

$$\begin{aligned}
\Phi_N'(t) &= 4\sum_{p=0}^{N}\cos 2(2p+1)\pi t = 4\Re\sum_{p=0}^{N} e^{2(2p+1)i\pi t}\\
&= 4\Re\frac{e^{2i\pi t} - e^{2(2N+3)i\pi t}}{1-e^{4i\pi t}}\\
&= 4\Re\left(e^{2(N+1)i\pi t}\frac{e^{-2(N+1)i\pi t} - e^{2(N+1)i\pi t}}{e^{-2i\pi t} - e^{2i\pi t}}\right)\\
&= 4\frac{\cos 2(N+1)\pi t\sin 2(N+1)\pi t}{\sin 2\pi t} = 2\frac{\sin 4(N+1)\pi t}{\sin 2\pi t}
\end{aligned}$$

Ainsi, on voit que Φ_N atteint un extremum local en $t_k = \frac{k}{4(N+1)}$, $k = 1,\dots,N$. Plus précisément, il s'agît d'un maximum local si k est impaire et d'un minimum local si k est pair.

De plus

$$\begin{aligned}
\Phi(t_{2k+1}) - \Phi(t_{2k-1}) &= 2\int_{\frac{2k-1}{4(N+1)}}^{\frac{2k+1}{4(N+1)}}\frac{\sin 4(N+1)\pi t}{\sin 2\pi t}\,\mathrm{d}t\\
&= 2\int_{-\frac{1}{4(N+1)}}^{\frac{1}{4(N+1)}}\frac{\sin 4(N+1)\pi t}{\sin 2\pi\left(t+\frac{k}{2(N+1)}\right)}\,\mathrm{d}t\\
&= \frac{1}{2(N+1)}\int_{-\pi}^{\pi}\frac{\sin s}{\sin\frac{s+2k\pi}{2(N+1)}}\,\mathrm{d}s\\
&= \frac{1}{2(N+1)}\left[\int_0^{\pi}\sin s\left[-\frac{1}{\sin\frac{2k\pi-s}{2(N+1)}} + \frac{1}{\sin\frac{2k\pi+s}{2(N+1)}}\right]\mathrm{d}s\right].
\end{aligned}$$

Comme $1/\sin t$ est décroissante pour $t\in[0,\pi/2]$ et $\frac{2k\pi-s}{2(N+1)}, \frac{2k\pi+s}{2(N+1)} \in [0,\pi/2]$ si $s\in[0,\pi]$ et $2k+1 \le N+1$ *i.e.* $t_{2k+1}\le 1/4$, on obtient $\Phi(t_{2k+1}) - \Phi(t_{2k-1}) \le 0$ pour ces k. En d'autres termes, les maxima locaux sont décroissants jusqu'à $1/4$ et

$$\sup_{t\in[0,1/2]} S_{2N+1}[H](t) = \frac{1}{2} + \sup_{t\in[0,1/2]}\Phi_N(t) = \frac{1}{2} + \Phi_N(t_1).$$

Mais

$$\Phi_N(t_1) = \frac{2}{\pi}\sum_{p=0}^{N}\frac{\sin\pi\frac{2p+1}{2(N+1)}}{2p+1} = 2\frac{1}{N+1}\sum_{p=0}^{N}\operatorname{sinc}\pi\frac{p+1/2}{N+1}$$

où on reconnaît une somme de Riemann de $\operatorname{sinc}\pi t = \frac{\sin\pi t}{\pi t}$ sur $[0,1]$. Ainsi

$$\Phi_N(t_1) \to_{N\to+\infty} \frac{2}{\pi}\int_0^1\frac{\sin\pi t}{t}\,\mathrm{d}t \simeq 1.179.$$

On peut également utiliser une inégalité classique de comparaison de somme de Riemann et d'intégrale

$$\left|\int_0^1 \varphi(t)\,\mathrm{d}t - \frac{1}{N+1}\sum_{k=0}^{N}\varphi\left(\frac{k+1/2}{N+1}\right)\right| \le \frac{1}{N+1}\sup_{t\in[0,1]}|\varphi'(t)|.$$

Quand $\varphi(t) = \dfrac{\sin \pi t}{\pi t}$, on a

$$\varphi'(t) = \frac{\pi t\cos \pi t - \sin \pi t}{\pi t^2}.$$

Il est facile de voir que cette fonction est bornée car elle se prolonge par continuité en 0 et tend vers 0 à l'infini. En écrivant

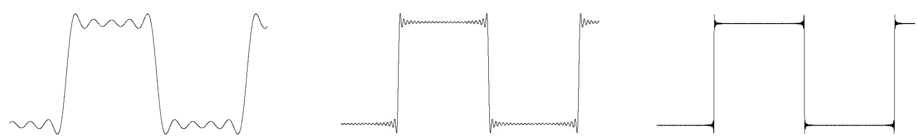
$$\varphi'(t) = \frac{1}{\pi t^2}\int_0^t s\sin \pi s\,\mathrm{d}s \le \frac{1}{t^2}\int_0^t s^2\,\mathrm{d}t \le \frac{t}{3}$$

on trouve $\psi(t) \le 1/3$ si $t \le 1$.Ainsi $\sup_{t\in[0,1]}|\varphi'(t)| \le 1/3$ donc

$$\left|\Phi_N(t_1) - \frac{2}{\pi}\int_0^1 \frac{\sin \pi t}{t}\,\mathrm{d}t\right| \le \frac{2}{3(N+1)}.$$

Ainsi la suite des premiers maximas $\Phi_N(t_1)$ est bornée et converge vers 1.179....

Ceci est illustré dans la figure suivante où on trace les sommes partielles de la fonction H.



On observe bien qu'il y a toujours un pic après le saut dont la hauteur est environ 20% supérieur au saut. Par ailleurs, on observe bien l'uniforme convergence lorsqu'on se place sur un intervalle fixe sur lequel H est continu.

2.5. Divergence dans $\mathcal{C}(\mathbb{T})$ et dans $L^1(\mathbb{T})$. Nous allons maintenant étudier deux cas de divergence de séries de Fourier: le cas des fonctions continues et le cas L^1. Pour cela, nous allons utiliser Banach-Steinhaus et le lemme suivant:

Lemme 2.12. *Soit $S_N : L^1(\mathbb{T}) \to L^1(\mathbb{T})$ donnée par $S_N(f) = D_N * f$. Alors*

$$\|S_N\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})} := \sup_{f\in\mathcal{C}(\mathbb{T}),\|f\|_\infty=1}\|S_N(f)\|_\infty = \|D_N\|_{L^1(\mathbb{T})}.$$

Si on considère $S_N : \mathcal{C}(\mathbb{T}) \to \mathcal{C}(\mathbb{T})$ donnée par $S_N(f) = D_N * f$. Alors

$$\|S_N\|_{\mathcal{C}(\mathbb{T})\to\mathcal{C}(\mathbb{T})} := \sup_{f\in\mathcal{C}(\mathbb{T}),\|f\|_\infty=1}\|S_N(f)\|_\infty = \|D_N\|_{L^1(\mathbb{T})}.$$

Soit $T_N : \mathcal{C}(\mathbb{T}) \to \mathbb{C}$ donnée par $T_N(f) = D_N * f(0)$. Alors

$$\|T_N\|_{\mathcal{C}(\mathbb{T})\to\mathbb{C}} := \sup_{f\in\mathcal{C}(\mathbb{T}),\|f\|\infty=1}|S_N(f)(0)| = \|D_N\|_{L^1(\mathbb{T})}.$$

Démonstration. On a clairement

$$|S_N(f)(t)| = \left|\int_0^1 D_N(s)f(t-s)\,\mathrm{d}s\right| \le \int_0^1 |D_N(s)|\,|f(s-t)|\,\mathrm{d}s \tag{2.10}$$

donc, avec Fubini

$$\int_0^1 |S_N(f)(t)|\,\mathrm{d}t \le \int_0^1 |D_N(s)| \int_0^1 |f(s-t)|\,\mathrm{d}t\,\mathrm{d}s = \|D_N\|_{L^1}\|f\|_1$$

Ainsi $\|S_N\|_{L^1\to L^1} \leq \|D_N\|_{L^1}$. Par ailleurs, en utilisant le noyau de Fejér F_N qui est de norme $\|F_N\|_1 = 1$, et

$$S_N(F_M) = D_N * F_M = F_M * D_N = \sigma_M(D_N) \to_{M\to\infty} D_N$$

dans L^1. Ainsi $\|S_N(F_M)\|_{L^1(\mathbb{T})} \to \|D_N\|_{L^1(\mathbb{T})}$. Mais

$$\|S_N\|_{L^1\to L^1} = \sup\{\|S_N(F)\|_{L^1(\mathbb{T})} : \|F\|_{L^1(\mathbb{T})} = 1\} \geq \|S_N(F_M)\|_{L^1(\mathbb{T})}$$

et, en faisant tendre $M \to \infty$, on trouve $\|S_N\|_{L^1\to L^1} \geq \|D_N\|_{L^1(\mathbb{T})}$.

Pour la norme de $S_N : \mathcal{C}(\mathbb{T}) \to \mathcal{C}(\mathbb{T})$ on déduit de (2.10) que $|S_N(f)(t)| \leq \|D_N\|_1 \|f\|_\infty$ d'où

$$|S_N(f)(0)| \leq \sup_{t\in\mathbb{R}} |S_N(f)(t)| \leq \|D_N\|_1 \|f\|_\infty.$$

Par suite $\|T_N\|_{\mathcal{C}(\mathbb{T})\to\mathbb{C}} \leq \|S_N\|_{\mathcal{C}(\mathbb{T})\to\mathcal{C}(\mathbb{T})} \leq \|D_N\|_{L^1(\mathbb{T})}$.

Il reste à trouver une suite $f_k \in \mathcal{C}(\mathbb{T})$ avec $\|f_k\|_\infty = 1$ telle que $|T_N(f_k)| \to \|D_N\|_{L^1(\mathbb{T})}$. Remarquons que

$$T_N(f) = \int_0^1 f(s)D_N(-s)\,\mathrm{d}s = \int_{-1/2}^{1/2} f(s)D_N(s)\,\mathrm{d}s$$

par parité et périodicité de D_N. Si on n'imposait pas la continuité de f, on prendrait simplement f paire, 1-périodique, définie sur $[0,1/2]$ par

$$\begin{aligned} f(s) &= \operatorname{sgn}(D_N)(s) \\ &= \begin{cases} 1 & \text{si } s \in [0, 1/(2N+1)] \\ 1 & \text{si } s \in [2j/(2N+1), (2j+1)/2(N+1)],\ j \in \mathbb{N} \\ -1 & \text{si } s \in [2(j+1)/(2N+1), (2j+2)/2(N+1)],\ j \in \mathbb{N} \end{cases}. \end{aligned}$$

On a alors

$$T_N(f) = \int_{-1/2}^{1/2} \operatorname{sgn}(D_N)(s)D_N(s)\,\mathrm{d}s = \int_{-1/2}^{1/2} |D_N(s)|\,\mathrm{d}s = \|D_N\|_{L^1(\mathbb{T})}.$$

Il suffit alors de définir pour $k \geq 10N$, f_k paire, 1-périodique, continue affine par morceaux, telle que $f_k = f$ sur les intervalles $[0, 1/(2N+1) - 1/k]$, $[j/(2N+1) + 1/k, j + 1/2(N+1) - 1/k]$, $j = \{1, \ldots, N\}$. Alors $f_k(s) \to \operatorname{sgn}(D_N)(s)$ pour tout $s \in \mathbb{R}$ et $|f_k| \leq 1$. Par convergence dominée $T_N(f_k) \to T_N(f)$. □

Il suffit alors d'appliquer le théorème de Banach-Steinhaus pour obtenir:

Corollaire 2.13. *Il existe une fonction $f \in L^1(\mathbb{T})$ dont la série de Fourier $S_N(f)$ ne converge pas vers f dans $L^1(\mathbb{T})$.*

Il existe une fonction continue f dont la série de Fourier $S_N(f)$ ne converge pas uniformément vers f.

Il existe une fonction continue f dont la série de Fourier en 0, $S_N(f)(0)$, ne converge pas vers $f(0)$.

3. Quelques applications des séries de Fourier

3.1. Résolution d'équations différentielles et aux dérivées partielles. De nombreuses équations différentielles et équations aux dérivées partielles peuvent se résoudre à l'aide de séries de Fourier. Nous allons ici nous concentrer sur léquation de la chaleur avec conditions de Dirichlet sur $[0,1] \times \mathbb{R}$

$$\begin{cases} 4\pi^2 \partial_t u(x,t) = \partial_x^2 u(x,t) & \text{pour } t > 0,\ 0 \leq x \leq 1 \\ u(x,0) = u_0(x) & \text{pour } 0 \leq x \leq 1 \\ u(t,0) = u(t,1) & \text{pour } t > 0 \end{cases}$$

où u_0 sera supposé de classe $\mathcal{C}^1$ (pour simplifier). Le facteur $4\pi^2$ est ici pour simplifier les calculs, on peut vérifier sans peine que si $\partial_t u(x,t) = \partial_x^2 u(x,t)$ alors $v(x,t) = u(x, t/\lambda)$ vérifie $\lambda\partial_t v(x,t) = \partial_x^2 v(x,t)$.

L'idée est de chercher une solution sous forme de série de Fourier en x.

$$u(x,t) = \sum_{k\in\mathbb{Z}} c_k(t) e^{2i\pi kx}$$

avec c_k des fonctions de classe $\mathcal{C}^1$ telles que $\sum_{k\in\mathbb{Z}} k^2|c_k(t)| < +\infty$ et $\sum_{k\in\mathbb{Z}} |c_k'(t)| < +\infty$ Cette condition garantie que u soit de classe $\mathcal{C}^2$ en x et que de classe $\mathcal{C}^1$ en t puisque les séries $\sum_{k\in\mathbb{Z}} \partial_t[c_k(t)e^{2i\pi kx}]$ et $\sum_{k\in\mathbb{Z}} \partial_x^j[c_k(t)e^{2i\pi kx}]$, $j=0,1,2$ sont alors normalement convergentes. De plus

$$\partial_t u = \sum_{k\in\mathbb{Z}} c_k'(t)e^{2i\pi kx} \quad \text{et} \quad \partial_x^2 u = -4\pi^2\sum_{k\in\mathbb{Z}} k^2 c_k e^{2i\pi kx}.$$

Ainsi, par unicité des coefficients de Fourier, $\partial_t u(x,t) = \partial_x^2 u(x,t)$ implique que pour tout k, $c_k'(t) = -k^2 c_k(t)$. Par ailleurs, En développant en série de Fourier u_0, $u_0 = \sum_{k\in\mathbb{Z}} c_k(u_0)e^{2i\pi kx}$. Ainsi $c_k(t)$ vérifie l'équation différentielle

$$\begin{cases} c_k'(t) = -k^2 c_k(t) \\ c_k(0) = c_k(u_0) \end{cases}$$

dont les solutions sont $c_k(t) = c_k(u_0)e^{-k^2t}$ et alors

$$u(x,t) = \sum_{k\in\mathbb{Z}} c_k(u_0)e^{-k^2t}e^{2i\pi kx}.$$

On vérifie sans peine que, du fait que u_0 soit $\mathcal{C}^1$, $c_k(u_0)$ est bornée (sommable) et alors

(1) $u(x,t) \to u_0(x)$ quand $t\to 0$: en effet $c_k(u_0)e^{-k^2t}e^{2i\pi kx} \to c_k(u_0)e^{2i\pi kx}$ et comme $\sum_{k\in\mathbb{Z}} |c_k(u_0)e^{2i\pi kx}| = \sum_{k\in\mathbb{Z}} |c_k(u_0)| < +\infty$, le théorème de convergence dominée implique
$$\sum_{k\in\mathbb{Z}} c_k(u_0)e^{-k^2t}e^{2i\pi kx} \to_{t\to 0} \sum_{k\in\mathbb{Z}} c_k(u_0)e^{2i\pi kx} = u_0(x);$$
(2) $\sum_{k\in\mathbb{Z}} |\partial_t^j[c_k(u_0)e^{-k^2t}e^{2i\pi kx}]| = \sum_{k\in\mathbb{Z}} |c_k(u_0)|k^{2j}e^{-k^2t} < +\infty$ puisque $k^{2j}e^{-k^2t}$ est borné si $t>0$. En particulier, u est de classe $\mathcal{C}^\infty$ en t.
(3) $\sum_{k\in\mathbb{Z}} |\partial_x^j[c_k(u_0)e^{-k^2t}e^{2i\pi kx}]| = \sum_{k\in\mathbb{Z}} |c_k(u_0)|(2\pi|k|)^j e^{-k^2t} < +\infty$ puisque $k^j e^{-k^2t}$ est borné si $t>0$. En particulier, u est de classe $\mathcal{C}^\infty$ en x.

Ainsi, la fonction $u(x,t)$ que nous avons trouvée est bien solution de l'équation de la chaleur et celle-ci est unique:

THÉORÈME 3.1. *Soit u_0 une fonction de classe $\mathcal{C}^1$, $\Omega = (0,1)\times(0,+\infty)$ et $\bar\Omega = [0,+\infty)\times[0,1]$. Alors l'unique fonction $u : [0,1]\times[0,+\infty) \to \mathbb{C}$ continue sur $\bar\Omega$, de classe $\mathcal{C}^2$ sur Ω solution de*

$$\begin{cases} 4\pi^2\partial_t u(x,t) = \partial_x^2 u(x,t) & (x,t)\in\Omega \\ u(x,0) = u_0(x) & \textit{pour } 0\le x\le 1 \\ u(t,0) = u(t,1) & \textit{pour } t>0 \end{cases}$$

est donnée par

$$u(x,t) = \sum_{k\in\mathbb{Z}} c_k(u_0)e^{-k^2t}e^{2i\pi kx}.$$

Cette fonction est de classe $\mathcal{C}^\infty$ sur Ω et la série ainsi que toutes ses dérivées en t et en x converge normalement sur $[0,1]$ pour $t>0$.

On peut maintenant remarquer que nous n'avons que très peu utilisé la régularité de u_0. D'une part, la convergence normale de sa série de Fourier est assurée par exemple si $u_0\in H^1$ et elle ne sert que pour obtenir $u(x,t)\to u_0(x)$ quand $t\to 0$.

D'autre part, pour $t>0$, la convergence uniforme des séries

$$\sum_{k\in\mathbb{Z}} |c_k k^j e^{-k^2t}e^{2i\pi kx}|$$

provient essentiellement du facteur rapidement décroissant donc il suffit que c_k soit à croissance au plus polynômiale. Par exemple

$$\theta(x,t) = \sum_{k\in\mathbb{Z}} c_k(u_0)e^{-\pi k^2 t}e^{2i\pi kx}$$

est solution de l'équation de la chaleur

$$\frac{\partial^2\theta}{\partial x^2} = 4\pi\frac{\partial\theta}{\partial t}.$$

Nous reviendrons sur θ lorsque nous aurons démontré la formule sommatoire de Poisson dont la démonstration relie séries de Fourier et transformée de Fourier.

3.2. Théorème de Lax-Milgram et problème de Sturm-Liouville.

THÉORÈME 3.2 (Théorème de Lax-Milgram). *Soit H un espace de Hilbert. Soit $a : H\times H\to\mathbb{C}$ tel que*

(i) *a est une forme bilinéaire;*
(ii) *a est continue: il existe $C>0$ tel que, pour tous $f,g\in H$, $|a(f,g)|\leq C\|f\|\|g\|$;*
(iii) *a est coercive: il existe $\alpha>0$ tel que $a(f,f)\geq\alpha\|f\|^2$.*

Soit L une forme linéaire continue sur H. Alors il existe un unique $g\in H$ tel que, pour tout $f\in H$,

$$a(f,g) = L(f).$$

DÉMONSTRATION. L'unicité de g est facile à voir: si on suppose que g_1, g_2 vérifient cette propriété, alors $a(f,g_1-g_2) = a(f,g_1)-a(f,g_2) = L(f)-L(f) = 0$ pour tout f. En particulier, pour $f = g_1-g_2$, $\alpha\|g_1-g_2\|^2 \leq a(g_1-g_2, g_1-g_2) = 0$ donc $g_1 = g_2$.

D'après le théorème de représentation de Riesz,
– il existe $\ell\in H$ tel que, pour tout $f\in H$, $L(f) = \langle f,\ell\rangle$;
– pour tout $g\in H$, il existe un unique $T_g\in H$ tel que, pour tout $f\in H$, $a(f,g) = \langle f, T_g\rangle$.

Notons que l'unicité de T_g implique que $T : g\to T_g$ est linéaire puisque

$$\begin{aligned}\langle f, T_{\lambda_1 g_1+\lambda_2 g_2}\rangle = a(f,\lambda_1 g_1+\lambda_2 g_2) &= \lambda_1 a(f,g_1)+\lambda_2 a(f,g_2)\\ &= \lambda_1\langle f,T_{g_1}\rangle+\lambda_2\langle f,T_{g_2}\rangle = \langle f,\lambda_1 T_{g_1}+\lambda_2 T_{g_2}\rangle.\end{aligned}$$

Par unicité, T est bien linéaire. On note maintenant Tg plutôt que T_g. Notons que de plus

$$\|Tg\|^2 = \langle Tg, Tg\rangle = a(Tg,g)\leq C\|Tg\|\|g\|.$$

Donc $\|Tg\|\leq C\|g\|$ (qui est aussi vérifié si $Tg=0$). Ainsi T est une application linéaire bornée $H\to H$. Notons que si $Tg=0$ alors

$$0 = \langle g, Tg\rangle = a(g,g)\geq\alpha\|g\|^2$$

donc $g=0$. Ainsi T est injective.

Montrons maintenant que l'image de T est dense. Soit $f\in(\operatorname{Im}T)^\perp$, c'est-à-dire que, pour tout $g\in H$, $\langle f, Tg\rangle = 0$. En particulier, si $g=f$,

$$0 = \langle f, Tf\rangle = a(f,f)\geq\alpha\|f\|^2$$

donc $f=0$. Ainsi $(\operatorname{Im}T)^\perp = 0$ et $H = \overline{\operatorname{Im}T}\oplus(\operatorname{Im}T)^\perp = \overline{\operatorname{Im}T}$.

Montrons maintenant que T est surjectif. Il suffit donc de montrer que $\operatorname{Im}T$ est fermé. Prenons donc une suite (Tg_k) de $\operatorname{Im}T$ qui converge et montrons que sa limite est dans $\operatorname{Im}T$. Pour cela remarquons qu'avec Cauchy-Schwarz,

$$\|g\|\|Tg\|\geq\langle g, Tg\rangle = a(g,g)\geq\alpha\|g\|^2$$

donc $\|Tg\|\geq\alpha\|g\|$. Ainsi, si (Tg_k) converge, c'est une suite de Cauchy, donc (g_k) aussi. Mais H est complet, donc (g_k) a une limite g. Comme T est continu, $Tg_k\to Tg\in\operatorname{Im}T$.

Ainsi T est une bijection continue. Le théorème de l'application ouverte nous dit que T^{-1} est également continue. On peut aussi remarquer que $\alpha\|T^{-1}f\|\leq\|TT^{-1}f\| = \|f\|$ *i.e.* $\|T^{-1}\|\leq 1/\alpha$.

Enfin, $a(f,g) = \langle f, Tg\rangle$, $L(f) = \langle f,\ell\rangle$. En posant $g = T^{-1}\ell$ on a bien, pour tout $f\in H$, $a(f,g) = \langle f, TT^{-1}\ell\rangle = \langle f,\ell\rangle = L(f)$. □

On considère maintenant l'équation de Sturm-Liouville sur $(0,1)$

$$(E) \qquad -(p(x)y')'(x) + q(x)y(x) = f(x) \qquad x \in (0,1)$$

avec les conditions au bord périodiques $y(0) = y(1)$, $y'(0) = y'(1)$. On suppose ici que $p \in \mathcal{C}^1(\mathbb{T})$, $q, f \in \mathcal{C}(\mathbb{T})$. Une fonction y qui vérifie (E) sera appelé une solution forte de (E).

Remarquons que si $\varphi \in \mathcal{C}^\infty(\mathbb{T})$ alors (E) implique

$$(3.11) \qquad \int_0^1 -(p(x)y')'(x)\varphi(x)\,\mathrm{d}x + \int_0^1 q(x)y(x)\varphi(x)\,\mathrm{d}x = \int_0^1 f(x)\varphi(x)\,\mathrm{d}x$$

d'où, après une intégration par parties

$$(F) \qquad \int_0^1 p(x)y'(x)\varphi'(x) + q(x)y(x)\varphi(x)\,\mathrm{d}x = \int_0^1 f(x)\varphi(x)\,\mathrm{d}x.$$

On a utilisé la périodicité de p, y et φ pour que

$$p(1)y'(1)\varphi(1) - p(0)y'(0)\varphi(0) = 0.$$

Inversement, si y est de classe $\mathcal{C}^2$ telle que (F) est vérifié pour tout $\varphi \in \mathcal{C}^\infty(\mathbb{T})$, alors on en déduit que y vérifie aussi (3.11). En posant $\psi(x) = -(p(x)y')'(x) + q(x)y(x) - f(x)$, on a alors $\psi \in \mathcal{C}(\mathbb{T})$ tel que pour tout $\varphi \in \mathcal{C}^\infty(\mathbb{T})$, $\displaystyle\int_0^1 \psi(x)\varphi(x)\,\mathrm{d}x = 0$ donc $\psi = 0$.

On dira que y est une solution faible de (E) si y vérifie (F) pour tout $\varphi \in \mathcal{C}^\infty(\mathbb{T})$ et on vient de voir que si y de classe $\mathcal{C}^2$ est solution faible, alors y est aussi solution forte.

Observons (F) un peu mieux. On peut voir que (F) est parfaitement définie en supposant moins de régularité: pour f, il suffit de supposer que $f \in L^2(\mathbb{T})$. On peut aussi supposer que $y \in H^1(\mathbb{T})$ et que $p, q \in L^\infty(\mathbb{T})$.

Ainsi, on pose $H = H^1(\mathbb{T})$. Pour $f \in L^2(\mathbb{T})$, on définit la forme linéaire L_f sur H par $L_f(\varphi) = \displaystyle\int_0^1 f(x)\varphi(x)\,\mathrm{d}x$. On remarque que $|L_f(\varphi)| \leq \|f\|_2\|\varphi\|_2 \leq \|f\|_2\|\varphi\|_{H^1}$. On définit ensuite la forme bilinéaire a sur $H \times H$ par

$$a(y,\varphi) = \int_0^1 p(x)y'(x)\varphi'(x) + q(x)y(x)\varphi(x)\,\mathrm{d}x.$$

On remarque que

$$\begin{aligned}
|a(y,\varphi)| &\leq \|p\|_\infty\|y'\|_2\|\varphi'\|_2 + \|q\|_\infty\|y\|_2\|\varphi\|_2 \\
&\leq \max(\|p\|_\infty, \|q\|_\infty)(\|y'\|_2 + \|y\|_2)(\|\varphi'\|_2 + \|\varphi\|_2) \\
&\leq \max(\|p\|_\infty, \|q\|_\infty)\|y\|_{H^1}\|\varphi\|_{H^1}.
\end{aligned}$$

Ainsi a est une forme bilinéaire continue sur $H \times H$. Enfin

$$a(y,y) = \int_0^1 p(x)|y'(x)|^2 + q(x)|y(x)|^2\,\mathrm{d}x \geq \alpha\|y\|_{H^1}$$

si $p(x), q(x) > \alpha$ presque partout. On peut alors appliquer le théorème de Lax-Milgram et en déduire le résultat suivant.

Théorème 3.3. *Soient $p, q \in L^\infty(\mathbb{T})$ et qu'il existe $\alpha > 0$ avec $p, q > \alpha$. Soit $f \in L^2(\mathbb{T})$. Alors l'équation*

$$(E) \qquad -(p(x)y')'(x) + q(x)y(x) = f(x) \qquad x \in (0,1)$$

avec conditions au bord périodique $y(0) = y(1)$, $y'(0) = y'(1)$ admet une unique solution faible $y \in H^1(\mathbb{T})$ i.e. pour tout $\varphi \in H^1(\mathbb{T})$

$$\int_0^1 p(x)y'(x)\varphi'(x) + q(x)y(x)\varphi(x)\,dx = \int_0^1 f(x)\varphi(x)\,dx.$$

Remarquons qu'une solution faible est une solution au sens des distributions (pour cela, on commence par prolonger p, q, f par 0 en-dehors de $(0,1)$ et on prolonge y par $y(0)$ sur $[-1,0]$ et sur $[1,2]$ puis sur $[-2,-1]$ et $[1,2]$ on la relie de façon $\mathcal{C}^\infty$ à 0 et ensuite on la prolonge par 0.)

Si f, q sont continues et p est $\mathcal{C}^1$, comme $y \in H^1(\mathbb{T}) \subset \mathcal{C}(\mathbb{T})$ on commence par remarquer que la distribution $(p(x)y')' = qy - f$ est en fait une fonction continue, donc py' est une fonction $\mathcal{C}^1$ donc en fait, comme p est $\mathcal{C}^1$ et $p \geq \alpha$, y' est $\mathcal{C}^1$ donc y est $\mathcal{C}^2$. Ainsi y est une solution forte.

3.3. Application en géométrie: inégalité isopérimétrique. Dans cette section on considère un domaine $\mathcal{D}$ du plan délimité par une courbe simple fermée Γ de classe $\mathcal{C}^1$.

Le but de cette section est de démontrer le résultat suivant:

THÉORÈME 3.4 (Inégalité isopérimétrique). *Soit $A(\mathcal{D})$ l'aire de $\mathcal{D}$ et $\ell(\Gamma)$ la longueur de Γ. Alors*

$$A(\mathcal{D}) \leq \frac{\ell(\Gamma)^2}{4\pi}$$

et on a égalité si et seulement si Γ est un cercle et A un disque.

Avant de démontrer le théorème, il nous faut préciser les quantités qui apparaissent.

Lorsque $\mathcal{D}$ est le disque de rayon $R > 0$, son aire est $A(\mathcal{D}) = \pi R^2$ alors que son périmètre est $\ell(\Gamma) = 2\pi R$. On trouve bien $\ell(\Gamma)^2 = 4\pi^2 R^2 = 4\pi A(\mathcal{D})$. Lorsque $\mathcal{D}$ est un carré de côté $R > 0$, (une courbe qui n'est que $\mathcal{C}^1$ par morceaux, mais cela ne change rien à l'inégalité isopérimétrique) son aire est $A(\mathcal{D}) = R^2$ alors que son périmètre est $\ell(\Gamma) = 4R$. On a bien $R^2 < 4R^2/\pi$.

De façon générale, la longueur d'une courbe paramétrée par

$$\gamma(t) = \big(x(t), y(t)\big), \qquad t \in [a,b]$$

est donnée par

$$\ell(\Gamma) = \int_a^b \|\gamma'(t)\| \, \mathrm{d}t = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} \, \mathrm{d}t.$$

Avec la formule de Green,

$$\int_\Gamma P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y = \int_\mathcal{D} \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \, \mathrm{d}x\,\mathrm{d}y$$

l'aire de $\mathcal{D}$ est donnée par

$$A(\mathcal{D}) = \int_\mathcal{D} \mathrm{d}x\,\mathrm{d}x = \frac{1}{2}\int_\Gamma x\,\mathrm{d}y - y\,\mathrm{d}x = \frac{1}{2}\int_a^b x(t)y'(t) - y(t)x'(t)\,\mathrm{d}t$$

où on prend $P(x,y) = -y/2$, $Q(x,y) = x/2$.

DÉMONSTRATION. Commençons par simplifier un peu le problème. Quitte à dilater la figure, $\big(x(t), y(t)\big) \to \big(\lambda x(t), \lambda y(t)\big)$, on peut supposer que $\ell(\Gamma) = 1$. Ensuite, on peut supposer qu'on parcourt la courbe à vitesse constant 1 en commençant à l'instant $t = 0$. Ceci est un peu plus subtile. En effet, on considère

$$\ell(s) = \int_a^s \sqrt{x'(t)^2 + y'(t)^2} \, \mathrm{d}t$$

et on remarque que ℓ est strictement croissante. Plus précisément, si ℓ est croissante, donc si elle n'est pas strictement croissante, elle est constante sur une intervalle $[\alpha, \beta] \subset [a,b]$ c'est-à-dire $x'(t) = y'(t) = 0$ sur $[\alpha, \beta]$ et $\big(x(t), y(t)\big)$ est constante sur cet intervalle. Mais alors on supprime cet intervalle de la paramétrisation.

Maintenant que ℓ est strictement croissante, elle est bijective $[a,b] \to [0,1]$. Comme ℓ est de classe $\mathcal{C}^1$, son inverse aussi, sauf aux points stationnaires.

Pour $s \in [0,1]$, on pose $\tilde{x}(s) = x\big(\ell^{-1}(s)\big)$ et $\tilde{y}(s) = y\big(\ell^{-1}(s)\big)$. On remarque que $\tilde{x}'(s) = \dfrac{x'\big(\ell^{-1}(s)\big)}{\ell'\big(\ell^{-1}(s)\big)}$ et $\tilde{y}'(s) = \dfrac{y'\big(\ell^{-1}(s)\big)}{\ell'\big(\ell^{-1}(s)\big)}$. Mais alors

$$\tilde{x}'(s)^2 + \tilde{y}'(s)^2 = \frac{x'\big(\ell^{-1}(s)\big)^2 + y'\big(\ell^{-1}(s)\big)^2}{\ell'\big(\ell^{-1}(s)\big)^2)} = 1.$$

puisque $\ell'(t) = \sqrt{x'(t)^2 + y'(t)^2}$.

On suppose donc à partir de maintenant que la courbe est paramétrée par $\big(x(t), y(t)\big)$, $t \in [0,1]$ avec $x'(s)^2 + y'(s)^2 = 1$.

Nous allons maintenant reformuler tout cela en termes de séries de Fourier. Écrivons les séries de Fourier de x et y

$$x(t) = \sum_{k\in\mathbb{Z}} c_k e^{2i\pi kt} \quad \text{et} \quad y(t) = \sum_{k\in\mathbb{Z}} c_k e^{2i\pi kt}$$

et ces séries convergent uniformément, donc dans $L^2(\mathbb{T})$, alors que

$$x'(t) = \sum_{k\in\mathbb{Z}} 2i\pi k c_k e^{2i\pi kt} \quad \text{et} \quad y'(t) = \sum_{k\in\mathbb{Z}} 2i\pi k d_k e^{2i\pi kt}$$

et ces séries convergent dans $L^2(\mathbb{T})$. De plus x, y étant à valeurs réelles, par unicité de Fourier, on voit immédiatement que $c_{-k} = \overline{c_k}$, $d_{-k} = \overline{d_k}$.

Par ailleurs, Parseval implique que

$$\sum_{k\in\mathbb{Z}} 4\pi^2 k^2 (|c_k|^2 + |d_k|^2) = \int_0^1 x'(s)^2 + y'(s)^2 \,\mathrm{d}s = \int_0^1 1 \,\mathrm{d}s = 1.$$

alors que

$$\begin{aligned} A(\mathcal{D}) &= \frac{1}{2}\int_0^1 x(t)y'(t) - y(t)x'(t)\,\mathrm{d}t = \frac{1}{2}\int_0^1 x(t)\overline{y'(t)} - y(t)\overline{x'(t)}\,\mathrm{d}t \\ &= \frac{1}{2}\sum_{k\in\mathbb{Z}} c_k \overline{2i\pi k d_k} - d_k \overline{2i\pi k c_k} = \pi \sum_{k\in\mathbb{Z}} k\big(d_k\overline{c_k} - c_k\overline{d_k}\big). \end{aligned}$$

On utilise ensuite le fait que

$$|d_k\overline{c_k} - c_k\overline{d_k}| \leq 2|c_k||d_k| \leq |c_k|^2 + |d_k|^2$$

d'où

$$\begin{aligned} A(\mathcal{D}) &\leq \pi \sum_{k\in\mathbb{Z}} |k||d_k\overline{c_k} - c_k\overline{d_k}| \leq \pi \sum_{k\in\mathbb{Z}} |k|\big(|c_k|^2 + |d_k|^2\big) \\ &\leq \frac{1}{4\pi}\sum_{k\in\mathbb{Z}} 4\pi^2|k|^2\big(|c_k|^2 + |d_k|^2\big) \\ &= \frac{1}{4\pi} = \frac{1}{4\pi}\ell(\Gamma). \end{aligned}$$

Regardons maintenant ce qui se passe pour que $A(\mathcal{D}) = \frac{1}{4\pi}\ell(\Gamma)$. Dans ce cas, la dernière inégalité est une inégalité, à savoir, pour $k \in \mathbb{Z}$,

$$\pi|k|^2\big(|c_k|^2 + |d_k|^2\big) = |k|\big(|c_k|^2 + |d_k|^2\big)$$

donc $|c_k|^2 + |d_k|^2 = 0$ si $k \notin \{-1, 0, 1\}$ soit $c_k = d_k = 0$ pour $|k| \geq 2$. Il en résulte que $x(t) = c_0 + c_1 e^{2i\pi kt} + c_{-1}e^{2i\pi kt} = c_0 + 2\Re(c_1 e^{2i\pi kt})$ et de même $y(t) = d_0 + 2\Re(d_1 e^{2i\pi kt})$.

Ensuite, en remontant, on a utilisé le fait que $2|c_k||d_k| \leq |c_k|^2 + |d_k|^2$. Donc, si on veut une égalité, il faut $2|c_k||d_k| = |c_k|^2 + |d_k|^2$ soit $|c_k| = |d_k|$. On écrit donc $d_1 = c_1 e^{i\theta}$ avec θ réel.

Enfin, la dernière inégalité utilisée est $|d_k\overline{c_k} - c_k\overline{d_k}| \leq 2|c_k||d_k|$. Il faut à nouveau avoir une égalité, (qui est évidente pour $|k| \geq 2$ puisque c'est $0 = 0$). Ainsi, si $d_1 = c_1 e^{i\theta}$ ceci se réduit à

$$|e^{i\theta} - e^{-i\theta}| = 2$$

soit $|\sin\theta| = 1$ et $\theta = \pm\pi/2$. Au final, on trouve

$$x(t) = c_0 + 2\cos\theta \quad \text{et} \quad y(t) = d_0 + 2\cos\Big(\theta \pm \frac{\pi}{2}\Big)$$

qui est bien la pramétrisation d'un cercle. □

3.4. Application en théorie des nombres: équirépartition. Une application simple de la classification des sous-groupes de $\mathbb{R}$ (ils sont soit denses, soit de la forme $a\mathbb{Z}$) est la suivante: pour $x \in \mathbb{R}$, on note $[x]$ la partie entière de x et $\{x\} = x - [x]$ la partie fractionnaire de x. Si $\alpha \notin \mathbb{Q}$, $(\{k\alpha\})_{k\in\mathbb{Z}}$ est dense dans $[0,1]$. Le but de cette section est de donner une version plus forte de ce résultat

THÉORÈME 3.5 (Weyl). *Si $\alpha \notin \mathbb{Q}$, $(\{k\alpha\})_{k\in\mathbb{Z}}$ est équi-répartie dans $[0,1]$: pour tous $0 \leq< a < b \leq 1$,*

$$\lim_{N\to+\infty} \frac{|\{k \in \{0,\ldots,N-1\} : a \leq \{k\alpha\} \leq b\}|}{N} = b - a. \tag{3.12}$$

En particulier, tout intervalle $[a,b]$ avec $b - a > 0$ contient une infinité de termes de la suite $(\{k\alpha\})_{k\in\mathbb{Z}}$ qui est donc dense dans $[0,1]$.

DÉMONSTRATION. Commençons par ré-écrire (3.12) sous la forme suivante : soit $f = \mathbf{1}_{[a,b]}$, alors

$$\lim_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) = \int_0^1 f(t)\,\mathrm{d}t. \tag{3.13}$$

Notons que si $b = 1$ et $a = 0$, cette identité est triviale, on peut donc supposer dans la suite que $b-a < 1$.

Prolongeons maintenant f en une fonction 1-périodique sur $\mathbb{R}$. Soit $0 < \varepsilon < \min(1-(b-a), b-a)$ et soient $f_1, f_2 \in \mathcal{C}(\mathbb{T})$ telles que $f_1 \leq f \leq f_2$ et

$$\int_0^1 f_1(t)\,\mathrm{d}t \geq b - a - \varepsilon \quad \text{alors que} \quad \int_0^1 f_2(t)\,\mathrm{d}t \leq b - a + \varepsilon.$$

Le plus simple est de prendre f_1, continue 1-périodiques et telles que $f_1 = f = 1$ sur $[a+\varepsilon/2, b-\varepsilon/2]$, $f_1 = 0$ sur $[0,1]\setminus[a,b]$ et f_1 affine sur $[a, a+\varepsilon/2]$ tout comme sur $[b-\varepsilon/2, b]$. Pour f_2, on pose pour $j \in \mathbb{Z}$, $f_2 = 1$ sur $[a+j, b+j]$ $f_2(a-\varepsilon/2+j) = f_2(b+\varepsilon/2+j) = 0$ et f_2 affine sur $[a-\varepsilon/2+j, a+j]$ tout comme sur $[b+j, b+\varepsilon/2+j]$ et enfin $f_2 = 0$ là où elle n'a pas encore été définie. Le lecteur dessinera f_1 et f_2 et remarquera qu'elles remplissent les critères du théorème de Dirichlet pour avoir convergence normale de leur série de Fourier.

Supposons maintenant qu'on sache démontrer (3.13) dès que f est une fonction continue, dont la série de Fourier converge normalement, alors

$$\limsup_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) \leq \limsup_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f_2(\{k\alpha\}) = \int_0^1 f_2(t)\,\mathrm{d}t \leq b - a + \varepsilon.$$

De même

$$\liminf_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) \geq \liminf_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f_1(\{k\alpha\}) = \int_0^1 f_1(t)\,\mathrm{d}t \geq b - a - \varepsilon.$$

En résumé

$$b - a - \varepsilon \leq \liminf_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) \leq \limsup_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) \leq b - a + \varepsilon.$$

Comme ε est arbitraire, on en déduit que

$$\liminf_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) = \limsup_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) = b - a$$

qui implique (3.12).

Il reste donc à démontrer (3.13) pour les fonctions continues, dont la série de Fourier converge normalement. Réduisons encore le problème. Supposons qu'on sache montrer que

$$\lim_{N\to+\infty} \frac{1}{N}\sum_{k=0}^{N-1} P(\{k\alpha\}) = \int_0^1 P(t)\,\mathrm{d}t \tag{3.14}$$

pour tout polynôme trigonométrique. Soit alors $\varepsilon > 0$ et supposons que f soit tel que sa série de Fourier converge normalement (nécessairement vers f). Il existe donc un polynôme trigonométrique P tel que $\|f - P\|_\infty \leq \varepsilon$. Mais alors, d'une part, il existe $N > 0$ tel que

$$\left|\frac{1}{N}\sum_{k=0}^{N-1} P(\{k\alpha\}) - \int_0^1 P(t)\,\mathrm{d}t\right| \leq \varepsilon$$

et d'autre part

$$\begin{aligned}\left|\frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\}) - \frac{1}{N}\sum_{k=0}^{N-1} P(\{k\alpha\})\right| &\leq \frac{1}{N}\sum_{k=0}^{N-1} |f(\{k\alpha\}) - P(\{k\alpha\})| \\ &\leq \|f-P\|_\infty \leq \varepsilon\end{aligned}$$

ainsi que

$$\left|\int_0^1 f(t)\,\mathrm{d}t - \int_0^1 P(t)\,\mathrm{d}t\right| \leq \int_0^1 |f(t) - P(t)|\,\mathrm{d}t \leq \|f-P\|_\infty \leq \varepsilon.$$

Mais alors

$$\begin{aligned}\left|\int_0^1 f(t)\,\mathrm{d}t - \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\})\right| &\leq \left|\int_0^1 f(t)\,\mathrm{d}t - \int_0^1 P(t)\,\mathrm{d}t\right| \\ &+ \left|\int_0^1 P(t)\,\mathrm{d}t - \frac{1}{N}\sum_{k=0}^{N-1} P(\{k\alpha\})\right| + \left|\frac{1}{N}\sum_{k=0}^{N-1} P(\{k\alpha\}) - \frac{1}{N}\sum_{k=0}^{N-1} f(\{k\alpha\})\right| \\ &\leq 3\varepsilon.\end{aligned}$$

ce qui montre bien (3.13). Il reste donc à montrer (3.14). Notons que cette identité est linéaire, il suffit donc de la montrer pour $P(t) = e^{2i\pi nt}$, $n \in \mathbb{Z}$. De plus, elle est trivialement vérifiée pour $n = 0$, *i.e.* pour $P(t) = 1$. Soit donc $n \in \mathbb{Z}$, $n \neq 0$ et remarquons que

$$\int_0^1 e^{2i\pi nt}\,\mathrm{d}t = 0.$$

Par ailleurs, comme $\{k\alpha\} = k\alpha - [k\alpha]$,

$$e^{2i\pi n\{k\alpha\}} = e^{2i\pi nk\alpha}e^{2i\pi n[k\alpha]} = e^{2i\pi nk\alpha}$$

puisque $n[k\alpha] \in \mathbb{Z}$. Ainsi, comme $\alpha \notin \mathbb{Q}$, $e^{2i\pi n\alpha} \neq 1$ donc

$$\frac{1}{N}\sum_{k=0}^{N-1} e^{2i\pi n\{k\alpha\}} = \frac{1}{N}\sum_{k=0}^{N-1} e^{2i\pi nk\alpha} = \frac{1}{N}\frac{1 - e^{2i\pi nN\alpha}}{1 - e^{2i\pi n\alpha}} \to 0$$

quand $N \to +\infty$ puisque $|1 - e^{2i\pi nN\alpha}| \leq 2$. On vient donc de démontrer (3.14) donc le théorème. □

CHAPTER 3

Opérateurs bornés, opérateurs compacts

1. Opérateurs bornés, critère de Schur, adjoint

Commençons par rappeler le fait fondamental suivant:

Théorème 1.1. *Soit X, Y deux espaces de Banach et $T : X \to Y$ un opérateur linéaire. On a équivalence entre*

1. *T est continu sur X;*
2. *T est continu en 0;*
3. *T est borné sur la boule unité;*
4. *Il existe $C > 0$ tel que, pour tout $x \in X$, $\|Tx\|_Y \leq C\|x\|_X$.*

En particulier, si X est de dimension finie, toutes les applications linéaires sur X sont continues.

On notera alors $\|T\|_{X\to Y} = \inf\{C > 0 \,:\, \forall x \in X,\ \|Tx\|_Y \leq C\|x\|_X\}$ ou plus simplement $\|T\|$ s'il n'y a aucune ambiguité possible. D'autres expressions sont possible

$$\|T\| = \sup\{\|Tx\| \,:\, \|x\| \leq 1\} = \sup\{\|Tx\| \,:\, \|x\| = 1\}$$

et on vérifie que c'est une norme sur $\mathcal{B}(X, Y)$ l'espace des applications linéaires continues (on dit aussi bornées) de X vers Y. On a vu qu'on obtient ainsi un espace de Banach.

Le cas qui nous intéresse le plus souvent est celui d'opérateurs $L^p(\Omega, \mu) \to L^q(\tilde{\Omega}, \nu)$ donnés par un noyau K:

$$T_K f(x) = \int_\Omega K(x, y) f(y) \,\mathrm{d}\mu(y). \tag{1.15}$$

En effet, cela couvre les cas suivants

1. $\Omega, \tilde{\Omega}$ sont des ensembles finis et μ, ν les mesures de comptage. Dans ce cas $L^p(\Omega, \mu) = (\mathbb{C}^m, \|\cdot\|_p)$ où $m = |\Omega|$ (le cardinal de Ω) et $L^q(\tilde{\Omega}, \nu) = (\mathbb{C}^n, \|\cdot\|_q)$ où $n = |\tilde{\Omega}|$. Dans ce cas, (1.15) se lit
$$T_K f(i) = \sum_{j=1}^m K(i, j) f(j)$$
et K est simplement la matrice de T (dans la base canonique).
2. $\Omega = \tilde{\Omega} = \mathbb{R}^d$ muni de la mesure de Lebesgue et K est de la forme $K(x, y) = k(x - y)$. Dans ce cas
$$T_K f(x) = \int_{\mathbb{R}^d} k(x - y) f(y) \,\mathrm{d}\mu(y) = k * f(x).$$
3. $\Omega = \tilde{\Omega} = \mathbb{R}^d$ muni de la mesure de Lebesgue et $K(x, y) = e^{-2i\pi\langle x, y\rangle}$ et dans ce cas, on on $T_K f = \widehat{f}$ est la transformée de Fourier.

Il est assez facile de trouver une condition sur K pour que T_K soit borné de $L^1 \to L^q$. En effet, l'inégalité de Minkovski montre que

$$\begin{aligned}
\left\|\int_\Omega K(x,y)f(y)\,\mathrm{d}y\right\|_{L^q(\tilde\Omega,\nu)} &\leq \int_\Omega \|K(x,y)f(y)\|_{L^q_x(\tilde\Omega,\nu)}\,\mathrm{d}\mu(y)\\
&= \int_\Omega \|K(x,y)\|_{L^q_x(\tilde\Omega,\nu)}|f(y)|\,\mathrm{d}\mu(y)\\
&\leq \sup_{y\in\Omega}\|K(x,y)\|_{L^q_x(\tilde\Omega,\nu)}\int_\Omega |f(y)|\,\mathrm{d}\mu(y)
\end{aligned}$$

(ici et dans tout le chapitre, sup signifie le sup essentiel). On a ainsi montré la proposition suivante:

Proposition 1.2. *Soient (Ω,μ), $(\tilde\Omega,\nu)$ deux espaces mesurés σ-finis et soit $1\leq q\leq\infty$. Soit $K:\tilde\Omega\times\Omega\to\mathbb{C}$. Supposons que*

$$M := \begin{cases} \displaystyle\sup_{y\in\Omega}\left(\int_{\tilde\Omega}|K(x,y)|^q\,\mathrm{d}\nu(x)\right)^{1/q} & \textit{si } q<+\infty\\ \sup_{(x,y)\in\Omega\times\tilde\Omega}|K(x,y)| & \textit{sinon}\end{cases} < +\infty.$$

Alors l'opérateur linéaire défini par

$$T_K f(x) = \int_\Omega K(x,y)f(y)\,\mathrm{d}\mu(y)$$

est borné $L^1(\Omega,\mu)\to L^q(\tilde\Omega,\nu)$ avec $\|T_K\|_{L^1(\Omega,\mu)\to L^q(\tilde\Omega,\nu)}\leq M$.

La situation duale, c'est-à-dire arriver dans L^∞, n'est pas plus compliquée: on prend $1\leq p\leq+\infty$ et p' tel que $\dfrac{1}{p}+\dfrac{1}{p'}=1$, l'exposant dual. Alors l'inégalité de Hölder nous donne

$$\sup_{x\in\tilde\Omega}\left|\int_\Omega K(x,y)f(y)\,\mathrm{d}\mu(y)\right| \leq \sup_{x\in\tilde\Omega}\|K(x,y\|_{L^{p'}_y(\Omega,\mu)}\|f\|_{L^p(\Omega,\mu)}.$$

On a ainsi obtenu le résultat suivant:

Proposition 1.3. *Soient (Ω,μ), $(\tilde\Omega,\nu)$ deux espaces mesurés σ-finis et soit $1\leq p\leq\infty$ et p' tel que $\dfrac{1}{p}+\dfrac{1}{p'}=1$, l'exposant dual. Soit $K:\Omega\times\tilde\Omega\to\mathbb{C}$. Supposons que*

$$M := \begin{cases} \displaystyle\sup_{x\in\tilde\Omega}\left(\int_{\Omega}|K(x,y)|^{p'}\,\mathrm{d}\mu(y)\right)^{1/p'} & \textit{si } p>1\\ \sup_{(x,y)\in\Omega\times\tilde\Omega}|K(x,y)| & \textit{si } p=1\end{cases} < +\infty.$$

Alors l'opérateur linéaire défini par

$$T_K f(x) = \int_\Omega K(x,y)f(y)\,\mathrm{d}\mu(y)$$

est borné $L^p(\Omega,\mu)\to L^\infty(\tilde\Omega,\nu)$ avec $\|T_K\|_{L^p(\Omega,\mu)\to L^\infty(\tilde\Omega,\nu)}\leq M$.

Pour le cas $L^p\to L^p$, on dispose d'un outil puissant (qui lui-même a été étendu de diverses façons que nous ne traiterons pas ici):

Théorème 1.4 (Test de Schur). *Soient (Ω,μ), $(\tilde\Omega,\nu)$ deux espaces mesurés σ-finis et soit $1\leq p\leq\infty$. Soit $K:\Omega\times\tilde\Omega\to\mathbb{C}$. Supposons que*

$$A := \sup_{x\in\tilde\Omega}\int_\Omega |K(x,y)|^p\,\mathrm{d}\mu(y) < +\infty$$

et

$$B := \sup_{y\in\Omega}\int_{\tilde\Omega} |K(x,y)|^p\,\mathrm{d}\nu(x) < +\infty$$

Alors l'opérateur linéaire défini par

$$T_K f(x) = \int_\Omega K(x,y) f(y)\,\mathrm{d}y$$

est borné $L^p(\Omega,\mu) \to L^p(\tilde\Omega,\nu)$ *avec* $\|T_K\|_{L^p(\Omega,\mu)\to L^p(\tilde\Omega,\nu)} \le A^{1/p'}B^{1/p}$. *Ici on prend la convention* $X^{1/\infty} = 1$.

Proof. Notons que si $A = 0$ ou $B = 0$ alors $K = 0$ presque partout et $T_K = 0$. On exclus donc ces cas dans la suite.

Les cas $p = 1$ et $p = \infty$ ont déjà été traités, on suppose donc que $1 < p < +\infty$. Nous prenons alors p' tel que $\dfrac{1}{p} + \dfrac{1}{p'} = 1$, l'exposant dual de p. Comme au premier semestre, il s'agit de montrer que, pour tout $f \in L^p(\Omega,\mu)$ et tout $g \in L^{p'}(\tilde\Omega,\nu)$

$$\left|\int_{\tilde\Omega} T_K f(x) g(x)\,\mathrm{d}\nu(y)\right| \le C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\nu)}$$

avec C à déterminer. En effet, cela implique que

$$\begin{aligned}\|T_K f\|_{L^p(\nu)} &= \sup_{\|g\|_{L^{p'}(\tilde\Omega,\nu)}=1} \left|\int_{\tilde\Omega} T_K f(x) g(x)\,\mathrm{d}\nu(y)\right| \\ &\le \sup_{\|g\|_{L^{p'}(\tilde\Omega,\nu)}=1} C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\nu)} = C\|f\|_{L^p(\mu)}.\end{aligned}$$

Ensuite, remarquons que

$$\left|\int_{\tilde\Omega} T_K f(x) g(x)\,\mathrm{d}\nu(y)\right| \le \int_{\tilde\Omega} |T_K f(x)|\,|g(x)|\,\mathrm{d}\nu(y)$$

et que

$$|T_K f(x)| = \left|\int_\Omega K(x,y) f(y)\,\mathrm{d}\mu(y)\right| \le \int_\Omega |K(x,y)|\,|f(y)|\,\mathrm{d}\mu(y).$$

Mais alors, avec Fubini,

$$\int_{\tilde\Omega} |T_K f(x)|\,|g(x)|\,\mathrm{d}\nu(y) = \int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x)$$

et l'ordre d'intégration n'importe pas dans cette intégrale double. Il suffit donc de montrer que

$$\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x) \le C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\nu)}. \tag{1.16}$$

L'étape d'après consiste à se ramener au cas $A = B = 1$. Pour cela, observons que si on pose $\mathrm{d}\tilde\mu = \dfrac{1}{A}\mathrm{d}\mu$ alors

$$\sup_{x\in\tilde\Omega}\int_\Omega |K(x,y)|^p\,\mathrm{d}\tilde\mu(y) = 1$$

tandis que

$$\|f\|_{L^p(\tilde\mu)} = \left(\int_\Omega |f(x)|^p\,\frac{1}{A}\mathrm{d}\mu(x)\right)^{1/p} = A^{-1/p}\|f\|_{L^p(\mu)}.$$

De même, si on pose $\mathrm{d}\tilde\nu = \dfrac{1}{B}\mathrm{d}\nu$ alors

$$\sup_{y\in\Omega}\int_{\tilde\Omega} |K(x,y)|^p\,\mathrm{d}\tilde\nu(x) = 1$$

alors que $\|g\|_{L^{p'}(\tilde\nu)} = B^{-1/p'}\|g\|_{L^{p'}(\nu)}$. Enfin

$$\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\tilde\mu(y)\,\mathrm{d}\tilde\nu(x) = A^{-1}B^{-1}\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x).$$

Ainsi, si on montre

$$\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\tilde\mu(y)\,\mathrm{d}\tilde\nu(x) \le \|f\|_{L^p(\tilde\mu)}\|g\|_{L^{p'}(\tilde\nu)} \tag{1.17}$$

alors

$$A^{-1}B^{-1}\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x) \le A^{-1/p}\|f\|_{L^p(\mu)}B^{-1/p'}\|g\|_{L^{p'}(\nu)}$$

qui est bien (1.16) avec $C = A^{1-1/p}B^{1-1/p'} = A^{1/p'}B^{1/p}$. En d'autre termes, on s'est ramené au cas $A = B = 1$ et on va maintenant démontrer (1.17) dans ce cas. Enfin, on remarque que (1.17) est homogène en f et en g, c'est-à-dire que, si elle est valable pour un couple (f, g) elle est aussi valable pour $(\lambda f, \mu g)$. En particulier, on remplaçant f et g par $f/\|f\|_{L^p(\tilde\mu)}$ et $g/\|g\|_{L^{p'}(\tilde\nu)}$, on se ramène à vouloir démontrer que

$$\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x) \le 1$$

si $\|f\|_{L^p(\tilde\mu)} = \|g\|_{L^{p'}(\tilde\nu)} = 1$ et

$$\sup_{x\in\tilde\Omega}\int_\Omega |K(x,y)|^p\,\mathrm{d}\mu(y) = \sup_{y\in\Omega}\int_{\tilde\Omega} |K(x,y)|^p\,\mathrm{d}\nu(x) = 1.$$

Pour cela, on utilise la convexité de l'exponentielle qui nous dit que, si $a, b > 0$, alors

$$e^{\frac1p\ln a + \frac1{p'}\ln b} \le \frac1p e^{\ln a} + \frac1{p'}e^{\ln b}$$

i.e. $a^{1/p}b^{1/p'} \le \dfrac1p a + \dfrac1{p'}b$ ou encore, pour tous $u, v > 0$

$$uv \le \frac1p u^p + \frac1{p'}v^{p'}.$$

En remarquant que cette inégalité est triviale si $a = 0$ ou $b = 0$, on en déduit que

$$|f(x)g(x)| \le \frac1p|f(x)|^p + \frac1{p'}|g(x)|^{p'}.$$

Il en résulte que

$$\begin{aligned}
&\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x)\\
&\qquad\le \frac1p\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|f(x)|^p\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x)\\
&\qquad\qquad + \frac1{p'}\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,|g(y)|^{p'}\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x)\\
&\qquad= \frac1p\int_{\tilde\Omega}\left(\int_\Omega |K(x,y)|\,\mathrm{d}\mu(y)\right)|f(x)|^p\,\mathrm{d}\nu(x)\\
&\qquad\qquad + \frac1{p'}\int_\Omega\left(\int_{\tilde\Omega} |K(x,y)|\,\mathrm{d}\nu(x)\right)|g(y)|^{p'}\,\mathrm{d}\mu(y)
\end{aligned}$$

avec Fubini (puisqu'on n'a que des termes positifs). La première intégrale est majorée par

$$\begin{aligned}\int_{\tilde\Omega}\left(\int_\Omega |K(x,y)|\,\mathrm{d}\mu(y)\right)\,|f(x)|^p\,\mathrm{d}\nu(x)&\\ \leq \left(\sup_{x\in\tilde\Omega}\int_\Omega |K(x,y)|\,\mathrm{d}\mu(y)\right)\int_{\tilde\Omega}|f(x)|^p\,\mathrm{d}\nu(x)&\\ =\int_{\tilde\Omega}|f(x)|^p\,\mathrm{d}\nu(x)=\|f\|_{L^p(\nu)}^p=1&\end{aligned}$$

alors que pour la seconde

$$\begin{aligned}\int_{\Omega}\left(\int_{\tilde\Omega} |K(x,y)|\,\mathrm{d}\nu(x)\right)\,|g(y)|^{p'}\,\mathrm{d}\mu(y)&\\ \leq \left(\sup_{y\in\Omega}\int_\Omega\int_{\tilde\Omega} |K(x,y)|\,\mathrm{d}\nu(x)\right)\int_{\Omega}|g(y)|^{p'}\,\mathrm{d}\mu(y)&\\ =\int_{\Omega}|g(y)|^{p'}\,\mathrm{d}\mu(y)=\|g\|_{L^{p'}(\nu)}^p=1.&\end{aligned}$$

Au final

$$\int_\Omega\int_{\tilde\Omega}|K(x,y)|\,|f(x)|\,|g(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x)\leq\frac{1}{p}+\frac{1}{p'}=1$$

comme voulu. $\square$

Remarque 1.5. Le test de Schur est parfois optimal. Ainsi, si μ,ν sont des mesures finies, $K\geq 0$ et que les conditions du test de Schur sont exacte au sens où

$$\int_\Omega K(x,y)\,\mathrm{d}\mu(y)=A\quad\text{et}\quad\int_{\tilde\Omega}K(x,y)\,\mathrm{d}\nu(x)=B$$

alors, en intégrant la première par rapport à ν et la seconde par rapport à μ, on trouve $A\nu(\tilde\Omega)=B\mu(\Omega)$. Remarquons que la première identité se lit $T_K\mathbf{1}_\Omega=A\mathbf{1}_{\tilde\Omega}$ mais $\|\mathbf{1}_\Omega\|_p=\mu(\Omega)^{1/p}$ et $\|\mathbf{1}_{\tilde\Omega}\|_{p'}=\nu(\tilde\Omega)^{1/p'}$. Ainsi on a $\|T_K\mathbf{1}_\Omega\|_{p'}=A^{1/p'}B^{1/p}\|\mathbf{1}_\Omega\|_p$.

À l'opposé, si K oscille, on ne s'attend pas à ce que le test de Schur soit optimal. Par exemple, si $\Omega=\tilde\Omega=\mathbb{R}^d$ et $K(x,y)=e^{-2i\pi\langle x,y\rangle}$ alors $T_Kf=\widehat{f}$, la transformée de Fourier de f. On sait avec Plancherel que $\|T_Kf\|_2=\|f\|_2$ alors que $A=B=+\infty$ dans le test de Schur.

2. Adjoint d'un opérateur

Soient X,Y deux espaces de Banach et X',Y' leurs espaces duaux. Pour $x\in X$ et $x'\in X'$, il sera plus pratique de noter $x'(x)=\langle x',x\rangle$. Rappelons que $\{\langle x',x\rangle,x'\in X'\}$ détermine de façon unique x.

Soit T un opérateur $X\to Y$ et $y'\in Y'$, on peut alors définir $x\to\langle y',Tx\rangle$ et on remarque que ceci définit une forme linéaire continue sur X:

$$|\langle y',Tx\rangle|\leq\|y'\|_{Y'}\|Tx\|_Y\leq\|T\|_{X\to Y}\|y'\|_{Y'}\|x\|_X.$$

On va noter cette forme linéaire $T'y'\in X'$. On remarque alors que $y'\to T'y'$ est linéaire puisque, pour tout x, $y'\to\langle y',Tx\rangle$ est linéaire et de plus, cette application est continue $Y'\to X'$ puisque

$$\|T'y'\|_{X'}=\sup_{x\in X,\|x\|_X=1}|T'y'(x)|=\sup_{\|x\|_X=1}|\langle y',Tx\rangle|\leq\|T\|_{X\to Y}\|y'\|_{Y'}.$$

Donc $\|T'\|_{Y'\to X'}\leq\|T\|_{X\to Y}$. Enfin, comme $\sup_{\|y'\|_{Y'}=1}|\langle y',Tx\rangle|=\|Tx\|$ on a

$$\sup_{\|x\|_X=1}\sup_{\|y'\|_{Y'}=1}|\langle y',Tx\rangle|=\sup_{\|x\|_X=1}\|Tx\|=\|T\|$$

donc $\|T'\|=\|T\|$.

Définition 2.1. Soient X, Y deux espaces de Banach et X', Y' leurs espaces duaux. Pour tout opérateur borné $T : X \to Y$, il existe un unique opérateur borné $T^* : Y' \to X'$ appelé l'adjoint de T tel que, pour tout $x \in X$ et tout $y' \in Y'$

$$\langle y', Tx\rangle = \langle T^*y', x\rangle.$$

De plus, $\|T'\|_{Y'\to X'} = \|T\|_{X\to Y}$.

Notons que si $X = Y = H$ un espace de Hilbert, la dualité est donnée par le produit scalaire et l'adjoint se définit de façon un peu plus directe:

Définition 2.2. Soit H un espace de Hilbert. Pour tout opérateur borné $T : H \to H$, il existe un unique opérateur borné $T^* : H \to H$ appelé l'adjoint de T tel que, pour tous $x, y \in H$

$$\langle y, Tx\rangle = \langle T^*y, x\rangle$$

ou, de façon équivalente, pour tous $x, y \in H$

$$\langle Tx, y\rangle = \langle x, T^*y\rangle.$$

Exemple 2.3. Soient $X = L^p(\mathbb{R}^d)$, $Y = L^q(\mathbb{R}^d)$ avec $1 < p, q < \infty$ et soient p', q' les exposants duaux $\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = 1$. Soit $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ une fonction localement bornée (*i.e.* pour tout $R > 0$, il existe C_R tel que $|K(x, y)| \le C_R$ si $|x|, |y| \le R$). On définit alors l'opérateur

$$T_K f(g) = \int_{\mathbb{R}^d} K(x, y)f(y)\,\mathrm{d}y \qquad f \in \mathcal{C}_c^\infty$$

et on suppose que T_K se prolonge en un opérateur continu $L^p(\Omega, \mu) \to L^q(\tilde\Omega, \nu)$. Alors T_K^* est continu $L^{q'}(\tilde\Omega, \mu) \to L^{p'}(\Omega, \nu)$ et

$$(T_K^*)g(y) = \int_{\mathbb{R}^d} K(x, y)g(x)\,\mathrm{d}x \qquad g \in \mathcal{C}_c^\infty.$$

Ainsi T_K^* est l'opérateur dont le noyau est $\mathcal{K}(x, y) = \overline{K(y, x)}$:

$$(T_K^*)g(x) = \int_{\mathbb{R}^d} \overline{K(y, x)}g(y)\,\mathrm{d}y \qquad g \in \mathcal{C}_c^\infty.$$

En effet, comme on sait que T_K et T_K^* sont continu, il suffit de vérifier $\langle T_K^*g, f\rangle = \langle g, T_K f\rangle$ pour $f, g \in \mathcal{C}_c^\infty$. Il faut ici prendre garde à l'identification d'un élément de $g \in L^{p'}$ comme élément du dual de L^p. En effet, on identifie g et la forme linéaire

$$f \to \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,\mathrm{d}x.$$

Ceci est cohérent avec le cas particulier de L^2. En effet, si X est un espace de Hilbert, la dualité est effectivement donnée par le produit scalaire (théorème de Riesz).

Mais alors

$$\begin{aligned}
\langle T_K^*g, f\rangle &= \langle g, T_K f\rangle := \int_{\mathbb{R}^d} \overline{g(x)}T_K f(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \overline{g(x)} \int_{\mathbb{R}^d} K(x, y)f(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K(x, y)\overline{g(x)}\,\mathrm{d}x\right) f(y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d} \left(\overline{\int_{\mathbb{R}^d} \overline{K(x, y)}g(x)\,\mathrm{d}x}\right) f(y)\,\mathrm{d}y = \langle T_{\mathcal{K}}g, f\rangle
\end{aligned}$$

avec Fubini qu'on peut appliquer car il existe $R > 0$ tel que $\operatorname{supp} f, \operatorname{supp} g \subset B(0, R)$ et alors

$$|K(x, y)g(x)f(y)| \le C_R\|f\|_\infty\|g\|_\infty \mathbf{1}_{B(0,R)\times B(0,R)}(x, y) \in L^1(\mathbb{R}^d \times \mathbb{R}^d).$$

Avant de poursuivre, rappelons que pour les opérateurs, on a trois modes de convergence:

(1) $T_n \to T$ fortement si $\|T_n - T\|_{X\to Y} \to 0$;

(2) $T_n \to T$ simplement si, pour tout $x \in X$, $T_n x \to Tx$, c'est-à-dire, pour tout $x \in X$, $\|T_n x - Tx\|_Y \to 0$;
(3) $T_n \to T$ faiblement si, pour tout $y \in Y'$, et tout $x \in X$, $\langle y, T_n x\rangle \to \langle y, Tx\rangle$.

Rappelons que si $X = H$ est un espace de Hilbert, il s'identifie à son dual via le produit scalaire et la convergence faible signifie: pour tous $x, y \in H$, $\langle y, T_n x\rangle \to \langle y, Tx\rangle$ ou, de façon équivalente, $\langle T_n x, y\rangle \to \langle Tx, y\rangle$. Évidemment la convergence forte implique la convergence simple qui implique la convergence faible.

Rappelons qu'on dit qu'un opérateur S sur un Hilbert H est positif si, pour tout $x \in H$, $\langle Sx, x\rangle \geq 0$. En particulier, $\langle Sx, x\rangle$ est réel donc

$$\langle S^* x, x\rangle = \overline{\langle x, S^* x\rangle} = \overline{\langle Sx, x\rangle} = \langle Sx, x\rangle \geq 0$$

donc S^* est également positif et alors l'opérateur auto-adjoint $A = \frac{S+S^*}{2}$ est positif. Notons aussi que si λ est une valeur propre d'un opérateur positif, alors $\lambda \geq 0$.

Les propriétés clés sont les suivantes

Proposition 2.4. *Pour H Hilbert, $S, T : H \to H$ linéaires bornés;*
(1) $\|T^*\| = \|T\|$;
(2) $(ST)^* = T^* S^*$ *et* $(T^*)^* = T$;
(3) *Si T est inversible, alors T^* aussi et* $(T^*)^{-1} = (T^{-1})^*$;
(4) *T^*T et TT^* sont des opérateurs positifs et*

$$\|T\| = \sqrt{\|TT^*\|} = \sqrt{\|T^*T\|};$$

(5) *Si $T_n \to T$ simplement, ou si $T_n \to T$ faiblement alors $T_n^* \to T^*$ faiblement.*
(6) $(\ker T)^\perp = \overline{\operatorname{Im} T^*}$ *et* $(\operatorname{Im} T)^\perp = \ker T^*$.
(7) *T est de rang fini si et seulement si T^* est de rang fini.*

Proof. La première a déjà été vue. Ensuite, pour tout $x, y \in H$

$$\langle y, STx\rangle = \langle y, S(Tx)\rangle = \langle S^* y, Tx\rangle = \langle T^* S^* y, x\rangle$$

et par unicité, $(ST)^* = T^* S^*$ puis

$$\langle (T^*)^* y, x\rangle = \langle y, T^* x\rangle = \overline{\langle T^* x, y\rangle} = \overline{\langle x, Ty\rangle} = \langle Ty, x\rangle$$

donc $(T^*)^* = T$.

(3): On a $TT^{-1} = T^{-1}T = I$, clairement $I^* = I$ donc $(T^{-1})^* T^* = T^*(T^{-1})^* = I$ ce qui montre que T^* est inversible (au sens algébrique), comme T^* est continue bijective, son inverse est continue et enfin $(T^{-1})^* = (T^*)^{-1}$.

(4): D'abord TT^* et T^*T sont positifs. En effet, on a $\langle TT^* x, x\rangle = \langle T^* x, T^* x\rangle = \|T^* x\|^2 \geq 0$ et de même $\langle T^* T x, x\rangle = \|Tx\|^2$. Appliquons l'inégalité de Cauchy-Schwarz à cette dernière identité, on obtient $\|Tx\|^2 \leq \|T^* Tx\|\|x\| \leq \|T^* T\|\|x\|^2$ et, en prenant le sup sur les x tels que $\|x\| \leq 1$, $\|T\|^2 \leq \|T^* T\|$. Inversement $\|T^* T\| \leq \|T^*\|\|T\| = \|T\|^2$.

(5): Découle directement de la définion. En effet, si $T_n \to T$ simplement alors, pour tout $x, y \in H$

$$\begin{aligned} |\langle T_n^* x, y\rangle - \langle T^* x, y\rangle| &= |\langle (T_n - T)^* x, y\rangle| = |\langle x, (T_n - T) y\rangle| \\ &\leq \|x\|\|T_n x - Tx\| \to 0 \end{aligned}$$

alors que si $T_n \to T$ faiblement, on a plus directement

$$\langle T_n^* x, y\rangle = \langle x, T_n y\rangle \to \langle x, Ty\rangle = \langle T^* x, y\rangle.$$

(6): $y \in (\operatorname{Im} T^*)^\perp$ signifie, pour tout $x \in H$, $0 = \langle y, T^* x\rangle = \langle Ty, x\rangle$ ce qui signifie $y \in \ker T$. En remplaçant T par T^* et en utilisant $(T^*)^* = T$ on trouve la deuxième identité.

(7): Si T est de rang fini, c'est-à-dire $\operatorname{Im} T$ est de dimension fini. En particulier, $\operatorname{Im} T$ est fermé donc $H = \ker T^* \oplus \operatorname{Im} T$. Ainsi, si $x \in H$, on peut écrire $x = a + b$ avec $a \in \ker T^*$ et $b \in \operatorname{Im} T$. Mais alors $T^* x = T^* b$, en d'autres termes, $\operatorname{Im} T^* = T^*(\operatorname{Im} T)$. Comme $\operatorname{Im} T$ est de dimension fini, $\operatorname{Im} T^*$ aussi. $\square$

Définition 2.5. Soit H un espace de Hilbert et $T : H \to H$ un opérateur borné sur H. On dira que

(i) T est une projection si $T^2 = T$;
(ii) T est auto-adjoint si $T^* = T$;
(iii) T est une projection orthogonale si T est une projection et que T est auto-adjoint;
(iv) T est normal si $T^*T = TT^*$;
(v) T est isométrique si $T^*T = I$;
(vi) T est unitaire si $T^*T = TT^* = I$ *i.e.* si T et T^* sont isométriques.

Commentons un peu ces définitions:
Tout d'abord, T^*T est auto-adjoint puisque $(T^*T)^* = T^*(T^*)^* = T^*T$.
Si T est isométrique, alors pour tout $x \in H$,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle x, x\rangle = \|x\|^2.$$

Inversement, si pour tout $x \in H$, $\|x\| = \|Tx\|$ alors $\langle T^*Tx, x\rangle = \langle x, x\rangle$ *i.e.* $\langle (T^*T - I)x, x\rangle = 0$. Soient alors $a, b \in H$ et $\lambda \in \mathbb{C}$, on trouve

$$\begin{aligned} 0 &= \langle (T^*T - I)(a + \lambda b), a + \lambda b\rangle \\ &= \langle (T^*T - I)a, a\rangle + \bar\lambda\langle (T^*T - I)a, b\rangle \\ &\quad +\lambda\langle (T^*T - I)b, a\rangle + |\lambda|^2\langle (T^*T - I)b, b\rangle. \end{aligned}$$

Mais, le premier et le quatrième terme sont nuls. Par ailleurs

$$\langle (T^*T - I)b, a\rangle = \overline{\langle a, (T^*T - I)b\rangle} = \overline{\langle (T^*T - I)^*a, b\rangle} = \overline{\langle (T^*T - I)a, b\rangle}$$

Ainsi on trouve

$$2\Re(\bar\lambda\langle (T^*T - I)a, b\rangle) = \bar\lambda\langle (T^*T - I)a, b\rangle + \lambda\overline{\langle (T^*T - I)^*a, b\rangle} = 0.$$

En particulier, en prenant $\lambda = 1$, on trouve $\Re(\langle (T^*T - I)a, b\rangle) = 0$ et en prenant $\lambda = i$, on trouve $\mathrm{Im}\,(\langle (T^*T - I)a, b\rangle) = 0$. Ainsi, pour tout $a, b \in H$, $\langle (T^*T - I)a, b\rangle = 0$ donc $T^*T - I = 0$. Les deux notions d'isométrie coincident donc.

Pour les projections orthogonale: si $E \subset H$ est un sous-espace vectoriel fermé et π_E la projection orthogonale sur E. Soient alors $x, x' \in H$, écrivons $x = a + b$, $x' = a' + b'$ avec $a, a' \in E$ et $b, b' \in H$. Alors $\pi_E(x) = a$ donc $\pi_E^2 = \pi_E a = a = \pi_E x$ et π_E est bien une projection. De plus

$$\begin{aligned} \langle \pi_E x, y\rangle = \langle a, a' + b'\rangle = \langle a, a'\rangle + \langle a, b'\rangle = \langle a, a'\rangle& \\ &= \langle a, a'\rangle + \langle b, a'\rangle = \langle x, a'\rangle = \langle x, \pi_E y\rangle. \end{aligned}$$

Ainsi $\pi_E^* = \pi_E$ et π_E est auto-adoint.

Inversement, si T est une projection et $T^* = T$ alors H est la somme directe de l'image et du noyau de T: $H = \mathrm{Im}\, T \oplus \ker T$ (rappelons que si T est une projection, $\mathrm{Im}\, T = \ker(I - T) = (I - T)^{-1}(0)$ est fermé). Il s'agit de vérifier que $\mathrm{Im}\, T \perp \ker T$. Mais, si $x \in \mathrm{Im}\, T$ et $y \in \ker T$ alors $Tx = x$ et

$$\langle x, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, Ty\rangle = \langle x, 0\rangle = 0.$$

3. Opérateurs compacts

Définition 3.1. Soit X, Y deux espaces de Banach et $T : X \to Y$ un opérateur borné. On dit que T est compact si l'image de la boule unité B_X de X par T, est d'adhérence compacte, $\overline{T(B_X)}$ est compact.

Exemple 3.2. Si T est un opérateur de rang fini, alors $T(B_X)$ est un ensemble borné dans un espace de dimension fini, son adhérence est donc compacte. Ainsi, *tout opérateur de rang fini est compact.*

Notons qu'un opérateur est compact si et seulement si il envoie tout ensemble borné sur un ensemble totalement borné, c'est-à-dire d'adhérence compact (on dit aussi relativement compact). En effet, si $A \subset X$ est borné, il existe $R > 0$ tel que $A \subset B(0, R) = RB_X$ donc $T(A) \subset RT(B_X)$ et alors $\overline{T(A)}$ est fermé dans le compact $R\overline{T(B_X)}$.

Comme les ensembles compacts sont les ensembles séquentiellement compacts, on obtient immédiatement la caractérisation suivante:

LEMME 3.3. *Soit $T : X \to Y$ un opérateur borné. Alors T est compact si et seulement si, pour toute suite (x_n) bornée, on peut extraire une sous-suite convergente.*

Le premier intérêt des opérateurs compacts est le suivant:

PROPOSITION 3.4. *Soit $T : X \to Y$ un opérateur borné. Si T est compact et si $x_n \rightharpoonup x$ faiblement, alors $Tx_n \to Tx$ fortement.*

DÉMONSTRATION. On a vu comme conséquence du théorème de Banach-Steinhaus que toute suite faiblement convergente était bornée. Par ailleurs, si $y \in Y^*$, $\langle y, Tx_n\rangle = \langle T^*y, x_n\rangle$ et comme $T^*y \in X^*$ on en déduit que $\langle T^*y, x_n\rangle$ converge vers $\langle T^*y, x\rangle = \langle y, Tx\rangle$. Ainsi $\langle y, Tx_n\rangle \to \langle y, Tx\rangle$ c'est-à-dire que Tx_n converge faiblement vers Tx.

Supposons que Tx_n ne converge pas fortement vers Tx. Alors il existe $\delta > 0$ et une sous-suite $(Tx_{n_k})_k$ telle que, pour tout k, $\|Tx_{n_k} - Tx\|_Y \geq \delta$. Notons que $(Tx_{n_k})_k$ converge encore faiblement vers Tx. Quitte à remplacer (x_n) par $(Tx_{n_k})_k$, on a donc une suite (x_n) bornée, telle que (Tx_n) converge faiblement vers Tx, mais, pour tout n, $\|Tx_n - Tx\|_Y \geq \delta$.

Or, par compacité de T, $(Tx_n)_n$ admet une sous-suite $(Tx_{n_k})_k$ qui converge fortement. En particulier, elle converge faiblement, donc sa limite est Tx. Mais, alors en passant à la limite dans $\|Tx_{n_k} - Tx\|_Y \geq \delta$, on trouve $\|Tx - Tx\|_Y \geq \delta$, ce qui est absurde. □

REMARQUE 3.5. Si X est un espace réflexif (le dual du dual s'identifie isométriquement à X), par exemple si X est un espace de Hilbert ou si $X = L^p(\mu)$, $1 < p < \infty$, alors la réciproque est vraie.

THÉORÈME 3.6. *Soient X, Y des espaces de Banach, T, et $(T_n)_n$ des opérateurs bornés $X \to Y$,*

(1) *Si pour tout n, T_n est compact et si $T_n \to T$ alors T est compact.*
(2) *Si T est compact et S borné, alors ST et TS sont compacts.*

DÉMONSTRATION. (1) Soit $(x_n)_{n\geq 0}$ une suite de $B(0,1)$. Alors de (T_0x_n), on peut extraire une sous-suite convergente $(T_0x_{n_k^0})_k$. Puis de $(T_1x_{n_k^0})_k$, on peut extraire une sous-suite convergente $(T_1x_{n_k^1})_k$. Une fois que n_k^j est obtenu, on construit ainsi n_k^j tel que $\{n_k^{j+1}, k \in \mathbb{N}\} \subset \{n_k^j, k \in \mathbb{N}\}$ et tel que $T_\ell(x_{n_k^{j+1}})$ converge pour $\ell \leq j+1$.

Mais alors $(T_kx_{n_k^k})$ est une suite convergente. Par ailleurs

$$\begin{aligned}\left\|Tx_{n_k^k} - Tx_{n_l^l}\right\| &\leq \left\|Tx_{n_k^k} - T_kx_{n_k^k}\right\| + \left\|T_kx_{n_k^k} - T_lx_{n_l^l}\right\| + \left\|T_lx_{n_l^l} - Tx_{n_l^l}\right\| \\ &\leq \|T - T_k\| + \left\|T_kx_{n_k^k} - T_lx_{n_l^l}\right\| + \|T_l - T\|\end{aligned}$$

ce qui montre que $\left(Tx_{n_k^k}\right)_k$ est une suite de Cauchy donc converge.

(2) Soit T est un opérateur compact, et (x_n) une suite bornée. Alors on peut extraire de (Tx_n) une suite $(Tx_{n_k})_k$ qui converge et donc $(STx_{n_k})_k$ converge aussi. Le lemme ci-dessus montre alors que ST est compact.

Par ailleurs, $S(x_n)$ est encore borné, donc on peut extraire une suite $(TSx_{n_k})_k$ de $(TSx_n)_n$ qui converge et TS est donc compact. □

THÉORÈME 3.7. *Soit H un espace de Hilbert séparable et T un opérateur borné $H \to H$. Alors T est compact si et seulement s'il existe une suite (T_n) d'opérateurs $H \to H$ de rang fini tel que $T_n \to T$.*

DÉMONSTRATION. On a déjà vu que si T_n est de rang fini, alors T_n est compact. En vertu du théorème précédent, il en découle que si $T_n \to T$ alors T est également compact. Montrons donc la réciproque.

Pour cela, rappelons qu'une famille orthonormée $(e_n)_{n\in\mathbb{N}}$ converge faiblement vers 0 puisque $\displaystyle\sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 \leq \|x\|^2 < +\infty$. Si $T : H \to H$ est compact, il en résulte que $Te_n \to T0 = 0$.

Nous allons raisonner par l'absurde et supposer que T est compact mais pas limite d'opérateurs de rang fini. Ainsi, il existe $\varepsilon_0 > 0$ tel que, pour tout opérateur R de rang fini, $\|T\| = \|T - R\| > \varepsilon_0$. En particulier, pour $R = 0$, on a $\|T - R\| > \varepsilon_0$. Il existe donc un vecteur $e_0 \in H$ de norme 1 tel que

$\|Te_0\| > \varepsilon$. Soit π_0 la projection orthogonale sur $\operatorname{Vect} e_0$ de sorte que $T\pi_0$ soit de rang fini. Par suite $\|T - T\pi_0\| > \varepsilon_0$. Il existe donc $u_1 \in H$ avec $\|u_1\| = 1$ tel que

$$\|Tu_1 - T\pi_0 u_1\| = \|T(I - \pi_0)u_1\| > \varepsilon_0 \geq \varepsilon_0\|(I - \pi_0)u_1\|.$$

Notons que $(I - \pi_0)u_1 \neq 0$ (sinon $T(I - \pi_0)u_1 = 0$) et que $(I - \pi_0)u_1$ est orthogonal à $\operatorname{Vect} e_0$. On pose alors $e_1 = \dfrac{(I - \pi_0)u_1}{\|(I - \pi_0)u_1\|}$ de sorte que $\|Te_1\| > \varepsilon_0$ et e_0, e_1 est orthonormé.

Une fois qu'on a construit $e_0, \ldots, e_n$, on pose π_n la projection orthogonale sur $E_n = \operatorname{Vect}\{e_0, \ldots, e_n\}$. Comme $T\pi_n$ est de rang fini, $\|T - T\pi_n\| > \varepsilon_0$. Il existe donc $u_{n+1} \in H$ avec $\|u_{n+1}\| = 1$ tel que

$$\|T(I - \pi_n)u_{n+1}\| > \varepsilon_0 \geq \varepsilon_0\|(I - \pi_n)u_{n+1}\|.$$

Ainsi $(I - \pi_n)u_{n+1} \neq 0$ et est orthogonal à E_n, donc à $e_0, \ldots, e_n$. On pose donc $e_{n+1} = \dfrac{(I - \pi_n)u_{n+1}}{\|(I - \pi_n)u_{n+1}\|}$ alors $e_0, \ldots, e_{n+1}$ est orthonormé et $\|Te_k\| > \varepsilon_0$ pour tout k. En particulier, $(Te_k)_k$ ne peut converger vers 0, ce qui est la contradiction souhaitée. □

4. Opérateurs de Hilbert-Schmidt

Dans toute cette section, H est un espace de Hilbert séparable.

Définition 4.1. Soit $T : H \to H$ une application linéaire. On dit que T est un opérateur de Hilbert-Schmidt s'il existe une base orthonormée $(e_n)_{n\geq 1}$ de H telle que

$$\|T\|_{HS}^2 = \sum_{n=1}^{+\infty} \|Te_n\|^2 < +\infty.$$

Lemme 4.2. *Si T est de Hilbert-Schmidt, $\|T\|_{HS}$ ne dépend pas de la base choisie. De plus, T est de Hilbert-Schmidt si et seulement si T^* est de Hilbert-Schmidt et $\|T\|_{HS} = \|T^*\|_{HS}$. Enfin, si T est de Hilbert-Schmidt, alors T est borné avec $\|T\| \leq \|T\|_{HS}$.*

Démonstration. En effet, soient $(e_n)_{n\geq 1}$ et $(f_n)_{n\geq 1}$ deux bases orthonormées de H. Alors

$$\|Te_n\|^2 = \sum_{k=1}^{+\infty} |\langle Te_n, f_k\rangle|^2 = \sum_{k=1}^{+\infty} |\langle e_n, T^*f_k\rangle|^2.$$

On somme alors en n et on utilise Fubini (que la somme soit finie ou non, puisqu'on ne somme que des termes positifs):

$$\begin{aligned}
\|T\|_{HS}^2 &= \sum_{n=1}^{+\infty} \|Te_n\|^2 \\
&= \sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty} |\langle e_n, T^*f_k\rangle|^2 = \sum_{k=1}^{+\infty}\sum_{n=1}^{+\infty} |\langle T^*f_k, e_n\rangle|^2 \\
&= \sum_{k=1}^{+\infty} \|T^*f_k\|^2 = \|T^*\|_{HS}^2.
\end{aligned}$$

Cela montre plusieurs choses:

– (e_n) n'apparaissant pas dans le membre de droite, la norme de Hilbert-Schmidt ne dépendant pas de la base choisie.

– T est de Hilbert-Schmidt si et seulement si T^* l'est également.

– Enfin, $\|T\|_{HS}^2 = \|T^*\|_{HS}^2$.

Finalement, soit $x \in H$ avec $\|x\| = 1$. On pose $e_1 = x$ et on prend une base orthonormée $(e_n)_{n\geq 2}$ de $\{x\}^\perp$ de sorte que $\{e_n\}_{n\geq 1}$ est une base orthonormée de H. Alors

$$\|Tx\|^2 = \|Te_1\|^2 \leq \sum_{n\geq 1} \|Te_n\|^2 = \|T\|_{HS}^2.$$

Ansi T est bornée. De plus, en prenant le supremum sur les x de norme 1, on trouve bien $\|T\| \leq \|T\|_{HS}$. $\square$

EXEMPLE 4.3. Soit E un sous-espace vectoriel fermé et π_E la projection orthogonale sur E. Alors π_E est de Hilbert-Schmidt si et seulement si E est de dimension finie.

En effet, il suffit de prendre une base orthonormée $(e_i)_{i\in I}$ de E et de la compléter par une base orhtonormée $(f_j)_{j\in J}$ de $E^\perp$. Alors

$$\|\pi_E\|_{HS}^2 = \sum_{i\in I} \|\pi_E e_i\|^2 + \sum_{j\in J} \|\pi_E f_j\|^2 = \sum_{i\in I} \|\pi_E e_i\|^2 = \sum_{i\in I} 1$$

qui est fini si et seulement si I est fini, c'est-à-dire que E est de dimension finie.

LEMME 4.4. *Soit T un opérateur de Hilbert-Schmidt et S un opérateur borné, alors ST et TS sont de Hilbert-Schmidt avec $\|ST\|_{HS} \leq \|S\|\,\|T\|_{HS}$ et $\|TS\|_{HS} \leq \|S\|\,\|T\|_{HS}$*

DÉMONSTRATION. Si T est de Hilbert-Schmidt, S borné et $(e_n)_{n\geq 1}$ une base orthonormée de H alors $\|STe_n\|^2 \leq \|S\|^2\|Te_n\|^2$. En sommant sur $n \geq 1$, on trouve bien $\|ST\|_{HS} \leq \|S\|\,\|T\|_{HS}$.

Ensuite, S^* est encore borné avec $\|S^*\| = \|S\|$, T^* est encore de Hilbert-Schmidt avec $\|T^*\|_{HS} = \|T\|_{HS}$ donc $(TS)^* = S^*T^*$ est de Hilbert-Schmidt et

$$\|TS\|_{HS} = \|(TS)^*\|_{HS} = \|S^*T^*\|_{HS} \leq \|S^*\|\,\|T^*\|_{HS} = \|S\|\,\|T\|_{HS}.$$

$\square$

EXEMPLE 4.5. Tout opérateur de rang fini est de Hilbert-Schmidt.

En effet, si T est de rang fini, son image E est de dimension finie. Si π_E est la projection orthogonale sur E, alors π_E est de Hilbert-Schmidt et comme $T = \pi_E T$, T aussi.

THÉORÈME 4.6. *Soit*

$$\mathcal{B}_2(H) = HS(H) = \{T : H \to H,\ T \text{ de Hilbert-Schmidt}\}$$

muni de la norme $\|T\|_{HS}$. Alors T est un espace de Hilbert.

DÉMONSTRATION. Fixons $(e_n)_{n\geq 1}$ une base orthonormée de H.

Montrons d'abord que HS est un espace vectoriel et que $\|T\|_{HS}$ est bien une norme.

i) On a évidemment $\|T\|_{HS} \geq 0$ et comme $\|T\| \leq \|T\|_{HS}$, si $\|T\|_{HS} = 0$ alors $|T\| = 0$ donc $T = 0$.

ii) Comme $\|\lambda Te_n\|^2 = |\lambda|^2\|Te_n\|^2$. En sommant sur $n \geq 1$ et en prenant la racine carrée, on trouve bien $\|\lambda T\|_{HS} = |\lambda|\,\|T\|_{HS}$. En particulier $\lambda T \in HS$.

iii) Si T, T' sont de Hilbert-Schmidt, alors $\|Te_n\|, \|T'e_n\|$ sont dans ℓ^2 et donc $\|Te_n\| + \|T'e_n\| \in \ell^2$. Comme $\|Te_n + T'e_n\| \leq \|Te_n\| + \|T'e_n\|$ on a $\|Te_n + T'e_n\| \in \ell^2$ avec $\big\|\|Te_n + T'e_n\|\big\|_{\ell^2} \leq \big\|\|Te_n\|\big\|_{\ell^2} + \big\|\|T'e_n\|\big\|_{\ell^2}$. Enfin

$$\|T + T'\|_{HS} = \big\|\|Te_n + T'e_n\|\big\|_{\ell^2} \leq \big\|\|Te_n\|\big\|_{\ell^2} + \big\|\|T'e_n\|\big\|_{\ell^2} = \|T\|_{HS} + \|T'\|_{HS}.$$

En particulier, $T + T' \in HS$.

On peut remarquer que la norme $\|\cdot\|_{HS}$ est issue du produit scalaire

$$\langle S, T\rangle_{HS} = \sum_{n\geq 1} \langle Se_n, Te_n\rangle.$$

Avec Cauchy-Schwarz dans H puis dans ℓ^2, cette somme est bien convergente

$$\begin{aligned}\sum_{n\geq 1} |\langle Se_n, Te_n\rangle| &\leq \sum_{n\geq 1} \|Se_n\|\,\|Te_n\| \\ &\leq \left(\sum_{n\geq 1} \|Se_n\|^2\right)^{1/2} \left(\sum_{n\geq 1} \|Te_n\|^2\right)^{1/2} = \|S\|_{HS}\|T\|_{HS}.\end{aligned}$$

Montrons maintenant que HS est complet. Soit (T_n) une suite de Cauchy pur $\|\cdot\|_{HS}$. Comme $\|S\| \leq \|S\|_{HS}$, (T_n) est aussi de Cauchy dans $\mathcal{B}(H)$. On a vu au premier semestre que $\mathcal{B}(H)$ est complet, donc (T_n) converge dans $\mathcal{B}(H)$. On note T sa limite pour la norme $\|\cdot\|$: $\|T_n - T\| \to 0$. Il faut montrer que $T \in HS$ et que $\|T_n - T\|_{HS} \to 0$.

Mais, pour tout $k \geq 1$, $T_n e_k \to T e_k$ quand $n \to +\infty$ et par ailleurs, si on pose $a_n = (\|T_n e_k\|)_{k\geq 1}$, alors $a_n \in \ell^2$ et $\|a_n - a_m\|_{\ell^2} = \|T_n - T_m\|_{HS}$. Donc $(a_n)_n$ est une suite de Cauchy dans ℓ^2, donc elle converge dans ℓ^2 vers une suite $a \in \ell^2$. Comme $T_n e_k \to T e_k$, nécessairement $a = (Te_k)_{k\geq 1}$ ce qui montre que $T \in HS$ et comme $\|T_n - T\|_{HS} = \|a_n - a\|_{\ell^2} \to 0$ on a bien $T_n \to T$ dans HS. □

Théorème 4.7. *Un opérateur de Hilbert-Schmidt est compact.*

Démonstration. Fixons $(e_n)_{n\geq 1}$ une base orthonormée de H. Soit π_N la projection orthogonale sur l'espace engendré par $\{e_1, \ldots, e_N\}$. Alors π_N est de rang fini donc $T\pi_N$ aussi. Mais

$$\|(T - T\pi_N)e_n\| = \begin{cases} 0 & \text{si } n \leq N \\ \|Te_n\| & \text{sinon} \end{cases}$$

donc

$$\|T - T\pi_N\|^2 \leq \|T - T\pi_N\|_{HS}^2 = \sum_{n>N} \|Te_n\|^2 \to 0$$

quand $N \to +\infty$. Comme $T\pi_N$ est de rang fini, T s'approche en norme $\mathcal{B}(H)$ par des opérateurs de rang fini, il est donc compact. □

Examinons enfin le cas particulier où $H = L^2(\Omega, \mu)$.

Théorème 4.8. *Soit (Ω, μ) un espace mesuré σ-fini. Soit T un opérateur sur $L^2(\Omega, \mu)$. Alors T est de Hilbert-Schmidt si et seulement s'il existe $K : \Omega \times \Omega \to \mathbb{C}$ avec*

$$\|K\|_{L^2(\mu)\otimes L^2(\mu)}^2 := \int_{\Omega\times\Omega} |K(x,y)|^2 \, d\mu(x) \, d\mu(y) < +\infty$$

tel que

$$Tf(x) = T_K f(x) := \int_\Omega K(x,y) f(y) \, d\mu(y).$$

De plus $\|T\|_{HS} = \|K\|_{L^2(\mu)\otimes L^2(\mu)}$.

Démonstration. Soit $(\varphi_k)_{k\geq 1}$ une base orthonormée de $L^2(\mu)$ et soit $f \in L^2(\mu)$.

Soit d'abord $K : \Omega \times \Omega \to \mathbb{C}$ avec $\|K\|_{L^2(\mu)\otimes L^2(\mu)} < +\infty$ et montrons que T_K est bien un opérateur de Hilbert-Schmidt, a forciori un opérateur borné.

Notons que $T_K \varphi_n(x) = \overline{\left\langle \overline{K(x,\cdot)}, \varphi_n \right\rangle_{L^2(\mu)}}$. donc

$$\|T_K \varphi_n\|^2 = \int_\Omega \left| \left\langle \overline{K(x,\cdot)}, \varphi_n \right\rangle_{L^2(\mu)} \right|^2 \mathrm{d}\mu(x).$$

Ainsi, avec Fubini

$$\begin{aligned} \sum_{n\geq 1} \|T_K \varphi_n\|^2 &= \sum_{n\geq 1} \int_\Omega \left| \left\langle \overline{K(x,\cdot)}, \varphi_n \right\rangle_{L^2(\mu)} \right|^2 \mathrm{d}\mu(x) \\ &= \int_\Omega \sum_{n\geq 1} \left| \left\langle \overline{K(x,\cdot)}, \varphi_n \right\rangle_{L^2(\mu)} \right|^2 \mathrm{d}\mu(x) \\ &= \int_\Omega \left\| \overline{K(x,\cdot)} \right\|_{L^2(\mu)} \mathrm{d}\mu(x) = \int_{\Omega\times\Omega} |K(x,y)|^2 \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y). \end{aligned}$$

Ceci montre qu'on a bien $T_K \in HS$ avec $\|T_K\|_{HS} = \|K\|_{L^2(\mu)\otimes L^2(\mu)}$.

Soit T un opérateur de Hilbert-Schmidt, Alors

$$f = \sum_{k\geq 1} \langle f, \varphi_k \rangle_{L^2(\mu)} \varphi_k$$

et cette série converge dans $L^2(\mu)$. Comme T est continue sur $L^2(\mu)$,

$$Tf = \sum_{k\geq 1} \langle f, \varphi_k\rangle_{L^2(\mu)} T\varphi_k.$$

Remarquons que

$$\begin{aligned}\langle f, \varphi_k\rangle_{L^2(\mu)} T\varphi_k(x) &= \int_\Omega f(y)\overline{\varphi_k(y)}\,\mathrm{d}\mu(y)\,T\varphi_k(x)\\ &= \int_\Omega T\varphi_k(x)\overline{\varphi_k(y)}f(y)\,\mathrm{d}\mu(y)\end{aligned}$$

donc formellement au moins

$$Tf(x) = \int_\Omega K(x,y)f(y)\,\mathrm{d}\mu(y)$$

où

$$K(x,y) = \sum_{k\geq 1} T\varphi_k(x)\overline{\varphi_k(y)}.$$

Notons que cette formule est valable au moins lorsque f est une combinaison linéaire finie des φ_k puisqu'alors la somme sur k est finie. Si on montre que $\|K\|_{L^2(\mu)\otimes L^2(\mu)} < +\infty$ alors T sera un opérateur de Hilbert-Schmidt, donc continu, donc le résultat s'étend à l'adhérence des combinaison linéaire finie des φ_k, c'est-à-dire à $L^2(\mu)$ en entier.

Mais

$$\begin{aligned}\int_\Omega |K(x,y)|^2\,\mathrm{d}\mu(y) &= \sum_{k\geq 1} |\langle K(x,\cdots), \varphi_k\rangle_{L^2(\mu)}|^2\\ &= \sum_{k\geq 1} |\langle \varphi_k, K(x,\cdots)\rangle_{L^2(\mu)}|^2 = \sum_{k\geq 1}|T\varphi_k(x)|^2\end{aligned}$$

donc

$$\begin{aligned}\|K\|^2_{L^2(\mu)\otimes L^2(\mu)} &= \int_{\Omega\times\Omega} |K(x,y)|^2\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) = \int_\Omega \sum_{k\geq 1} |T\varphi_k(x)|^2\,\mathrm{d}\mu(x)\\ &= \sum_{k\geq 1}\int_\Omega |T\varphi_k(x)|^2\,\mathrm{d}\mu(x)\\ &= \sum_{k\geq 1} \|T\varphi_k\|_{L^2(\mu)} = \|T\|^2_{HS} < +\infty.\end{aligned}$$

Le théorème est donc démontré. □

Exemple 4.9. Évidemment, ℓ^2 est un cas particulier. Si on prend une suite à deux indices (une matrice infinie) $a = (a_{n,m})_{n,m\in\mathbb{N}}$, on peut définir l'opérateur T_a sur ℓ^2 par

$$(T_a x)_n = \sum_{m\in\mathbb{N}} a_{n,m}x_m.$$

Alors cet opérateur est un opérateur de Hilbert-Schmidt si et seulement si $\sum_{m\in\mathbb{N}} |a_{n,m}|^2 < +\infty$.

Notons que $a_{n,m} = \langle Te_m, e_n\rangle$ où $(e_n)_{n\in\mathbb{N}}$ est la base canonique de ℓ^2.

5. Théorie spectrale

5.1. Spectre et valeurs propres. Rappelons le résultat fondamental d'algèbre linéaire suivant

Théorème 5.1 (Théorème du rang). *Soient E un espace vectoriel de dimension finie, et $T : E \to E$ une application linéatire. Alors*

$$\dim E = \dim\ker T + \dim \operatorname{Im} T.$$

En particulier, on a équivalence entre

(i) *T est injectif i.e.* $\ker T = \{0\}$*;*
(ii) *T est surjectif i.e.* $\operatorname{Im} T = E$*;*

(iii) *T est bijectif;*
(iv) *T a un inverse à gauche: il existe $S_g : E \to E$ linéaire tel que $S_g T = I$;*
(v) *T a un inverse à gauche: il existe $S_d : E \to E$ linéaire tel que $TS_d = I$.*

De plus, dans ce cas, $S_g = S_d$ est noté T^{-1}.

Notons que l'existence d'un inverse à gauche implique l'injectivité de T: si $Tx = 0$ alors $x = S_g Tx = 0$. Réciproquement, en dimension finie si T est injectif, alors si $y \in \operatorname{Im} T$, *i.e.* s'il existe $x \in E$ tel que $y = Tx$, cet x est unique, ce qui permet de définir $S_g y$ comme l'unique x tel que $y = Tx$ donc $S_g Tx = x$. Notez que nous n'avons pour l'instant défini S_g que sur $\operatorname{Im} T$ et il est facile de voir que cette application est linéaire. Pour définir S_g sur E en entier, quand E est de dimension finie, il suffit de prendre un supplémentaire $\tilde{E}$, $E = \tilde{E} \oplus \operatorname{Im} T$ et définir $S_g x = 0$ si $x \in \tilde{E}$. Nous ne cherchons pas ici à étendre cette définition au cas de la dimension infini, et encore moins à montrer la continuité de S_g. (Le lecteur attentif pourra étudier le cas où $\operatorname{Im} T$ est fermé dans un espace de Hilbert). Notez que ce raisonnement est valable même si $T : E \to F$ avec F un second espace vectoriel pas forcément de même dimension et on obtient ainsi $S_g : F \to E$.

L'existence d'un inverse à droite implique la surjectivité de T: $x = T(S_d x) \in \operatorname{Im} T$. Réciproquement, en dimension finie, supposons que $T : E \to F$ est surjectif et écrivons $E = \tilde{E} \oplus \ker T$. Alors T est bijectif sur $\tilde{E}$ puisque $\tilde{E} \cap \ker T = \{0\}$ donc T est injectif et il est évidemment toujours surjectif puisque si $y \in F$, il existe $x \in E$ tel que $y = Tx$. En écrivant $x = x_1 + x_2$, $x_1 \in \tilde{E}$ et $x_2 \in \ker T$ on voit que $y = Tx_1$. Comme cette écriture est unique, on peur définir $S_d y = x_1$ et on a bien $TS_d y = y$.

Notons que si $T : E \to E$ admet une inverse à droite S_d et une inverse à gauche S_g (et ces deux propriétés sont équivalentes en dimension finie) alors

$$S_g = S_g I = S_g(TS_d) = (S_g T)S_d = IS_d = S_d$$

i.e. ces deux inverses sont égales. Ceci est vrai même en dimension infinie. Ce qui n'est pas vrai en dimension infinie est que l'existence d'une inverse à droite soit équivalente à l'existence d'une inverse à gauche.

Par exemple, soit $D_+ : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur de décalage à droite défini par

$$D_+(a_0, a_1, \ldots) = (0, a_0, a_1, \ldots).$$

Cet opérateur est clairement injectif mais pas surjectif donc n'admet pas d'inverse à droite. Son inverse à gauche est l'opérateur de décalage à gauche $D_- : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur de décalage défini par

$$D_-(b_0, b_1, b_2, \ldots) = (b_1, b_2, \ldots).$$

On a $D_- D_+ = I$ mais $D_+ D_- = I - P_0$ où $P_0(a_0, a_1, \ldots) = (a_0, 0, \ldots)$ *i.e.*

$$D_+ D_-(a_0, a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

Définition 5.2. Soit E un espace de Banach et $T : E \to E$ une application linéaire bornée. On dit que T est inversible s'il existe une application linéaire $S : E \to E$ telle que $ST = TS = I$.

Ainsi T est bijective, on note $S = T^{-1}$ et cette application est automatiquement bornée (théorème de l'application ouverte).

Le lemme suivant permet de facilement construire des opérateurs inversibles:

Lemme 5.3 (Neumann). *Soit E un espace de Banach et $T : E \to E$ une application linéaire bornée avec $\|T\| < 1$. Alors*

(i) *$I - T$ est inversible;*
(ii) $(I - T)^{-1} = \displaystyle\sum_{k=0}^{+\infty} T^k$;
(iii) *l'application $B(E) \to B(E)$, $T \to (I - T)^{-1}$ est continue sur $\{T \in B(E), \|T\| < 1\}$;*
(iv) $\left\|(I - T)^{-1}\right\| \leq \dfrac{1}{1 - \|T\|}$.

Démonstration. L'ennoncé contient déjà l'idée de la démonstration puisqu'on reconnaît le développement en série entière de $(1-t)^{-1}$ dans l'expression de $(I-T)^{-1}$.

On note $P_n(t) = \sum_{k=0}^{n} t^k = \dfrac{1-t^{n+1}}{1-t}$ donc $(1-t)P_n(t) = P_n(t)(1-t) = 1 + t^{n+1}$. Mais, ceci est une identité entre polynômes et on on peut appliquer un polynôme à un opérateur: $P(T) = \sum_{finie} a_k T^k$ si $P(t) = \sum_{finie} a_k t^k$. Il en résulte que

$$(1-T)P_n(T) = P_n(T)(I-T) = I + T^{n+1}. \tag{5.18}$$

D'une part $\left\|T^{n+1}\right\| \leq \|T\|^{n+1} \to 0$ puisque $\|T\| < 1$. D'autre part, $\mathcal{B}(E)$ est complet et la série $\sum_{k=0}^{\infty} \left\|T^k\right\| \leq \sum_{k=0}^{\infty} \|T\|^k < +\infty$ converge puisque $\|T\| < 1$, donc $\sum_{k=0}^{+\infty} T^k = \lim P_n(T)$ est une série normalement convergente donc convergente. On peut passer à la limite dans (5.18) ce qui donne

$$(I-T)\left(\sum_{k=0}^{\infty} T^k\right) = \left(\sum_{k=0}^{\infty} T^k\right)(I-T) = I$$

ce qui nous dit que $I-T$ est bien inversible et que son inverse est bien $\sum_{k=0}^{\infty} T^k$. Il reste à voir que

$$\|P_n(T)\| = \left\|\sum_{k=0}^{n} T^k\right\| \leq \sum_{k=0}^{n} \left\|T^k\right\| = \frac{1-\|T\|^{n+1}}{1-\|T\|} \leq \frac{1}{1-\|T\|}.$$

En passant à la limite sur n, on obtient bien $\left\|(I-T)^{-1}\right\| \leq \dfrac{1}{1-\|T\|}$. □

Corollaire 5.4. *Si* $|\lambda| > \|T\|$ *alors* $\lambda I - T$ *est inversible,*

$$(\lambda I - T)^{-1} = \sum_{k=0}^{+\infty} \frac{T^k}{\lambda^{k+1}}$$

et $\left\|(T-\lambda I)^{-1}\right\| \leq \frac{1}{|\lambda| - \|T\|}$.

Démonstration. Remarquons que $\lambda I - T = \lambda\left(I - \frac{1}{\lambda}T\right)$ et comme $\left\|\frac{1}{\lambda}T\right\| = \frac{1}{|\lambda|}\|T\| < 1$ donc $I - \frac{1}{\lambda}T$ est inversible d'inverse

$$\left(I - \frac{1}{\lambda}T\right)^{-1} = \sum_{k=0}^{+\infty} \frac{T^k}{\lambda^k}$$

De plus

$$\left\|(T-\lambda I)^{-1}\right\| = \frac{1}{|\lambda|}\left\|\left(I - \frac{1}{\lambda}T\right)^{-1}\right\| \leq \frac{1}{|\lambda|}\frac{1}{1 - \left\|\frac{1}{\lambda}T\right\|} = \frac{1}{|\lambda| - \|T\|}.$$

□

Corollaire 5.5. *L'ensemble des applications inversibles est un ouvert de* $B(E)$ *et l'application* $T \to T^{-1}$ *est continue sur cet ouvert.*

Démonstration. Tout d'abord supposons que T_0 soit inversible et soit T tel que $\|T - T_0\| < r$ et essayons de voir que si r est assez petit, alors T est également inversible.

Observons que, comme T_0 est inversible, T est inversible si et seulement si $T_0^{-1}T$ est inversible. Mais pour cela, on écrit $T_0^{-1}T = I - (I - T_0^{-1}T)$ qui permet d'appliquer le lemme de Neumann si $\left\|I - T_0^{-1}T\right\| < 1$. Mais $I - T_0^{-1}T = T_0^{-1}(T_0 - T)$ donc

$$\left\|I - T_0^{-1}T\right\| \leq \left\|T_0^{-1}(T_0 - T)\right\| \leq \left\|T_0^{-1}\right\|\|T_0 - T\|$$

et on voit que T est inversible dès que $\|T_0 - T\| < \dfrac{1}{\left\|T_0^{-1}\right\|}$.

De plus, comme $T = T_0(T_0^{-1}T)$ on aura alors

$$T^{-1} = (T_0^{-1}T)^{-1}T_0^{-1} = \big(I - (I - T_0^{-1}T)\big)^{-1}T_0^{-1} = \sum_{n=0}^{+\infty}(I - T_0^{-1}T)^k T_0^{-1}.$$

Mais alors

$$T^{-1} - T_0^{-1} = \sum_{n=1}^{+\infty}(I - T_0^{-1}T)^k T_0^{-1}$$

donc

$$\begin{aligned}\big\|T^{-1} - T_0^{-1}\big\| &\leq \sum_{n=1}^{+\infty}\big\|(I - T_0^{-1}T)^k T_0^{-1}\big\| \leq \sum_{n=1}^{+\infty}\big\|(I - T_0^{-1}T)\big\|^k \|T_0\| \\ &= \sum_{n=1}^{+\infty}\big\|T_0^{-1}(T_0 - T)\big\|^k \|T_0\| \leq \sum_{n=1}^{+\infty}\|(T_0 - T)\|^k \|T_0\|\big\|T_0^{-1}\big\| \\ &= \frac{\|T_0 - T\|}{1 - \|T_0 - T\|}\|T_0\|\big\|T_0^{-1}\big\| \to 0\end{aligned}$$

quand $T \to T_0$.

□

Remarque 5.6. Si E est de dimension fini n, il y a une démonstration plus directe. Identifions $B(E)$ et $\mathcal{M}_n(\mathbb{C})$. Alors l'ensemble des matrices inversibles est $\{A \in \mathcal{M}_n(\mathbb{C}) \,:\, \det A \neq 0\}$. En remarquant que le déterminant est un polynôme (dont les variables sont les coefficients de A) donc une fonction continue, l'ensemble des matrices inversibles est l'image réciproque de l'ouvert $\mathbb{C} \setminus \{0\}$ par det et est donc ouvert.

Ensuite, $A^{-1} = \dfrac{1}{\det A}Com(A)^t$. Comme les entrées de $Com(A)$ sont des déterminants de matrices extraites de A, c'est un polynôme en A donc A^{-1} est une fraction rationnelle en A donc est continue sur son domaine, et même $\mathcal{C}^\infty$.

La démonstration ci-dessus dit en fait un peu plus puisqu'on a écrit

$$\begin{aligned}T^{-1} &= \sum_{n=0}^{+\infty}(I - T_0^{-1}T)^k T_0^{-1} = \sum_{n=0}^{+\infty}(T_0^{-1})^k(T_0 - T)^k T_0^{-1} \\ &= T_0^{-1} + T_0^{-1}(T_0 - T)T_0^{-1} + o(\|T - T_0\|)\end{aligned}$$

ou encore

$$(T_0 + A)^{-1} = T_0^{-1} - T_0^{-1}AT_0^{-1} + o(\|A\|).$$

Ainsi l'application $Inv \,:\, T \to T^{-1}$ est différentiable et sa différentielle en T_0^{-1} est $\mathrm{d}Inv_{T_0}(A) = -T_0^{-1}AT_0^{-1}$. C'est l'analogue non commutatif de $(1/x)' = -1/x^2$.

Dans toute la suite de cette section, on suppose que E est un espace de Banach sur $\mathbb{C}$. Certains résultats utilisent la théorie des fonctions holomorphes et ne sont pas valables pour les espaces sur $\mathbb{R}$.

Définition 5.7. Soit $T : E \to E$ un opérateur borné. On définit:
– une *valeur propre* $\lambda \in \mathbb{C}$ de T s'il existe $x \in E$ tel que $Tx = \lambda x$. Le *spectre ponctuel* de T, noté $\sigma_p(T)$ est l'ensemble des valeurs propres de T.
– le *spectre* de T, noté $\sigma(T)$ est l'ensemble des λ pour lesquels $T - \lambda I$ *n'est pas* inversible.
– la *résolvante* de T est l'application $R_T : \mathbb{C} \setminus \sigma(T) \to B(E)$, $\lambda \to (T - \lambda I)^{-1}$.

Remarquons que si λ est valeur propre, alors l'espace propre correspondant $\ker(T - \lambda I) \neq \{0\}$ donc $T - \lambda I$ n'est pas injectif, donc pas inversible, ainsi $\lambda \in \sigma(T)$. Si E est de dimension fini, le fait que $T - \lambda I$ ne soit pas inversible est équivalent au fait qu'il ne soit pas injectif donc $\sigma(T) = \sigma_p(T)$. Il n'en va pas de même en dimension infinie.

Théorème 5.8. *Soit E un espace de Banach et $T : E \to E$ une application linéaire bornée. Alors*
(i) *le spectre ponctuel est une partie du spetre: $\sigma_p(T) \subset \sigma(T)$;*

(ii) *le spectre est borné* $\sigma(T) \subset \{z \in \mathbb{C} : |z| \leq \|T\|\}$*;*
(iii) *le spectre* $\sigma(T)$ *est fermé et* $\mathbb{C} \setminus \sigma(T)$ *est un ouvert;*
(iv) *l'application* $\lambda \to R(\lambda)$ *est analytique sur cet ouvert (dévelopable en série entière avec des coefficients opérateurs);*
(v) *le spectre* $\sigma(T)$ *est compact;*
(vi) *le spectre n'est pas vide.*

Démonstration. Nous avons démontré (i) juste avant le théorème et vu (ii) dans le corollaire 5.4. Enfin (ii)-(iii) montrent que $\sigma(T)$ est fermé borné dans $\mathbb{C}$ donc compact. Ainsi (v) sera démontré une fois que nous aurons montré les assertions (iii)-(iv), qui se démontrent simultanément à l'aide du Lemme de Neumann.

Soit maintenant $\lambda_0 \in \mathbb{C} \setminus \sigma(T)$ *i.e.* $T - \lambda_0 I$ est inversible. On prend $\lambda \in \mathbb{C}$ et on écrit $z = \lambda - \lambda_0$ qui sera petit. Puis

$$\begin{aligned} \lambda I - T &= \lambda_0 I - T + zI = \lambda_0 I - T + z(\lambda_0 I - T)(\lambda_0 I - T)^{-1} \\ &= (\lambda_0 I - T)\big(I + z(\lambda_0 I - T)^{-1}\big). \end{aligned}$$

Mais maintenant, en posant $N = \big\|(\lambda_0 I - T)^{-1}\big\|$, on a pour $|z| < 1/N$, $\big\|z(\lambda_0 I - T)^{-1}\big\| < 1$ donc $T_z := \big(I + z(\lambda_0 I - T)^{-1}\big)$ est inversible et son inverse est donnée par

$$T_z^{-1} = \sum_{k=0}^{+\infty}(-z)^k\big((\lambda_0 I - T)^{-1}\big)^k = \sum_{k=0}^{+\infty}(\lambda_0 - \lambda)^k\big((\lambda_0 I - T)^{-1}\big)^k.$$

Finalement, le calcul ci-dessus montre que $\lambda I - T$ est inversible et que son inverse est

$$(\lambda I - T)^{-1} = T_z^{-1}(\lambda_0 I - T)^{-1} = \sum_{k=0}^{+\infty}(\lambda_0 - \lambda)^k\big((\lambda_0 I - T)^{-1}\big)^{k+1}$$

qui est bien de la forme voulue.

□

Exemple 5.9. Considérons $E = \mathbb{R}^2$ et T dont la matrice est $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Alors $\sigma(T) = 0$ et $R_T(\lambda) = -\dfrac{1}{\lambda^2}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Considérons $E = L^2(0,1)$ et $Tf(t) = e^{it}f(t)$. Alors $\sigma_p(T) = \emptyset$ et $\sigma(T) = \mathbb{T} = \{z \in \mathbb{C}, |z| = 1\}$.

En effet, si on avait $Tf = \lambda f$ alors, pour tout $t \in (0,1)$, $(\lambda - e^{it})f(t) = 0$. Mais $\lambda - e^{it}$ ne s'annule que sur un ensemble de mesure nulle, donc $f = 0$ p.p. Il n'y a donc pas de vecteur propre.

Pour le spectre, si $\lambda \in \mathbb{C} \setminus \mathbb{T}$, alors $e^{it} - \lambda$ ne s'annule pas sur $[0,1]$ donc $m : t \to \frac{1}{\lambda - e^{it}}$ est une fonction bornée. Il en résulte que $M : g \to m(t)g(t)$ est borné $L^2(0,1) \to L^2(0,1)$ et très clairement, cet opérateur est l'inverse (algébrique) de $\lambda I - T$. Lorsque $|\lambda| = 1$, cet opérateur est toujous l'inverse algébrique de $\lambda I - T$, mais il n'est pas borné puique $M1 = m \notin L^2(0,1)$.

Définition 5.10. Soit T un opérateur borné sur un espace de Banach E, son rayon spectral est

$$\rho(T) = \sup_{\lambda \in \sigma(T)} |\lambda|.$$

En particulier, $\rho(T) \leq \|T\|$.

Théorème 5.11. *Soit T un opérateur borné sur un espace de Banach E alors* $\lim \|T^n\|^{1/n}$ *existe et*

$$\rho(T) = \lim \|T^n\|^{1/n}.$$

5.2. L'opérateur de Voltera. On considère ici l'opérateur de Voltera sur $L^2(0,1)$ défini par

$$Vf(x) = \int_0^x f(t)\,\mathrm{d}t = \int_0^1 \mathbf{1}_{[0,x]}(t)f(t)\,\mathrm{d}t = \langle f, \mathbf{1}_{[0,x]}\rangle.$$

Ainsi Vf est bien défini pour tout $x \in [0,1]$ et c'est un opérateur à noyau $K(x,t) = \mathbf{1}_{[0,x]}(t)$. On remarque que

$$\int_0^1 \int_0^1 |K(x,t)|^2\,\mathrm{d}t\,\mathrm{d}x = \int_0^1 \int_0^x \mathrm{d}t\,\mathrm{d}x = \int_0^1 x\,\mathrm{d}x = \frac{1}{2} < +\infty$$

donc V est un opérateur de Hilbert-Schmidt, en particulier c'est un opérateur compact (et donc aussi borné). De plus $\|V\|_{HS} = 2^{-1/2}$.

Son adjoint a pour noyau $\tilde{K}(x,t) = K(t,x) = \mathbf{1}_{[0,t]}(x) = \mathbf{1}_{[x,1]}(t)$. En d'autre termes

$$V^*g(x) = \int_0^1 \mathbf{1}_{[x,1]}(t)g(t)\,\mathrm{d}t = \int_x^1 g(t)\,\mathrm{d}t.$$

On voit donc que V n'est pas auto-adjoint puisque $V\mathbf{1} = \mathbf{1}$ et $V^*\mathbf{1} = \mathbf{1} - x$.

Supposons que $\lambda \in \mathbb{C}$ et que $\lambda f(x) = Vf(x) = \displaystyle\int_0^x f(t)\,\mathrm{d}t$. Tout d'abord, si $\lambda = 0$ alors pour presque tout $x \in [0,1]$, $Vf(x) = 0$ donc pour presque tous x, y,

$$0 = Vf(y) - Vf(x) = \int_x^y f(t)\,\mathrm{d}t = \int_0^1 \mathbf{1}_{[x,y]}f(t)\,\mathrm{d}t.$$

Par suite f est orthogonal à toutes les fonctions étagées et, par densité de celles-ci dans $L^2(0,1)$, $f = 0$. Ainsi 0 n'est pas valeur propre.

Observons ensuite que, avec Cauchy-Schwarz, $|Vf(y) - Vf(x)| \le |y-x|^{1/2}\|f\|_2$. En particulier, Vf est continue. Mais alors, si $\lambda \neq 0$ et $Vf = \lambda f$ *i.e.* $f = \frac{1}{\lambda}Vf$ donc f est continue. Comme Vf est une primitive de f, Vf est alors $\mathcal{C}^1$ donc f aussi. Une récurrence immédiate montre alors que f est de classe $\mathcal{C}^\infty$. En dérivant, on trouve $f = (Vf)' = (\lambda f)' = \lambda f'$ et comme $f(0) = \frac{1}{\lambda}Vf(0) = 0$ on en déduit immédiatement que $f = 0$.

Nous venons donc de montrer que $\sigma_p(V) = \emptyset$.

Remarquons ensuite que

$$\begin{aligned} V^2 f(x) &= \int_0^1 K(x,y)\int_0^1 K(y,t)f(t)\,\mathrm{d}t\,\mathrm{d}y \\ &= \int_0^1 \left(\int_0^1 K(x,y)K(y,t)\,\mathrm{d}y\right) f(t)\,\mathrm{d}t \end{aligned}$$

avec Fubini (exercice, le justifier). Mais alors que le noyau de cet opérateur est

$$\begin{aligned} K_2(x,t) = \int_0^1 K(x,y)K(y,t)\,\mathrm{d}y &= \int_0^1 \mathbf{1}_{[0,x]}(y)\mathbf{1}_{[0,y]}(t)\,\mathrm{d}y \\ &= \int_t^x \mathrm{d}y \mathbf{1}_{[0,x]}(t) = \mathbf{1}_{[0,x]}(t)(x-t) \end{aligned}$$

(remarquons que $\mathbf{1}_{[0,x]}(y)\mathbf{1}_{[0,y]}(t) = 0$ sauf si $y < x$ et $t < y$ donc $t < x$ et vaut 1 sinon). Plus généralement, V^n est un opérateur dont le noyau K_n vérifie

$$K_{n+1}(x,t) = \int_0^1 K(x,y)K_n(y,t)\,\mathrm{d}y$$

Par récurrence immédiate avec l'argument de support ci-dessus montre que $K_n(x,t) = 0$ si $t > x$. De plus, si $0 \le K_n(x,t) \le \frac{(x-t)^n}{n!}\mathbf{1}_{[0,x]}(t)$ (qui est vérifiée pour $n = 0$ et $n = 1$) alors clairement $K_{n+1} \ge 0$ et

$$K_{n+1}(x,t) \le \int_t^x \frac{(x-t)^n}{n!}\,\mathrm{d}y\mathbf{1}_{[0,x]}(t) = \frac{(x-t)^{n+1}}{(n+1)!}\mathbf{1}_{[0,x]}(t).$$

Il en résulte que

$$\|V^n\| \leq \|V^n\|_{HS} = \frac{1}{n!}\left(\int_0^1\int_0^x (x-t)^{2n}\,\mathrm{d}t\,\mathrm{d}x\right)^{1/2}$$

$$= \frac{1}{n!}\left(\int_0^1 \frac{x^{2n+1}}{2n+1}\,\mathrm{d}x\right)^{1/2} = \frac{1}{n!\sqrt{(2n+1)(2n+2)}}.$$

En particulier, $\|V^n\|^{1/n} \to 0$ donc $\rho(V) = 0$. Comme le spectre est non vide, il est réduit à $\{0\}$.

Proposition 5.12. *L'opérateur de Voltera est un opérateur compact (non auto-adjoint) sans valeur propre et de spectre réduit à* $\{0\}$.

Notons qu'une fois qu'on a montré que le spectre est réduit à $\{0\}$, la seule valeur propre possible est 0, une partie du raisonnement concernant le spectre ponctuel est donc inutile.

5.3. Théorème spectral des opérateurs compacts auto-adjoints. Le but de cette section est de démontrer que les opérateurs compacts auto-adjoints sont diagonalisables dans une base orthonormée de vecteurs propres. C'est donc le pendant infini-dimensionel du théorème de diagonalisation des matrices symétriques réelles.

Dans toute cette section, on considérera donc un espace de Hilbert H, un opérateur $T : H \to H$ dont on supposera qu'il est *compact* et *auto-adjoint*, $T^* = T$. La diagonalisation de T va résulter de plusieurs observations.

Lemme 5.13. *Soit T un opérateur auto-adjoint T sur H et $E \subset H$ un sous-espace stable par T: $T(E) \subset E$. Alors $E^\perp$ est également stable par T.*

Démonstration. En effet, soit $x \in E^\perp$ et $y \in E$. On a

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, Ty\rangle = 0$$

puisque $Ty \in E$. Ainsi on a bien $Tx \in E^\perp$. □

Lemme 5.14. *Si $\lambda \in \mathbb{C}$ est une valeur propre de l'opérateur auto-adjoint T, alors λ est réelle.*

Démonstration. En effet, si $x \neq 0$ est vecteur propre pour la valeur propre λ, alors

$$\begin{aligned}\lambda\|x\|^2 &= \langle \lambda x, x\rangle = \langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, Tx\rangle \\ &= \langle x, \lambda x\rangle = \bar{\lambda}\|x\|^2.\end{aligned}$$

Comme $x \neq 0$, $\|x\| \neq 0$ donc, en simplifiant par $\|x\|^2$, on trouve $\lambda = \bar{\lambda}$, c'est-à-dire λ réel. □

Si $\lambda \in \mathbb{C}$ est une valeur propre de T, on notera

$$E_\lambda = \{x \in H : Tx = \lambda x\} = \ker(T - \lambda I)$$

le sous-espace propre associé. Notons que E_λ est fermé dans H puisque T est supposé borné (et même compact). On va noter π_λ la projection orthogonale sur E_λ. Rappelons que $I - \pi_\lambda$ est la projection orthogonale sur $E_\lambda^\perp$.

Lemme 5.15. *Si $\lambda \neq \mu \in \mathbb{R}$ sont deux valeurs propres distinctes de l'opérateur auto-adjoint T, alors les espaces propres correspondant E_λ et E_μ sont orthogonaux.*

En particulier, si $\mu \neq 0$, E_μ est aussi le sous-espace propre associé à μ de l'opérateur auto-adjoint $(I - \pi_\lambda)T(I - \pi_\lambda)$.

Démonstration. Soient $x \in E_\lambda$ et $y \in E_\mu$ *i.e.* $Tx = \lambda x$ et $Ty = \mu y$. Alors

$$\begin{aligned}\lambda\langle x, y\rangle &= \langle \lambda x, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, Ty\rangle \\ &= \langle x, \mu y\rangle = \bar{\mu}\langle x, y\rangle = \mu\langle x, y\rangle\end{aligned}$$

puisque μ est réel. Comme $\lambda \neq \mu$, ceci implique $\langle x, y\rangle = 0$. Ainsi E_λ et E_μ sont bien orthogonaux.

Pour la seconde assertion, d'une part, $y \in E_\lambda^\perp$ donc $(I - \pi_\lambda)y = y$ donc

$$(I - \pi_\lambda)T(I - \pi_\lambda)y = (I - \pi_\lambda)Ty = \mu(I - \pi_\lambda)y = \mu y.$$

Inversément, si $(I-\pi_\lambda)T(I-\pi_\lambda)y=\mu y$ alors $y\in\mathrm{Im}\,(I-\pi_\lambda)=E_\lambda^\perp$ donc $(I-\pi_\lambda)y=y$. De plus, E_λ étant stable par T, $E_\lambda^\perp$ aussi, donc $Ty\in E_\lambda^\perp$ et $(I-\pi_\lambda)Ty=Ty$. Au final

$$\mu y=(I-\pi_\lambda)T(I-\pi_\lambda)y=(I-\pi_\lambda)Ty=Ty$$

et y est valeur propre de T. □

Remarque 5.16. Ces deux lemmes nous indiquent une stratégie de diagonalisation d'un opérateur auto-adjoint T.

On commence par observer que $E_0=\ker T$ est stable par T donc $(\ker T)^\perp$ aussi. On remplace donc T par $(I-\pi_0)T(I-\pi_0)$, *i.e.* la restriction de T à $(\ker T)^\perp$. Cela revient à supposer que T est injectif.

On recherche alors une valeur propre λ. Une fois celle-ci trouvée, on remplace T par l'opérateur $(I-\pi_\lambda)T(I-\pi_\lambda)$, c'est-à-dire l'opérateur T restreint à $(E_\lambda)^\perp$. Cet opérateur est encore auto-adjoint car une projection orthogonale est auto-adjointe. En dimension fini, il suffit donc de montrer que tout opérateur auto-adjoint non nul a au moins une valeur propre non-nulle. Le processus s'arrêtera nécessairement puisqu'on enlève au moins 1 à la dimension de l'espace à chaque étape et on obtiendra alors la décomposition de H en sous-espaces propres *i.e.* la diagonalisation de T. En dimension infinie, la situation est plus compliquée, comme le montre le lemme suivant qui nous dit qu'il faudra une infinité d'étapes:

Lemme 5.17. *Si $\lambda\in\mathbb{C}\setminus\{0\}$ une valeur propre de l'opérateur compact T, alors l'espace propre correspondant E_λ est de dimension finie.*

Démonstration. En effet, considérons l'opérateur $\pi_\lambda T$ comme opérateur $T_\lambda: E_\lambda\to E_\lambda$. D'une part, comme λ est valeur propre, cet opérateur est λI_{E_λ} (où I_{E_λ} est l'identié sur E_λ). D'autre part, $\pi_\lambda T$ est compact puisque T est compact. A forciori, T_λ est compact et, comme $\lambda\neq 0$, I_{E_λ}. Cela signifiie que la boule unité B_λ de E_λ est compacte puisque $B_\lambda=I_{E_\lambda}(B_\lambda)$. D'après le théorème de Riesz, ceci implique que E_λ est de dimension finie. □

Démonstration alternative. Soit $(e_k)_{k\in I}$ une base orthonormée de E_λ. Alors $Te_k=\lambda e_k$. Notons que pour $k\neq l$, $\|Te_k-Te_\ell\|=|\lambda|\sqrt{2}\neq 0$. Ainsi, si I était infini, on ne pourrait pas extraire de sous-suite convergente de $(Te_k)_{k\in I}$. Cela contre-dirait alors la compactité de T. □

Montrons maintenant que T a au moins une valeur propre non nulle:

Lemme 5.18. *Si l'opérateur T est compact auto-adjoint, alors soit $\|T\|$, soit $-\|T\|$ est valeur propre de T.*

En particulier, $\rho(T)=\|T\|$.

Démonstration. Le résultat est évident si $T=0$, on suppose donc que $\|T\|\neq 0$. Notons $\bar B=\{x\in H, \|x\|\leq 1\}$; posons $\alpha=\|T\|$ et remarquons que

$$\begin{aligned}\alpha^2 &= \sup\{\|Tx\|^2 : \|x\|=1\}=\sup\{\langle Tx,Tx\rangle : \|x\|=1\}\\ &= \sup\{\langle T^2x,x\rangle : \|x\|=1\}\end{aligned}$$

puisque T est auto-adjoint. Soit (x_n) une suite de vecteurs de norme 1, $\|x_n\|=1$ tel que $\langle T^2x_n,x_n\rangle\to\alpha^2$. Alors

$$\begin{aligned}\left\|T^2x_n-\alpha^2x_n\right\|^2 &= \left\|T^2x_n\right\|^2-2\alpha^2\Re\langle T^2x_n,x_n\rangle+\alpha^4\|x_n\|^4\\ &\leq \|T\|^4\|x_n\|^2-2\alpha^2\Re\langle T^2x_n,x_n\rangle+\alpha^4\|x_n\|^4\\ &= 2\alpha^4-2\alpha^2\Re\langle T^2x_n,x_n\rangle\to 0.\end{aligned}$$

Par ailleurs, T est compact, donc T^2 aussi, donc $T^2x_n\subset T^2(\bar B)$ qui est compact. Quitte à remplacer x_n par une suite extraire, on peut donc supposer que T^2x_n converge. Comme $\left\|T^2x_n-\alpha^2x_n\right\|^2\to 0$ et $\alpha\neq 0$, il en résulte que $x_n=\frac{1}{\alpha}\left(T^2x_n+\alpha^2x_n-T^2x_n\right)$ convege. On note x sa limite, alors $T^2x_n\to T^2x$ donc $T^2x=\alpha^2x$. Notons que, comme $\|x_n\|=1$, $\|x\|=1$ donc $x\neq 0$.

Pour conclure, on écrit $T^2x=\alpha^2x$ sous la forme

$$0=T^2x-\alpha^2x=(T+\alpha I)(T-\alpha I)x.$$

Si $(T - \alpha I)x = 0$, c'est-à-dire $Tx = \alpha x$, alors x est vecteur propre pour la valeur propre α, sinon $y = (T - \alpha I)x \neq 0$ et alors $(T + \alpha I)y = 0$, c'est-à-dire que y est vecteur propre pour la valeur propre $-\alpha$. □

Remarque 5.19. Avec la remarque précédente, on retrouve le fait que tout opérateur auto-adjoint sur un espace de Hilbert de dimension fini est diagonalisable dans une base orthonormée (on prend une base orthonormé de chaque espace propre, leur réunion est alors une base ortonormée de l'espace en entier).

Remarque 5.20. On n'a utilisé l'hypothèse que T est auto-adjoint que dans la dernière étape. Avant cela, on a montré que $\|T\|^2$ est valeur propre de l'opérateur T^*T (et de TT^*).

Nous sommes maintenant en mesure de démontrer le théorème spectral.

Théorème 5.21 (Théorème spectral). *Soit T un opérateur compact auto-adjoint sur un espace de Hilbert H. Alors il existe une suite $(\lambda_n)_{n\in I} \subset \mathbb{R}^*$ décroisssante et soit finie (et alors I est fini) soit tendant vers 0 (alors $I = \mathbb{N}$) telle que*

$$H = E_0 \oplus \bigoplus_{n\in I} E_{\lambda_n}.$$

De façon équivalente, en notant π_λ la projection orthogonale sur E_λ,

$$T = \sum_{n\in I} \lambda_n \pi_{\lambda_n}.$$

De plus, chaque E_{λ_n} étant de dimension finie, il admet une base orthonormée $(e_{n,k})_{k=1,\dots,d_n}$ et alors T s'écrit

$$Tx = \sum_{n\in I} \lambda_n \sum_{k=1}^{d_n} \langle x, e_{n,k}\rangle e_{n,k}.$$

Remarque 5.22. Ce théorème est ennoncé de façon un peu inutilement compliquée mais sous sa forme exacte qui prend en compte que $\ker T$ puisse être de dimension infinie et que T puisse être de rang fini.

La formulation (volontairement vague ici) usuelle de ce théorème est plus ou moins celle-ci: il existe une suite $(\lambda_n)_{n\in\mathbb{N}}$ telle que $|\lambda_n|$ est décroissante et tend vers 0 et une base orthonormée $(e_n)_{n\geq 0}$ de H telle que

$$Tx = \sum_{n\in\mathbb{N}} \lambda_n \langle x, e_n\rangle e_n.$$

Cette formulation est fausse si $\ker T$ et $\operatorname{Im} T$ sont tous deux de dimension infini puisqu'il faut ajouter la base de $\ker T$ après la réunion des bases des espaces propres. Cela suppose aussi que H soit séparable.

Pour éviter cela, on peut supposer que T est injectif (et s'y ramener dans le cas général). Si H est de dimension infini, cela implique que le rang de T soit infini et que la suite des valeurs propres est infinie.

Démonstration. L'écriture

$$H = E_0 \oplus \bigoplus_{n\in I} E_{\lambda_n}. \tag{5.19}$$

signifie que tout $x \in H$ s'écrit

$$x = \pi_0 x + \sum_{n\in I} \pi_{\lambda_n} x$$

et cette série converge dans H (quand la somme est infinie, et la somme est inconditionnellement convergente, c'est-à-dire que l'ordre des E_{λ_n} n'a pas d'importance). Comme T est continu, il en résulte que

$$Tx = T\pi_0 x + \sum_{n\in I} T\pi_{\lambda_n} x = \sum_{n\in I} \lambda_n \pi_{\lambda_n} x.$$

En prenant une base orthonormée de chaque E_{λ_n}, on obtient alors

$$Tx = \sum_{n\in I} \lambda_n \sum_{k=1}^{d_n} \langle x, e_{n,k}\rangle e_{n,k}.$$

Il suffit donc de montrer (5.19). Mais pour cela, il suffit de considérer π_0 la projection orthogonale sur $E_0 = \ker T$. Alors $H = E_0 \oplus E_0^\perp$ et ces deux espaces sont stables par T (puisque E_0 l'est évidemment). On remarque ensuite que $(I-\pi_0)T(I-\pi_0)$ est encore auto-adjoint (calcul direct) et compact (car T l'est, donc si $S = (I-\pi_0)$, ST est compact puis STS aussi). Mais alors, en considérant cet opérateur comme un opérateur $E_0^\perp \to E_0^\perp$, on obtient, on obtient un opérateur auto-adjoint injectif.

Il suffit donc de montrer la version suivante du théorème spectral. □

Théorème 5.23 (Théorème spectral 2). *Soit T un opérateur compact auto-adjoint injectif sur un espace de Hilbert H.*

Alors

– soit H est de dimension fini et il existe alors une base orthonormée $(e_n)_{n\in\{1,\ldots,\dim H\}}$ de H formée de vecteurs propres de T

– soit H est de dimension infini et il existe une suite (λ_n) avec $|\lambda_n|$ décroissante et $\lambda_n \to 0$ et une une base orthonormée $(e_n)_{n\in\mathbb{N}}$et de H telle que

$$Tx = \sum_{n=0}^{+\infty} \lambda_n \langle x, e_n\rangle e_n.$$

En particulier, H est séparable.

Remarque 5.24. Si H est de dimension infini et séparable, le théorème reste vrai même si T n'est pas injectif si on se passe de la conclusion $|\lambda_n|$ décroissante:

Soit T un opérateur compact auto-adjoint injectif sur un espace de Hilbert H séparable, de dimension infinie. Il existe une suite (λ_n) avec $\lambda_n \to 0$ et une une base orthonormée $(e_n)_{n\in\mathbb{N}}$et de H telle que

$$Tx = \sum_{n=0}^{+\infty} \lambda_n \langle x, e_n\rangle e_n.$$

En effet, on prend $(e_n^1)_{n\in\mathcal{N}_1}$ ($\mathcal{N}_1 = \{1,\ldots,N_1\}$ si $\mathcal{N}_1$ fini sinon $\mathcal{N}_1 = \mathbb{N}$ et on pose $N_1 = +\infty$) une base de vecteur propre de l'opérateur T restreint à $(\ker T)^\perp$ et $(e_n^2)_{n\in\mathcal{N}_2}$ ($\mathcal{N}_2 = \{1,\ldots,N_2\}$ si $\mathcal{N}_2$ fini sinon $\mathcal{N}_2 = \mathbb{N}$ et on pose $N_2 = +\infty$) une base de $\ker T$. On réunit alors ces deux bases en insérant les éléments de la deuxième une place sur 2. Formellement, on remarque que $\max(N_1, N_2) = +\infty$ et on définit, pour $n \geq 1$,

$$e_n = \begin{cases} e_p^1 & \text{si } n = 2p-1 \text{ et } p \leq N_1 \\ e_p^2 & \text{si } n = 2p \text{ et } p \leq N_2 \\ e_{n-N_2}^1 & \text{si } N_1 \geq N_2 \text{ et } n > 2N_2 \\ e_{n-N_1}^2 & \text{si } N_2 > N_1 \text{ et } n > 2N_2 \end{cases}.$$

Par exemple si $N_1 = +\infty$ et $N_2 = 2$, c'est la suite $(e_1^1, e_1^2, e_2^1, e_2^2, e_3^1, \ldots)$ et si $N_1 = 2$, $N_2 = +\infty$, c'est la suite $(e_1^1, e_1^2, e_2^1, e_2^2, e_3^2, \ldots)$.

Démonstration. Nous avons vu que soit $\|T\|$, soit $-\|T\|$ est valeur propre et que les espaces propres correspondants sont de dimension fini.

On pose alors

(i) $F_1 = E_{\|T\|} \oplus E_{-\|T\|}$ (rappelons que ces deux espaces sont orthogonaux, l'un des deux pouvant être nul, mais pas les deux);
(ii) $\pi_1^\pm$ la projection orthogonale sur $E_{\pm\|T\|}$ et $\Pi_1 = \pi_1^+ + \pi_1^-$ la projection orthogonale sur F_1;
(iii) $m_1 = \dim E_{\|T\|}$, $m_2 = \dim E_{-\|T\|}$ et $n_1 = m_1 + m_2 \geq 1$;
(iv) $e_1, \ldots, e_{m_1}$ une base orthonormée de $E_{\|T\|}$ (éventuellement vide) et $e_{m_1+1}, \ldots, e_{n_1}$ une base orthonormée de $E_{-\|T\|}$ de sorte que $e_1, \ldots, e_{n_1}$ soit une base orthonormée de F_1;
(v) $\lambda_1 = \cdots = \lambda_{m_1} = \|T\|$ et $\lambda_{n_1+1} = \cdots = \lambda_{n_1} = -\|T\|$.

En particulier, si $x \in F_1$ alors $x = \sum_{k=1}^{n_1} \langle x, e_k \rangle e_k$ donc $Tx = \sum_{k=1}^{n_1} \langle x, e_i \rangle Te_i = \sum_{k=1}^{n_1} \lambda_k \langle x, e_k \rangle e_k$. En particulier F_1 est stable par T donc $F_1^\perp$ aussi. On pose alors $T_0 = T$ et $T_1 = (I - \Pi_1)T_0(I - \Pi_1)$. Notons aussi que $\Pi_1 = \pi_1^+ + \pi_1^-$, $\pi_1^- \pi_1^+ = 0$ donc $(I - \Pi_1) = (I - \pi_1^-)(I - \pi_1^+)$. On constate que T_1 est un opérateur compact auto-adjoint sur H, qu'on peut voir comme l'opérateur T, restreint à $F_1^\perp$. De plus T_1 est soit nul (dans ce cas $H = F_1$ est de dimension finie), soit injectif puisque T est injectif. De plus, tout sous-espace propre de T_1 est aussi sous-espace propre de T.

Il faut ensuite remarquer que $\|T_1\| < \|T\|$. En effet, si ce n'était pas le cas, il existerait un vecteur propre x avec $\|x\| = 1$ de T_1 pour la valeur propre $\|T\|$ (ou $-\|T\|$, l'argument sera la même). Mais alors $x \in \mathrm{Im}\,(I - \Pi_1)$ donc

$$\|T\|x = T_1 x = (I - \Pi_1)T(I - \Pi_1)x = (I - \Pi_1)Tx$$

donc $Tx = \|T\|x + \Pi_1 Tx$. Comme $x \in \mathrm{Im}\,(I - \Pi_1)$, cette décomposition est orthogonale donc

$$\|T\|^2 \geq \|Tx\|^2 = \|T\|^2\|x\| + \|\Pi_1 Tx\|^2 = \|T\|^2 + \|\Pi_1 Tx\|^2$$

d'où $\|\Pi_1 Tx\| = 0$ et finalement $\Pi_1 Tx = 0$ et $Tx = \|T\|x$. Ainsi $x \in E_{\|T\|}$ ce qui contredit $x \in \mathrm{Im}\,(I - \Pi_1)$.

On recommence donc la première étape avec T_1... on fait donc une récurrence, une fois que T_n est construit, soit $\|T_n\|$ soit $-\|T_n\|$ est valeur propre de T_n. On pose alors

(i) $F_{n+1} = E_{\|T_n\|} \oplus E_{-\|T_n\|}$;
(ii) $\pi_{n+1}^\pm$ la projection orthogonale sur $E_{\pm\|T_n\|}$ et Π_{n+1} la projection orthogonale sur F_{n+1};
(iii) $m_{2n+1} = \dim E_{\|T_n\|}$, $m_{2n+2} = \dim E_{-\|T\|}$ et $d_{n+1} = m_{2n+1} + m_{2n+2} \geq 1$;
(iv) $D_n = \sum_{j=1}^{n} d_j$, $(e_j)_{j=D_n+1,\dots,D_n+m_{2n+1}}$ une base orthonormée de $E_{\|T_n\|}$ et $(e_j)_{j=D_n+m_{2n+1}+1,\dots,D_{n+1}}$ une base orthonormée de $E_{-\|T_n\|}$ de sorte que $e_{D_n+1}, \dots, e_{D_{n+1}}$ soit une base orthonormée de F_{n+1};
(v) $\lambda_{d_n+1} = \cdots = \lambda_{d_n+m_{2n+1}} = \|T_{n+1}\|$ et $\lambda_{d_n+m_{2n+1}+1} = \cdots = \lambda_{d_n+d_{n+1}} = -\|T_{n+1}\|$.

On pose alors $T_{n+1} = (I - \Pi_{n+1})T_n(I - \Pi_{n+1})$.

Si H est de dimension fini, ce processus s'arrête si $D_{n+1} = \dim H$ et alors $T_{n+1} = 0$.

Si H est de dimension infini, ce processus ne s'arrête pas et $0 < \|T_{n+1}\| < \|T_n\|$.

À ce stade, on constate

$$\begin{aligned} T_n &= (I - \Pi_n)T_n(I - \Pi_n) = (I - \Pi_n)(I - P_{n-1})T_{n-2}(I - \Pi_{n-1})(I - \Pi_n) \\ &= \left(\prod_{k=1}^{n}(I - \Pi_k)\right) T \left(\prod_{k=1}^{n}(I - \Pi_k)\right) \end{aligned}$$

et on remarque que les Π_k étant des projections sur des espaces 2 à 2 orthogonaux, on a

$$\prod_{k=1}^{n}(I - \Pi_k) = I - \sum_{k=1}^{n+1} \Pi_k = I - P_n.$$

où $H_n := \bigoplus_{k=1}^n F_k$ et P_n désigne la projection orthogonale sur H_n. Ainsi $T_n = (I - P_n)T(I - P_n)$. Notons que H_n est stable par T (c'est une somme directe d'espaces propres) donc $H_n^\perp$ aussi. En particulier, si $x \in H_n$ et $y \in H_n^\perp$, alors Tx et y sont orthogonaux de même que x et Ty le sont. Il en résulte que $(I - P_n)TP_n = P_n T(I - P_n) = 0$ donc

$$T = (P_n + I - P_n)T(P_n + I - P_n) = P_n T P_n + (I - P_n)T(I - P_n) = P_n T P_n + T_n.$$

Par ailleurs, si $x \in H$,

$$P_n x = \sum_{k=1}^{d_n} \langle x, e_k \rangle e_k$$

et

$$TP_n x = P_n T P_n x = \sum_{k=1}^{d_n} \lambda_k \langle x, e_k \rangle e_k.$$

On remarque aussi que $\lambda_{d_n} = \|T_n\|$. Mais

$$|\lambda_n| = |\lambda_n|\|e_n\| = \|\lambda_n e_n\| = \|Te_n\|.$$

Mais (e_n) est une suite(infinie) orthonormée donc l'inégalité de Bessel nous dit que, pour tout $x \in H$,

$$\sum_n |\langle x, e_n\rangle|^2 \leq \|x\|^2 < +\infty.$$

En particulier, $\langle x, e_n\rangle \to 0$ c'est-à-dire e_n tend faiblement vers 0. Comme T est compact, $Te_n \to 0$ *i.e.* $\lambda_n \to 0$. Il en résulte que $\|T - P_nTP_n\| = \|T_n\| \to 0$ soit

$$Tx = \sum_{k=1}^{+\infty} \lambda_k \langle x, e_k\rangle e_k \tag{5.20}$$

comme annoncé.

Reste à voir que $(e_n)_n$ est une base hilbertienne de H (donc que H est séparable). Pour cela, il suffit de voir que l'espace engendré par $(e_n)_n$ est dense, ou encore que son orthogonal est 0. Mais si x est orthogonal à tous les e_k, alors (5.20) montre que $Tx = 0$ et comme T est supposé injectif, $x = 0$. □

Lemme 5.25. *Soient S, T deux opérateurs continus sur H qui commutent $ST = TS$, alors les espaces propres de T sont stables par S.*

En particulier, si S et T sont compacts et auto-adjoints, sur un Hilbert séparable, il existe une base orthonormée commune de vecteurs propres.

Démonstration. En effet, si x est vecteur propre de T pour une valeur propre λ, $Tx = \lambda x$, alors $TSx = STx = S(\lambda x) = \lambda Sx$. Ainsi, soit $Sx = 0$ soit Sx est un vecteur propre pour la valeur propre λ.

Pour la deuxième partie, on écrit

$$H = E_0^T \oplus \bigoplus_{i\in I} E_{\lambda_i}^T$$

où E_λ^T désigne l'espace propre de T pour la valeur propre λ. Alors, comme chaque E_λ est stable par S, en considérant S comme un opérateur $E_\lambda^T \to E_\lambda^T$ (*i.e.* comme $P_{E_\lambda^T} S P_{E_\lambda^T}$), S est un opérateur compact auto-adjoint sur E_λ^T. Il suffit donc de diagonaliser cet opérateur et de prendre une base orthonormée de vecteur propres pour chaque E_λ^T. En réunissant toutes ces bases orthonormées, on obtient une base orthonormée de H.

Alternativement, on peut écrire

$$E_\lambda^T = \bigoplus_{\mu\in I_\lambda} E_\mu^S \cap E_\lambda^T$$

où I_λ désigne l'ensemble des valeurs propres de S restreint à E_λ (cet ensemble est fini si $\lambda \neq 0$) et E_μ^S est l'espace propre de S pour la valeur propre μ. Alors

$$H = \left(\bigoplus_{\mu\in I_0} E_\mu^S \cap E_0^T\right) \oplus \bigoplus_{i\in I} \left(\bigoplus_{\mu\in I_{\lambda_i}} E_\mu^S \cap E_{\lambda_i}^T\right)$$

qui est bien la décomposition en espaces propres communs. □

Rappelons qu'un opérateur est dit *normal* s'il commute avec son adjoint $TT^* = T^*T$ (par exemple s'il est unitaire au quel cas $TT^* = T^*T = I$).

Remarquons que si T est normal alors, pour tout $x \in H$

$$\begin{aligned} \|Tx\|^2 &= \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, TT^*x\rangle = \overline{\langle TT^*x, x\rangle} \\ &= \overline{\langle T^*x, T^*x\rangle} = \|T^*x\| \end{aligned}$$

(ce qui est plus fort que $\|T\| = \|T^*\|$). De plus, si T est normal, alors $T - \lambda I$ est encore normal. En effet, son adjoint est $T^* - \bar\lambda I$ donc

$$\begin{aligned} (T - \lambda I)(T^* - \bar\lambda I) &= TT^* - \lambda T - \bar\lambda T^* - |\lambda|^2 I = T^*T - \lambda T - \bar\lambda T^* - |\lambda|^2 I \\ &= (T^* - \bar\lambda I)(T - \lambda I). \end{aligned}$$

Mais alors, si x est un vecteur propre de T pour la valeur propre λ,

$$0 = \|Tx - \lambda x\|^2 = \left\|T^* x - \bar{\lambda} x\right\|^2$$

donc x est aussi vecteur propre de T^* pour la valeur propre $\bar{\lambda}$.

COROLLAIRE 5.26 (Théorème spectral 3). *Soit T un opérateur compact normal sur un espace de Hilbert H séparable, de dimension infinie. Il existe une suite (λ_n) avec $\lambda_n \to 0$ et une une base orthonormée $(e_n)_{n\in\mathbb{N}}$et de H telle que*

$$Tx = \sum_{n=0}^{+\infty} \lambda_n \langle x, e_n \rangle e_n.$$

DÉMONSTRATION. On décompose $T = A + iB$ avec $A = \frac{T+T^*}{2}$ et $B = \frac{T-T^*}{2i}$ et on remarque que

(i) $A^* = A$ et $B^* = B$ *i.e.* A et B sont auto-adjoints (c'est cette propriété qui détermine A et B);
(ii) A et B sont compact puisque T est compact donc T^* aussi;
(iii) A et B commutent: comme $TT^* = T^*T$, $AB = \frac{T^2-(T^*)^2}{4i} = BA$.

On applique alors le lemme ci-dessus qui nous fournit une base orthonormée commune (e_n) de vecteurs propres: $Ae_n = \alpha_n e_n$, $Be_n = \beta_n e_n$ avec $\alpha_n, \beta_n \to 0$ et alors $Te_n = \frac{\alpha_n + i\beta_n}{2} e_n$ fournit la décomposition cherchée. $\square$