Quelques théorèmes classiques d'analyse fonctionnelle et de dynamique symbolique

Pierre Rousselin

28 Juillet 2014

Mémoire de M1 de Mathématiques Fondamentales de l'Université Paris 13
Sous la direction de Julien Barral

Résumé

Le présent travail est le carnet de bord de mon voyage à la découverte des systèmes dynamiques symboliques, à l'invitation de Julien Barral, et guidé par l'ouvrage de Michel Zinsmeister [Zin96]. On y trouvera les démonstrations des théorèmes suivants :

- le Théorème de récurrence de Poincaré ;
- le Théorème ergodique maximal ;
- le Théorème ergodique de Birkhoff ;
- dans le cas d'un espace métrique compact et d'une transformation continue, la compacité et la métrisabilité de l'ensemble des mesures de probabilité, puis de l'ensemble des mesures invariantes ;
- dans le même cadre, l'existence et la caractérisation des mesures ergodiques ;
- dans le cadre symbolique, le principe variationnel pour la pression et
- le théorème de Perron-Frobenius-Ruelle.

Peu à peu, il est apparu qu'il était nécessaire de faire un détour par l'analyse fonctionnelle pour tenter de comprendre les outils très puissants qu'elle fournit, lesquels étaient très souvent employés dans les démonstrations des résultats ci-dessus. Une annexe assez importante a donc vu le jour, principalement guidée par l'Analyse Fonctionnelle de Walter Rudin ([RA95]). Nous y démontrons :

- le Théorème des prébases d'Alexander et le Théorème de Tychonov sur les produits d'espaces compacts ;
- le Théorème de point fixe de Schauder-Tychonov ;
- les Théorèmes de prolongement et de séparation de Hahn-Banach ;
- le Théorème de Krein-Milman et
- le Théorème de Banach-Alaoglu.

Les Théorèmes de Brouwer, d'Ascoli, de Radon-Nikodym, de Riesz, de Carathéodory et quelques autres résultats de théorie de la mesure sont cités par commodité et utilisés mais ne sont pas démontrés.

Je tiens à remercier Julien Barral de bien avoir voulu encadrer ce mémoire. Plus généralement, je remercie les enseignants-chercheurs du M1 de mathématiques de Paris 13 pour avoir fait une place à l'étudiant atypique que je suis. Merci également à ma famille, mes amis et mes collègues pour le soutien immense qu'ils m'ont apporté pendant cette année de reprise d'études. Parmi eux, je me dois de remercier particulièrement mon ami Pierre Clare qui m'a conseillé, suivi et soutenu depuis le premier partiel jusqu'à la relecture de ce travail.

Merci à Mélanie.

Table des matières

1 Transformations préservant la mesure, ergodicité

1.1 Transformations préservant la probabilité

Définition 1.1.1. On appelle transformation préservant la probabilité (TPP) un quadruplet $(X, \mathcal{B}, \mu, T)$ où $(X, \mathcal{B}, \mu)$ est un espace de probabilité et $T : X \to X$ est une transformation mesurable préservant la mesure μ, i.e.

$$\forall A \in \mathcal{B}, \quad \mu\left(T^{-1}(A)\right) = \mu(A)$$

On appellera parfois X l'*espace des phases* et les éléments $x \in X$ des *états.*
La transformation T doit être vue comme l'évolution d'un système après une unité de temps. Pour un état initial x_0, $x_1 := T(x_0)$ représente l'état du système après une unité de temps, $x_2 = T^2(x_0) = T \circ T(x_0)$ est l'état du système après deux unités de temps, ...
Ainsi, s'intéresser à l'évolution temporelle du système à partir d'un état initial x_0 revient à étudier la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_{n+1} = T(x_n)$.
Poursuivons cette analogie. Soit $E \in \mathcal{B}$, une partie mesurable de l'espace des phases.

- L'événement $T^{-n}(E)$ est l'ensemble des états qui seront dans E au temps n.
- L'événement $E_N := \bigcup\limits_{n \geq N} T^{-n}(E)$ est l'ensemble des états qui seront dans E à un temps supérieur ou égal à N.
  On peut déjà remarquer que :
  - $E_{N+1} \subset E_N$
  - $E_{N+1} = T^{-1}(E_N)$ donc, comme T préserve la mesure, $\mu(E_{N+1}) = \mu(E_N)$
  - En particulier $E_0 = E \cup T^{-1}(E) \cup T^{-2}(E) \cup ...$ est l'ensemble des états qui atteignent E.
- L'événement $E_\infty := \bigcap\limits_{N\in\mathbb{N}} E_N$ est l'ensemble des états qui passeront infiniment souvent dans E.
  Vu les remarques précédentes, on a :
  - $T^{-1}(E_\infty) = E_\infty$.
  - Comme la suite d'événements $(E_N)_{N\in\mathbb{N}}$ est décroissante,

$$\begin{aligned}\mu(E_\infty) = \mu\left(\bigcap_{N\in\mathbb{N}} E_N\right) &= \lim_{N\to+\infty} \mu(E_N) \\ &= \mu(E_0) = \mu\left(E \cup T^{-1}(E) \cup T^{-2}(E) \cup ...\right).\end{aligned}$$

  L'événement E_∞ étant évidemment inclus dans E_0, on obtient $\mu(E_0 \setminus E_\infty) = 0$, et nous venons de démontrer le :

Théorème 1.1.1 (Théorème de récurrence de Poincaré (1890))**.** *Soit $E \in \mathcal{B}$. Presque tout état qui atteint E y retournera infiniment souvent. En particulier, pour presque tout état initial $x_0 \in E$, la suite (x_n) retourne infiniment souvent dans E.*

Ce beau théorème soulève quelques questions :
Premièrement, si $\mu(E) > 0$, on sait maintenant que presque tout x de E y retournera infiniment souvent. Qu'en est-il des états qui ne sont pas dans E ?
Deuxièmement, il ne dit rien sur la *fréquence* des retours en E.
Le premier point motivera la définition de transformation ergodique, qui est au coeur de ce mémoire.
Pour cette classe de transformation, le second point a été résolu par Birkhoff en 1931 : pour presque tout état initial x, la fréquence (donc la moyenne temporelle) de passage de (x_n) dans E est égal à la mesure de E.

1.2 Systèmes dynamiques symboliques

Nous définissons maintenant le cadre principal de ce travail.

Soit $\Sigma = \{\alpha_1; \alpha_2; ...; \alpha_d\}$ un ensemble fini, appelé *alphabet* et dont les éléments sont appelés des *lettres*, et $X = \Sigma^{\mathbb{N}}$ l'ensemble des suites d'éléments de Σ. On appellera de telles suites des *mots infinis*.
Sur X, on considère la distance ultramétrique d définie par :

$$\forall a = a_0a_1a_2..., \quad b = b_0b_1b_2... \in X, \quad d(a,b) = 2^{-\inf\{k\geq 0 \ | \ a_k \neq b_k\}}$$

Définition 1.2.1. Soient n un entier naturel et a_0, a_1, ..., a_n des lettres. On note $[a_0a_1...a_n]$ l'ensemble de mots défini par :

$$[a_0a_1...a_n] = \{x \in X \ | \ \forall 0 \leq i \leq n, x_i = a_i\}$$

On appelle *cylindre* de précision $n+1$ un tel ensemble. Par commodité, nous incluons X et $\emptyset$ dans l'ensemble des cylindre.

La proposition suivante ne fait qu'énoncer certaines propriétés évidentes mais utiles.

Proposition 1.2.1. *1.* $[a_0a_1...a_n] = \overline{B}\left(a, 2^{-(n+1)}\right) = B(a, 2^{-n})$, *où a désigne tout mot commençant par $a_0a_1...a_n$.*

2. *Soit $\varepsilon > 0$ et n_ε l'entier tel que $2^{-(n_\varepsilon+1)} < \varepsilon \leq 2^{-n_\varepsilon}$. Soit $a = a_0a_1...a_n....$ Alors,*

$$B(a,\varepsilon) = \overline{B}\left(a, 2^{-(n_\varepsilon+1)}\right) = B\left(a, 2^{-n_\varepsilon}\right)$$

3. *Toutes les boules sont à la fois ouvertes et fermées et coïncident avec les cylindres.*
4. *Les cylindres forment une base dénombrable de voisinages. X est séparable et une partie dénombrable dense est donnée par le choix d'un mot par cylindre.*
5. *Deux cylindres sont soit emboîtés, soit disjoints.*

Remarque 1.2.1. Le point 3. de la dernière proposition a une conséquence remarquable : les fonctions indicatrices des boules, c'est-à-dire des cylindres sont continues.

Proposition 1.2.2. *Muni de sa topologie métrique, X est compact.*

Démonstration. Soit $\left(a^{(n)}\right)_{n\in\mathbb{N}}$ une suite de mots. Comme l'alphabet Σ est fini, d'après le principe des tiroirs, on peut en extraire une sous-suite $\left(a^{\varphi_0(n)}\right)$ de mots commençant par la même lettre ℓ_0.
En réitérant ce procédé, on peut pour tout naturel s, aboutir à une suite extraite

$$\left(a^{\varphi_0\circ\varphi_1\circ\ldots\circ\varphi_s(n)}\right)$$

de mots commençant tous par le même préfixe $\ell_0\ell_1\ldots\ell_s$. Il suffit alors de poser $\psi(n) = \varphi_0\circ\varphi_1\circ\ldots\circ\varphi_n(n)$ et la suite extraite $\left(a^{\psi(n)}\right)$ converge vers le mot $a = \ell_0\ell_1\ldots\ell_n\ldots$. □

Maintenant que nous avons défini la topologie de X, il faut nous poser la question de la mesure.

Remarque 1.2.2. Comme les cylindres forment une base dénombrable de voisinages ouverts, la tribu borélienne de X est la tribu engendrée par les cylindres.

Pour se donner une mesure de probabilité borélienne μ sur X, il est nécessaire d'avoir :

$$\sum_{i=1}^{d}\mu\left([\alpha_i]\right) = 1 \quad \text{et,}$$

$$\forall n\in\mathbb{N},\ \forall(a_0,a_1,\ldots,a_n)\in\Sigma^{n+1}, \sum_{i=1}^{d}\mu\left([a_0a_1\ldots a_n\alpha_i]\right) = \mu\left([a_0a_1\ldots a_n]\right) \qquad (1)$$

L'objet de la discussion qui va suivre est de montrer que la condition (1) est suffisante.
Nous aurons besoin du théorème de Carathéodory, dont la démonstration peut être trouvée, par exemple dans le manuel de théorie de l'intégration [BP00].

Définition 1.2.2. Une famille $\mathcal{C}$ de parties de X est une *algèbre d'ensembles* ssi $X\in\mathcal{C}$ et $\mathcal{C}$ est stable par *union finie, intersection finie, et passage au complémentaire.*

Théorème 1.2.3 (Théorème d'extension de Carathéodory)**.** *Soit $\mathcal{C}$ une algèbre d'ensembles de X. Soit $\mu:\mathcal{C}\longrightarrow\mathbb{R}^+$ telle que :*

1. *$\mu(\emptyset) = 0$ et*
2. *Pour toute union dénombrable disjointe $\bigsqcup_{i\in\mathbb{N}} A_i$, d'éléments de $\mathcal{C}$ telle que $\bigsqcup_{i\in\mathbb{N}} A_i\in\mathcal{C}$, on ait*

$$\mu\left(\bigsqcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}}\mu\left(A_i\right)$$

Alors, μ se prolonge à la tribu engendrée $\sigma(\mathcal{C})$. De plus, si μ est finie ou σ-finie, ce prolongement est unique.

Proposition 1.2.4. *Toute fonction d'ensembles μ définie sur les cylindres et vérifiant la condition* (1) *s'étend de façon unique à une mesure de probabilité borélienne sur X.*

Notons $\mathcal{A}$ la famille des cylindres et $\mathcal{C}$ l'algèbre qu'ils engendrent. Le théorème de Carathéodory nous fournira directement le résultat, une fois établis les deux lemmes qui vont suivre.

Lemme 1.2.5. *La famille $\mathcal{A}$ est telle que :*

1. $\emptyset \in \mathcal{A}$*.*

2. $\mathcal{A}$ *est stable par intersection finie.*

3. Le complémentaire de tout élément de $\mathcal{A}$ *est une union finie disjointe d'éléments de* $\mathcal{A}$*.*

4. Pour toute union finie disjointe $A = \bigsqcup_{i=1}^{r} A_i$*, d'éléments de* $\mathcal{A}$*, telle que* $A \in \mathcal{A}$*, on a*

$$\mu(A) = \sum_{i=1}^{r} \mu(A_i)$$

Autrement dit, $\mathcal{A}$ *est une* semi-algèbre, sur laquelle μ est additive.

Démonstration. Les points 1. et 2. sont évidents. Pour démontrer le point 3, considérons le cylindre $A = [a_0 a_1 ... a_n]$. Nous pouvons écrire son complémentaire sous la forme

$$A^c = \left(\bigsqcup_{\alpha_i \neq a_0} [\alpha_i] \right) \bigsqcup \left(\bigsqcup_{\alpha_i \neq a_1} [a_0 \alpha_i] \right) \bigsqcup ... \bigsqcup \left(\bigsqcup_{\alpha_i \neq a_n} [a_0 a_1 ... a_{n-1} \alpha_i] \right)$$

Pour le point 4, soit $A = [a_0 a_1 ... a_n]$ et une union disjointe d'éléments de $\mathcal{A}$, $\bigsqcup_{i \in I}^{r} A_i = A$, où I est fini de cardinal r.
Si $r = 1$, il n'y a rien à démontrer. Si $r \geq 2$, alors aucun des A_i n'est inclus dans A mais chacun d'entre eux est inclus dans un et un seul des cylindres $[a_0 a_1 ... a_n \alpha_j]$. Partitionnons l'ensemble d'indice I en réunion des

$$I_j = \{ i \in I \quad | \quad A_i \subset [a_0 a_1 ... a_n \alpha_j] \}$$

$$A = \bigsqcup_{j=1}^{d} \left(\bigsqcup_{i \in I_j} A_i \right)$$

Comme $[a_0 a_1 ... a_n \alpha_j]$ est inclus dans A mais disjoint de tous les $\bigsqcup_{i \in I_{j'}} A_i$ pour $j' \neq j$, on en déduit :

$$\forall 1 \leq j \leq d, \quad \left(\bigsqcup_{i \in I_j} A_i \right) = [a_0 a_1 ... a_n \alpha_j]$$

Par conséquent, pour tout $1 \leq j \leq d$, $1 \leq \text{Card}(I_j) < r$.
Si tous les I_j sont de cardinal 1, la partition de A n'est autre que

$$[a_0 a_1 ... a_n] = \bigsqcup_{j=1}^{d} [a_0 a_1 ... a_n \alpha_j],$$

et la propriété d'additivité est vraie par hypothèse.
Sinon, on raisonne par récurrence forte sur la taille r des partitions, pour obtenir

$$\forall 1 \leq j \leq d, \quad \mu([a_0 a_1 ... a_n \alpha_j]) = \mu \left(\bigsqcup_{i \in I_j} A_i \right) = \sum_{i \in I_j} \mu(A_i)$$

Finalement,

$$\mu(A)=\sum_{j=1}^{d}\mu\left([a_0a_1...a_n\alpha_j]\right)=\sum_{j=1}^{d}\sum_{i\in I_j}\mu\left(A_i\right)=\sum_{i\in I}\mu\left(A_i\right).$$

$\square$

Lemme 1.2.6. *1. L'algèbre $\mathcal{C}$ est l'ensemble des unions disjointes finies de cylindres.*

2. On peut étendre sans ambigüité μ à $\mathcal{C}$ en posant, pour tout union disjointe finie de cylindres,

$$A=\bigsqcup_{i=1}^{r}A_i,\quad \mu\left(A\right)=\sum_{i=1}^{r}\mu\left(A_i\right)$$

3. Pour toute union dénombrable disjointe $\bigsqcup_{i\in\mathbb{N}}A_i$, d'éléments de $\mathcal{C}$ telle que $A=\bigsqcup_{i\in\mathbb{N}}A_i\in\mathcal{C}$, on a

$$\mu\left(\bigsqcup_{i\in\mathbb{N}}A_i\right)=\sum_{i\in\mathbb{N}}\mu\left(A_i\right)$$

Démonstration. 1. Ce premier point est évident grâce aux propriétés de $\mathcal{A}$.

2. Soit $A=\bigsqcup_{i=1}^{r}A_i=\bigsqcup_{j=1}^{s}A'_j$. Alors pour tout $1\leq i\leq r$, on a

$$A_i=\bigsqcup_{j=1}^{s}\left(A_i\cap A'_j\right)\qquad\text{donc, d'après le lemme précédent,}\qquad \mu\left(A_i\right)=\sum_{j=1}^{s}\mu\left(A_i\cap A'_j\right)$$

De la même manière, pour tout $1\leq j\leq s$,

$$\mu\left(A'_j\right)=\sum_{i=1}^{r}\mu\left(A_i\cap A'_j\right)$$

$$\text{Donc,}\quad \sum_{i=1}^{r}\mu\left(A_i\right)=\sum_{i=1}^{r}\sum_{j=1}^{s}\mu\left(A_i\cap A'_j\right)=\sum_{j=1}^{s}\sum_{i=1}^{r}\mu\left(A_i\cap A'_j\right)=\sum_{j=1}^{s}\mu\left(A'_j\right)$$

3. Comme A est dans $\mathcal{C}$, c'est un fermé de X, comme réunion finie de fermés, donc un compact. Les A_i sont ouverts comme union d'ouverts donc par compacité, on peut extraire du recouvrement de A par les A_i un recouvrement fini. Les A_i étant disjoints, ceux qui sont non vides sont en nombre fini. La σ-additivité revient donc à montrer l'additivité finie dans $\mathcal{C}$, qui est maintenant évidente.

$\square$

Pour finir cette première visite de $X=\Sigma^{\mathbb{N}}$, étudions la transformation T la plus naturelle sur X : le *décallage*, ou *shift*, défini par :

$$\forall x=x_0x_1...x_n...\in X,\qquad T(x)=x_1x_2...x_n...$$

Proposition 1.2.7. *1. La transformation T est continue sur X.*

2. Une mesure μ est invariante par rapport à T si et seulement si :

$$\forall A \in \mathcal{A}, \qquad \mu(A) = \sum_{\ell \in \Sigma} \mu(\ell A)$$

Démonstration. 1. Le shift est 2-lipschitzien. En effet, si $d(x,y) = 2^{-k-1}$, alors $d(Tx, Ty) = 2^{-k}$.

2. La condition donnée est clairement nécessaire. Elle est de plus suffisante car dans ce cas $\mathcal{S}$ est un π-système (i.e. est un ensemble de parties stable par intersections finies) qui engendre la tribu borélienne $\mathcal{B}(X)$ de X, et sur lequel coïncident les mesures μ et μ_T. (Voir, par exemple, [Wil91]).

□

1.3 Transformations ergodiques, mesures ergodiques

Revenons au théorème de récurrence de Poincaré.
Nous avions $T^{-1}(E_\infty) = E_\infty$ et $\mu(E_\infty) = \mu(E_0) \geq \mu(E)$.
Si E est de mesure nulle, il en sera de même pour E_0 (union dénombrable d'événements de mesure nulle) et donc pour E_∞.
Lorsque E est de mesure non nulle, il arrive que presque tous les états de X (et pas seulement ceux de E ou de E_0) visitent E infiniment souvent, i.e. $\mu(E_\infty) = 1$.

Définition 1.3.1. Lorsque ce sera le cas pour tous les événements E de mesure non nulle, nous dirons que la transformation T est *ergodique*.

Théorème 1.3.1. *Les propriétés suivantes sont équivalentes :*

1. *T est ergodique.*
2. $\forall E \in \mathcal{B},\ \ \mu(E) > 0 \Rightarrow \mu\left(\bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} T^{-n}(E)\right) = 1.$
3. $\forall E \in \mathcal{B},\ \ \mu(E) > 0 \Rightarrow \mu\left(\bigcup_{n\geq 1} T^{-n}(E)\right) = 1.$
4. $\forall A, B \in \mathcal{B},\ \ \mu(A) > 0, \mu(B) > 0 \Rightarrow \exists n > 0, \mu\left(T^{-n}(A) \cap B\right) > 0.$
5. $\forall f : X \longrightarrow \mathbb{R},$ *mesurable*, $f \circ T = f$ *μ-p.p* $\Rightarrow \exists a \in \overline{\mathbb{R}},\ \ f = a$ *μ-p.p.*
6. $\forall f \in L^2(\mu),\ \ f \circ T = f \Rightarrow \exists a \in \mathbb{R},\ \ f = a$ *μ-p.p.*
7. $\forall E \in \mathcal{B},\ \ \mu\left(T^{-1}(E)\Delta E\right) = 0 \Rightarrow \mu(E) = 0$ *ou* 1.
8. $\forall E \in \mathcal{B},\ \ T^{-1}(E) = E \Rightarrow \mu(E) = 0$ *ou* 1.

Démonstration.

- 1) $\Longleftrightarrow$ 2) : C'est la définition que nous avons choisie.
- 2) $\Rightarrow$ 3) : Presque tous les états $x \in X$ visitent infiniment souvent E, donc presque tous les états visitent E à un temps ≥ 1.

- 3) $\Rightarrow$ 4) : Comme $\mu\left(\bigcup_{n\geq 1} T^{-n}(A)\right) = 1$, on a $\mu\left(\bigcup_{n\geq 1} T^{-n}(A)\cap B\right) = \mu(B) > 0$.

  Or,

$$\mu\left(\bigcup_{n\geq 1} T^{-n}(A)\cap B\right) \leq \sum_{n\geq 1} \mu\left(T^{-n}(A)\cap B\right)$$

  de sorte que pour au moins un entier $n \geq 1$, on a

$$\mu\left(T^{-n}(A)\cap B\right) > 0$$

- 4) $\Rightarrow$ 5) : Soit $k \in \mathbb{R}$. Nous allons montrer que l'une des mesures, $\mu\left(f < k\right)$ ou $\mu\left(f > k\right)$ est nulle.

  Supposons que ce n'est pas le cas et notons $A = \{f > k\}$, $B = \{f < k\}$ et $X' = \{f \circ T = f\}$. Par hypothèse, $\mu\left(X'\right) = 1$, $\mu(A) > 0$ et $\mu(B) > 0$. Appliquant 3) à A et $B \cap X'$, on voit qu'il existe un entier $n \geq 1$ tel que $\mu\left(T^{-n}(A)\cap B \cap X'\right) > 0$.

  Or, si $x \in X' \cap T^{-n}(A)$, on a $f(x) = f\left(T(x)\right) = f\left(T^2(x)\right) = ... = f\left(T^n(x)\right)$ et $f\left(T^n(x)\right) > k$, donc $f(x) > k$. Par conséquent, l'intersection $T^{-n}(A)\cap B\cap X'$ est vide, ce qui est contradictoire.

  On a donc prouvé que pour chaque réel k, on a $\mu\left(f < k\right) = 0$ ou $\mu\left(f > k\right) = 0$.

  Posons $a = \sup\left\{k \in \mathbb{R} : \mu\left(f < k\right) = 0\right\}$. Si la partie $\left\{k \in \mathbb{R} : \mu\left(f < k\right) = 0\right\}$ n'est pas majorée, alors f est presque sûrement égale à $+\infty$ et le résultat est démontré.

  De même, considérons $b = \inf\left\{k \in \mathbb{R} : \mu\left(f > k\right) = 0\right\}$.

  Notons que $a \leq b$, sinon pour tout $\alpha \in]b; a[$, on a $\mu\left(f = \alpha\right) = 1$, ce qui n'est clairement pas possible.

  Puis, si $a < b$, alors pour $\alpha \in]a; b[$, on aura $\mu\left(f < \alpha\right) > 0$ et $\mu\left(f > \alpha\right) > 0$, ce qui n'est pas possible d'après la discussion précédente. On a donc $a = b$ et pour tout $r \in \mathbb{N}^*$, $\mu\left(f \in \left]a - \frac{1}{r}; a + \frac{1}{r}\right[\right) = 1$, donc $\mu\left(f = a\right) = 1$.

- 5) $\Rightarrow$ 6) : évident. Notons qu'on aurait pu choisir n'importe quel L^p, pour $p \in [1; +\infty]$.
- 6) $\Rightarrow$ 7) : L'indicatrice 1_E est dans $L^2(\mu)$ par finitude de μ. Si $\mu\left(T^{-1}(E)\Delta E\right) = 0$, alors les fonctions 1_E et $1_E \circ T$ sont égales presque partout, donc en appliquant 6), 1_E est constante presque partout, ce qui est équivalent à $\mu(E) = 0$ ou 1.
- 7) $\Rightarrow$ 8) : évident. Si $T^{-1}(E) = E$, alors $T^{-1}(E)\Delta E = \emptyset$.
- 8) $\Rightarrow$ 2) : Notons encore $E_\infty = \bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} T^{-n}(E)$.

  On sait déjà que $T^{-1}\left(E_\infty\right) = E_\infty$, donc $\mu\left(E_\infty\right) = 0$ ou 1 et comme $\mu\left(E_\infty\right) \geq \mu(E)$ d'après le théorème de récurrence de poincaré, on en déduit que $\mu\left(E_\infty\right) = 1$.

$\square$

Remarque 1.3.1. Habituellement, la définition de l'ergodicité est la propriété 8). Il nous paraissait plus intuitif de partir du point de vue de la récurrence. Comme nous le voyons, ces deux définitions coïncident.

Remarque 1.3.2. Lorsqu'on change de point de vue et que l'on fixe la transformation T, on parle de *mesure ergodique* par rapport à T.

2 Théorèmes ergodiques

2.1 La tribu des invariants

Soit $(X, \mathcal{B}, \mu, T)$ une TPP. Notons :

$$\mathfrak{Inv}T = \left\{ B \in \mathcal{B} \mid T^{-1}(B) = B \right\}$$

Proposition 2.1.1. *1. $\mathfrak{Inv}T$ est une sous-tribu de $\mathcal{B}$.*

2. La transformation T est ergodique ssi $\mathfrak{Inv}T$ est n'est composée que d'événements de mesure 0 ou 1.

3. Soit $f : X \longrightarrow \mathbb{R}$ une fonction mesurable. Alors f est $\mathfrak{Inv}T$-mesurable ssi $f \circ T = f$.

Démonstration. Les points 1. et 2. sont évidents. Démontrons le point 3.

$\Leftarrow$: Soient f une fonction T-invariante et A un borélien de $\mathbb{R}$. On a

$$T^{-1}\left(f^{-1}(A)\right) = (f \circ T)^{-1}(A) = f^{-1}(A)$$

La tribu $\sigma(f)$ est donc incluse dans $\mathfrak{Inv}T$ et f est $\mathfrak{Inv}T$-mesurable.

$\Rightarrow$: Soit f une fonction $\mathfrak{Inv}T$-mesurable. Par définition, pour tout borélien A de $\mathbb{R}$, on a $f^{-1}(A) \in \mathfrak{Inv}T$, c'est-à-dire $T^{-1}\left(f^{-1}(A)\right) = f^{-1}(A)$. En particulier, pour tout $c \in \mathbb{R}$, on a

$$\{f \circ T < c\} = T^{-1}(\{f < c\}) = \{f < c\}$$

On en déduit que $\{f \circ T < c, f > c\} = \emptyset$, et finalement que

$$\{f \circ T < f\} = \bigcup_{c \in \mathbb{Q}} \{f \circ T < c, f > c\} = \emptyset$$

On raisonne de même pour l'autre inégalité et on obtient le résultat voulu. □

2.2 Théorème ergodique maximal

Le théorème suivant nous servira à établir le théorème de Birkhoff. Il a été démontré par Yosida et Kakutani en 1939 [YK+39]. Sa forme et sa démonstration nous viennent de Garcia [Gar65].

Dans cette partie, l'espace mesuré $(X, \mathcal{B}, \mu)$ est supposé de mesure finie.

Remarque 2.2.1. Soit T une transformation préservant la mesure μ. Le point de vue adopté est de considérer l'opérateur linéaire $U : L^1(X) \longrightarrow L^1(X)$ défini par $Uf = f \circ T$. Comme T conserve la mesure,

$$\|Uf\|_1 = \int |Uf| \, d\mu = \int |f \circ T| \, d\mu = \int |f| \, d\mu_T = \int |f| \, d\mu = \|f\|_1$$

L'opérateur U est donc une isométrie.

Le théorème ergodique maximal et ses corollaires sont en fait vérifiés pour une classe bien plus importante d'opérateurs. Nous les énoncerons et les démontrerons dans ce cadre plus général. Avant cela, nous aurons besoin de quelques définitions.

Définition 2.2.1. Soit U un endomorphisme de $L^1(X)$.

- On dit que U est *positif* si, pour tout $f \in L^1(X)$, $f \geq 0 \Rightarrow Uf \geq 0$.
- Soit A une partie mesurable de X. On dira que la partie A est invariante par l'opérateur U ssi pour toutes f et g dans $L^1(X)$

  $$f_{|A} = g_{|A} \ \ \mu\text{-p.p} \Rightarrow (Uf)_{|A} = (Ug)_{|A} \ \ \mu\text{-p.p}$$

  Dans ce cas, on peut définir sans ambigüité un endomorphisme de $L^1(A)$ qu'on note U_A et qu'on définit par :

  $$\forall f \in L^1(A),\ \ U_A f = \left(U\tilde{f}\right)_{|A},\ \text{ où } \tilde{f} \text{ est toute fonction dans } L^1(X) \text{ telle que } \tilde{f}_{|A} = f$$

Théorème 2.2.1 (Théorème ergodique maximal)**.** *Soit $U : L^1(\mu) \longrightarrow L^1(\mu)$ un opérateur linéaire positif continu, de norme $\|U\| \leq 1$.*
Soit $f \in L^1(\mu)$. Pour n entier ≥ 1, on note $S_n^U f = f + Uf + U^2 f + ... + U^{n-1} f$, $M_n^U f = max\left(0, S_1^U f, S_2^U f, ..., S_n^U f\right)$ et
$M^U f = \sup\left\{0, S_1^U f, S_2^U f, ..., S_n^U f, ...\right\}$. *Alors,*

$$\forall n \geq 1, \ \int_{\{M_n^U f > 0\}} f d\mu \geq 0 \quad \text{ et } \quad \int_{\{M^U f > 0\}} f d\mu \geq 0$$

Démonstration. • Pour la première inégalité :
Soient n et N deux entiers ≥ 1 avec $N \geq n$. Comme $S_n^U f \leq M_N^U f$, on a, par positivité de l'opérateur U,

$$U\left(S_n^U f\right) \leq U\left(M_N^U f\right) \quad \text{donc} \quad \forall 1 \leq n \leq N, \quad U\left(M_N^U f\right) + f \geq U\left(S_n^U f\right) + f = S_{n+1}^U f$$

On a donc, $U\left(M_N^U f\right) + f \geq \max\left(S_2^U f, S_3^U f, ..., S_N^U f\right)$ et comme $S_1^U f = f$ et $U\left(M_N^U f\right) \geq 0$,

$$U\left(M_N^U f\right) + f \geq \max\left(S_1^U f, S_2^U f, S_3^U f, ..., S_N^U f\right)$$

Finalement, sur l'ensemble mesurable $\left\{M_N^U f > 0\right\}$, on a

$$M_N^U f = \max\left(S_1^U f, S_2^U f, S_3^U f, ..., S_N^U f\right),$$

donc

$$U\left(M_N^U f\right) + f \geq M_N^U f$$

ou encore :

$$\text{sur} \quad \left\{M_N^U f > 0\right\}, \quad f \geq M_N^U f - U\left(M_N^U f\right)$$

En intégrant cette inégalité sur $\left\{M_N^U f > 0\right\}$, il vient :

$$\int_{\{M_N^U f > 0\}} f d\mu \geq \int_{\{M_N^U f > 0\}} M_N^U f d\mu - \int_{\{M_N^U f > 0\}} U\left(M_N^U f\right) d\mu$$

Comme $M_N^U f \geq 0$, $\int_{\{M_N^U f>0\}} M_N^U f d\mu = \int_X M_N^U f d\mu$. Par ailleurs, comme $U\left(M_N^U f\right) \geq 0$,

$$\int_{\{M_N^U f>0\}} U\left(M_N^U f\right) d\mu \leq \int_X U\left(M_N^U f\right) d\mu \leq \left\|U\left(M_N^U f\right)\right\|$$
$$\leq \|U\| \left\|M_N^U f\right\| \leq 1 \times \left\|M_N^U f\right\| = \int_X M_N^U f d\mu.$$

On en déduit l'inégalité cherchée.

- Pour la seconde inégalité :
On sait maintenant que pour chaque $n \geq 1$, $\int f\mathbf{1}_{\{M_n^U f>0\}} d\mu \geq 0$. Comme la suite de mesurables $\left(\{M_n^U f > 0\}\right)$ est croissante, on a la convergence simple

$$\lim_{N\to+\infty} f\mathbf{1}_{\{M_n^U f>0\}} = f\mathbf{1}_{\cup_{n\geq 1}\{M_n^U f>0\}} = f\mathbf{1}_{\{M^U f>0\}}$$

La fonction f étant intégrable, on peut conclure par convergence dominée.

$\square$

Corollaire 2.2.2. *Soient $f \in L^1(\mu)$, et α et β deux réels. Notons $B_\alpha = \left\{\sup_{n\geq 1} \frac{S_n^U f}{n} > \alpha\right\}$ et $C_\beta = \left\{\inf_{n\geq 1} \frac{S_n^U f}{n} < \beta\right\}$. Alors, on a :*

1. $\int_{B_\alpha} f d\mu \geq \alpha\mu\left(B_\alpha\right)$
2. $\int_{C_\beta} f d\mu \leq \beta\mu\left(C_\beta\right)$
3. *Pour tout ensemble mesurable A, invariant par l'opérateur U,* $\int_{B_\alpha\cap A} f d\mu \geq \alpha\mu\left(B_\alpha \cap A\right)$ *et* $\int_{C_\beta\cap A} f d\mu \leq \beta\mu\left(C_\beta \cap A\right)$
4. *En particulier, pour tout ensemble mesurable A, invariant par l'opérateur U, contenu dans B_α et dans C_β, on a :*

$$\alpha\mu(A) \leq \int_A f d\mu \leq \beta\mu(A)$$

Démonstration. 1. On pose $g = f - \alpha$. La fonction g est encore dans $\mathrm{L}^1(\mu)$ par finitude de μ.

Par linéarité des S_n^U, on a $\frac{S_n^U f}{n} > \alpha \Leftrightarrow S_n^U g > 0$, si bien que

$$x \in B_\alpha \Longleftrightarrow \exists n \geq 1,\ \ S_n^U g(x) > 0 \Longleftrightarrow x \in \left\{M^U g > 0\right\}$$

On applique le théorème maximal à g pour obtenir

$$\int_{B_\alpha} g d\mu \geq 0$$

Revenant à f, on obtient $\int_{B_\alpha} f d\mu \geq \int_{B_\alpha} \alpha d\mu$ d'où le résultat.

2. Commençons par remarquer que, par linéarité des S_n^U,

$$\left\{\inf_{n\geq 1}\frac{S_n^U f}{n}<\beta\right\}=\left\{-\inf_{n\geq 1}\frac{S_n^U f}{n}>-\beta\right\}=\left\{\sup_{n\geq 1}-\frac{S_n^U f}{n}>-\beta\right\}=\left\{\sup_{n\geq 1}\frac{S_n^U(-f)}{n}>-\beta\right\}$$

Il ne reste qu'à appliquer le point précédent à $-f$ et $-\beta$ pour obtenir $\int_{C_\beta}(-f)d\mu \geq -\beta\mu\left(C_\beta\right)$, puis le résultat.

3. Comme A est invariant par l'opérateur U, il l'est aussi par tous les S_n^U et on a

$$\begin{aligned}B_\alpha\cap A&=\left\{x\in A\ \mid\ \sup_{n\geq 1}\frac{\left(S_n^U f\right)(x)}{n}>\alpha\right\}=\left\{x\in A\ \mid\ \sup_{n\geq 1}\frac{\left(S_n^U f\right)_{|A}(x)}{n}>\alpha\right\}\\&=\left\{x\in A\ \mid\ \sup_{n\geq 1}\frac{\left(S_n^{U_A} f_{|A}\right)(x)}{n}>\alpha\right\}\end{aligned}$$

Il suffit maintenant d'appliquer les points 1. et 2. à l'espace mesuré A, pour l'opérateur U_A et la fonction $f_{|A}$.

4. Le point 3. donne directement le résultat.

□

2.3 Théorème ergodique ponctuel de Birkhoff

Théorème 2.3.1 (Théorème ergodique ponctuel de Birkhoff, 1931 [Bir31])**.** *Soit* $(X,\mathcal{B},\mu)$ *un espace probabilisé et* T *une transformation préservant la mesure. Pour toute fonction* $f\in L^1(\mu)$*, on a :*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f\circ T^k=\mathbb{E}\left[f\mid \mathfrak{Inv}T\right]\quad \mu\text{-}p.p$$

Démonstration. Nous allons nous servir à plusieurs reprises des points 3. et 4. du corollaire 2.3.2. L'opérateur U est défini par $Uf=f\circ T$. C'est un opérateur de norme 1, positif, et si un mesurable A est invariant par T, alors il l'est par l'opérateur U.

Les opérateurs S_n, ainsi que les ensembles B_α et C_β sont définis comme au paragraphe précédent.

On considère les fonction $f^*=\limsup_{n\to\infty}\dfrac{S_nf}{n}$ et $f_*=\liminf_{n\to\infty}\dfrac{S_nf}{n}$.

- Commençons par montrer que les fonction f^* et f_* sont T-invariantes.
  En effet, pour tout $n\geq 1$ et $x\in X$, on a

  $$S_n\left(f\circ T\right)(x)=S_{n+1}f(x)-f(x)\quad\text{donc}\quad \frac{S_n\left(f\circ T\right)(x)}{n}=\frac{S_{n+1}f(x)}{n+1}\times\frac{n+1}{n}-\frac{f(x)}{n}$$

  Comme f est intégrable, elle est finie presque partout et les suites $\left(\dfrac{f(x)}{n}\right)$ convergent vers 0 pour presque tout x. On en déduit que les lim sup des deux membres sont égales et donc que $f^*\circ T=f^*$. La démonstration est la même pour f_*.

- Nous allons maintenant montrer que $f^* = f_*$ μ-p.p. Pour cela, considérons, pour $\alpha < \beta$, l'ensemble
$$E_{\alpha,\beta} = \{f^* > \alpha, f_* < \beta\}$$
On remarque que $E_{\alpha,\beta}$ est inclus dans B_α et dans C_α car
$$f^*(x) > \alpha \Rightarrow \forall n \geq 1 \sup_{k \geq n} \frac{S_k f(x)}{k} > \alpha \Rightarrow \sup_{k \geq 1} \frac{S_k f(x)}{k} > \alpha$$
L'ensemble $E_{\alpha,\beta}$ est T-invariant car f^* et f_* sont T-invariantes donc $\mathfrak{Inv}(T)$-mesurables. On peut appliquer le dernier point du corollaire du paragraphe précédent pour obtenir :
$$\alpha\mu\left(E_{\alpha,\beta}\right) \leq \beta\mu\left(E_{\alpha,\beta}\right)$$
Par conséquent, si $\alpha > \beta$, $\mu\left(E_{\alpha,\beta}\right) = 0$. Il ne reste plus pour conclure qu'à remarquer que
$$\mu\left(f^* \neq f_*\right) = \mu\left(\bigcup_{\alpha > \beta, (\alpha,\beta) \in \mathbb{Q}^2} \{f^* > \alpha, f_* < \beta\}\right) = 0$$
Ses lim sup et lim inf étant presque partout égales, la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\right)$ converge presque partout vers $f^* = f_*$.
- Nous pouvons maintenant montrer que $f^* \in \mathrm{L}^1(\mu)$.
En effet, par invariance de T :
$$\begin{aligned}\int \left|\frac{S_n f}{n}\right| d\mu \leq & \frac{1}{n}\left(\int |f|\, d\mu + \int |f \circ T|\, d\mu + \ldots + \int \left|f \circ T^{n-1}\right| d\mu\right) \\ = & \frac{1}{n}\left(n \int |f| d\mu\right) = \int |f| d\mu\end{aligned}$$
Puis, par Fatou,
$$\int |f^*|\, d\mu = \int \liminf_{n \to +\infty} \left|\frac{S_n f}{n}\right| d\mu \leq \liminf_{n \to +\infty} \int \left|\frac{S_n f}{n}\right| d\mu \leq \int |f| d\mu < \infty$$
- Montrons que $\int f d\mu = \int f^* d\mu$. Pour cela, fixons un entier $n \geq 1$ et considérons, pour $k \in \mathbb{Z}$, la partition de X formée par les ensembles $D_{n,k} = \left\{\frac{k}{n} \leq f^* < \frac{k+1}{n}\right\}$.
Pour $\varepsilon > 0$ suffisamment petit, on a $D_{n,k} \subset B_{\frac{k}{n} - \varepsilon}$ et, comme $D_{n,k}$ est T-invariant, par application du point 3. du corollaire 2.2.2,
$$\int_{D_{n,k}} f d\mu \geq \left(\frac{k}{n} - \varepsilon\right) \mu\left(D_{n,k}\right)$$
Ceci étant vrai pour tout $\varepsilon > 0$ assez proche de 0, on obtient :
$$\int_{D_{n,k}} f d\mu \geq \frac{k}{n} \mu\left(D_{n,k}\right)$$

Par ailleurs, la définition de $D_{n,k}$ implique la majoration

$$\int_{D_{n,k}} f^* d\mu \leq \frac{k+1}{n} \mu\left(D_{n,k}\right) \leq \frac{k}{n} \mu\left(D_{n,k}\right) + \frac{1}{n} \mu\left(D_{n,k}\right) \leq \int_{D_{n,k}} f d\mu + \frac{1}{n} \mu\left(D_{n,k}\right)$$

On peut sommer sur tous les $k \in \mathbb{Z}$ pour obtenir :

$$\int_X f^* d\mu \leq \int_X f d\mu + \frac{1}{n} \mu\left(X\right)$$

Ceci étant vrai pour tous les entiers $n \geq 1$, finalement, comme $\mu\left(X\right) < \infty$,

$$\int_X f^* d\mu \leq \int_X f d\mu$$

On peut appliquer ce dernier résultat à $-f$:

$$\int_X (-f)^* d\mu \leq -\int_X f d\mu \quad \text{donc} \quad -\int_X f_* d\mu \leq -\int_X f d\mu$$

ce qui implique

$$\int_X f_* d\mu \geq \int_X f d\mu$$

On conclut car $f^* = f_*$ μ-p.p.

- Il reste à prouver que $f^* = \mathbb{E}\left[f \mid \mathfrak{Inv}T\right]$ μ-p.p. Comme on sait déjà que f^* est $\mathfrak{Inv}T$-mesurable, pour terminer, il ne reste qu'à montrer que pour tout événement $A \in \mathfrak{Inv}T$, on a $\int_A f d\mu = \int_A f^* d\mu$.

  Cela est clairement vrai si $\mu\left(A\right) = 0$. Sinon, on se place dans $\left(A, \mathcal{C}, \nu, T_{|A}\right)$ où $\mathcal{C}$ est la trace de $\mathcal{B}$ sur l'ensemble mesurable A et $\nu = \frac{\mu}{\mu(A)}$. Alors, le résultat précédent appliqué à $f_{|A}$ nous donne l'égalité voulue, de sorte qu'on a bien $f^* = \mathbb{E}\left[f \mid \mathfrak{Inv}T\right]$ μ-p.p.

□

Corollaire 2.3.2. *Dans le cas où T est ergodique,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \int_X f d\mu \ \ \mu\text{-p.p}$$

Démonstration. Si T est ergodique, alors la tribu $\mathfrak{Inv}T$ n'est composée que d'événements de mesure 0 ou 1, elle est donc indépendante de toute autre tribu et en particulier de la tribu engendrée par f, donc

$$\mathbb{E}\left[f \mid \mathfrak{Inv}T\right] = \mathbb{E}\left[f\right] = \int_X f d\mu \ \ \mu\text{-p.p}$$

□

Corollaire 2.3.3. *Toujours dans le cas où T est ergodique, on considère un événement $E \in \mathcal{B}$. Alors pour presque tout $x \in X$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_E \left(T^k(x)\right) = \mu(E)$$

Autrement dit, pour presque tout état initial x_0, la fréquence (temporelle) de passage dans E de la suite (x_n) converge vers la mesure (spatiale) de E

Corollaire 2.3.4. *Soit $p \in [1; +\infty[$. On suppose que $f \in L^p$. Alors, dans le théorème ergodique ponctuel de Birkhoff, la convergence a également lieu dans L^p.*

Pour démontrer ce résultat, nous aurons besoin de généraliser un peu le théorème de convergence dominée de Lebesgue.

Lemme 2.3.5 (Convergence dominée généralisée)**.** *Soit $p \in [1; +\infty[$. Soient (f_n) et (g_n) deux suites de fonctions telles que :*
$f_n \to f p.p.$, $|f_n|^p \leq g_n$, les g_n sont dans L^1, $g_n \to g$, p.p. et $\int g_n d\mu \to \int g d\mu$.
Alors, les f_n sont dans L^p et $f_n \to f$ dans L^p.

Démonstration du lemme. En passant à la limite dans $|f_n|^p \leq g_n$, on obtient déjà $|f|^p \leq |g|$.
Puis, en utilisant la convexité de $(x \longmapsto |x|^p)$,

$$|f_n - f|^p \leq (|f_n| + |f|)^p \leq 2^p \left(\frac{|f_n| + |f|}{2}\right)^p \leq 2^p \left(\frac{|f_n|^p + |f|^p}{2}\right) \leq 2^{p-1} (g_n + g)$$

On peut appliquer le lemme de Fatou à la fonction positive $2^{p-1}(g_n^p + g^p) - |f_n - f|^p$.

$$\int \liminf_{n^{\circ} \to\infty} \left(2^{p-1}(g_n + g) - |f_n - f|^p\right) d\mu \leq \liminf_{n\to+\infty} \int \left(2^{p-1}(g_n + g) - |f_n - f|^p\right) d\mu$$

La convergence ponctuelle des g_n et des f_n implique que le membre de gauche est égal à $2^p \int g d\mu$.
La convergence des intégrales des g_n implique que le membre de droite est égal à

$$2^p \left(\int g d\mu - \limsup_{n\to\infty} \int (|f_n - f|^p)\, d\mu\right).$$

On en déduit que

$$\limsup_{n\to\infty} \int |f_n - f|^p \, d\mu \leq 0 \quad \text{donc} \quad \lim_{n\to\infty} \int |f_n - f|^p \, d\mu = 0$$

□

Démonstration du corollaire 2.3.4. On pose, pour $n \geq 1$, $f_n = \dfrac{1}{n} S_n f$ et $g_n = \dfrac{1}{n} S_n |f|^p$.
On a, par convexité de $(x \longmapsto |x|^p)$, $|f_n|^p \leq g_n$ et, par invariance de T, pour tout $n \in \mathbb{N}$, $\int g_n d\mu = \int |f|^p \, d\mu$. En appliquant le théorème de Birkhoff à la fonction $|f|^p \in \mathrm{L}^1$, on voit que (g_n) converge presque partout vers $\mathbb{E}[|f|^p \mid \mathfrak{Inv}T]$ qui a même intégrale que $|f|^p$.
Les conditions du lemme précédent sont remplies et ce corollaire est démontré. □

Remarque 2.3.1. En revanche, même si $f \in L^{\infty}(X)$, la convergence de $\dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1} f \circ T^k$ peut ne pas avoir lieu dans $L1^{\infty}(X)$.

Démonstration. Un contre-exemple est donné par le schéma de Bernoulli $X = \{0;1\}^{\mathbb{N}}$, muni de $\mu = {\mu_0}^{\otimes\mathbb{N}}$, où μ_0 est uniforme sur $\{0;1\}$, $f(a_0a_1a_2...) = a_0$ et $T(a_0a_1a_2...) = a_1a_2...$. En effet, T est ergodique et $\mathbb{E}f = \dfrac{1}{2}$. Si l'on avait convergence dans L^{∞}, on aurait pour un certain n_0,

$$\left|\frac{1}{n_0}\sum_{k=0}^{n_0-1} f \circ T^k - \frac{1}{2}\right| \leq \frac{1}{4} \quad \text{sauf éventuellement sur un ensemble de mesure nulle}$$

Or, le cylindre $[0^{n_0}]$ est de mesure $\dfrac{1}{2^{n_0}}$ et sur celui-ci, $\dfrac{1}{n_0}\displaystyle\sum_{k=0}^{n_0-1} f \circ T^k = 0$. □

3 Topologie des mesures boréliennes de probabilités sur un espace métrique compact

Dans tout cette partie, $(X, d, \mathcal{B})$ désigne un espace métrique compact muni de sa tribu borélienne.

3.1 Séparabilité de $\mathcal{C}(X)$

Les idées de ce paragraphe sont tirées du polycopié de cours de Bernard Maurey [Mau01].

On note $\mathcal{C}(X)$ l'ensemble des fonctions continues de X dans $\mathbb{R}$, muni de la norme $\| \ \|_\infty$. Rappelons que cette norme en fait un espace de Banach.

Proposition 3.1.1. *Soit $\mathcal{R}$ un recouvrement fini de X par des ouverts. Il existe dans $\mathcal{C}(X)$, une partition de l'unité adaptée à $\mathcal{R}$, i.e. une famille $(\varphi_R)_{R\in\mathcal{R}} \subset \mathcal{C}(X)$ telle que :*

$$\forall R \in \mathcal{R}, \quad 0 \leq \varphi_R \leq 1, \quad \varphi_{|R^c} = 0, \quad \varphi_{|R} > 0, \quad \text{et} \quad \sum_{R\in\mathcal{R}} \varphi_R = 1$$

Démonstration. Posons, pour $R \in \mathcal{R}$ et $x \in X$, $\psi_R(x) = d(x, R^c)$.

La fonction ψ_R est bien continue, nulle sur R^c et, comme R^c est fermé, elle est strictement positive sur R. La partition $\mathcal{R}$ recouvrant X, la fonction $\psi := \sum_{R\in\mathcal{R}} \psi_R$ est strictement positive sur X.

En définissant $\varphi_R = \dfrac{\psi_R}{\psi}$, cette proposition est démontrée. □

Théorème 3.1.2. *L'espace $\mathcal{C}(X)$ est séparable.*

Démonstration. Soit $f \in \mathcal{C}(X)$ et $\delta \in \mathbb{R}^+$, on note

$$\omega_f(\delta) = \sup \{ |f(x) - f(t)| \quad | \quad d(s,t) \leq \delta \}$$

le module de continuité de f en δ. Comme X est compact et f est continue, $\omega_f(\delta)$ est bien défini. De plus, dans un compact la continuité équivaut à la continuité uniforme, de sorte que,

$$\forall f \in \mathcal{C}(X), \quad \lim_{\delta\to 0} \omega_f(\delta) = 0$$

Pour tout $n \in \mathbb{N}$, on peut choisir un recouvrement fini $\mathcal{R}_n$ de X par des boules ouvertes de diamètres égaux à 2^{-n}, et une partition de l'unité associée, $(\varphi_R)_{R\in\mathcal{R}_n}$.

Notons F_n l'espace vectoriel engendré par $(\varphi_R)_{R\in\mathcal{R}_n}$. C'est un espace vectoriel de dimension fini et,

$$\forall f \in \mathcal{C}(X), \quad d(f, F_n) \leq \omega_f(2^{-n})$$

En effet, choisissons arbitrairement dans chaque ouvert R de $\mathcal{R}_n$ un élément x_R et considérons la fonction continue g, définie par

$$g(x) = \sum_{R\in\mathcal{R}_n} f(x_R)\, \varphi_R(x)$$

On a alors, pour tout x dans X,

$$|f(x)-g(x)| = \left| \sum_{R\in\mathcal{R}_n} f(x)\varphi_R(x) - \sum_{R\in\mathcal{R}_n} f(x_R)\varphi_R(x) \right| \leq \sum_{R\in\mathcal{R}_n} |f(x)-f(x_R)|\varphi_R(x)$$

Lorsqu'un terme de la somme de droite n'est pas nul, alors $x\in R$. Comme $x_R\in R$ et que le diamètre de R vaut 2^{-n}, on a :

$$|f(x)-f(x_R)| \leq \omega_f\left(2^{-n}\right)$$

Finalement, on obtient, pour tout $x\in X$,

$$|f(x)-g(x)| \leq \sum_{R\in\mathcal{R}_n} \omega_f\left(2^{-n}\right)\varphi_R(x) = \omega_f\left(2^{-n}\right)$$

Donc

$$d(f,F_n) \leq \|f-g\|_\infty \leq \omega_f\left(2^{-n}\right)$$

On en déduit que $\bigcup_{n\in\mathbb{N}} F_n$ est dense dans $\mathcal{C}(X)$. Or, chaque F_n est un espace vectoriel normé de dimension fini donc est séparable. Leur réunion dénombrable l'est aussi, donc il existe une suite $(g_k)_{k\in\mathbb{N}}$ d'éléments de $\mathcal{C}(X)$ tels que $\overline{\{g_k \mid k\in\mathbb{N}\}} \supset \bigcup_{n\in\mathbb{N}} F_n$ puis,

$$\overline{\{g_k \mid k\in\mathbb{N}\}} \supset \overline{\bigcup_{n\in\mathbb{N}} F_n} = \mathcal{C}(X)$$

□

3.2 Topologie de $\mathcal{M}(X)$ et de $\mathcal{M}(X,T)$

Notons $\mathcal{S}(X)$ l'ensemble des mesures boréliennes signées sur X (voir la définition C.2.1). Comme X est métrique compact, le Théorème de représentation de Riesz, couplé au théorème C.3.2 permet d'identifier $\mathcal{S}(X)$ au dual $\mathcal{C}(X)^*$ de $\mathcal{C}(X)$.

Nous pouvons donc munir $\mathcal{S}(X)$ de la topologie $*$-faible, c'est-à-dire de la topologie la moins fine rendant les applications

$$\gamma_f : f \longmapsto \int f d\mu$$

continues. Notons

$$B = \left\{ \mu\in\mathcal{S}(X) \mid \forall f\in\mathcal{C}(X),\ \|f\|_\infty \leq 1 \Rightarrow \left|\int_X f d\mu\right| \leq 1 \right\}$$

D'après le Théorème de Banach-Alaoglu, B est compact pour la topologie $*$-faible. De plus, comme $\mathcal{C}(X)$ est séparable, son corollaire B.6.5 montre que la topologie $*$-faible est métrisable sur B.

Notons $\mathcal{M}(X)$ l'ensemble des mesures de probabilités boréliennes de X. Il est clair que $\mathcal{M}(X)\subset B$, ce qui en fait un espace métrisable.

De plus, $\mathcal{M}(X)$ est $*$-faiblement fermé dans $\mathcal{S}(X)$ car

$$\mathcal{M}(X) = \bigcap_{f\in\mathcal{C}(X),\ f\geq 0} {\gamma_f}^{-1}\left([0;+\infty[\right) \cap \gamma_1^{-1}\left(\{1\}\right)$$

Par conséquent, $\mathcal{M}(X)$ est compact, car c'est un fermé inclus dans le compact B.

De plus, il est facile de vérifier que $\mathcal{M}(X)$ est convexe. En conclusion,

Théorème 3.2.1. *L'ensemble $\mathcal{M}(X)$ des mesures boréliennes de probabilité, muni de la topologie initiale associées aux applications $\gamma_f : \mu \mapsto \int f d\mu$, où $f \in \mathcal{C}(X)$ est un espace compact, convexe, métrisable.*

Jusqu'à la fin de ce travail, il sera toujours sous-entendu que $\mathcal{M}(X)$ est muni de cette topologie.

Soit maintenant $T : X \to X$, une application continue. Nous pouvons définir l'application T^* sur $\mathcal{S}(X)$ par

$$\forall f \in \mathcal{C}(X), \quad \int f d(T^*\mu) = \int f \circ T d\mu$$

D'après la propriété caractéristique de la topologie initiale, T^* est continue si et seulement si les formes linéaires $\gamma_f \circ T^*$ sont continues. Or, pour tout $\mu \in \mathcal{S}(X)$,

$$\gamma_f \circ T^*(\mu) = \int f d(T^*\mu) = \int f \circ T d\mu = \gamma_{f\circ T}(\mu)$$

Comme T est continue, $f \circ T \in \mathcal{C}(X)$ donc $\gamma_{f\circ T}$ est continue sur $\mathcal{S}(X)$. Ceci a lieu pour toute fonction $f \in \mathcal{C}(X)$ donc T^* est continue sur $\mathcal{S}(X)$. De plus, il est facile de vérifier que T^* est linéaire sur $\mathcal{S}(X)$.

Par conséquent, l'ensemble des mesures signées qui sont fixes par T^*, qui n'est autre que le sous-espace propre de T^* pour la valeur propre 1, est un sous-espace vectoriel fermé de $\mathcal{S}(X)$.

Par ailleurs, il est immédiat que $\mathcal{M}(X)$ est stable par T^*. Comme T^* est continue et que $\mathcal{M}(X)$ est compact, convexe et non vide, on peut appliquer le Théorème de Schauder-Tychonov qui prouve que T^* admet au moins un point fixe dans $\mathcal{M}(X)$.

Notons $\mathcal{M}(X,T)$ l'ensemble de ces points fixes, c'est-à-dire l'ensemble des probabilités boréliennes sur X qui sont invariantes par T. Nous venons de voir que cet ensemble est non vide. Il est aussi compact, car c'est l'intersection d'une partie fermée avec le compact $\mathcal{M}(X)$ et enfin, il est convexe comme intersection du convexe $\mathcal{M}(X)$ avec un sous-espace vectoriel. En conclusion,

Théorème 3.2.2. *L'espace $\mathcal{M}(X,T)$ des mesures de probabilités boréliennes invariantes par T est une partie non vide, compacte et convexe de $\mathcal{M}(X)$.*

3.3 Caractérisation des mesures ergodiques dans $\mathcal{M}(X,T)$

Notons $\mathcal{E}(X,T)$ l'ensemble des mesures de probabilité borélienne sur X, T-invariantes et ergodiques. Nous allons caractériser les mesures ergodiques parmi les mesures invariantes à l'aide de la notion de point extrémal (voir la définition B.5.1).

Théorème 3.3.1. *La partie $\mathcal{E}(X,T)$ est l'ensemble des points extrémaux de $\mathcal{M}(X,T)$.*

Démonstration. Soit μ une mesure invariante non ergodique. Dans ce cas, il existe un borélien E tel que

$$0 < \mu(E) < 1 \quad \text{et} \quad T^{-1}(E) = E$$

et les mêmes relations sont vérifiées par le complémentaire E^c.

Posons $\mu_1 = \dfrac{1}{\mu(E)} 1_E \mu$ et $\mu_2 = \dfrac{1}{\mu(E^c)} 1_{E^c} \mu$. Alors μ_1 et μ_2 sont encore des mesures de probabilité. Elles sont de plus invariantes, car, pour tout borélien A,

$$\begin{aligned}\mu_1\left(T^{-1}A\right) &= \frac{1}{\mu(E)}\mu\left(T^{-1}(A)\cap E\right) = \frac{1}{\mu(E)}\mu\left(T^{-1}(A)\cap T^{-1}(E)\right)\\ &= \frac{1}{\mu(E)}\mu\left(T^{-1}(A\cap E)\right) = \frac{1}{\mu(E)}\mu(A\cap E)\end{aligned}$$

donc $\mu_1(T^{-1}(A)) = \mu_1(A)$.
Comme $\mu = \mu(E)\mu_1 + \mu(E^c)\mu_2$, la mesure μ n'est pas extrémale dans $\mathcal{M}(X,T)$.

Supposons maintenant que $\mu \in \mathcal{E}(X,T)$. Soient $0 < t < 1$ et $\mu_1, \mu_2 \in \mathcal{M}(X,T)$ tels que

$$\mu = t\mu_1 + (1-t)\mu_2$$

Le fait que $0 < t < 1$ implique que μ_1 et μ_2 sont absolument continues par rapport à μ. D'après le Théorème de Radon-Nikodym, il existe deux fonctions positives et mesurables f_1 et f_2 telles que

$$d\mu_1 = f_1 d\mu \quad \text{et} \quad d\mu_2 = f_2 d\mu$$

Posons $E = \{f_1 < 1\}$. Alors,

$$\mu_1(E) = \mu_1\left(E\cap T^{-1}(E)\right) + \mu_1\left(E\cap\left(T^{-1}(E)\right)^c\right) \quad \text{et}$$

$$\mu_1\left(T^{-1}(E)\right) = \mu_1\left(E\cap T^{-1}(E)\right) + \mu_1\left(E^c\cap T^{-1}(E)\right)$$

Comme μ_1 est invariante, $\mu_1(T^{-1}(E)) = \mu_1(E)$, donc après simplification,

$$\mu_1\left(E\cap\left(T^{-1}(E)\right)^c\right) = \mu_1\left(E^c\cap T^{-1}(E)\right)$$

Ce calcul est en fait vrai pour toute mesure invariante (et tout mesurable E), donc on a aussi

$$\mu\left(E\cap\left(T^{-1}(E)\right)^c\right) = \mu\left(E^c\cap T^{-1}(E)\right)$$

Supposons que $\mu\left(E\cap(T^{-1}(E))^c\right) > 0$. Alors, par définition de E,

$$\mu_1\left(E\cap\left(T^{-1}(E)\right)^c\right) = \int_{E\cap(T^{-1}(E))^c} f_1 d\mu < \mu\left(E\cap\left(T^{-1}(E)\right)^c\right)$$

Mais, par ailleurs,

$$\mu_1\left(E^c\cap T^{-1}(E)\right) = \int_{E^c\cap T^{-1}(E)} f_1 d\mu \geq \mu\left(E^c\cap T^{-1}(E)\right)$$

On obtient donc

$$\mu\left(E^c\cap T^{-1}(E)\right) < \mu\left(E\cap\left(T^{-1}(E)\right)^c\right)$$

ce qui est contradictoire. Ceci prouve que $\mu\left(E\cap(T^{-1}(E))^c\right) = 0$ et finalement, que $\mu(E\Delta T^{-1}(E)) = 0$.

La mesure μ étant ergodique, d'après le théorème 1.3.1, cela impose que $\mu(E) = 0$ ou 1 et comme $\mu_1(X) = 1$, il est impossible que $\mu(E) = 1$. Finalement, $\mu(\{f_1 < 1\}) = 0$. On montrerait exactement de la même manière que $\mu(\{f_1 > 1\}) = 0$. Par conséquent, $f_1 = 1$, μ-presque partout, donc $\mu_1 = \mu$. De façon identique, on obtient $\mu_2 = \mu$, ce qui achève de démontrer cette implication. $\square$

En utilisant le Théorème de Krein-Milman dans le compact convexe non vide $\mathcal{M}(X,T)$, on obtient le corollaire suivant.

Corollaire 3.3.2. *L'ensemble $\mathcal{E}(X,T)$ des mesures ergodiques T-invariantes sur X, où T est continue et X est métrique compact est non vide. De plus, toute partie extrémale de $\mathcal{M}(X,T)$ contient au moins une mesure ergodique.*

4 Entropie en dynamique symbolique

4.1 Entropie d'une partition par rapport à une mesure

L'entropie a été introduite dans l'étude des systèmes dynamiques par Kolmogorov et Sinai dans les années 1950. Nous devons rester brefs sur son introduction et passerons sous silence beaucoup de ses propriétés.

Les références [Wal00], [Zin96] ou encore [ELW11] sont très riches et claires sur ce sujet.

Nous nous contentons ici de donner une construction, d'abord générale, de l'entropie d'une partition par rapport à une mesure, puis dans le cas de la dynamique symbolique, de l'entropie d'une transformation préservant la mesure, en suivant la monographie [Zin96].

On note $h : [0;1] \to \mathbb{R}^+$ la fonction définie par $h(x) = -x\ln(x)$ si $x > 0$ et $h(0) = 0$.

Proposition 4.1.1. *La fonction h est continue et strictement concave sur $[0;1]$, i.e.*

$$\forall x, y \in [0;1], \qquad \forall t \in]0;1[, \qquad th(x) + (1-t)h(y) \leq h\left(tx + (1-t)y\right)$$

avec égalité si et seulement si $x = y$.

Son maximum est $\frac{1}{e}$, atteint en $\frac{1}{e}$.

Démonstration. La continuité en 0 est un fait bien connu. De plus, la fonction h est dérivable sur $]0;1]$ et sa dérivée $h'(x) = -\ln(x) - 1$ y est strictement décroissante. Elle est donc strictement concave sur $]0;1]$, et comme $h(0) = 0$, cela s'étend sur $[0:1]$.

Le dernier point se vérifie simplement. $\square$

Soit $\Delta_k = \left\{(p_1, p_2, ..., p_k) \in {\mathbb{R}^+}^k \quad | \quad \sum_{i=1}^k p_i = 1\right\}$. On définit la fonction H sur $\bigcup_{k=1}^\infty \Delta_k$ par :

$$\forall k \geq 1, \qquad \forall (p_1, ..., p_k) \in \Delta_k, \qquad H\left(p_1, p_2, ..., p_k\right) = \sum_{i=1}^k h\left(p_i\right)$$

Proposition 4.1.2. *La fonction H est positive. À k fixé, son minimum est égal à 0 et est atteint lorsque l'un des p_i vaut 1 et son maximum, qui vaut $\ln(k)$, est atteint lorsque tous les p_i sont égaux à $\frac{1}{k}$.*

Démonstration. Il est clair que H est positive. Elle vaut 0 lorsque l'un des p_i vaut 1 car alors, tous ses autres arguments valent 0. Si aucun des p_i n'est égal à 1, au moins deux d'entre eux sont compris dans l'intervalle $]0;1[$, et alors $H(p_1, ..., p_k) > 0$.

Pour le maximum, on utilise la concavité de h avec les coefficients tous égaux à $\frac{1}{k}$. Il en résulte :

$$\sum_{i=1}^k \frac{1}{k} h\left(p_i\right) \leq h\left(\sum_{i=1}^k \frac{1}{k} p_i\right) = h\left(\frac{1}{k}\right) = \frac{1}{k}\ln\left(k\right)$$

D'où l'on déduit que $H\left(p_1, p_2, ..., p_k\right) \leq \ln(k)$.

Comme la concavité est stricte, $H\left(p_1, p_2, ..., p_k\right) = \ln(k) \Leftrightarrow p_1 = p_2 = ... = p_k = \frac{1}{k}$. $\square$

Dans tout ce qui suit, $(X, \mathcal{B}, \mu)$ est un espace de probabilité. Lorsque nous parlerons d'une partition de X, il sera toujours sous-entendu que ses éléments sont mesurables.

Définition 4.1.1. Soit $\mathcal{A}$ une partition finie de X. L'entropie de $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ par rapport à μ est le réel :

$$H_\mu(\mathcal{A}) = H(\mu(A_1), \mu(A_2), ..., \mu(A_k))$$

Remarque 4.1.1. Intuitivement, l'entropie de la partition $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ est l'information moyenne apportée par la connaissance, pour les x dans X, de l'élément de $\mathcal{A}$ qui le contient. Ainsi, si l'un des A_i est de probabilité 1, la partition $\mathcal{A}$ n'offre aucune information, tandis que si les A_i sont équiprobables, la quantité d'information moyenne donnée par "$x \in A_i$" est maximale.

Définition 4.1.2. Soient $\mathcal{A}$ et $\mathcal{C}$ deux partitions finies de X. Soit B un événement de X. Si $\mu(B) > 0$, on note μ_B la probabilité conditionnelle sachant B et l'entropie conditionnelle de $\mathcal{A}$ sachant l'événement B est :

$$H_\mu(\mathcal{A} \mid B) = H_{\mu_B}(\mathcal{A}) = -\sum_{A \in \mathcal{A}} \frac{\mu(A \cap B)}{\mu(B)} \ln\left(\frac{\mu(A \cap B)}{\mu(B)}\right)$$

Si $\mu(B) = 0$, on fixe arbitrairement $H_\mu(\mathcal{A} \mid B) = 0$.

L'entropie conditionnelle de la partition $\mathcal{A}$ sachant la partition $\mathcal{C}$ est :

$$H_\mu(\mathcal{A} \mid \mathcal{C}) = \sum_{C \in \mathcal{C}} \mu(C) H_\mu(\mathcal{A} \mid C) = -\sum_{A \in \mathcal{A},\ C \in \mathcal{C},\ \mu(C)>0} \mu(A \cap C) \ln\left(\frac{\mu(A \cap C)}{\mu(C)}\right)$$

Définition 4.1.3. Soient $\mathcal{A}$ et $\mathcal{C}$ deux partitions de X. La partition $\mathcal{A} \vee \mathcal{C}$ est définie par :

$$\mathcal{A} \vee \mathcal{C} = \{A \cap C \mid A \in \mathcal{A},\ C \in \mathcal{C}\}$$

Proposition 4.1.3 (Sous-additivité de l'entropie)**.** *Soient $\mathcal{A}$ et $\mathcal{B}$ deux partitions finies de X. Alors,*

1. $H(\mathcal{A} \vee \mathcal{C}) = H(\mathcal{A}) + H(\mathcal{C} \mid \mathcal{A})$ *;*

2. $H(\mathcal{A}) \geq H(\mathcal{A} \mid \mathcal{C})$ *et*

3. $H(\mathcal{A} \vee \mathcal{C}) \leq H(\mathcal{A}) + H(\mathcal{C})$.

Démonstration. 1. L'entropie de $\mathcal{C}$ sachant $\mathcal{A}$ vaut :

$$\begin{aligned}
H(\mathcal{C} \mid \mathcal{A}) &= -\sum_{C \in \mathcal{C},\ A \in \mathcal{A},\ \mu(A)>0} \mu(C \cap A) \ln\left(\frac{\mu(C \cap A)}{\mu(A)}\right) \\
&= -\sum_{C \in \mathcal{C},\ A \in \mathcal{A},\ \mu(A)>0} \mu(C \cap A) \ln(\mu(C \cap A)) + \sum_{C \in \mathcal{C},\ A \in \mathcal{A},\ \mu(A)>0} \mu(A \cap C) \ln(\mu(A)) \\
&= -\sum_{C \in \mathcal{C},\ A \in \mathcal{A}} \mu(C \cap A) \ln(\mu(C \cap A)) + \sum_{A \in \mathcal{A},\ \mu(A)>0} \ln(\mu(A)) \sum_{C \in \mathcal{C}} \mu(A \cap C) \\
&= H_\mu(\mathcal{A} \vee \mathcal{C}) + \sum_{A \in \mathcal{A},\ \mu(A)>0} \ln(\mu(A)) \mu(A) \\
&= H_\mu(\mathcal{A} \vee \mathcal{C}) - H_\mu(\mathcal{A})
\end{aligned}$$

2. On observe que

$$
\begin{aligned}
H_\mu(\mathcal{A} \mid \mathcal{C}) &= -\sum_{A\in\mathcal{A},\ C\in\mathcal{C},\ \mu(C)>0} \mu(C)\frac{\mu(A\cap C)}{\mu(C)} \ln\left(\frac{\mu(A\cap C)}{\mu(C)}\right) \\
&= \sum_{A\in\mathcal{A}} \sum_{C\in\mathcal{C},\ \mu(C)>0} \mu(C)\, h\left(\frac{\mu(A\cap C)}{\mu(C)}\right)
\end{aligned}
$$

La concavité de la fonction h donne alors :

$$
\sum_{C\in\mathcal{C},\ \mu(C)>0} \mu(C)\, h\left(\frac{\mu(A\cap C)}{\mu(C)}\right) \le h\left(\sum_{C\in\mathcal{C},\ \mu(C)>0} \mu(A\cap C)\right) = h(\mu(A))
$$

On en déduit que :

$$
H_\mu(\mathcal{A} \mid \mathcal{C}) \le \sum_{A\in\mathcal{A}} h(\mu(A)) = H_\mu(\mathcal{A})
$$

3. L'inégalité 3. se déduit des deux propriétés précédentes.

□

Donnons une dernière propriété qui nous servira par la suite.

Proposition 4.1.4. *Soient μ et ν deux mesures de probabilité sur X et $t \in [0;1]$. Alors, pour toute partition finie $\mathcal{A}$ de X, on a :*

$$
tH_\mu(\mathcal{A}) + (1-t)H_\nu(\mathcal{A}) \le H_{t\mu+(1-t)\nu}(\mathcal{A})
$$

Démonstration. Pour chaque événement $A \in \mathcal{A}$, on a, par concavité de la fonction h,

$$
h(t\mu(A) + (1-t)\nu(A)) \ge th(\mu(A)) + (1-t)h(\nu(A))
$$

Le résultat s'en déduit en faisant la somme sur tous les événements A de $\mathcal{A}$. □

4.2 Entropie d'une transformation

Dorénavant, nous nous placerons dans le cadre de la dynamique symbolique. Rappelons que l'espace X est l'ensemble des mots infinis sur l'alphabet fini Σ, que cet espace est muni d'une distance qui en fait un espace compact et que sa tribu borélienne $\mathcal{B}$ est la tribu engendrée par les cylindres. La transformation T est le shift, et il est continu sur X. On note $\mathcal{A}_n$ la partition formée par les cylindres de précision $n+1$. Nous allons définir dans ce cas particulier l'entropie de T par rapport à une mesure invariante par T.

Comme l'illustre la proposition suivante, le cadre symbolique offre des facilités qui n'ont pas toujours lieu dans le cas général.

Proposition 4.2.1. *Soit $(\mu_k)_{k\in\mathbb{N}}$ une suite de mesures de probabilités qui converge *-faiblement vers μ. Alors, pour tout $n \in \mathbb{N}$,*

$$
\lim_{k\to+\infty} H_{\mu_k}(\mathcal{A}_n) = H_\mu(\mathcal{A}_n)
$$

Démonstration. Rappelons que les cylindres sont à la fois ouverts et fermés (voir la remarque 1.2.1) donc que leurs indicatrices sont continues. On aura donc, pour tout $A \in \mathcal{A}_n$,

$$\lim_{k\to+\infty} \mu_k(A) = \mu(A)$$

La proposition s'en déduit en faisant tendre k vers $+\infty$ dans la somme finie :

$$H_{\mu_k}(\mathcal{A}_n) = \sum_{A\in\mathcal{A}_n} h(\mu_k(A))$$

□

Nous aurons également besoin du petit lemme d'analyse qui suit.

Lemme 4.2.2. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels positifs telle que*

$$\forall n, p \in \mathbb{N}, \qquad u_{n+p} \leq u_n + u_p \qquad (1)$$

Alors, la suite $\left(\frac{u_n}{n}\right)_{n\in\mathbb{N}^*}$ converge vers $\ell = \inf_{n\in\mathbb{N}^*} \frac{u_n}{n}$.

Démonstration. Soit $\varepsilon > 0$. Par définition de la borne inférieure, il existe un entier n_0 tel que

$$\ell \leq \frac{u_{n_0}}{n_0} \leq \ell + \frac{\varepsilon}{2}$$

Soit $n \geq n_0$. La division euclidienne de n par n_0 donne $n = qn_0 + r$, avec $0 \leq r \leq q - 1$. La propriété (1) appliquée récursivement donne

$$u_n \leq qu_{n_0} + u_r$$

Puis, comme $n \geq qn_0$,

$$\ell \leq \frac{u_n}{n} \leq \frac{qu_{n_0}}{qn_0} + \frac{u_r}{qn_0} \leq \ell + \frac{\varepsilon}{2} + \frac{\max_{0\leq i\leq n_0-1} u_i}{n - n_0}$$

Pour n assez grand, on a

$$\frac{\max_{0\leq i\leq n_0-1} u_i}{n - n_0} \leq \frac{\varepsilon}{2}$$

ce qui démontre ce lemme. □

Proposition 4.2.3. *Soit μ une mesure invariante par T et n et p deux entiers naturels.*

1. $H_\mu(T^{-1}(\mathcal{A}_n)) = H_\mu(\mathcal{A}_n)$;
2. $H_\mu(\mathcal{A}_{n+p}) \leq H_\mu(\mathcal{A}_n) + H_\mu(\mathcal{A}_p)$ *et*
3. *le réel*

$$h_\mu(T) = \lim_{n\to+\infty} \frac{H_\mu(A_n)}{n} = \inf_{n\in\mathbb{N}^*} \frac{H_\mu(A_n)}{n}$$

est bien défini. C'est l'entropie *de T par rapport à μ.*

Démonstration. Le premier point est évident car μ est invariante par rapport à T. Le deuxième se déduit de la sous-additivité vue à la proposition 4.1.3 et du fait que

$$\mathcal{A}_{n+p} = \mathcal{A}_n \vee T^{-n}\mathcal{A}_p$$

Enfin, le dernier point n'est que l'application du lemme 4.2.2. □

Rappelons qu'une fonction d'un espace topologique, à valeurs dans $\mathbb{R}$ est dite semi-continue supérieurement en x, lorsque :

$$\forall \varepsilon > 0, \quad \exists U \text{ voisinage de } x, \ \forall y \in U, \ f(y) < f(x) + \varepsilon$$

Dans un espace métrique, cela est équivalent à la semi-continuité supérieure séquencielle, i.e. pour toute suite (u_n) qui converge vers x,

$$\limsup_{n\to+\infty} f(u_n) \leq f(x)$$

Proposition 4.2.4. *L'application de $\mathcal{M}(X,T)$ dans $\mathbb{R}^+$, qui à μ associe $h_\mu(T)$ est semi-continue supérieurement et affine.*

Démonstration. Soit $(\mu_k)_{k\in\mathbb{N}} \subset \mathcal{M}(X,T)$ telle que, pour la topologie faible-*, $\mu_k \underset{k\to+\infty}{\rightharpoonup} \mu$. On sait que $\mu \in \mathcal{M}(X,T)$ car cette partie est fermée.

Il nous faut montrer que pour tout $\varepsilon > 0$ et pour k assez grand, on a

$$h_{\mu_k}(T) < h_\mu(T) + \varepsilon$$

Par définition de la borne inférieure, il existe un entier n tel que

$$\frac{H_\mu(\mathcal{A}_n)}{n} < h_\mu(T) + \frac{\varepsilon}{2}$$

Or, d'après la proposition 4.2.1,

$$\lim_{k\to\infty} H_{\mu_k}(\mathcal{A}_n) = H_\mu(\mathcal{A}_n)$$

Il existe donc un entier K tel que pour tout $k \geq K$,

$$\frac{H_{\mu_k}(\mathcal{A}_n)}{n} \leq \frac{H_\mu(\mathcal{A}_n)}{n} + \frac{\varepsilon}{2} \leq h_\mu(T) + \varepsilon$$

Alors, pour tout $k \geq K$,

$$h_{\mu_k}(T) = \inf_{m\in\mathbb{N}*} \frac{H_{\mu_k}(\mathcal{A}_m)}{m} \leq \frac{H_{\mu_k}(\mathcal{A}_n)}{n} < h_\mu(T) + \varepsilon$$

ce qui conclut la démonstration de la semi-continuité supérieure de l'application $\mu \longmapsto h_\mu(T)$.

Démontrons maintenant que cette application est affine. Soient $\mu, \nu \in \mathcal{M}(X,T)$ et $t \in [0;1]$.
Soit $m = t\mu + (1-t)\nu$. Nous allons démontrer que pour tout $n \in \mathbb{N}^*$,

$$0 \leq H_m(\mathcal{A}_n) - tH_\mu(\mathcal{A}_n) - (1-t)H_\nu(\mathcal{A}_n) \leq \frac{2}{e}$$

Une fois cette inégalité établie, après une division par n et un passage à la limite, la proposition sera démontrée.

L'inégalité de gauche a déjà été établie à la proposition 4.1.4

Pour démontrer l'inégalité de droite, on considère un élément A de $\mathcal{A}_n$ tel que $\mu(A) > 0$ et $\nu(A) > 0$.

$$\begin{aligned}
& m(A)\ln(m(A)) - t\mu(A)\ln(\mu(A)) - (1-t)\nu(A)\ln(\nu(A)) \\
= \; & t\mu(A)\left[\ln(t\mu(A) - (1-t)\nu(A)) - \ln(\mu(A))\right] + (1-t)\nu(A)\left[\ln(t\mu(A) - (1-t)\nu(A)) - \ln(\nu(A))\right] \\
= \; & t\mu(A)\ln(t)\left[1 + \ln\left(1 + \frac{(1-t)\nu(A)}{t\mu(A)}\right)\right] + (1-t)\nu(A)\ln(1-t)\left[1 + \ln\left(1 + \frac{t\mu(A)}{(1-t)\nu(A)}\right)\right] \\
\leq \; & t\ln(t)\mu(A) + (1-t)\ln(1-t)\,\nu(A)
\end{aligned}$$

Et cette inégalité subsiste dans le cas où $\mu(A) = 0$ ou $\nu(A) = 0$.

On obtient alors, en passant à l'opposé et en sommant sur les $A \in \mathcal{A}_n$,

$$H_m(\mathcal{A}_n) - tH_\mu(\mathcal{A}_n) - (1-t)H_\nu(\mathcal{A}_n) \leq h(t) + h(1-t)$$

En majorant la fonction h par son maximum, égal à $\dfrac{1}{e}$ d'après la proposition 4.1.1, on obtient l'inégalité cherchée. $\square$

5 Applications en dynamique symbolique

5.1 Pression topologique et principe variationnel

Théorème 5.1.1 (Définition de la pression topologique)**.** *Soit $\varphi \in \mathcal{C}(X)$. On peut lui associer le réel positif* $\mathrm{P}(\varphi)$*, appelé pression topologique de φ, défini par :*

$$\mathrm{P}(\varphi) = \lim_{n \to +\infty} \frac{1}{n} \ln \left(\sum_{C \in \mathcal{A}_n} e^{\sup_{x \in C} S_n \varphi(x)} \right)$$

Démonstration. Soit $N \geq 1$ un entier naturel. Pour $A = [\ell_0 \ell_1 ... \ell_N]$ dans $\mathcal{A}_N$, on note x_A l'un des mots du compact X pour lequel

$$S_N \varphi(x_A) = \sup_{x \in A} S_N \varphi(x)$$

Si $n \leq N$, on note $A_{|n} = [\ell_0 \ell_1 ... \ell_n]$ le cylindre de $\mathcal{A}_n$ qui contient A.

Enfin, pour $n \in \mathbb{N}$, on note $V_n = \sum_{C \in \mathcal{A}_n} e^{S_n \varphi(x_C)}$. Soient n et p deux entiers naturels non nuls. Nous allons montrer que

$$V_{n+p} \leq V_n V_p$$

Soit $A \in \mathcal{A}_{n+p}$. Alors

$$S_{n+p} \varphi(x_A) = S_n \varphi(x_A) + \sum_{k=0}^{p-1} \varphi \circ T^k (T^n(x_A)) = S_n \varphi(x_A) + S_p \varphi(T^n(x_A))$$

Comme $A \subset A_{|n}$, on a

$$S_n \varphi(x_A) \leq S_n \varphi\left(x_{A_{|n}}\right)$$

Puis, comme $x_A \in A$, $T^n(x_A) \in T^n A$, où $T^n(A) \in \mathcal{A}_p$, de sorte que :

$$S_p \varphi(T^n(x_A)) \leq S_p \varphi\left(x_{T^n(A)}\right)$$

Nous pouvons maintenant découper $\mathcal{A}_{n+p}$ de façon adaptée :

$$\begin{aligned}
V_{n+p} &= \sum_{A \in \mathcal{A}_{n+p}} e^{S_{n+p}\varphi(x_A)} \\
&\leq \sum_{A \in \mathcal{A}_{n+p}} e^{S_n \varphi\left(x_{A_{|n}}\right) + S_p \varphi\left(x_{T^n(A)}\right)} \\
&\leq \sum_{C \in \mathcal{A}_n} \sum_{A \in \mathcal{A}_{n+p},\ A_{|n} = C} e^{S_n \varphi(x_C)} e^{S_p \varphi\left(x_{T^n(A)}\right)} \\
&\leq \sum_{C \in \mathcal{A}_n} e^{S_n \varphi(x_C)} \sum_{A \in \mathcal{A}_{n+p},\ A_{|n} = C} e^{S_p \varphi\left(x_{T^n(A)}\right)}
\end{aligned}$$

Finalement, comme à C fixé dans $\mathcal{A}_n$, l'ensemble $\left\{ A \in \mathcal{A}_{n+p} \mid A_{|n} = C \right\}$ est en bijection avec $\mathcal{A}_p$ via la fonction d'ensembles T^n, on obtient,

$$V_{n+p} \leq \sum_{C \in \mathcal{A}_n} e^{S_n \varphi(x_C)} \sum_{D \in \mathcal{A}_p} e^{S_p \varphi(x_D)} = V_n V_p$$

Comme les V_i sont strictement positifs, on peut poser pour $n \geq 1$, $u_n = \ln(V_n)$ et $u_0 = 0$. On a alors, pour tous les entiers naturels n et p,

$$u_{n+p} \leq u_n + u_p$$

et le lemme 4.2.2 s'applique et implique la convergence de la suite $\left(\frac{\ln(V_n)}{n}\right)$. □

Avant de passer à l'énoncé et la démonstration du théorème principal de ce chapitre, nous donnons quelques rappels sur le module de continuité.

Définition 5.1.1. Soit (X, d) un espace métrique et f une fonction à valeurs réelles sur X. Le *module de continuité* de f est la fonction $\omega_f : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$ définie par

$$\omega_f(t) = \sup\{|f(x) - f(y)| \quad | \quad d(x, y) \leq t\}$$

Dans le cadre de notre étude, la distance ne prend que les valeurs 2^{-n}, donc il suffit de s'intéresser à la suite décroissante $\left(\omega_f\left(2^{-k-1}\right)\right)_{k \geq -1}$.

Et comme X est métrique compact, la continuité de f est équivalente à sa continuité uniforme, donc f est continue si et seulement si

$$\lim_{k \to +\infty} \omega_f\left(2^{-k-1}\right) = 0$$

Théorème 5.1.2 (Principe variationnel pour la pression)**.** *Soit $\varphi \in \mathcal{C}(X)$. Alors*

$$\mathrm{P}(\varphi) = \sup\left\{h_\mu(T) + \int_X \varphi d\mu \quad | \quad \mu \in \mathcal{M}(X, T)\right\}$$

Commençons par montrer que $\mathrm{P}(\varphi) \geq \sup\{h_\mu(T) + \int_X \varphi d\mu \quad | \quad \mu \in \mathcal{M}(X, T)\}$. Pour cela, nous aurons besoin du lemme de concavité suivant :

Lemme 5.1.3. *Soient $n \in \mathbb{N}^*$, p_1, p_2, ..., p_n des réels positifs tels que $\sum_{i=1}^n p_i = 1$ et a_1, a_2, ..., a_n des réels quelconques. Alors,*

$$\sum_{i=1}^n p_i(a_i - \ln(p_i)) \leq \ln\left(\sum_{i=1}^n e^{a_i}\right)$$

avec égalité si et seulement si pour tout $i = 1, ..., n$,

$$p_i = \frac{e^{a_i}}{\sum_{i=1}^n e^{a_i}}$$

Démonstration du lemme. Commençons par remarquer que l'on peut, dans le membre de gauche, omettre les indices i tels que $p_i = 0$. Ainsi, il suffit de démontrer ce lemme dans le cas où tous les p_i sont non nuls.

On utilise alors la concavité de la fonction ln :

$$\ln\left(\sum_{i=1}^n p_i \frac{e^{a_i}}{p_i}\right) \geq \sum_{i=1}^n p_i \ln\left(\frac{e^{a_i}}{p_i}\right)$$

ce qui démontre l'inégalité.

Pour le cas d'égalité, il est impossible si l'un des p_i est nul. Si tous les p_i sont non nuls, par stricte concavité de la fonction ln, il ne peut advenir que si tous les réels $\frac{e^{a_i}}{p_i}$ sont égaux, i.e.

$$\forall i = 1, ..., n \qquad p_i = p_1 \frac{e^{a_i}}{e^{a_1}}$$

Comme $\sum_{i=1}^{n} p_i = 1$, on obtient $p_1 = \frac{e^{a_1}}{\sum_{i=1}^{n} e^{a_i}}$, puis, pour $i = 1, ..., n$, $p_i = \frac{e^{a_i}}{\sum_{j=1}^{n} e^{a_j}}$. □

Démonstration de la première inégalité. Soit $\mu \in \mathcal{M}(X, T)$.

Nous appliquons le lemme 5.1.3 avec pour ensemble d'indices l'ensemble de cylindres $\mathcal{A}_n$, $p_A = \mu(A)$ et $a_A = \sup_{x \in A} S_n\varphi(x)$. Il vient :

$$H_\mu(\mathcal{A}_n) + \sum_{A \in \mathcal{A}_n} p_A a_A \leq \ln\left(\sum_{A \in \mathcal{A}_n} e^{a_A}\right)$$

Après division par n, le membre de droite converge vers $\mathrm{P}(\varphi)$ lorsque n tend vers $+\infty$ et la première partie du membre de gauche tend vers $h_\mu(T)$. Il ne reste qu'à étudier $\sum_{A \in \mathcal{A}_n} p_A a_A$.

Or, par invariance de μ et du fait que $\mathcal{A}_n$ est une partition de X, on a :

$$n \int_X \varphi d\mu = \int S_n\varphi d\mu = \sum_{A \in \mathcal{A}_n} \int_A S_n\varphi d\mu = \sum_{A \in \mathcal{A}_n} \int_A a_A d\mu + \sum_{A \in \mathcal{A}_n} \int_A (S_n\varphi - a_A)\, d\mu$$

On obtient donc :

$$\frac{1}{n} \sum_{A \in \mathcal{A}_n} \mu(A)\, a_A = \int_X \varphi d\mu - \frac{1}{n} \sum_{A \in \mathcal{A}_n} \int_A (S_n\varphi - a_A)\, d\mu$$

Soit, pour chaque cylindre $A \in \mathcal{A}_n$, $x_A \in A$ tel que $S_n\varphi(x_A) = \sup_{x \in A} S_n\varphi(x)$. On a alors, en introduisant le module de continuité ω_φ de la fonction continue φ,

$$\left| \int_A \sum_{k=0}^{n-1} \left(\varphi \circ T^k(x_A) - \varphi \circ T^k(x)\right) \mu(dx) \right| \leq \sum_{k=0}^{n-1} \mu(A) \omega_\varphi\left(2^{-n-1+k}\right)$$

Par conséquent,

$$\left| \frac{1}{n} \sum_{A \in \mathcal{A}_n} \int_A S_n\varphi - a_A d\mu \right| \leq \frac{1}{n} \sum_{k=1}^{n} \omega_\varphi\left(2^{-k-1}\right)$$

Comme φ est continue sur X, $\lim_{n \to +\infty} \omega_\varphi\left(2^{-n-1}\right) = 0$ et il en est de même de la somme de Césaro :

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \omega_\varphi\left(2^{-k-1}\right) = 0$$

Ceci achève de démontrer que

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{A \in \mathcal{A}_n} p_A a_A = \int_X \varphi d\mu$$

et cette première inégalité.

□

L'autre inégalité est un peu plus délicate. Nous aurons besoin d'un nouveau lemme qui sert à fabriquer des mesures invariantes.

Lemme 5.1.4. *Soit $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}(X)$. Pour tout $n \in \mathbb{N}$, on définit la mesure de probabilité ν_n par*

$$\nu_n = \frac{1}{n}\left(\mu_n + \mu_n \circ T + \mu_n \circ T^2 + ... + \mu_n \circ T^{n-1}\right) = \frac{1}{n} S_n^* \mu_n$$

Alors la suite (ν_n) admet des valeurs d'adhérences et ces valeurs d'adhérence sont toutes des mesures invariantes *par T.*

Démonstration. Comme $\mathcal{M}(X)$ est métrique compact, la suite (ν_n) admet des suites extraites convergentes. Soit (ν_{n_k}) l'une d'entre elle et ν sa limite.

On a alors, pour tout entier k,

$$\nu_{n_k} \circ T - \nu_{n_k} = \frac{1}{n_k}\left(T^* S_{n_k}^* \mu_{n_k} - S_{n_k}^* \mu_{n_k}\right) = \frac{1}{n_k}\left(\mu \circ T^{n_k} - \mu\right)$$

Or, pour toute fonction $f \in \mathcal{C}(X)$,

$$\left|\int_X f d\left(\frac{1}{n_k}\left(\mu \circ T^{n_k} - \mu\right)\right)\right| \leq \frac{2\left\|f\right\|_\infty}{n_k} \underset{k\to+\infty}{\longrightarrow} 0$$

Donc la suite $(\nu_{n_k} \circ T - \nu_{n_k})$ converge *-faiblement vers 0. Par ailleurs, comme le shift T est continu, son adjoint l'est aussi et $(\nu_{n_k} \circ T - \nu_{n_k})$ converge également vers $\nu \circ T - \nu$. On en déduit donc que toute valeur d'adhérence ν de la suite (ν_n) vérifie

$$\nu \circ T = \nu$$

□

Enfin, pour pouvoir découper de façon efficace nos partitions en sous-partitions, nous aurons besoin d'un lemme d'arithmétique :

Lemme 5.1.5. *Soient $0 < q < n$ des entiers. Pour tout $j \in \{0; 1; ...; q-1\}$, on définit a_j et r_j comme le quotient et le reste de la division euclidienne de $n - j$ par q, i.e.*

$$n - j = a_j q + r_j \qquad \textit{avec} \qquad 0 \leq r_j < q$$

1. *On a la partition*

$$\{0; 1; ...; n-1\} = \bigsqcup_{\alpha\in\{0;1;...;a_j-1\}} \{j + \alpha q \ \ | \ \ 0 \leq \beta \leq q-1\} \sqcup S_j$$

où $\mathrm{Card}\, S_j \leq 2q$.

2. *Lorsque j parcourt $\{0; ...; q-1\}$ et α parcourt $\{0; 1; ...; a_j - 1\}$, les entiers $j + \alpha q$ sont tous distincts et inférieurs ou égaux à $n - q$*

Démonstration. 1. La partie restante est

$$S_j = \{0; 1; ...; j-1; j+a_jq; ...; j+a_jq+r_j-1\}$$

Il devient dés lors évident, vu le choix de j et la définition de r_j, que

$$\text{Card}\, S_j \leq j + r_j - 1 \leq 2q$$

2. Pour tout j dans $\{0; ...; q-1\}$ et α dans $\{0; 1; ...; a_j - 1\}$,

$$j + \alpha q \leq j + a_j q - q \leq n - r_j - q \leq n - q$$

De plus, $j + \alpha q = j' + \alpha' q$ entraîne que q divise $j - j'$, donc que $j = j'$.

□

Nous pouvons enfin nous atteler à la démonstration, très subtile, de l'autre inégalité du principe variationnel.

Démonstration de la seconde inégalité. On choisit, encore une fois, pour tout entier n et tout cylindre $A \in \mathcal{A}_n$, un mot x_A tel que $S_n\varphi\,(x_A) = \sup_{x\in A} S_n\varphi(x)$.

On définit alors, pour tout entier $n \geq 1$,

$$\mu_n = \frac{1}{\sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)}} \sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)}\delta_{x_A} \qquad \text{et} \qquad \nu_n = \frac{1}{n} S_n^* \mu_n$$

Notons $Z_n = \sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)}$ la constante de normalisation de μ_n et calculons :

$$H_{\mu_n}(\mathcal{A}_n) = -\frac{1}{Z_n} \sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)} \left(\ln\left(e^{S_n\varphi(x_A)}\right) - \ln(Z_n)\right) = -\frac{1}{Z_n} \sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)} S_n\varphi\,(x_A) + \ln(Z_n)$$

Par ailleurs, par définition de μ_n,

$$\int S_n\varphi d\mu_n = \frac{1}{Z_n} \sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)} S_n\varphi\,(x_A)$$

On obtient donc, après division par n, l'égalité suivante :

$$\frac{1}{n} H_{\mu_n}(\mathcal{A}_n) + \int \varphi d\left(\frac{1}{n} S_n^* \mu_n\right) = \frac{1}{n} \ln\left(\sum_{A\in\mathcal{A}_n} e^{S_n\varphi(x_A)}\right)$$

Nous pouvons, d'après le lemme 5.1.4, trouver une suite strictement croissante d'entiers $(n_k)_{k\in\mathbb{N}}$ telle que la suite de probabilités (ν_{n_k}) converge *-faiblement vers une mesure de probabilité invariante ν.

On aura alors :

$$\lim_{k\to+\infty} \int \varphi d\left(\frac{1}{n_k} S_{n_k}^* \mu_{n_k}\right) = \lim_{k\to+\infty} \int \varphi d\nu_k = \int \varphi d\nu$$

Et, bien sûr,

$$\lim_{k\to\infty} \frac{1}{n_k} \ln \left(\sum_{A\in\mathcal{A}_{n_k}} e^{S_{n_k}\varphi(x_A)} \right) = \mathrm{P}\left(\varphi\right)$$

Il ne reste donc qu'à étudier le comportement de $H_{\mu_n}\left(\mathcal{A}_n\right)$ et pour cela, il semble essentiel de faire apparaître les mesures ν_n. C'est là que notre découpage intervient. Soient $0 < q < n$ et $0 \leq j \leq q-1$. En utilisant le lemme d'arithmétique 5.1.5, il vient :

$$\mathcal{A}_n = \bigvee_{\alpha=0}^{a_j-1} T^{-j-\alpha q}\left(\mathcal{A}_q\right) \vee \mathcal{C}_{S_j}$$

où $\mathcal{C}_{S_j}$ est l'information manquante :

$$\mathcal{C}_{S_j} = \left\{ \{x \in X \quad | \quad \forall i \in S_j, \;\; x_i = \ell_i\} \quad | \quad (\ell_i)_{i\in S_j} \in \Sigma^{S_j} \right\}$$

Par sous-additivité, on obtient :

$$H_{\mu_n}\left(\mathcal{A}_n\right) \leq \sum_{\alpha=0}^{a_j-1} H_{T^{j+\alpha q *}\mu_n}\left(\mathcal{A}_q\right) + H_{\mu_n}\left(\mathcal{C}_{S_j}\right) \qquad (I_j)$$

On peut majorer le terme résiduel en utilisant la proposition 4.1.2 :

$$H_{\mu_n}\left(\mathcal{C}_{S_j}\right) \leq \ln\left(\mathrm{Card}\left(\mathcal{C}_{S_j}\right)\right) \leq \ln\left(\left(\mathrm{Card}\,\Sigma\right)^{\mathrm{Card}\,S_j}\right) \leq 2q \ln\left(\mathrm{Card}\,\Sigma\right)$$

D'après la deuxième partie du lemme 5.1.5, en faisant la somme des inégalités (I_j) pour $j = 0, 1, .., q-1$, on obtient :

$$qH_{\mu_n}\left(\mathcal{A}_n\right) \leq \sum_{i=0}^{n-1} H_{T^{i*}\mu_n}\left(\mathcal{A}_q\right) + 2q^2 \ln\left(\mathrm{Card}\,\Sigma\right)$$

On divise ensuite par nq et l'on utilise l'inégalité de concavité 4.1.4 pour obtenir :

$$\frac{1}{n} H_{\mu_n}\left(\mathcal{A}_n\right) \leq \frac{1}{q} H_{\nu_n}\left(\mathcal{A}_q\right) + 2\frac{q}{n} \ln\left(\mathrm{Card}\,\Sigma\right)$$

Il suffit ensuite de remplacer n par n_k dans l'inégalité précédente et de faire tendre k vers $+\infty$ en utilisant la propriété 4.2.1 pour avoir :

$$\limsup_{k\to\infty} \frac{1}{n_k} H_{\mu_{n_k}}\left(\mathcal{A}_{n_k}\right) \leq \frac{1}{q} H_\nu\left(\mathcal{A}_q\right)$$

Enfin, en faisant tendre q vers $+\infty$, il vient :

$$\limsup_{k\to\infty} \frac{1}{n_k} H_{\mu_{n_k}}\left(\mathcal{A}_{n_k}\right) \leq h_\nu\left(T\right)$$

Et cela achève la démonstration de cette inégalité. □

Avant de donner un corollaire très important de ce théorème, nous rappelons quelques lemmes de topologie.

Lemme 5.1.6. *Soient X est un espace topologique compact et $f : X \to \mathbb{R}^+$ une fonction semi-continue supérieurement. Alors f est majorée et atteint sa borne supérieure.*

Démonstration. La semi-continuité supérieure entraîne que, pour tout $\alpha \in \mathbb{R}$, l'ensemble $f^{-1}(]-\infty;\alpha[)$ est ouvert. En effet, soit $x_0 \in f^{-1}(]-\infty;\alpha[)$. On pose $\varepsilon = \dfrac{\alpha - f(x_0)}{2}$ et on utilise la semi-continuité supérieure en x_0 : il existe un voisinage U de x_0, tel que pour tout x dans U,

$$f(x) < f(x_0) + \varepsilon \leq f(x_0) + \frac{\alpha - f(x_0)}{2} \leq \frac{\alpha + f(x_0)}{2} < \alpha$$

On peut alors écrire K comme une réunion d'ouverts :

$$K \subset \bigcup_{M \in \mathbb{R}} f^{-1}(]-\infty; M[)$$

d'où l'on extrait un recouvrement fini

$$K \subset \bigcup_{i=1}^{n} f^{-1}(]-\infty; M_i[)$$

ce qui prouve que f est majorée par $\max_{i\in\{1;\dots;n\}} M_i$.

Notons $s = \sup_{x\in X} f(x)$. Supposons que s ne soit pas atteint. Alors,

$$K \subset \bigcup_{\alpha < s} f^{-1}(]-\infty; \alpha[)$$

On peut alors en extraire un sous-recouvrement fini,

$$K \subset \bigcup_{i=1}^{m} f^{-1}(]-\infty; \alpha_i[)$$

et alors, $\sup_{x\in X} f(x) < \max_{i\in\{1;\dots;m\}} \alpha_i < s$, ce qui est contradictoire. □

Lemme 5.1.7. *Soit X un espace vectoriel topologique, et K une partie convexe et compacte de X. Soit f une fonction affine, semi-continue supérieurement, de K dans $\mathbb{R}$, et $s = \sup_{x\in K} f(x)$.*

Alors, l'ensemble $S = f^{-1}(s)$ est une partie extrémale et non vide de K. En particulier, elle contient au moins un point extrémal de K.

Démonstration. Le fait que S soit non vide a déjà été démontré dans le lemme précédent.

Soient x et y deux éléments de K et $t \in]0;1[$.

Supposons que $tx + (1-t)y$ soit dans S, c'est-à-dire $f(tx + (1-t)y) = s$.

Comme f est affine, on a alors $tf(x) + (1-t)f(y) = s$ et comme $f(x)$ et $f(y)$ sont inférieurs ou égaux à s, nécessairement, $f(x) = f(y) = s$, donc x et y sont dans S, ce qui prouve que S est une partie extrémale de K.

Donc d'après le Théorème de Krein-Milman, elle contient au moins un point extrémal de K. □

Corollaire 5.1.8. *Soit $\varphi \in \mathcal{C}(X)$. L'ensemble des mesures $\nu \in \mathcal{M}(X,T)$ telles que*

$$\mathrm{P}(\varphi) = h_\nu(T) + \int_X \varphi d\nu$$

est non vide et contient au moins une mesure ergodique.

Démonstration. D'après la proposition 4.2.4, l'application

$$\mu \longmapsto h_\nu(T) + \int_X \varphi d\nu$$

est affine et semi-continue supérieurement, du compact convexe non vide $\mathcal{M}(X,T)$ dans $\mathbb{R}^+$.

En appliquant le lemme précédent, on voit que sa borne supérieure est atteinte en au moins un point extrémal de $\mathcal{M}(X,T)$.

Or, d'après le théorème 3.3.1, les points extrémaux de $\mathcal{M}(X,T)$ sont précisément les mesures ergodiques par rapport à T. □

5.2 Théorème de Perron-Frobenius-Ruelle

Définition 5.2.1. Soit $\varphi \in \mathcal{C}(X)$. On définit l'opérateur de Ruelle associé à φ par :

$$\forall f \in \mathcal{C}(X), \ \forall x \in X, \ \mathcal{L}_\varphi f(x) = \sum_{\ell \in \Sigma} e^{\varphi(\ell x)} f(\ell x)$$

Notons que cela définit également un opérateur sur l'ensemble $\mathcal{S}(X)$ des mesures signées sur X par :

$$\forall \mu \in \mathcal{S}(X), \ \forall f \in \mathcal{C}(X), \ \mathcal{L}_\varphi^* \mu(f) = \int_X f d\left(\mathcal{L}_\varphi^* \mu\right) = \int_X \mathcal{L}_\varphi f d\mu$$

Proposition 5.2.1. *1. Les opérateurs $\mathcal{L}_\varphi$ et $\mathcal{L}_\varphi^*$ sont continus et positifs.*
2. Pour tous $f \in \mathcal{C}(X)$, $x \in X$ et $n \geq 1$,

$$\mathcal{L}_\varphi^n f(x) = \sum_{A \in \Sigma^n} e^{S_n\varphi(Ax)} f(Ax) = \sum_{y \in T^{-n}(x)} e^{S_n\varphi(y)} f(y)$$

Démonstration. Comme φ est continue sur X, elle est bornée, et l'on voit que

$$|\mathcal{L}_\varphi f(x)| \leq \|f\|_\infty \operatorname{Card}(\Sigma) e^{\|\varphi\|_\infty}$$

donc $\mathcal{L}_\varphi$ est continu et son adjoint l'est aussi. La positivité est claire (mais essentielle) et la formule sur les itérés s'obtient simplement par récurrence. □

Sous certaines conditions de régularité supplémentaires sur φ, il est possible de contrôler les fonctions $S_n\varphi$ de façon uniforme. Pour cela on introduit l'espace $\mathcal{B}_0$.

Définition 5.2.2. On note $\mathcal{B}_0(X)$ l'ensemble des fonctions $f : X \to \mathbb{R}$ telles que :

$$\|f\|_{\mathcal{B}_0} := \sum_{k \geq -1} \omega_f\left(2^{-k-1}\right) < +\infty$$

Cet espace contient les fonctions lipschitziennes et même, plus généralement, toutes les fonctions hölderiennes.

Théorème 5.2.2 (Théorème de Perron-Frobenius-Ruelle)**.** *Soit $\varphi \in \mathcal{B}_0(X)$ et $\mathcal{L}_\varphi$ l'opérateur de Ruelle qui lui est associé.*

1. *Il existe un unique réel $\beta > 0$ tel que l'adjoint $\mathcal{L}_\varphi^*$ admette une mesure de probabilité propre, .i e.*
$$\exists \mu \in \mathcal{M}(X), \;\; \mathcal{L}_\varphi^*(\mu) = \beta\mu$$
De plus, cette mesure de probabilité propre est unique.
2. *Le réel β est également valeur propre simple de $\mathcal{L}_\varphi$ et il existe une unique fonction propre $h > 0$ dans $\mathcal{C}(X)$ telle que*
$$\mathcal{L}_\varphi(h) = \beta h \quad et \quad \int_X h d\mu = 1$$
3. *Pour toute fonction $v \in \mathcal{C}(X)$, on a, uniformément sur X,*
$$\beta^{-n}\mathcal{L}_\varphi^n(v) \underset{n\to+\infty}{\longrightarrow} h\int_X v d\mu$$
4. *Enfin, $\beta = e^{\mathrm{P}(\varphi)}$, où $\mathrm{P}(\varphi)$ est la pression topologique de φ.*

Le reste de cette partie est consacré à la démonstration de ce théorème.

a) Commençons par montrer l'existence dans 1.
Pour cela, nous allons appliquer le théorème de Schauder-Tychonov dans le compact convexe $\mathcal{M}(X)$.
Avant cela, remarquons que, comme X est compact et φ est continue, on a, pour tout $x \in X$,
$$\mathcal{L}_\varphi(1)(x) = \sum_{\ell\in\Sigma} e^{\varphi \ell x} \geq \mathrm{Card}\,(\Sigma)\, e^{-\|\varphi\|_\infty} := b > 0$$
En particulier, pour toute mesure de probabilité ν sur X, on a
$$\int_X \mathcal{L}_\varphi(1) d\nu \geq b > 0$$
Définissons sur $\mathcal{M}(X)$ la fonction
$$\Phi : \nu \longmapsto \frac{\mathcal{L}_\varphi^*(\nu)}{\int_X \mathcal{L}_\varphi(1)d\nu}$$
Par positivité de l'opérateur $\mathcal{L}_\varphi$, $\Phi(\nu)$ est une mesure positive, et c'est encore une mesure de probabilité car
$$\int_X 1d\,(\Phi(\nu)) = \frac{1}{\int_X \mathcal{L}_\varphi(1)d\nu}\int_X 1d\left(\mathcal{L}_\varphi^*(\nu)\right) = \frac{1}{\int_X \mathcal{L}_\varphi(1)d\nu}\int_X \mathcal{L}_\varphi(1)d\nu = 1$$
Ainsi, Φ est continue du compact convexe $\mathcal{M}(X)$ dans lui-même, donc d'après le théorème de Tychonov, elle y admet un point fixe μ. Il suffit alors de poser
$$\beta = \int_X \mathcal{L}_\varphi(1)d\mu$$
pour obtenir $\mathcal{L}_\varphi^*(\mu) = \beta\mu$, avec $\beta > 0$.

b) Pour obtenir l'existence de la fonction propre h, il suffirait de pouvoir utiliser de nouveau le théorème de Schauder-Tychonov à l'opérateur $\mathcal{T} := \beta^{-1}\mathcal{L}_\varphi$. Pour cela, il faut travailler dans une partie convexe compact de $\mathcal{C}(X)$. Les parties compactes de $\mathcal{C}(X)$ sont caractérisées par le Théorème d'Ascoli. Il nous faut donc trouver un ensemble convexe et assez petit de fonctions continues qui soit en plus stable par $\mathcal{T}$. Le lemme suivant va nous permettre de construire une telle partie.

Lemme 5.2.3. *Soit $C : X \times X \to \mathbb{R}^+$ une fonction telle que*

$$C(x,y) \underset{d(x,y)\to 0}{\longrightarrow} 0 \quad \textit{et} \quad \exists M > 0,\ \forall x, y \in X,\ C(x,y) \leq M$$

Alors,

$$K = \left\{ f \in \mathcal{C}(X) \ \mid \ f \geq 0,\ \int_X f d\mu = 1,\ \forall x, y \in X,\ f(x) \leq f(y) e^{C(x,y)} \right\}$$

est une partie compacte, convexe et non vide de $\mathcal{C}(X)$.

Démonstration du lemme. La partie K est non vide car la fonction constante égale à 1 lui appartient.
De plus, K est bornée. En effet, si $f \in K$, il existe $y_0 \in X$ tel que $f(y_0) \leq 1$, sinon l'intégrale $\int_X f d\mu$ ne serait pas égale à 1 et la dernière condition impose que pour tout $x \in X$, $0 \leq f(x) \leq e^{C(x,y_0)} f(y_0) \leq e^M$ et donc que $\|f\|_\infty \leq e^M$.
La partie K est équicontinue : pour tout f dans K, on a

$$f(x) - f(y) \leq f(y)\left(e^{C(x,y)} - 1\right) \leq e^M \left(e^{C(x,y)} - 1\right)$$

et, de façon symétrique, $f(y) - f(x) \leq e^M \left(e^{C(y,x)} - 1\right)$, d'où

$$\forall f \in K,\ \forall x, y \in X,\ |f(x) - f(y)| \leq e^M \left[\max\left(e^{C(x,y)}, e^{C(y,x)}\right) - 1\right]$$

et cette dernière expression tend bien vers 0 lorsque $d(x,y)$ tend vers 0.
D'après le Théorème d'Ascoli, K est relativement compacte dans $\mathcal{C}(X)$ et il est clair que les trois conditions définissant K sont fermées. Enfin, la convexité de K se vérifie simplement. □

Les deux premières conditions définissant l'ensemble K sont bien préservées par $\mathcal{T}$: la première car $\mathcal{L}_\varphi$ est un opérateur positif et $\beta > 0$, la deuxième car μ est un vecteur propre à gauche de $\mathcal{L}_\varphi$ pour β. Il reste à trouver une bonne fonction C pour que la dernière condition soit préservée par $\mathcal{T}$, ou, ce qui revient au même, par $\mathcal{L}_\varphi$.
On doit avoir en particulier, pour tout f dans K et tous x et y dans X,

$$\mathcal{L}_\varphi(f)(x) \leq e^{C(x,y)} \mathcal{L}_\varphi(f)(y), \ \text{ i.e. } \sum_{\ell \in \Sigma} e^{\varphi(\ell x)} f(\ell x) \leq e^{C(x,y)} \sum_{\ell \in \Sigma} e^{\varphi(\ell y)} f(\ell y) \quad (1)$$

Mais comme f est dans K, on a, pour tout x, y dans X,

$$\sum_{\ell \in \Sigma} e^{\varphi(\ell x)} f(\ell x) \leq \sum_{\ell \in \Sigma} e^{\varphi(\ell x)} f(\ell y) e^{C(\ell x, \ell y)} \leq \sum_{\ell \in \Sigma} e^{\varphi(\ell y)} f(\ell y) e^{\varphi(\ell x) - \varphi(\ell y) + C(\ell x, \ell y)}$$

où l'on voit qu'une condition suffisante pour avoir l'inégalité (1) est :

$$\forall x, y \in X, \ \ \forall \ell \in \Sigma, \ \ \varphi(\ell x) - \varphi(\ell y) + C(\ell x, \ell y) \leq C(x, y)$$

Cette dernière condition, appliquée récursivement, donne alors :

$$\forall x, y \in X, \ \ \forall n \geq 1, \forall A \in \Sigma^n, \ \ C(x, y) \geq S_n\varphi(Ax) - S_n\varphi(Ay) + C(Ax, Ay) \quad (*)$$

Comme C est positive, une condition nécessaire pour que $(*)$ ait lieu, est que

$$\forall x, y \in X, \ \ C(x, y) \geq \sup_{n \geq 1} \sup_{A \in \Sigma^n} (S_n\varphi(Ax) - S_n\varphi(Ay))$$

Posons maintenant, pour x, y dans X,

$$C_\varphi(x, y) = \sup_{n \geq 1} \sup_{A \in \Sigma^n} (S_n\varphi(Ax) - S_n\varphi(Ay))$$

La condition de continuité $\varphi \in \mathcal{B}_0(X)$ apparaît de façon naturelle, pour pouvoir contrôler la fonction C_φ.

Lemme 5.2.4. *Si $\varphi \in \mathcal{B}_0(X)$, alors C_φ vérifie la condition $(*)$, ainsi que les hypothèses du lemme 5.2.3.*

Démonstration. En effet, soit $k \in \mathbb{N}$ et x, y tels que $d(x, y) \leq 2^{-k-1}$. On a alors, pour tout $n \geq 1$, pour tout $A \in \Sigma^n$,

$$\begin{aligned} S_n\varphi(Ax) - S_n\varphi(Ay) &\leq \omega_\varphi\left(2^{-k-1-n}\right) + \omega_\varphi\left(2^{-k-n}\right) + \ldots + \omega_\varphi\left(2^{-k-2}\right) \\ &= \sum_{i=k+1}^{n+k} \omega_\varphi\left(2^{-i-1}\right) \leq \sum_{i=k+1}^{+\infty} \omega_\varphi\left(2^{-i-1}\right) \end{aligned}$$

Ainsi, C_φ est bornée par $\|\varphi\|_{\mathcal{B}_0}$ et $C(x, y)$ tend vers 0 lorsque $d(x, y)$ tend vers 0.
Il reste à démontrer que pour tous $\ell \in \Sigma$, $x, y \in X$, on a

$$\varphi(\ell x) - \varphi(\ell y) + C_\varphi(\ell x, \ell y) \leq C_\varphi(x, y)$$

Soient $n \geq 1$ et $A \in \Sigma^n$, on a

$$S_n\varphi(A\ell x) - S_n\varphi(A\ell y) + \varphi(\ell x) - \varphi(\ell y) = S_{n+1}\varphi((A\ell)x) - S_{n+1}\varphi((A\ell)y) \leq C_\varphi(x, y)$$

et il suffit de passer au sup, pour les A de Σ^n, puis les entiers $n \geq 1$. □

Notons

$$K_\varphi = \left\{ f \in \mathcal{C}(X) \ \ | \ \ f \geq 0, \ \ \int_X f d\mu = 1, \ \ \forall x, y \in X, \ \ f(x) \leq f(y) e^{C_\varphi(x,y)} \right\}$$

D'après les deux lemmes précédents, K_φ est un compact convexe de l'espace vectoriel normé $\mathcal{C}(X)$. De plus, il est stable par $\mathcal{T}$, donc toutes les conditions sont réunies pour appliquer le théorème de Schauder-Tychonov : il existe une fonction $h \in K_\varphi$ telle que $\mathcal{L}_\varphi(h) = \beta h$. En particulier, comme $h \in K_\varphi$, on a $h > 0$ et $\int_X h d\mu = 1$.

c) Passons à l'unicité de h. Soit g une autre fonction propre pour la valeur propre $\beta > 0$. Quitte à prendre son opposée si besoin, on peut supposer qu'il existe $z \in X$ tel que $g(z) > 0$.
Ainsi, l'ensemble de réels $\{u \geq 0 \quad | \quad h \geq ug\}$ est bien majoré. Soit t sa borne supérieure. On a $t > 0$, sinon on aurait $h(z) = 0$.
De plus, $h \geq tg$, sinon, par compacité de X, il existerait un réel α tel que $h - tg \leq \alpha < 0$, donc pour $\varepsilon > 0$ assez petit, on aurait

$$h - (t - \varepsilon)g = h - tg + \varepsilon g \leq \alpha + \varepsilon g < 0$$

Enfin, il existe $x_0 \in X$ tel que $h(x_0) - tg(x_0) = 0$, sinon, il existerait $\alpha' > 0$ tel que pour tout $x \in X$,
$h(x) - tg(x) \geq \alpha' > 0$, ce qui entraînerait, pour $\varepsilon > 0$ assez petit,

$$h(x) - (t + \varepsilon)g(x) \geq \alpha' - \varepsilon g(x) > 0$$

Comme h et g sont des fonctions propres pour β, on a alors, pour tout $n \in \mathbb{N}$,

$$0 = \beta^n \left(h(x_0) - tg(x_0)\right) = \sum_{y \in T^{-n}(x_0)} e^{S_n\varphi(y)} \left(h - tg\right)(y)$$

Mais cela implique que la fonction $h - tg$ est nulle sur $\bigcup_{n \geq 1} T^{-n}(x_0)$ et ce dernier ensemble est dense dans X, car il contient un mot de chacun de ses cylindres. Finalement, par continuité de $h - tg$, on a $h = tg$ sur X, donc la valeur propre β est simple à droite.

d) Il est pratique de "normaliser" φ, et pour cela, d'introduire :

$$\psi = \varphi - \ln h \circ T + \ln h - \ln \beta$$

de façon à avoir

$$\mathcal{L}_\psi = \beta^{-1} M_h^{-1} \mathcal{L}_\varphi M_h$$

Pour cet opérateur $\mathcal{L}_\psi$, la fonction constante égale à 1 est vecteur propre (à droite) pour la valeur propre 1, et la mesure de probabilité $\nu := h\mu$ est propre à gauche, encore pour la valeur propre 1 :

$$\mathcal{L}_\psi(1) = 1 \quad \text{et} \quad \mathcal{L}_\psi^*(h\mu) = h\mu$$

Remarquons que $S_n\psi = S_n\varphi - \ln h \circ T^n + \ln h - n \ln \beta$, donc que, pour tout $n \geq 1$ et $A \in \Sigma^n$,

$$S_n\psi(Ax) - S_n\psi(Ay) = S_n\varphi(Ax) - S_n\varphi(Ay) + \ln h(Ax) - \ln h(Ay) - \ln h(x) + \ln h(y)$$

Donc

$$S_n\psi(Ax) - S_n\psi(Ay) \leq C_\varphi(x, y) + 2\omega_{\ln h}(d(x, y))$$

d'où l'on déduit que $C_\psi(x, y)$ est bornée et tend vers 0 lorsque $d(x, y)$ tend vers 0.
Nous allons maintenant démontrer que, pour toute fonction $v \in \mathcal{C}(X)$, on a la convergence uniforme

$$\mathcal{L}_\psi^n v \underset{n \to \infty}{\longrightarrow} \int_X v d\nu$$

Quitte à lui soustraire son minimum, on peut supposer que $v \geq 0$. Cela ne pose pas de problème car X est compact, $\mathcal{L}_\psi$ est l'identité sur les constantes et ν est une mesure de probabilité.
Remarquons qu'alors, $\|v\|_\infty = \sup_{x \in X} v(x)$

e) On commence par montrer que la suite $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ est relativement compacte. Pour cela, on utilise encore le théorème d'Ascoli.
D'abord, la norme de l'opérateur $\mathcal{L}_\psi$ est inférieure ou égale à 1 : en effet, pour tout $f \in \mathcal{C}(x)$, on a :

$$|\mathcal{L}_\psi f(x)| = \left|\sum_{\ell\in\Sigma} e^{\ell x} f(\ell x)\right| \leq \|f\|_\infty \sum_{\ell\in\Sigma} e^{\ell x} = \|f\|_\infty \mathcal{L}_\psi(1)(x) = \|f\|_\infty$$

On a donc, pour tout $n \in \mathbb{N}$, $\left\|\mathcal{L}_\psi^n v\right\|_\infty \leq \|v\|_\infty$ et la suite $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ est bien bornée.
Pour ce qui est de l'équicontinuité, on a, pour tout $n \in \mathbb{N}$, et x, y dans X tels que $d(x,y) \leq 2^{-k-1}$,

$$\left|\mathcal{L}_\psi^n v(x) - \mathcal{L}_\psi^n(y)\right| \leq \left|\sum_{A\in\Sigma^n} e^{S_n\psi(Ax)}\left(v\left(Ax\right) - v\left(Ay\right)\right)\right| + \left|\sum_{A\in\Sigma^n} v\left(Ay\right)\left(e^{S_n\psi(Ax)} - e^{S_n\psi(Ay)}\right)\right|$$

Majorons la première des deux sommes :

$$\begin{aligned}\left|\sum_{A\in\Sigma^n} e^{S_n\psi(Ax)}\left(v\left(Ax\right) - v\left(Ay\right)\right)\right| &\leq \left(\sum_{A\in\Sigma^n} e^{S_n\psi(Ax)}\right) \sup_{A\in\Sigma^n} \left|v\left(Ax\right) - v\left(Ay\right)\right| \\ &\leq \mathcal{L}_\psi^n(1)(x)\omega_v\left(2^{-n-k-1}\right) \leq \omega_v\left(2^{-k-1}\right)\end{aligned}$$

Pour la seconde somme, on doit se servir de la fonction C_ψ :

$$\left|\sum_{A\in\Sigma^n} v\left(Ay\right)\left(e^{S_n\psi(Ax)} - e^{S_n\psi(Ay)}\right)\right| \leq \|v\|_\infty \sum_{A\in\Sigma^n} e^{S_n\psi(Ay)}\left|e^{S_n\psi(Ax)-S_n\psi(Ay)} - 1\right|$$

d'où l'on déduit,

$$\left|e^{S_n\psi(Ax)-S_n\psi(Ay)} - 1\right| \leq \max\left(e^{C_\psi(x,y)} - 1, 1 - e^{-C_\psi(y,x)}\right)$$

Finalement, la majoration obtenue est :

$$\left|\mathcal{L}_\psi^n v(x) - \mathcal{L}_\psi^n(y)\right| \leq \omega_v\left(2^{-k-1}\right) + \|v\|_\infty \max\left(e^{C_\psi(x,y)} - 1, 1 - e^{-C_\psi(y,x)}\right)$$

et le majorant tend vers 0 lorsque $d(x,y)$ tend vers 0 uniformément en n, ce qui achève de montrer l'équicontinuité de la suite $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ et donc sa compacité relative.

f) Soit $\left(\mathcal{L}_\psi^{n_k} v\right)_{k\in\mathbb{N}}$ une suite extraite de $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ qui converge vers une fonction v^*. Nous allons montrer que v^* est constante.

Notons que par continuité de $\| \ \|_\infty$, $\|v^*\|_\infty = \lim_{k\to+\infty} \left\|\mathcal{L}_\psi^{n_k} v\right\|_\infty$.
Mais comme la suite $\left(\left\|\mathcal{L}_\psi^{n} v\right\|_\infty\right)_{n\in\mathbb{N}}$ est décroissante, elle converge et c'est donc vers $\|v^*\|_\infty$.
Or, pour tout $k \geq 0$, on a

$$v^* \geq \mathcal{L}_\psi^{n_k}(v) - \varepsilon_k \qquad \text{avec } \varepsilon_k \underset{k\to+\infty}{\longrightarrow} 0$$

En appliquant à cette inégalité l'opérateur positif $\mathcal{L}_\psi^N$, pour un N quelconque, il vient :

$$\mathcal{L}_\psi^N(v^*) \geq \mathcal{L}_\psi^{N+n_k} v - \varepsilon_k$$

Puis, comme $\sup_{x\in X} \mathcal{L}_\psi^{N+n_k} v(x) \geq \sup_{x\in X} v^*(x)$, on obtient, pour tout N,

$$\sup_{x\in X} \mathcal{L}_\psi^N(v^*) \geq \sup_{x\in X} v^*(x) =: s$$

Or, comme X est compact, il existe x_0 dans X tel que $v^*(x_0) = s$ et $x_N \in X$ tel que

$$\mathcal{L}_\psi^N(v^*)(x_N) = \sup_{x\in X} \mathcal{L}_\psi^N(v^*)$$

On a donc, pour tout N,

$$s = v^*(x_0) \leq \sum_{x\in T^{-N}(x_N)} e^{S_N\psi(x)} v^*(x) \leq \sum_{x\in T^{-N}(x_N)} e^{S_N\psi(x)} v^*(x_0) \leq v^*(x_0)$$

ce qui impose, pour tout $x \in T^{-N}(x_N)$, $v^*(x) = v^*(x_0) = s$ et l'on conclut par densité de $\bigcup_{N\in\mathbb{N}} T^{-N}(x_N)$ et que v^*, continue, est constante et égale à s.
La convergence étant uniforme sur le compact X, on a également,

$$\int_X \mathcal{L}_\psi^{n_k} v d\nu \underset{k\to\infty}{\longrightarrow} \int_X s d\nu = s$$

Or, comme ν est propre à gauche pour la valeur propre 1, on a

$$\int_X \mathcal{L}_\psi^{n_k} v d\nu = \int_X v d\left(\left(\mathcal{L}_\psi^{n_k}\right)^*(\nu)\right) = \int_X v d\nu$$

Donc $s = \int_X v d\nu$. Et ce dernier résultat, comme tout ce qui précède, ne dépend pas de la suite extraite convergente choisie : toute suite extraite convergente de $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ converge vers la fonction constante égale à $\int_X v d\nu$.
Ainsi, la suite $\left(\mathcal{L}_\psi^n v\right)_{n\in\mathbb{N}}$ est relativement compacte et n'a qu'une seule valeur d'adhérence. Elle converge vers $\int_X v d\nu$.
Ceci nous permet également de montrer que ν est unique : si ν' est une mesure de probabilité telle que $\mathcal{L}_\psi^*(\nu') = \nu'$, on a également, pour toute fonction $v \in \mathcal{C}(X)$

$$\mathcal{L}_\psi^n v \underset{n\to+\infty}{\longrightarrow} \int_X v d\nu'$$

donc pour toute fonction $v \in \mathcal{C}(X)$, $\int_X v d\nu = \int_X v d\nu'$, ce qui implique $\nu = \nu'$

g) Ceci implique également l'unicité de μ, car les mesures de probabilité propres de $\mathcal{L}_\varphi^*$ pour β sont en bijection avec celles de $\mathcal{L}_\psi^*$ pour la valeur propre 1, via

$$\mu \longmapsto h\mu \qquad \text{où } \ h > 0$$

Enfin, pour toute fonction $v \in \mathcal{C}(X)$,

$$\beta^{-n}\mathcal{L}_\varphi^n(v) = M_h \mathcal{L}_\psi^n M_h^{-1}(v)$$

et le membre de droite converge uniformément vers

$$h \int_X v \frac{d\nu}{h} = h \int_X v d\mu$$

Il ne reste qu'à montrer l'égalité $\beta = e^{\mathrm{P}(\varphi)}$, qui implique l'unicité de β. Pour cela, on applique ce qui précède à la fonction constante égale à 1. Ainsi, pour tout $x \in X$,

$$\mathcal{L}_\varphi^n(1)(x) = \sum_{A \in \Sigma^n} e^{S_n\varphi(Ax)} \sim_{n\to+\infty} \beta^n h(x) \qquad \text{avec } \ h(x) > 0$$

Par ailleurs, la pression $\mathrm{P}\left(\varphi\right)$ est caractérisée par

$$\sum_{A \in \Sigma^n} e^{\sup_{x\in X} S_n\varphi(Ax)} = \sum_{A \in \Sigma^n} e^{S_n\varphi(Ax_A)} \sim_{n\to+\infty} e^{n\,\mathrm{P}(\varphi)}$$

où x_A est tel que $S_n\varphi\left(Ax_A\right) = \sup_{x\in X} S_n\varphi\left(Ax\right)$.
Il est clair que $\mathcal{L}_\varphi^n(1)(x) \leq \sum_{A\in\Sigma_n} e^{\sup_{x\in X} S_n\varphi(Ax)}$, donc que $\ln\left(\beta\right) \leq \mathrm{P}\left(\varphi\right)$.
D'autre part,

$$\begin{aligned}\sum_{A \in \Sigma^n} e^{S_n\varphi(Ax_A)} &= \sum_{A \in \Sigma^n} e^{S_n\varphi(Ax_A) - S_n\varphi(Ax)} e^{S_n\varphi(Ax)} \\ &\leq \sum_{A \in \Sigma^n} e^{C_\varphi(x_A, x)} e^{S_n\varphi(Ax)} \leq e^M \mathcal{L}_\varphi^n(1)(x)\end{aligned}$$

D'où l'on déduit que

$$\sum_{A \in \Sigma^n} e^{S_n\varphi(Ax_A)} = O_{n\to+\infty}\left(\mathcal{L}_\varphi^n(1)(x)\right)$$

ce qui impose $\mathrm{P}\left(\varphi\right) \leq \ln\left(\beta\right)$ et conclut cette démonstration.

A Quelques rappels de topologie générale

A.1 Lemme de Zorn

Le lemme de Zorn est un outil extrêmement puissant lorsque l'on a besoin de raisonner par induction mais que les ensembles considérés sont trop gros pour pouvoir utiliser le raisonnement par récurrence.

Nous le rappelons car il sera utilisé trois fois dans ces appendices.

Définition A.1.1. Soit $\mathcal{E}$ un ensemble muni d'une relation d'ordre $\preceq$ (a priori partielle). On dit que $\mathcal{E}$ est *inductif* lorsque toute partie $\mathcal{F}$ de $\mathcal{E}$ qui est strictement ordonnée par $\preceq$, admet un majorant dans $\mathcal{E}$, c'est-à-dire qu'il existe un élément $M \in E$ tel que pour tout $F \in \mathcal{F}$, $F \preceq M$.

Le lemme de Zorn affirme alors :

Théorème A.1.1 (Lemme de Zorn)**.** *Tout ensemble non vide et inductif admet des éléments maximaux.*

Rappelons qu'un élément S de $\mathcal{E}$ est dit maximal si

$$\forall E \in \mathcal{E}, \;\; S \preceq E \Rightarrow S = E$$

Il est un peu abusif de parler de théorème (ou même de lemme) car le lemme de Zorn est équivalent à l'axiome du choix (voir [Sch95], par exemple, pour une démonstration).

A.2 Topologie et systèmes fondamentaux de voisinages

Définition A.2.1. Soit X un ensemble. Une *topologie* sur X est une partie $\mathcal{O}$ de $\mathcal{P}(X)$ contenant X, $\emptyset$, stable par intersection finie et réunion quelconque. Les éléments de X sont appelés ouverts de la topologie $\mathcal{O}$. Le complémentaire d'un ouvert est appelé un fermé. Un ensemble muni d'une topologie est appelé espace topologique.

Exemple A.2.1. On peut toujours définir au moins deux topologies sur un ensemble X. La topologie discrète est définie par $\mathcal{O} = \mathcal{P}(X)$. C'est la plus grande (on dit aussi la "plus fine", la "plus forte") des topologies sur X.

À l'inverse, la topologie grossière est $\mathcal{O} = \{\emptyset, X\}$. C'est la plus petite (la "moins fine", la "plus faible") des topologies sur X.

Définition A.2.2. Soient X un espace topologique et $x \in X$. On appelle voisinage de x, toute partie de X contenant un ouvert qui contient x. On appelle système fondamental de voisinage de x, toute partie $\mathcal{V}_x$ de $\mathcal{P}(X)$, formée de voisinages de x, et telle que tout voisinage de x contienne un élément de $\mathcal{V}_x$.

Remarquons que tout point $x \in X$ admet un système fondamental de voisinages ouverts.

Proposition A.2.1. *Une partie U de X est un ouvert si et seulement si c'est un voisinage de chacun de ses points.*

Démonstration. En effet, pour tout $u \in U$, il existe un ouvert V_u contenant u et inclus dans U, de sorte que $U = \bigcup_{u \in U} V_u$ est un ouvert. $\square$

Remarque A.2.1. Ainsi, une topologie sur un ensemble X est entièrement déterminée par la donnée de systèmes fondamentaux de voisinages pour chaque point de X.

Définition A.2.3. Soient X et Y deux espaces topologiques.

1. Soit (x_n) une suite d'éléments de X. On dit que la suite (x_n) converge vers un élément ℓ de X lorsque tout voisinage de ℓ contient tous les termes de la suite (x_n) à partir d'un certain rang.
2. On dit qu'une application $f : X \to Y$ est continue au point $x \in X$ lorsque pour tout voisinage V de $f(x)$, il existe un voisinage U de x tel que $f(U) \subset V$.
3. On dit que f est continue sur X si elle est continue en tout point de X.

Proposition A.2.2. *Soit $f : X \to Y$. Alors f est continue sur X si et seulement si pour tout ouvert V de Y, $f^{-1}(V)$ est un ouvert de X.*

Démonstration. Posons $U = f^{-1}(V)$. Soit $x \in U$ et $y = f(x)$. L'élément y appartient à V qui est ouvert, donc V est un voisinage de y. Il existe donc un voisinage W_x de x tel que $f(W_x) \subset V$, donc $W_x \subset U$. L'ensemble U contient un voisinage de x, donc est lui-même un voisinage de x, pour chacun de ses éléments x : c'est un ouvert. La réciproque est évidente. $\square$

Définition A.2.4. Soit X un ensemble muni d'une topologie. On dit que la topologie de X est séparée lorsque deux points distincts de X admettent toujours des voisinages disjoints.

A.3 Topologies engendrées par des familles de parties, topologies initiales

Proposition A.3.1. *Toute intersection de topologies sur un ensemble X est encore une topologie sur X.*

Bien qu'évidente, cette proposition est très importante pour la suite.

Considérons une famille de parties $\mathcal{A} \subset \mathcal{P}(X)$. L'ensemble des topologies qui contiennent cette famille est non vide, car il contient la topologie discrète. Il admet donc un plus petit élément d'après la proposition précédente.

Définition A.3.1. Soit X un ensemble et $\mathcal{A} \subset \mathcal{P}(X)$. On appelle topologie engendrée par $\mathcal{A}$, la plus petite topologie sur X contenant $\mathcal{A}$.

Proposition A.3.2. *Soit X un ensemble et $\mathcal{A} \subset \mathcal{P}(X)$, alors les ouverts de la topologie engendrée par $\mathcal{A}$ sont les ensembles X, $\emptyset$, ou toute* réunion quelconque d'intersections finies d'éléments de $\mathcal{A}$. *En particulier, chaque élément x de X admet un système fondamental de voisinages formé d'intersections finies d'éléments de $\mathcal{A}$.*

Démonstration. Commençons par supposer que $\mathcal{A}$ est stable par intersection finie. Posons

$$\mathcal{O} = \{\emptyset, X\} \cup \left\{ \bigcup_{A \in I} A \quad | \quad I \subset \mathcal{P}(\mathcal{A}) \right\}$$

Il est clair que $\mathcal{O}$ est stable par réunion quelconque. Montrons qu'il est stable par intersection finie. Pour cela, considérons les ensembles

$$V_1 = \bigcup_{A_1 \in I_1} A_1, \;\; V_2 = \bigcup_{A_2 \in I_2} A_2, \;\; \ldots \;\; V_n = \bigcup_{A_n \in I_n} A_n \qquad I_1, I_2, \ldots, I_n \subset \mathcal{P}(\mathcal{A})$$

et $V = V_1 \cap V_2 \cap \ldots \cap V_n$. On a alors

$$V = \bigcup_{A_1 \in I_1} A_1 \cap \bigcup_{A_2 \in I_2} A_2 \cap \ldots \cap \bigcup_{A_n \in I_n} A_n = \bigcup_{(A_1, A_2, \ldots, A_n) \in I_1 \times I_2 \times \ldots \times I_n} A_1 \cap A_2 \cap \ldots \cap A_n$$

Comme $\mathcal{A}$ est supposé stable par intersection finie, chaque élément de cette réunion est dans $\mathcal{A}$, donc V est dans $\mathcal{O}$, qui est alors clairement la plus petite topologie contenant $\mathcal{A}$.

Pour le cas général, soit $\mathcal{A}'$ l'ensemble des intersections finies d'éléments de $\mathcal{A}$. Les familles $\mathcal{A}'$ et $\mathcal{A}$ engendrent la même topologie et on peut appliquer ce qui précède à $\mathcal{A}'$. □

Définition A.3.2. Lorsque $\mathcal{A}$ engendre une topologie $\mathcal{O}$, on dit que $\mathcal{A}$ est une *prébase* de la topologie $\mathcal{O}$. Si de plus, $\mathcal{A}$ est stable par intersections finies, on dit qu'il en est une *base*.

Exemple A.3.1. Si X est un ensemble muni d'une distance, sa topologie métrique est celle engendrée par les boules ouvertes. Il n'est pas nécessaire ici de travailler avec les intersections finies de boules ouvertes, car une intersection finie de boules ouverte est soit vide, soit contient une boule ouverte. Ainsi, les boules ouvertes centrées en $x \in X$ forment un système fondamental de voisinages de x.

Voyons maintenant un procédé classique d'obtention d'une topologie, qui nous servira beaucoup par la suite.

Définition A.3.3. Soient X un ensemble, $(Y_i)_{i \in I}$ une famille d'espaces topologiques, et $(f_i)_{i \in I}$ une famille d'applications, où $f_i : X \to Y_i$. On appelle *topologie initiale* définie par les $(f_i)_{i \in I}$, la topologie la moins fine rendant toutes les applications f_i continues.

Remarque A.3.1. C'est la topologie engendrée par $\bigcup_{i \in I} \left\{ f_i^{-1}(V) \quad | \quad V \text{ ouvert de } Y_i \right\}$ donc la proposition précédente s'applique pour donner des systèmes fondamentaux de voisinages très pratiques.

Proposition A.3.3 (Propriété fondamentale de la topologie initiale)**.** *Avec les notations précédentes, soit Z un espace topologique et $f : Z \to X$. f est continue pour la topologie initiale définie par la famille $(f_i)_{i \in I}$ si, et seulement si, toutes les applications $f_i \circ f$ le sont.*

Démonstration. Soit $z \in Z$. Posons $x = f(z)$ et, pour tout $i \in I$, $y_i = f_i(x)$. Soit W un voisinage de x dans X. Il existe alors i_1, i_2, ..., i_n dans I, V_{i_1} voisinage de y_{i_1} dans Y_{i_1}, ..., V_{i_n} voisinage ouvert de y_{i_n} dans Y_{i_n} tels que

$$f_{i_1}^{-1}(V_{i_1}) \cap f_{i_2}^{-1}(V_{i_2}) \cap ... \cap f_{i_n}^{-1}(V_{i_n}) \subset W$$

Comme chaque $f_{i_k} \circ f$ est supposée continue en z, il existe U_{i_k}, voisinage de z tel que $f_{i_k} \circ f(U_{i_k}) \subset V_{i_k}$, c'est-à-dire $f(U_{i_k}) \subset f_{i_k}^{-1}(V_{i_k})$. Pour conclure, on pose $U = U_{i_1} \cap U_{i_2} \cap ... \cap U_{i_n}$ et on a $f(U) \subset W$. □

Proposition A.3.4. *Avec les notations de la définition précédente, si les Y_i sont séparés et les f_i séparent les points de X (i.e. pour tout $x \neq y$ dans X, il existe $i \in I$ tel que $f_i(x) \neq f_i(y)$), alors la topologie initiale de X associée aux f_i est* séparée.

Démonstration. Soient $x \neq y$ dans X et i tel que $f_i(x) \neq f_i(y)$. Comme Y_i est séparé, il existe des voisinages ouverts disjoints U_i et V_i de $f_i(x)$ et $f_i(y)$, et par définition de la topologie de X, $f_i^{-1}(U_i)$ et $f_i^{-1}(V_i)$ sont des voisinages ouverts disjoints de x et de y dans X. □

Exemple A.3.2 (Topologie induite)*.* Soit X un espace topologique et $A \subset X$. La topologie induite sur A est la topologie initiale associée à l'injection de A dans X. Les ouverts de A sont donc tous les ensembles $A \cap U$, où U est un ouvert de X. Remarquons que si A est ouvert, ses ouverts sont des ouverts de X et si A est fermé, ses fermés sont des fermés de X.

Exemple A.3.3 (Topologie d'un produit fini)*.* Soient X et Y deux espaces topologiques.

Soient $\pi_1 : X \times Y \to X$ et $\pi_2 : X \times Y \to Y$. La topologie produit sur $X \times Y$ est la topologie initiale associée à ces deux projections canoniques.

En particulier, toute application à valeurs dans $X \times Y$ est continue si et seulement si ses composantes le sont.

Toute application $f : X \times Y \to Z$ est continue en (x, y) si et seulement si, pour tout voisinage W de $f(x, y)$, il existe un voisinage U de x dans X et un voisinage V de y dans Y tels que $f(U \times V) \subset W$. En particulier, si f est continue, toutes les applications partielles $(x \mapsto f(x, y))$ et $(y \mapsto f(x, y))$ sont continues.

Exemple A.3.4 (Topologie d'un produit infini)*.* Soit I un ensemble non vide et $(X_i)_{i \in I}$ une famille d'espaces topologiques. Les logiciens vous confirmeront que l'ensemble $\bigcup_{i \in I} X_i$ est bien défini. Le produit cartésien de la famille $(X_i)_{i \in I}$ est l'ensemble

$$\prod_{i \in I} X_i = \left\{ f : I \to \bigcup_{i \in I} X_i \quad | \quad \forall i \in I, f(i) \in X_i \right\}$$

Notons $X = \prod_{i \in I} X_i$. Sur X, les projections $\pi_i : X \to X_i$ sont bien définies par $\pi_i(f) = f(i)$ et sont surjectives.

Par définition, la topologie produit sur X est la moins fine rendant toutes les projections π_i continues. Elle est donc engendrée par $\bigcup_{i \in I} \pi_i^{-1}(\mathcal{O}_i)$, où $\mathcal{O}_i$ est la topologie de X_i.

En particulier, un voisinage d'un élément f de X contient toujours un voisinage du type $\pi_{i_1}^{-1}(V_{i_1}) \cap ... \cap \pi_{i_n}^{-1}(V_{i_n})$ où les V_{i_k} sont des voisinages de $f(i_k)$ dans X_{i_k}.

Cette topologie est séparée si les X_i le sont, et c'est la topologie de la convergence simple : une suite $(f_n)_{n\in N}$ converge dans X si et seulement si pour tout i dans I, $(f_n(i))$ converge vers $f(i)$.

Proposition A.3.5. *Soit X un espace topologique. La diagonale de $X \times X$, muni de la topologie produit est fermée si, et seulement si, X est séparé.*

Démonstration. Supposons X séparé. Soit $\Delta = \{(x,x) \in X^2\}$. Si $(x,y) \notin \Delta$, alors $x \neq y$ et comme X est séparé, il existe un voisinage U de x et un voisinage V de y tels que $U \cap V = \emptyset$. Alors $U \times V$ est un voisinage de (x,y) et celui-ci ne contient aucun élément de Δ (sinon U et V ne seraient pas disjoints).

Réciproquement, supposons Δ fermée. Soient x et y deux éléments distincts de X. Comme Δ est fermée, il existe un voisinage W de (x,y) disjoint de Δ, puis, par définition de la topologie produit, il existe un voisinage U de x et V de y tels que $U \times V \subset W$. La partie $U \times V$ étant disjointe de Δ, U et V sont disjoints. □

A.4 Compacité

Définition A.4.1. Soit X un espace topologique. Notons $\mathcal{O}$ l'ensemble de ses ouvert. On appelle recouvrement de X par des ouverts toute partie $R \subset \mathcal{O}$ telle que $X = \bigcup_{u\in R} u$.

On appelle sous-recouvrement d'un recouvrement R une partie S de R telle que $X = \bigcup_{u\in S} u$

On dit que X est *compact* lorsqu'il est séparé et que de tout recouvrement de X par des ouverts, on peut extraire un sous-recouvrement *fini*.

On dit qu'une partie K d'un espace topologique X est compacte lorsqu'elle est compacte pour la topologie induite.

Proposition A.4.1. *Soit X un espace topologique* séparé.

1. *L'espace X est compact si et seulement si, pour toute famille $(F_i)_{i\in I}$ de fermés de X telle que $\bigcap_{i\in I} F_i = \emptyset$, il existe un sous-ensemble* fini $\{i_1, i_2, ..., i_n\}$ *de I tel que $\bigcap_{k=1}^{n} F_{i_k} = \emptyset$.*
2. *Soient K un compact de X et $y \notin K$. Alors il existe deux ouverts disjoints U et V tels que $K \subset U$ et $y \in V$.*
3. *Soient K et L deux compacts disjoints de X. Alors il existe deux ouverts disjoints U et V tels que $K \subset U$ et $L \subset V$.*
4. *Toute partie compacte de X est fermée.*
5. *Si X est un espace topologique compact, alors toute partie fermée de X est encore compacte.*
6. *Toute intersection de compacts de X est encore compacte.*
7. *Soit I un ensemble non vide et $(K_i)_{i\in I}$ une famille de compacts* non vides, totalement ordonnée. *Alors l'intersection $\bigcap_{i\in I} K_i$ est non vide.*

Démonstration. 1. Il s'agit seulement de passer au complémentaire.

2. Soit $x \in K$. Comme $y \notin K$, on a $x \neq y$, donc comme X est séparé, il existe des voisinages ouverts disjoins U_x de x et V_x de y. On a donc $K \subset \bigcup_{x \in K} U_x$. Comme K est compact, on peut en extraire un sous-recouvrement fini $K \subset \bigcup_{k=1}^{n} U_{x_k}$. Il suffit maintenant de poser $U = \bigcup_{k=1}^{n} U_{x_k}$ et $V = \bigcap_{k=1}^{n} V_{x_k}$ qui est ouvert car intersection finie d'ouverts.
3. Soit $y \in L$. Comme $y \notin K$, on peut appliquer le point précédent : il existe deux ouverts disjoints U_y et V_y tels que $K \subset U_y$ et $y \in V_y$. On utilise ensuite la compacité de L, $L \subset \bigcup_{k=1}^{n} V_y$. En posant $V = \bigcup_{k=1}^{n} V_{y_k}$ et $U = \bigcap_{k=1}^{n} U_{y_k}$, on obtient ce premier résultat de séparation.
4. Le point 2. montre que le complémentaire de K est voisinage de chacun de ses points.
5. Soit F un fermé de X. Comme X est supposé séparé, toute partie de X munie de sa topologie induite est encore séparée. En particulier, F est séparé.

   Soit $(F_i)_{i \in I}$ une famille de fermés de F telle que $\bigcap_{i \in I} F_i = \emptyset$. Comme F est fermé, ce sont aussi des fermés de X et par compacité, il existe une partie finie $\{i_1, i_2, ..., i_n\}$ de I telle que $\bigcap_{k=1}^{n} F_{i_k} = \emptyset$, de sorte que F est encore compact, d'après le point 1.
6. Soit $(K_i)_{i \in I}$ une famille de compacts de X. Soit $K = \bigcap_{i \in I} K_i$. D'après le point 4, K est fermé comme intersection de fermés. D'après le point 5, comme K est fermé dans l'un des K_i, c'est encore un compact.
7. Supposons que $\bigcap_{i \in I} K_i = \emptyset$. Soit $i_0 \in I$. Alors, $\bigcap_{i \in I} K_i \cap K_{i_0} = \emptyset$ et comme les $K_i \cap K_{i_0}$ sont fermés dans le compact K_{i_0}, il existe des éléments $i_1, i_2, ..., i_n$ dans I tels que $\bigcap_{k=1}^{n} K_{i_k} \cap K_{i_0} = \emptyset$, mais comme la famille $(K_i)_{i \in I}$ est totalement ordonnée, l'intersection $K_{i_0} \cap K_{i_1} \cap ... \cap K_{i_n}$ est égale au plus petit des K_{i_k}, donc est non vide.

□

Le théorème suivant permet d'affaiblir considérablement la condition de compacité, il nous dit que dans un espace séparé, il suffit de vérifier la condition de recouvrement pour une prébase de la topologie.

Théorème A.4.2 (Théorème des prébases d'Alexander)**.** *Soit $\mathcal{A}$ une famille de parties d'un espace topologique séparé X qui engendre sa topologie. Alors, X est compact si et seulement si de tout recouvrement de X par des éléments de $\mathcal{A}$, on peut extraire un recouvrement fini.*

Démonstration. Soit $\mathcal{O}$ la topologie sur X.

Nous raisonnons par l'absurde en supposant que X n'est pas compact.

Soit $\mathcal{R}$, l'ensemble des recouvrements de X qui n'admettent pas de sous-recouvrement fini. Par hypothèse, $\mathcal{R}$ est non vide. Nous allons montrer qu'il est inductif pour l'inclusion. Soit $\mathcal{L}$ une partie totalement ordonnée de $\mathcal{R}$, et soit $M = \bigcup_{L \in \mathcal{L}} L$.

Il est clair que M est un recouvrement de X. S'il admettait un sous-recouvrement fini $\{m_1, m_2, ..., m_n\}$, alors, comme $\mathcal{L}$ est totalement ordonnée et les m_i en nombre fini, ils appartiendrait tous à l'un des recouvrements L_0 de $\mathcal{L}$, lequel admettrait un sous-recouvrement fini, ce qui est contradictoire. Le recouvrement M est donc un majorant de $\mathcal{L}$ dans $\mathcal{R}$, ce qui fait de $\mathcal{R}$ un ensemble inductif. Par le lemme de Zorn, $\mathcal{R}$ admet un élément maximal S. Soit $S' = \mathcal{A} \cap S$. Nous souhaitons montrer que S' est un recouvrement de X.

Supposons que ce n'est pas le cas : il existe alors $x \in X$ qui n'est pas recouvert par S', donc qui n'est inclus dans aucun élément de S'. Mais comme S recouvre x, il existe un ouvert $W \in S$, tel que W soit un voisinage de x. Comme $\mathcal{A}$ engendre $\mathcal{O}$, cela implique qu'il existe des ouverts $V_1, V_2, ..., V_q$ dans $\mathcal{A}$, tels que

$$x \in V_1 \cap V_2 \cap ... \cap V_q \subset W$$

En particulier, aucun des V_i n'est dans S', donc dans S. Mais alors, chaque $S \cup \{V_i\}$ n'est plus dans $\mathcal{R}$, par maximalité de S, donc admet un sous-recouvrement fini. Il existe donc, pour chaque i, un ouvert Y_i qui est réunion finie d'éléments de S et tel que $X = Y_i \cup V_i$. On obtient alors :

$$X = Y_1 \cup Y_2 \cup ... \cup Y_q \cup \left(\bigcap_{i=1}^{q} V_i \right) \subset Y_1 \cup Y_2 \cup ... \cup Y_q \cup W$$

D'où une première contradiction, car on obtient un sous-recouvrement fini de S.

On en déduit que S' est un recouvrement de X, mais, par hypothèse, il a alors un sous-recouvrement fini qui est lui-même un sous-recouvrement fini de S, ce qui est la contradiction finale recherchée.

□

Nous pouvons maintenant énoncer et démontrer le théorème de Tychonov, qui joue un rôle crucial dans la démonstration du théorème de Banach-Alaoglu (lequel est fondamental pour nous).

Théorème A.4.3 (Théorème de Tychonov)**.** *Tout produit d'espaces compacts est encore compact pour la topologie produit.*

Démonstration. Reprenons les notations de l'exemple A.3.4. Nous supposons cette fois que chaque X_i est compact.

Soit $\mathcal{R}$ un recouvrement de X par des ouverts "cylindriques", c'est-à-dire $\mathcal{R} \subset \mathcal{A}$. Si l'on parvient à extraire de $\mathcal{R}$ un sous-recouvrement fini, le théorème d'Alexander nous permettra de conclure.

Partitionnons $\mathcal{R}$ sous la forme $\mathcal{R} = \bigcup_{i \in I} \mathcal{R}_i$, où $\mathcal{R}_i = \mathcal{R} \cap \pi_i^{-1}(\mathcal{O}_i)$. Les éléments de $\mathcal{R}_i$ sont des cylindres dont la base est un ouvert de X_i. Notons $\mathcal{S}_i = \pi_i(\mathcal{R}_i)$ la famille d'ouverts correspondante dans X_i.

Supposons qu'aucun $\mathcal{R}_i$ ne recouvre X. Alors, pour tout i, il existe un élément $f_i \in X$ tel que $f_i(i)$ n'est pas recouvert par $\mathcal{S}_i$. Soit alors g, définie par $g(i) = f_i(i)$, pour tout $i \in I$. On a bien $g \in X$, car chaque f_i est dans X, donc pour tout $i \in I$, $g(i) = f_i(i) \in X_i$. Mais g n'est recouvert par aucun des $\mathcal{R}_i$, donc n'est pas recouvert par leur réunion $\mathcal{R}$, ce qui est contradictoire.

Il existe donc $i_0 \in I$, tel que $X = \bigcup_{U \in \mathcal{R}_{i_0}} U$.

Comme π_{i_0} est surjective, on obtient $X_{i_0} = \bigcup_{V \in \mathcal{S}_{i_0}} V$. Mais alors, comme X_{i_0} est compact, on peut extraire un sous-recouvrement fini $\mathcal{F}_{i_0}$ de $\mathcal{S}_{i_0}$, de sorte que

$$X_{i_0} = \bigcup_{V \in \mathcal{F}_{i_0}} V$$

En notant $\mathcal{F} = \pi_{i_0}^{-1}(\mathcal{F}_{i_0})$ et en appliquant $\pi_{i_0}^{-1}$ à cette dernière égalité, on obtient

$$X = \bigcup_{U \in \mathcal{F}} U$$

Le recouvrement $\mathcal{F}$ est alors fini, est un sous-recouvrement de $\mathcal{R}_{i_0}$, donc de $\mathcal{R}$ et par le théorème d'Alexander, X est compact.

□

Pour finir cette partie sur la compacité, on rappelle une version du Théorème d'Ascoli. Sa démonstration se trouve, par exemple, dans [Mau01]. Rappelons qu'une partie d'un espace est dite relativement compacte lorsque son adhérence est compacte.

Théorème A.4.4. *Soit X un espace métrique compact et $\mathcal{C}(X)$ l'ensemble des fonctions continues à valeurs réelles sur X, muni de $\| \ \|_\infty$. Alors, une partie A de $\mathcal{C}(X)$ est relativement compacte si et seulement si, elle est bornée et équicontinue, i.e.*

$$\exists M > 0, \ \ \forall f \in A, \ \ \|f\|_\infty < M \quad \textit{et}$$

$$\forall \varepsilon > 0, \ \ \exists \eta, \ \ \forall x, y \in X, \ \ d(x, y) < \eta \Rightarrow |f(x) - f(y)| < \varepsilon$$

B Quelques résultats d'analyse fonctionnelle

B.1 Espaces vectoriels topologiques

Ce qui suit est très largement inspiré de l'ouvrage de référence de Walter Rudin [RA95]. Le lecteur qui voudrait davantage de détails ou qui préfèrerait, légitimement, l'original à la copie est invité à s'y référer.

Jusqu'à la fin de cet appendice, $\mathbb{K}$ désignera le corps des réels ou des complexes.

Définition B.1.1. Soit X un espace vectoriel sur $\mathbb{K}$. On dit que X est un espace vectoriel topologique lorsque :

1. Le singleton $\{0\}$ est fermé dans X.
2. L'addition des vecteurs est continue de $X \times X$ dans X.
3. Le produit des scalaires et des vecteurs est continue de $\mathbb{K} \times X$ dans X.

Proposition B.1.1. *Soit X un espace vectoriel topologique. Soient $a \in X$ et λ un scalaire non nul.*

1. *La translation $t_a : x \longmapsto x + a$ est un homéomorphisme de X dans X.*
2. *L'homothétie $h_\lambda : x \longmapsto \lambda x$ est un homéomorphisme de X dans X.*
3. *Si U est un ouvert de X, il en est de même de $a + U$ et λU.*
4. *Si F est un fermé de X, $a + F$ et λF sont aussi fermés.*
5. *Tout voisinage W de a peut s'écrire sous la forme $W = a + V$, où V est un voisinage de 0.*
6. *Tout système fondamental de voisinages de 0 définit, par translation un système de voisinage de a.*

Définition B.1.2. On appelle base de voisinages de X, tout système fondamental de voisinages de 0.

Nous allons maintenant établir un théorème très utile de séparation d'un fermé et d'un compact, pour lequel nous aurons besoin du lemme suivant :

Lemme B.1.2. *Soient X un espace vectoriel topologique et W un voisinage de 0. Alors, il existe un voisinage U de 0, symétrique (i.e. tel que $U = -U$) et tel que $U + U \subset W$.*

Démonstration. En effet, comme la somme est continue en 0, il existe deux voisinages ouverts de 0, V_1 et V_2 tels que $V_1 + V_2 \subset W$. Comme le produit par le scalaire -1 est un homéomorphisme, les parties $-V_1$ et $-V_2$ sont encore des voisinages ouverts de 0. En posant $U = V_1 \cap (-V_1) \cap V_2 \cap (-V_2)$, on obtient bien le voisinage cherché. □

Théorème B.1.3. *Soient X un espace vectoriel topologique et K et F deux parties disjointes de X, avec F fermée et K compacte. Alors, il existe un voisinage ouvert V de 0 tel que*

$$(K + V) \cap (F + V) = \emptyset$$

Démonstration. Soit x dans K. Comme F est fermé, il existe un voisinage W_x de 0 tel que $x + W_x$ soit disjoint de F. On applique le lemme précédent deux fois, pour obtenir un voisinage V_x de 0, symétrique, tel que

$$V_x + V_x + V_x + V_x \subset W_x$$

En particulier, $x + V_x + V_x + V_x$ est disjoint de F, donc $(x + V_x + V_x) \cap (F - V_x) = \emptyset$, puis comme V_x est symétrique, on a obtenu une famille de voisinages $(V_x)_{x \in K}$ telle que

$$(x + V_x) \cap (F + V_x) = \emptyset$$

On utilise maintenant la compacité de K :

$$K \subset \bigcup_{x \in K} (x + V_x) \subset \bigcup_{i=1}^{n} (x_i + V_{x_i})$$

Nous pouvons maintenant poser $V = \bigcap_{i=1}^{n} V_{x_i}$ et remarquer que :

$$K + V \subset \bigcup_{i=1}^{n} (x_i + V_{x_i} + V) \subset \bigcup_{i=1}^{n} (x_i + V_{x_i} + V_{x_i})$$

Chaque terme $(x_i + V_{x_i} + V_{x_i})$ de cette dernière réunion est disjoint de $F + V_{x_i} \supset F + V$ donc $K + V$ est disjoint de $F + V$. □

Corollaire B.1.4. *Les espaces vectoriels topologiques (tels que nous les avons définis) sont séparés.*

Démonstration. Soient deux éléments distincts x et y de X, espace vectoriel topologique. On applique la propriété précédente à $\{x\}$ compacte et $\{y\}$ fermée (c'est l'image du fermé $\{0\}$ par l'homéomorphisme t_y). □

Définition B.1.3. Soient X un espace vectoriel topologique et $A \subset X$. On dit que A est :

— équilibrée lorsque pour tout scalaire α, $|\alpha| \leq 1 \Rightarrow \alpha A \subset A$;

— bornée lorsque, pout tout voisinage V de 0, il existe un réel $t > 0$ tel que $A \subset tV$;

— absorbante lorsque, pour tout $x \in X$, il existe $t > 0$ tel que $x \in tA$.

Proposition B.1.5. *Soit X un espace vectoriel topologique. Alors :*

1. *tout voisinage de* 0 *est absorbant ;*
2. *X admet une base de voisinages équilibrés.*

Démonstration.

Soit V un voisinage de 0 dans X et $x \in X$. Comme $0x = 0$, par continuité du produit par un scalaire, il existe un réel $\eta > 0$ tel que $|\alpha| \leq \eta$ implique $\alpha x \in V$, d'où $x \in \frac{1}{\eta} V$.

Soit V un voisinage de 0 dans X. Par continuité du produit des scalaires et des vecteurs, on peut trouver un réel $\delta > 0$ et un voisinage U de 0 dans X tels que

$$|\alpha| < \delta \quad \text{et} \quad x \in U \Rightarrow \alpha x \in U$$

Posons $W = \bigcup_{|\alpha| \leq 1} \alpha U$. Il est alors clair que W est un voisinage de 0 inclus dans V et que W est équilibré. □

Proposition B.1.6. *Soient X un espace vectoriel topologique de X, A et B deux parties de X et α et β deux scalaires.*

1. $\overline{A} + \overline{B} \subset \overline{A+B}$ *et* $A^o + B^o \subset (A+B)^o$.
2. $\overline{\alpha A} = \alpha \overline{A}$ *et si* $\alpha \neq 0$, $(\alpha A)^o = \alpha A^o$.
3. *Si A est équilibré, il en est de même pour $\overline{A}$.*
4. *Si A est équilibré et A^o contient 0, alors A^o est équilibré.*
5. *Si A est un sous-espace vectoriel de X, alors $\overline{A}$ est un sous-espace vectoriel de X.*
6. *Si A est un sous-espace vectoriel strict de X, alors $A^o = \emptyset$.*

Démonstration. 1. Soient $x \in \overline{A}$, $y \in \overline{B}$. On veut prouver que tout voisinage W de $x + y$ coupe $A + B$.

Par continuité de la somme, il existe un voisinage U de x et un voisinage V de y tels que $U + V \subset W$. Comme $x \in \overline{A}$, U coupe A donc il existe $a \in A \cap U$. De même, il existe $b \in B \cap V$. On en déduit que $a + b \in U + V \subset W$ et $a + b \in A + B$ donc que $W \cap (A + B)$ est non vide.

2. La propriété est évidente pour $\alpha = 0$. Supposons $\alpha \neq 0$. Alors, comme l'homothétie $(x \mapsto \alpha x)$ est ouverte et fermée, $\alpha\overline{A}$ est fermée et αA^o est ouvert. On en déduit que $\overline{\alpha A} \subset \alpha\overline{A}$ et $\alpha A^o \subset (\alpha A)^o$. Il suffit pour conclure de reprendre le même raisonnement en remplaçant α par α^{-1} et A par αA.
3. Soit α tel que $|\alpha| \leq 1$. Alors $\alpha\overline{A} = \overline{\alpha A} \subset \overline{A}$.
4. Si α est non nul et tel que $|\alpha| \leq 1$, la démonstration est la même que la précédente. Si $\alpha = 0$, on a $0 \in A^o$ par hypothèse.
5. On a, pour tous scalaires α et β,

$$\alpha\overline{A} + \beta\overline{A} = \overline{\alpha A} + \overline{\beta A} \subset \overline{\alpha A + \beta A} \subset \overline{A}$$

6. Si A un sous espace vectoriel strict, il ne peut contenir aucun voisinage de 0 car ceux-ci sont tous absorbants.

□

Proposition B.1.7. *1. Soit $\varphi : X \to Y$ une application linéaire entre deux espaces vectoriels topologiques. Alors φ est continue sur X si et seulement si elle est continue en* 0.

2. Soit φ une forme linéaire sur un espace vectoriel topologique X. Alors φ est continue si et seulement si elle est borné sur un certain voisinage de 0.

3. *Soit F un espace vectoriel normé de dimension finie et X un espace topologique. Toute application linéaire $\varphi : F \to X$ est continue.*

Démonstration. 1. Soit $x \in X$ et $y \in Y$. Considérons un voisinage $y+U$ de y dans Y, où U est un voisinage de 0. Alors, comme φ est supposée continue en 0, il existe un voisinage V de 0 dans X tel que $\varphi(V) \subset U$. Puis on voit, par linéarité, que $\varphi(x+V) \subset y+U$, donc que φ est continue en x.

2. Si φ est continue en 0, alors pour tout réel $r > 0$, il existe un voisinage V de 0 tel que $|\varphi(V)| < r$.

   Réciproquement, si φ est bornée par un réel positif $M > 0$ sur un voisinage V, alors pour tout $r > 0$, elle est bornée par r sur le voisinage $\frac{r}{M}V$, donc continue en 0.

3. Notons $(e_1, e_2, ..., e_n)$ une base de f et pour chaque i, e_i^* la forme linéaire qui à un vecteur x associe sa i-ième coordonnée x_i dans cette base.

   Les e_i^* sont continues sur F (par exemple pour la norme infinie associée à cette base, mais sur F, toutes les normes sont équivalentes).

   On a alors, pour tout $x \in F$,

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} e_i^*(x) e_i \quad \text{donc} \quad \varphi(x) = \sum_{i=1}^{n} e_i^*(x) \varphi(e_i)$$

   Et on peut conclure à la continuité de φ, par continuité des e_i^*, du produit des scalaires et des vecteurs dans X et de la somme dans X.

□

Théorème B.1.8. *Soient X un espace vectoriel topologique et F un sous-espace de dimension n de X, alors*

1. *tout isomorphisme de $\mathbb{K}^n$ dans F est un homéomorphisme, et*
2. *F est fermé.*

Ce théorème affirme donc que tout sous-espace vectoriel de dimension fini de X est *normable* et qu'il n'existe qu'une seule topologie faisant d'un espace vectoriel de dimension finie un espace vectoriel topologique.

La démonstration suivante, encore tirée de [RA95], utilise un argument astucieux de connexité.

Démonstration. 1. Notons $\varphi : \mathbb{K}^n \xrightarrow{\sim} F \subset X$. Nous savons déjà que φ est continue et il reste à prouver que φ^{-1} l'est également.

   Choisissons une norme arbitraire sur $\mathbb{K}^n$ et notons S sa sphère unité et B sa boule unité fermée.

   Comme S est compacte et φ est continue, $K = \varphi(S)$ est compact dans X. L'origine $0 \notin K$ car φ est injective. D'après le théorème B.1.3 et la proposition B.1.5, il existe donc un voisinage équilibré V de 0 tel que $V \cap K = \emptyset$. Posons

$$E = \varphi^{-1}(V) = \varphi^{-1}(V \cap F)$$

L'ensemble E est bien un voisinage de 0 et comme $V \cap F$ est équilibré et φ^{-1} est linéaire, il est encore équilibré. Dans $\mathbb{K}^n$, il est donc connexe, car étoilé : pour tout $x \in E$, le segment paramétré $\gamma : [0;1] \to E$ défini par $\gamma(t) = tx$ relie continûment x à 0.

Le voisinage E est connexe, disjoint de S et contient 0, donc il est inclus dans la boule unité ouverte B de $\mathbb{K}^n$.

On en déduit que pour tout $x \in V$, $\|\varphi^{-1}(x)\| \leq 1$. Par conséquent, chaque composante de $\varphi^{-1} : F \to \mathbb{K}^n$ est bornée sur V et puisque ce sont des formes linéaires sur F, elles sont continues. Enfin, chacune de ses composantes étant continue, $\varphi^{-1} : F \to \mathbb{K}^n$ est elle-même continue.

2. Démontrons que F est fermé dans X. Comme V est absorbant et équilibré, on peut écrire X comme la réunion :

$$X = \bigcup_{n \in \mathbb{N}^*} nV \text{ donc } \overline{F} = \bigcup_{n \in \mathbb{N}^*} nV \cap \overline{F} \subset \bigcup_{n \in \mathbb{N}^*} n\overline{V \cap F}$$

Or, $\overline{V \cap F} \subset \overline{\varphi(B)} \subset \overline{\varphi\left(\overline{B}\right)} = \varphi\left(\overline{B}\right)$, car $\overline{B}$ est compact donc son image l'est aussi.

Finalement, $\overline{V \cap F} \subset F$, donc $\overline{F} \subset F$.

□

B.2 Convexité, convexité locale, jauges et semi-normes

Définition B.2.1. Soit X un espace vectoriel topologique. Une partie C de X est dite *convexe* lorsque pour tout $t \in [0,1]$, $tC + (1-t)C \subset C$.

L'espace X est dit *localement convexe* lorsqu'il a des bases de voisinages convexes.

La plupart des espaces de l'analyse fonctionnelle sont localement convexes.

Proposition B.2.1. *Soit X un espace vectoriel topologique et $C \subset X$.*

1. *C est convexe si et seulement si, pour tous réels positifs α_1, α_2, ..., α_n tels que $\alpha_1 + ... + \alpha_n = 1$, on a $\sum_{i=1}^n \alpha_i C \subset C$.*
2. *Si C est une partie contenant 0, alors pour tous réels positifs α_1, α_2, ..., α_n tels que $\alpha_1 + ... + \alpha_n \leq 1$, on a $\sum_{i=1}^n \alpha_i C \subset C$.*
3. *Si C est une partie convexe contenant 0, alors pour tout $t \in [0,1]$, $tC \subset C$.*
4. *Si C est convexe, alors $\overline{C}$ et C^o sont convexes.*

Démonstration. 1. Soient $\alpha_1, ..., \alpha_n, \alpha_{n+1}$ des réels positifs de somme 1, et $x_1, ..., x_n, x_{n+1}$ des éléments de C. Si $\alpha_1 + ... + \alpha_n = 0$, alors $\alpha_{n+1} = 1$ et $\alpha_1 x_1 + ... + \alpha_{n+1} x_{n+1} = x_{n+1} \in C$. Sinon,

$$\alpha_1 x_1 + ... + \alpha_n x_n + \alpha_{n+1} x_{n+1} = \left(\sum_{i=1}^n \alpha_i\right) \left(\frac{\alpha_1}{\sum_{i=1}^n \alpha_i} x_1 + ... + \frac{\alpha_1}{\sum_{i=1}^n \alpha_i} x_n\right) + \alpha_{n+1} x_{n+1}$$

et on peut conclure par récurrence.

2. Il suffit d'appliquer la propriété précédente avec $x_{n+1} = 0$ et $\alpha_{n+1} = 1 - (\sum_{i=1} \alpha_i)$.

3. Il suffit d'appliquer la propriété précédente au cas $n = 1$.
4. Pour tout $t \in [0, 1]$, on a

$$t\overline{C} + (1-t)\overline{C} = \overline{tC} + \overline{(1-t)C} \subset \overline{tC + (1-t)C} \subset \overline{C}$$

La démonstration est identique pour l'intérieur C^o.

□

Proposition B.2.2. *Soit X un espace topologique localement convexe. Alors X admet une base de voisinages ouverts, convexes et équilibrés.*

Démonstration. Soit U un voisinage convexe de 0 dans X. Soit $A = \bigcap_{|\alpha|=1} \alpha U \subset U$.

D'après la proposition B.1.5, il existe un voisinage équilibré $W \subset U$. Pour $|\alpha| = 1$, on a $\alpha^{-1}W \subset W$, donc $W \subset \alpha U$. On a donc $W \subset A$, de sorte que l'intérieur de A est non vide et contient 0. De plus, A est convexe, comme intersection de convexes.

Montrons que A est équilibré : soit β un scalaire de module inférieur ou égal à 1. Notons $\beta = r\beta'$, avec $r \in [0, 1]$ et $|\beta'| = 1$. On a alors

$$\beta A = r\beta' A = \bigcup_{|\alpha|=1} r\beta'\alpha U = \bigcup_{|\alpha|=1} r\alpha U \subset \bigcup_{|\alpha|=1} \alpha U = A$$

car αU est convexe et contient 0.

L'intérieur A^o de A est donc un voisinage ouvert, équilibré et convexe de 0 inclus dans U. □

Nous allons maintenant voir une façon très classique de définir une topologie d'espace vectoriel localement convexe à l'aide d'une famille séparante de semi-normes.

De plus, nous verrons qu'on obtient ainsi tous les espaces vectoriels localement convexes. Les jauges des convexes équilibrés, cas particulier de semi-normes, nous serviront dans les démonstrations des théorèmes de Hahn-Banach, de Schauder-Tychonoff et de Krein-Milman.

Définition B.2.2. Soit X un espace vectoriel et $p : X \to \mathbb{R}^+$. On dit que X est une semi-norme lorsque :

1. pour tous vecteurs x et y de X, $p(x + y) \leq p(x) + p(y)$, et
2. pour tout scalaire α et tout vecteur x, $p(\alpha x) = |\alpha|\, p(x)$.

On appelle p-semi-boule (ouverte) de centre x et de rayon $r > 0$, l'ensemble

$$\mathrm{sb}_p(x, r) = \{y \in X \quad | \quad p(y - x) > 0\}$$

On dit qu'une famille $\mathcal{P}$ de semi-normes est séparante lorsque pour tout $x \neq 0$, il existe une semi-norme $p \in \mathcal{P}$ telle que $p(x) \neq 0$.

Proposition B.2.3. *Soit $\mathcal{P}$ une famille séparante de semi-normes sur un espace vectoriel X. La topologie engendrée par les p-semi-boules ouvertes, pour $p \in \mathcal{P}$, définit sur X une topologie d'espace vectoriel topologique.*

Démonstration. 1. La somme est continue. En effet, soient a et b dans X et W un voisinage de W. Alors, il existe $p_1, p_2, ..., p_n \in \mathcal{P}$ et $\varepsilon > 0$ tels que

$$\mathrm{sb}_{p_1}(a+b, \varepsilon) \cap ... \cap \mathrm{sb}_{p_n}(a+b, \varepsilon) \subset W$$

Posons

$$U = \mathrm{sb}_{p_1}\left(a, \frac{\varepsilon}{2}\right) \cap ... \cap \mathrm{sb}_{p_n}\left(a, \frac{\varepsilon}{2}\right) \quad \text{et} \quad V = \mathrm{sb}_{p_1}\left(b, \frac{\varepsilon}{2}\right) \cap ... \cap \mathrm{sb}_{p_n}\left(b, \frac{\varepsilon}{2}\right)$$

On a alors, pour tous x dans U, y dans V et i entre 1 et n,

$$p_i\left((x+y) - (a+b)\right) = p_i\left((x-a) + (y-b)\right) \leq p_i(x-a) + p_i(y-b) < \varepsilon$$

de sorte que $U + V \subset W$.

2. Soient $\lambda \in \mathbb{K}$ et $a \in X$. Pour tous $p \in \mathcal{P}$, $x \in X$ et $\mu \in \mathbb{K}$, on a

$$\begin{aligned} p(\mu x - \lambda a) &= p(\mu x - \mu a + \mu a - \lambda a) \leq |\mu|\, p(x-a) + |\mu - \lambda|\, p(a) \\ &\leq (|\lambda| + |\mu - \lambda|)\, p(x-a) + |\mu - \lambda|\, p(a) \end{aligned}$$

Soit $\varepsilon > 0$. Si λ et $p(a)$ sont non nuls, on pose $\eta = \min\left(\frac{\varepsilon}{3p(a)}, \frac{\sqrt{\varepsilon}}{3}\right)$ et $\delta = \min\left(\frac{\varepsilon}{3|\lambda|}, \sqrt{\varepsilon}\right)$. Alors,

$$|\mu - \lambda| < \eta \quad \text{et} \quad p(x-a) < \delta \Rightarrow p(\mu x - \lambda a) < \varepsilon$$

On en déduit que le produit est continu.

3. L'ensemble $\{0\}$ est bien fermé : si $x \neq 0$, par hypothèse, il existe $p \in \mathcal{P}$ tel que $p(x) > 0$ donc $\mathrm{sb}_p\left(x, \frac{p(x)}{2}\right)$ est un voisinage de x ne contenant pas 0.

□

Proposition B.2.4. *Toute semi-norme p sur un espace vectoriel X vérifie les propriétés suivantes :*

1. pour tous vecteurs x et y, on a $|p(x) - p(y)| \leq p(x-y)$,

2. les p-semi-boules sont convexes.

Démonstration.

L'inégalité triangulaire donne $p(x)+p(y-x) \geq p(y)$, d'où $p(y)-p(x) \leq p(y-x) = p(x-y)$. On procède de même pour majorer $p(x) - p(y)$.

Si y_1 et y_2 sont dans la semi-boule $\mathrm{sb}_p(x, r)$, et $t \in [0,1]$ alors

$$\begin{aligned} p(x - (ty_1 + (1-t)y_2)) &= p(t(x-y_1) + (1-t)(x-y_2)) \\ &\leq tp(x-y_1) + (1-t)p(x-y_2) < t + (1-t)r = r \end{aligned}$$

□

Proposition B.2.5. *Soit X un espace vectoriel muni d'une famille séparante de semi-normes $\mathcal{P}$ et de la topologie associée. Alors X est un espace vectoriel topologique localement convexe dans lequel les semi-normes sont continues.*

Proposition B.2.6. *Soit X un espace vectoriel topologique et V un voisinage convexe de 0 dans X. Notons, pour $x \in X$*

$$j_V(x) = \inf\{t \geq 0 \quad | \quad x \in tV\}$$

Alors l'application j_V, appelée jauge *du convexe V est bien définie et vérifie :*

$$\forall x, y \in X, \;\; \forall t \geq 0, \;\; j_V(x+y) \leq j_V(x) + j_V(y) \quad \textit{et} \quad j_V(tx) = tj_V(x)$$

Si, de plus, V est équilibré, alors j_V est une semi-norme sur X.

Démonstration. Soient x, y dans V, $\varepsilon > 0$. On considère les réels $t = j_V(x) + \varepsilon$ et $s = j_v(y) + \varepsilon$. Alors $\frac{x}{t}$ et $\frac{y}{s}$ sont dans le convexe V, et il en est de même de

$$\frac{x+y}{t+s} = \frac{t}{t+s}\frac{x}{t} + \frac{s}{s+t}ys$$

Ainsi, $j_V(x+y) \leq s+t \leq j_V(x)+j_V(y)+2\varepsilon$ pour tout $\varepsilon > 0$, donc $j_V(x+y) \leq j_V(x)+j_V(y)$.

Les autres propriétés sont évidentes. □

Proposition B.2.7. *Tout espace vectoriel topologique localement convexe est "semi-normable" par la famille $\mathcal{P} = \{j_V \quad | \quad V \in \mathcal{B}\}$, où $\mathcal{B}$ est une base de voisinages ouverts convexes équilibrés.*

En particulier, dans un espace vectoriel localement convexe, les jauges des voisinages ouverts convexes équilibrés de 0 sont continues.

B.3 De Brouwer à Schauder-Tychonoff

Dans des espaces vectoriels de dimension finie, nous disposons du prodigieux théorème de point fixe de Brouwer.

Théorème B.3.1 (Théorème du point fixe de Brouwer)**.** *Soient X un espace vectoriel normé de dimension finie, K une partie convexe, compacte, non vide de X et une application continue $f : K \to K$. Alors, f admet un point fixe dans K.*

Ses démonstrations les plus classiques utilisent la géométrie différentielle (via le théorème de Stokes dans [Laf12]) ou la topologie algébrique. Elles nous emmèneraient trop loin de notre sujet principal et nous admettons ce théorème. Nous allons nous servir des parties précédentes pour le généraliser à tous les espaces vectoriels topologiques localement convexes, tout en gardant à l'esprit que le théorème de Brouwer en reste l'ingrédient principal.

Théorème B.3.2 (Théorème du point fixe de Schauder-Tychonoff)**.** *Soient X un espace vectoriel topologique localement convexe, K une partie convexe, compacte, non vide de X et une application continue $f : K \to K$. Alors, f admet un point fixe dans K.*

Démonstration. Nous raisonnons par l'absurde. Supposons que f n'admette pas de point fixe. Alors, son graphe $G = \{(x, f(x)) : x \in K\}$ ne rencontre jamais la diagonale $\Delta = \{(x,x) : x \in X\}$. Comme chacune de ses composantes est continue, l'application $x \mapsto$

$(x, f(x))$ est continue du compact K dans $X \times X$, donc son image G est compacte. De plus, comme X est séparé, la diagonale Δ est fermée.

Nous pouvons appliquer le théorème B.1.3 dans l'espace vectoriel topologique $X \times X$: il existe un voisinage W de $(0,0)$ dans $X \times X$ tel que $K \cap (\Delta + W) = \emptyset$.

Quitte à rétrécir W, on peut le prendre sous la forme $V_1 \times V_2$, où V_1 et V_2 sont des voisinages de 0 dans X. On peut même choisir pour V_1 et V_2 des voisinages ouverts, convexes et équilibrés de 0, puis, posant $V = V_1 \cap V_2$, on obtient $K \cap (\Delta + V \times V) = \emptyset$, où V est un voisinage ouvert convexe et équilibré de 0. En particulier, pour tout $x \in K$, $f(x) \notin x + V$, sinon on aurait $(x, f(x)) \in (x, x) + \{0\} \times V \subset (x, x) + V \times V$.

On utilise encore la compacité de K pour le recouvrir par l'ensemble fini d'ouverts :

$$K \subset (x_1 + V) \cup (x_2 + V) \cup ... \cup (x_n + V) \quad x_1, x_2, ..., x_n \in K$$

Considérons maintenant la jauge j de V. D'après la proposition B.2.7, j est continue sur X. On introduit, pour $i = 1, 2, ..., n$, les applications continues $\alpha_i : x \longmapsto \max\{0, 1 - j(x - x_i)\}$. On a alors, pour tout $x \in X$, $0 \leq \alpha_i(x) \leq 1$ et $\alpha_i(x) > 0 \Leftrightarrow x \in x_i + V$. Ainsi, la somme $\alpha_1(x) + \alpha_2(x) + ... + \alpha_n(x)$ n'est jamais nulle si x est dans K. On peut donc définir sur K la fonction continue :

$$g : x \longmapsto \sum_{i=1}^{n} \frac{\alpha_i(x) x_i}{\alpha_1(x) + \alpha_2(x) + ... + \alpha_n(x)}$$

Introduisons maintenant l'enveloppe convexe H de $\{x_1; x_2; ...; x_n\}$. Comme K est convexe, $H \subset K$, et g est à valeurs dans H qui est compacte, convexe et non vide. Remarquons également que, pour tout $x \in K$,

$$x - g(x) = \sum_{i=1}^{n} \frac{\alpha_i(x)\,(x - x_i)}{\alpha_1(x) + \alpha_2(x) + ... + \alpha_n(x)}$$

Comme $\alpha_i(x) = 0$ si $x \notin x_i + V$, les éléments non nuls de cette somme forment une combinaison convexe d'éléments de V, donc est dans V.

Ainsi, on a décomposé tout élément x de K, sous la forme

$$x = g(x) + (x - g(x)) \quad (1)$$

où $g(x) \in H$ et $x - g(x) \in V$.

L'enveloppe H est de plus incluse dans le sous-espace vectoriel E de X engendré par x_1, x_2, ... x_n. La topologie induite sur E par X est la topologie d'un espace vectoriel normé de dimension finie d'après le théorème B.1.7.

On peut donc appliquer le théorème de Brouwer à la fonction continue $g \circ f_{|H}$.
Il existe un élément $x^* \in H \subset K$ tel que $g(f(x^*)) = x^*$. Mais alors on peut appliquer l'égalité (1) à $f(x^*)$ pour trouver $f(x^*) = x^* + (f(x^*) - g(f(x^*)))$ où $(f(x^*) - g(f(x^*)))$ est dans V, ce qui est contradictoire. □

B.4 Formes linéaires, théorèmes de Hahn-Banach

Théorème B.4.1 (Théorème de prolongement de Hahn-Banach)**.** *Soient X un espace vectoriel* réel*, M un sous-espace vectoriel de X, $\varphi : M \to \mathbb{R}$, une forme linéaire et $p : X \to \mathbb{R}^+$ telles que*

- $\forall, x, y \in X,\ \ \forall t \geq 0,\ \ p(x+y) \leq p(x) + p(y) \quad et \quad p(tx) = tp(x)$
- $\forall x \in M,\ \ \varphi(x) \leq p(x)$

Alors, il existe une forme linéaire $\tilde{\varphi} : X \to \mathbb{R}$ telle que pour tout $x \in X$, $\tilde{\varphi}(x) \leq p(x)$.

Démonstration. Si $M = X$, il n'y a rien à démontrer. Sinon, il existe $x_1 \in X \setminus M$ et nous allons commencer par étendre φ à $M \oplus \mathbb{R}x_1$. Remarquons que :

$$\forall x, y \in M,\ \ \varphi(x) + \varphi(y) \leq \varphi(x+y) \leq p(x - x_1) + p(x_1 + y)$$

On en déduit que

$$\forall x, y \in M,\ \ \varphi(x) - p(x - x_1) \leq p(y + x_1) - \varphi(y)\ \ (*)$$

En particulier, le membre de gauche est majoré (par exemple par $p(x_1)$), notons $\alpha = \sup\{\varphi(x) - p(x - x_1)\ \ |\ \ x \in M\}$. On a alors, par définition de α,

$$\forall x \in M,\ \ \varphi(x) - \alpha \leq p(x - x_1)\ \ (1)$$

Puis, en utilisant l'inégalité $(*)$,

$$\forall y \in M,\ \ \varphi(y) + \alpha \leq p(y + x_1)\ \ (2)$$

Soit alors φ_1 égale à φ sur M et telle que $\varphi_1(x_1) = \alpha$.

Si $t > 0$, on a $\varphi_1\left(\frac{1}{t}y + x_1\right) \leq p\left(\frac{1}{t}y + x_1\right)$ d'après l'inégalité (1) et en multipliant par t, on obtient bien
$\varphi_1(y + tx_1) \leq p(y + tx_1)$, qui est l'inégalité voulue.

Si $t < 0$, on a, en posant $t' = -t$ et en utilisant cette fois l'inégalité (2), $\varphi_1(x - t'\alpha) \leq p(x - t'x_1)$.

Finalement, nous avons bien réussi à étendre φ à $M \oplus \mathbb{R}x_1$, nous allons pouvoir raisonner par récurrence transfinie.

Considérons l'ensemble $\mathcal{E}$ des couples (N, ψ) tels que N soit un sous-espace vectoriel de X contenant M, $\psi : N \to \mathbb{R}$ soit linéaire, prolonge φ et $\psi \leq p$ sur N. Munissons $\mathcal{E}$ de l'ordre partiel $\preceq$ défini par

$$(N, \psi) \preceq (N', \psi') \Leftrightarrow N \subset N' \quad \text{et} \quad \psi'_{|N} = \psi$$

Soit $\mathcal{F} = \{(N_i, \psi_i)\ \ |\ \ i \in I\}$ une partie totalement ordonnée de $\mathcal{E}$. Soient

$$N = \bigcup_{i \in I} N_i \quad \text{et} \quad \psi(x) = \psi_i(x)\ \text{ pour tout }\ x \in N_i$$

Alors, N est un sous-espace vectoriel de X car les N_i sont totalement ordonnés pour l'inclusion et pour la même raison, ψ est définie sans ambigüité.

Ainsi, $\mathcal{E}$ est inductif et par le lemme de Zorn, il admet un élément maximal $\left(\tilde{M}, \tilde{\varphi}\right)$. On a nécessairement $\tilde{M} = X$, sinon la première partie de la démonstration permettrait de prolonger encore $\tilde{\varphi}$, ce qui contredirait sa maximalité.

$\square$

Corollaire B.4.2. *Soit X un espace vectoriel, M un sous-espace vectoriel de X, $\varphi : M \to \mathbb{R}$, une forme linéaire et p une semi-norme telle que $\varphi \leq p$ sur M.*

Alors, il existe une forme linéaire $\tilde{\varphi} : X \to \mathbb{K}$ telle que $|\tilde{\varphi}| \leq p$ sur X.

Démonstration. Dans le cas où X est un espace vectoriel réel, il suffit d'appliquer le théorème précédent en remarquant qu'alors, $\tilde{\varphi}(-x) \leq p(-x)$, donc, comme $\tilde{\varphi}$ est linéaire et p est une semi-norme, $\tilde{\varphi}(x) \geq -p(x)$

Notation 1. Pour un complexe $z \in \mathbb{C}$, on note $\Re z$ sa partie réelle.

Dans le cas où X est un espace vectoriel complexe, nous allons utiliser le lemme suivant.

Lemme B.4.3. *Soient X un espace vectoriel complexe et φ une forme linéaire complexe sur X. Posons $u = \Re\varphi$. Alors, X est également un espace vectoriel réel, sur lequel u est une forme linéaire réelle, et u caractérise φ grâce à l'égalité :*

$$\forall x \in X, \ \ \varphi(x) = u(x) - iu(ix)$$

Réciproquement, toute forme linéaire réelle u sur X définit une unique forme linéaire complexe à l'aide de l'égalité précédente.

Il suffit pour le démontrer d'utiliser l'identité, pour tout $z \in \mathbb{C}$, $z = \Re z - i\Re(iz)$. Le reste de la preuve n'est qu'une vérification facile.

Pour en finir avec le cas complexe, soit $u = \Re\varphi$ et $\tilde{u}$ défini sur X à l'aide du cas réel. Il existe alors une unique forme linéaire complexe $\tilde{\varphi}$ de partie réelle $\tilde{u}$. Il reste à démontrer que pour tout $x \in X$, $|\tilde{\varphi}(x)| \leq p(x)$. Pour cela, on considère un complexe α de module 1 tel que $\alpha\tilde{\varphi}(x) \in \mathbb{R}$.

On a alors

$$|\tilde{\varphi}(x)| = |\alpha\tilde{\varphi}(x)| = |\tilde{\varphi}(\alpha x)| = |\tilde{u}(\alpha x)| \leq p(\alpha x) = p(x)$$

□

Définition B.4.1. Soit X un espace vectoriel topologique. On appelle *dual* de X, l'ensemble, noté X^* des formes linéaires continues sur X.

Lemme B.4.4. *Soit X un espace vectoriel topologique. Toute forme linéaire ψ non nulle sur X est ouverte (i.e. l'image de tout ouvert U de X par ψ est un ouvert de $\mathbb{K}$).*

Démonstration. Soient U un ouvert de X, $x_0 \in U$ et $y_0 = \psi(u_0)$. Comme U est supposé ouvert, il existe un voisinage ouvert équilibré V de 0 tel que $x_0 + V \subset U$. Par linéarité, $\psi(x_0) + \psi(V) \subset \psi(U)$.

Comme ψ est supposée non nulle, il existe $v_0 \in V$ tel que $\psi(v_0) \neq 0$, car tout voisinage de 0 est absorbant. Le voisinage V étant équilibré, on a $Dv_0 \subset V$, où D désigne le disque unité ouvert de $\mathbb{K}$. On en déduit que

$$\psi(x_0) + D\psi(v_0) \subset \psi(U)$$

Puisque $D\psi(v_0)$ est le disque ouvert de rayon $|\psi(v_0)|$ dans $\mathbb{K}$, on a bien trouvé un voisinage de y_0 inclus dans $\psi(U)$. □

Théorème B.4.5 (Premier théorème de séparation de Hahn-Banach)**.** *Soit X un espace vectoriel topologique, U un* ouvert convexe *de X et C une partie* convexe *de X, disjointe de U. Alors, il existe $\varphi \in X^*$ et un réel s tel que*

$$\forall x \in U, \;\; \forall y \in C, \;\; \Re\varphi x < s \leq \Re\varphi y$$

Démonstration. Vu le lemme B.4.3, il suffit de démontrer le théorème dans le cas où $\mathbb{K} = \mathbb{R}$.

Fixons $u_0 \in U$, $c_0 \in C$, posons $x_0 = c_0 - u_0 \neq 0$ et considérons la partie $V = U - C + x_0$. On vérifie facilement que V est un voisinage ouvert convexe de 0. Soit j la jauge de V. On a alors $V = \{x \in X \;\;|\;\; j(x) < 1\}$ et comme $x_0 \notin V$, $j(x_0) \geq 1$.

On définit alors φ sur $\mathbb{R}x_0$ par $\varphi(x_0) = 1$ et l'on a alors $\varphi \leq j$ sur le sous-espace $\mathbb{R}x_0$. Les conditions du théorème d'extension sont réunies et l'on peut prolonger φ en une forme linéaire $\tilde{\varphi}$ sur X tout entier, qui satisfait, pour tout $x \in X$,

$$\tilde{\varphi}(x) \leq j(x)$$

La forme linéaire $\tilde{\varphi}$ est continue : en effet, $-V$ est aussi un voisinage de 0, et pour $x \in -V$, on a $\tilde{\varphi}(-x) \leq j(-x) < 1$, donc $\tilde{\varphi}(x) > -1$. On a donc, pour tout $x \in V \cap (-V)$, $|\tilde{\varphi}(x)| < 1$, de sorte que $\tilde{\varphi}$, bornée sur un voisinage de 0, est bien continue.

Puis, remarquons que pour tous $u \in U$, $c \in C$, on a $u - c + x_0 \in V$ donc

$$\tilde{\varphi}(u) - \tilde{\varphi}(c) + 1 = \tilde{\varphi}\left(u - c + x_0\right) \leq j\left(u - c + x_0\right) < 1$$

On en déduit que pour tout $u \in U$ et pour tout $c \in C$, $\tilde{\varphi}(u) < \tilde{\varphi}(c)$.

Les images $I = \tilde{\varphi}\left(U\right)$ et $J = \tilde{\varphi}\left(C\right)$ sont donc des parties convexes de $\mathbb{R}$, c'est-à-dire des intervalles et I et J sont disjoints, avec I à gauche de J.

De plus, d'après le lemme précédent, I est un intervalle ouvert. Si l'on note s sa borne supérieure, on obtient $I \subset]-\infty; s[$ et $J \subset [s; +\infty[$, ce qui conclut la preuve. □

Théorème B.4.6 (Second théorème de séparation de Hahn-Banach)**.** *Soit X un espace vectoriel topologique localement convexe, K un compact convexe de X et F un* fermé convexe *de X. Alors, il existe $\varphi \in X^*$ et deux réels a et b tels que*

$$\forall x \in K, \;\; \forall y \in F, \;\; \Re\varphi(x) \leq a < b < \Re\varphi(y)$$

Démonstration. D'après le théorème B.1.3 et la proposition B.2.2, il existe un voisinage V, ouvert convexe de 0 tel que $(K + V) \cap F = \emptyset$.

On peut alors appliquer le théorème précédent avec l'ouvert convexe $K + V$ et le convexe F, pour trouver $\tilde{\varphi} \in X^*$ telle que $\tilde{\varphi}\left(K + V\right)$ soit un intervalle ouvert disjoint de l'intervalle $\tilde{\varphi}\left(F\right)$.

Comme $\tilde{\varphi}$ est continue, $\tilde{\varphi}\left(K\right)$ est un intervalle compact inclus strictement dans l'intervalle ouvert $\tilde{\varphi}\left(K + V\right)$. Notons a la borne supérieure de $\tilde{\varphi}\left(K\right)$ et b la borne supérieure de $\tilde{\varphi}\left(K + V\right)$, le théorème est alors démontré. □

Corollaire B.4.7. *Soit X un espace vectoriel localement convexe. Alors X^* sépare les points de X.*

Démonstration. Soient x et y deux vecteurs distincts de X. Il suffit d'appliquer le théorème précédent au compact $\{x\}$ et au fermé $\{y\}$. □

B.5 Théorème de Krein-Milman

Définition B.5.1. Soit A une partie d'un espace vectoriel X. Soit $S \subset A$. On dit que S est une *partie extrémale* de A lorsque

$$\forall x, y \in A,\ \ \forall t \in]0;1[, tx + (1-t)y \in S \Rightarrow x \in S \quad \text{et} \quad y \in S$$

Autrement dit, si l'un des points x ou y n'est pas dans S, les points qui sont à la fois dans le segment ouvert $]x;y[$ et dans A ne sont pas dans S.

Exemple B.5.1. Dans un triangle, les sommets sont des points extrêmaux, les côtés sont des parties extrémales.

Théorème B.5.1 (Théorème de Krein-Milman)**.** *Soient X un espace vectoriel topologique localement convexe et K une partie* compacte *et non vide de X. Alors :*

1. *Toute partie extrémale compacte et non vide de K contient des points extrêmaux de K. En particulier, l'ensemble E des points extrêmaux de K est non vide.*
2. *$K \subset \overline{\text{co}}\,(E)$.*
3. *Si, de plus, K est convexe, alors $K = \overline{\text{co}}\,(E)$.*

Démonstration. 1. Soit $\mathcal{P}$ l'ensemble des parties *extrémales, compactes et non vides* de K. Comme K est extrémal dans lui-même, $\mathcal{P}$ est non vide.

L'idée la plus importante de la preuve est l'utilisation du dual topologique de X pour obtenir, à partir d'un élément de $\mathcal{P}$, un nouvel élément plus petit.

Si $S \in \mathcal{P}$ et $\varphi \in X^*$, alors, comme $\Re\varphi$ est continue sur le compact S, elle y atteint un maximum et on note S_φ l'ensemble des points de S où ce maximum est atteint. Alors S_φ est encore compact, extrémal et non vide.

En effet, c'est l'image réciproque dans S d'un singleton par l'application continue $\Re\varphi$ donc c'est un fermé de S et par conséquent un compact.

Montrons qu'il est extrémal. Soient x, y dans K et $t \in]0;1[$ tels que $z = tx + (1-t)y$ soit dans S_φ. Alors z est en particulier dans S, donc x et y aussi. Comme $\Re\varphi(z)$ est le maximum de $\Re\varphi$ sur S, on en déduit que $\Re\varphi(x) = \Re\varphi(y) = \Re\varphi(z)$, c'est-à-dire que x et y sont dans S_φ.

Cette idée va nous permettre de démontrer l'existence de points extrêmaux par un argument de récurrence transfinie.

Soit $S \in \mathcal{P}$ et $\mathcal{P}_S$ l'ensemble des éléments de $\mathcal{P}$ qui sont inclus dans S. L'ensemble $\mathcal{P}_S$ est non vide et est ordonné par $\supset$. Soit $\mathcal{Q} \subset \mathcal{P}_S$ une partie totalement ordonnée, et $M = \bigcap_{S \in Q} Q$. Alors M est non vide, car une intersection décroissante de compacts non vides ne l'est jamais, M est compacte, et il est facile de vérifier que M est encore extrémale, ainsi M majore $\mathcal{Q}$ pour $\supset$, de sorte que $\mathcal{P}_S$ est inductif.

Par le lemme de Zorn, il admet des éléments maximaux pour $\supset$. Soit P un tel élément. Si P n'est pas réduit à un point, alors il existe deux points distincts x et y dans P. Mais, comme X est localement convexe, X^* sépare les points de X, donc il existe $\psi \in X^*$ telle que $\Re\psi(x) < \Re\psi(y)$ de sorte que $P_\psi \subsetneq P$, ce qui contredit la maximalité de P pour $\supset$.

2. Raisonnons par l'absurde et supposons qu'il existe $x_0 \in K$ tel que $x_0 \notin \overline{\mathrm{co}}\,(E)$. Le théorème de séparation de Hahn-Banach, dans sa version de séparation d'un compact convexe et d'un fermé convexe dans un espace localement convexe affirme alors qu'il existe $\varphi \in X^*$ tel que $\sup_{x \in \overline{\mathrm{co}}(E)} \varphi(x) < \varphi(x_0)$.
   Une telle forme φ atteint donc son maximum sur K en-dehors de $\overline{\mathrm{co}}\,(E)$. En définissant K_φ comme dans la partie précédente, on obtient donc une partie compacte extrémale non vide de K qui est disjointe de $\overline{\mathrm{co}}\,(E)$. Or K_φ contient des points extrêmaux de K, c'est-à-dire des éléments de E, ce qui est contradictoire.
3. Si K est convexe, alors $\mathrm{co}\,(E) \subset K$, et comme K est fermé, $\overline{\mathrm{co}}\,(E) \subset K$.

□

B.6 Topologie $*$-faible, théorème de Banach-Alaoglu, métrisabilité

Définition B.6.1. Soit X un espace vectoriel topologique et X^* son dual. La topologie $*$-faible sur X^* est la topologie initiale associée aux formes linéaires :

$$\pi_x : \varphi \mapsto \varphi(x)$$

Proposition B.6.1. *La topologie $*$-faible sur X^* est aussi la topologie associée à la famille de semi-normes $p_x = |\pi_x|$. Par conséquent, elle fait de X^* un espace vectoriel topologique localement convexe.*

Théorème B.6.2 (Théorème de Banach-Alaoglu)**.** *Soit X un espace vectoriel topologique et X^* son dual topologique muni de la topologie $*$-faible.*

Soit V un voisinage de 0 dans X. Alors la partie de X^*

$$P_V = \{\varphi \in X^* \quad | \quad \forall x \in V, |\varphi(x)| \leq 1\}$$

est compacte.

Démonstration. Soit $x \in X$. Comme V est un voisinage de 0, il est absorbant et le réel $\lambda_x = \inf\{\lambda > 0 \quad | \quad x \in \lambda V\}$ est bien défini.

On a alors, pour tout φ dans P_V, $|\varphi(x)| \leq \lambda_x$. Si l'on note $\pi_x : \varphi \mapsto \varphi(x)$ et D_x le disque fermé de centre 0 et de rayon λ_x dans le corps $\mathbb{K}$ des scalaires, on obtient :

$$P_V \subset \bigcap_{x \in X} \pi_x^{-1}\,(D_x)$$

On voit apparaître l'argument principal de cette démonstration : on peut aussi voir P_V comme un sous-ensemble de l'ensemble de fonctions

$$T_V = \{f : X \to \mathbb{K} \quad | \quad \forall x \in \mathbb{K}, f(x) \in D_x\}$$

qui n'est autre que le produit cartésien des D_x.

Comme chaque D_x est compact, le théorème de Tychonoff permet d'établir que T_V est compact.

On peut ainsi munir P_V de deux topologies a priori différentes :

- La topologie $*$-faible induite par l'inclusion $P_V \subset X^*$, c'est-à-dire la topologie la moins fine rendant les applications π_x continues sur X^*.
- La topologie produit induite par l'inclusion $P_V \subset T_V$, c'est-à-dire la topologie la moins fine rendant les applications π_x continues sur T_V.

Mais en fait, ces topologies proviennent toutes les deux de la topologie produit sur $\mathbb{K}^X$. Il est alors évident qu'elles coïncident sur

$$P_V \subset T_V \cap X^* \subset \mathbb{K}^X$$

Il ne reste plus qu'à montrer que P_V est fermé dans $\mathbb{K}^X$.

Or la condition de linéarité est fermée : f est linéaire si et seulement si

$$\forall \alpha, \beta \in \mathbb{K}, \forall x, y \in X, \pi_{\alpha x + \beta y}(f) = \alpha \pi_x(f) + \beta \pi_y(f)$$

Donc l'ensemble des fonctions linéaires de X dans $\mathbb{K}$ est égal à

$$\bigcap_{(\alpha,\beta,x,y) \in X^2 \times \mathbb{K}^2} \left(\pi_{\alpha x + \beta y} - \alpha \pi_x - \beta \pi_y\right)^{-1} (\{0\})$$

qui est fermé comme intersection de fermés.

La condition d'être bornée par 1 sur V est également fermée. Si l'on note $\overline{D} = \{\alpha \in \mathbb{K} \mid |\alpha| \leq 1\}$, l'ensemble des telles fonctions est

$$\bigcap_{x \in V} \pi_x^{-1} \left(\overline{D}\right)$$

Ainsi, P_V est fermé dans $\mathbb{K}^X$, et inclus dans le compact T_V, il est donc fermé. □

Les propositions générales suivantes vont nous permettre d'utiliser ce puissant théorème pour métriser des parties d'espaces vectoriels topologiques.

Proposition B.6.3. *Soit X un ensemble. Si τ_1 et τ_2 sont deux topologies sur X telles que (X, τ_1) est séparé, (X, τ_2) est compact et $\tau_1 \subset \tau_2$, alors $\tau_1 = \tau_2$.*

Autrement dit, on ne peut pas appauvrir une topologie compacte sans en perdre la séparation.

Démonstration. Nous démontrons que τ_1 et τ_2 ont les mêmes fermés.

Soit F un fermé pour τ_2. Alors F est compact pour τ_2. Soit $\mathcal{R}$ un recouvrement de F par des ouverts de τ_1. Comme $\tau_1 \subset \tau_2$, $\mathcal{R}$ est aussi un recouvrement de F par des ouverts de τ_2. On peut en extraire un sous-recouvrement fini, puis comme τ_1 est séparé, F est compact, donc fermé. □

Proposition B.6.4. *Soient (X, τ) un espace topologique compact et $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions* continues *de X dans $\mathbb{K}$ qui sépare les points de X. Alors X est métrisable.*

Démonstration. Quitte à les diviser par $\sup_{x\in X}|f_n(x)|$, on peut supposer que :

$$\forall x \in X, \forall n \in \mathbb{N}, |f_n(x)| \leq 1$$

Posons alors, pour x et y dans X,

$$d(x,y) = \sum_{n\in\mathbb{N}} 2^{-n} |f_n(x) - f_n(y)|$$

Puisque les f_n séparent les points, d définit bien une distance sur X.

Cette série est uniformément convergente, donc par continuité des fonctions $(x,y) \mapsto |f_n(x) - f_n(y)|$, la distance d est continue. Par conséquent, la boule ouverte $B(x,r)$, image réciproque par l'application $d(x,.)$ de l'ouvert $[0;r[$ de $\mathbb{R}^+$ est un ouvert de τ.

La topologie τ_d définie par cette métrique est donc plus faible que τ, elle est séparée comme toute topologie métrique et τ est compacte, donc d'après la proposition précédente, $\tau_d = \tau$. □

Corollaire B.6.5. *Soit X un espace topologique* séparable *et K une partie *-faiblement compacte de X^*. Alors K est métrisable. En particulier, avec les notations du théorème de Banach-Alaoglu, P_V est métrisable.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une famille dénombrable dense de X. On définit $f_n : X^* \to \mathbb{K}$ par $f_n(\varphi) = \varphi(x_n)$.

Par définition de la topologie $*$-faible, les f_n sont continues. De plus, elles séparent les points de X^*. En effet, si φ et ψ sont telles que pour tout $n \in \mathbb{N}$, $\varphi(x_n) = \psi(x_n)$, ces deux applications continues coïncident sur une partie dense de X, donc sont égales.

On peut alors appliquer la proposition précédente à K. □

C Mesures boréliennes et théorèmes de représentations

Les résultats qui vont suivre sont très classiques (ils sont entièrement démontrés, par exemple, dans le chapitre 6 de [Rud75]). Pour cette raison, nous nous contenterons d'énoncer les définitions et les théorèmes, à l'exception du théorème C.3.2 qui est plus spécifique à notre étude.

C.1 Théorème de Radon-Nikodym

Soit $\mathcal{F}$ une tribu sur un ensemble X.

Définition C.1.1. Soient μ et ν deux mesures positives sur $(X, \mathcal{F})$. On dit que :

1. la mesure μ est σ-finie, lorsqu'il existe une suite $(X_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ telle que

$$\forall n \in \mathbb{N}, \quad \mu(X_n) < \infty \quad \text{et} \quad \bigcup_{n\in\mathbb{N}} X_n = X$$

2. la mesure ν est *absolument continue* par rapport à μ lorsque, pour tout $E \in \mathcal{F}$,

$$\mu(E) = 0 \Rightarrow \nu(E) = 0$$

Dans ce cas, on note $\nu \ll \mu$.

Théorème C.1.1 (Théorème de Radon-Nikodym)**.** *Soient μ et ν deux mesures positives σ-finies sur $(X, \mathcal{F})$, telles que $\nu << \mu$.*

Alors, il existe une fonction h, mesurable, positive telle que $d\nu = hd\mu$, i.e.

$$\forall E \in \mathcal{F}, \quad \nu(A) = \int 1_A g d\mu$$

De plus, h est unique, à un ensemble μ-négligeable près.

Remarque C.1.1. Il existe bien d'autres versions plus générales de ce théorème, mais nous n'avons besoin dans ce mémoire que de celle-ci.

C.2 Mesures signées

Soit $\mathcal{F}$ une tribu sur un ensemble X.

Définition C.2.1. Une mesure signée μ sur $(X, \mathcal{F})$ est une fonction de $\mathcal{F}$ dans $\mathbb{R}$ telle que, pour tout mesurable $E \in \mathcal{F}$ et toute partition dénombrale $(E_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ de E, on ait

$$\mu(E) = \sum_{n=0}^{+\infty} \mu(E_n)$$

où la convergence de cette série est absolue.

On note ici $\mathcal{S}(X, \mathcal{F})$ l'ensemble des mesures signées sur $(X, \mathcal{F})$.

On définit également la *variation totale* $|\mu|$ de μ comme la fonction d'ensembles égale à :

$$|\mu|(E) = \sup \sum_{n=0}^{+\infty} |\mu(E_n)|$$

où la borne supérieure est prise sur l'ensemble des partitions $(E_n)_{n\in\mathbb{N}}$ de E.

Théorème C.2.1. *La variation totale $|\mu|$ d'une mesure signée μ est une mesure positive et finie sur $(X, \mathcal{F})$.*

Notation 2. Pour $\mu \in \mathcal{S}(X, \mathcal{F})$, on note $\|\mu\|_{\mathcal{S}} = |\mu|(X)$.

Théorème C.2.2. *L'espace $\mathcal{S}(X, \mathcal{F})$ est un espace vectoriel, sur lequel $\| \ \|_{\mathcal{S}}$ est une norme.*

C.3 Théorème de représentation de Riesz

Dans cette partie, X est muni d'une topologie et $\mathcal{B}$ est la tribu borélienne associée à cette topologie.

Définition C.3.1. 1. On dit qu'une mesure (positive ou signée) est *borélienne* lorsqu'elle est définie sur $\mathcal{B}$.

2. Une mesure borélienne positive est dite *régulière* lorsque, pour tout $E \in \mathcal{B}$,

$$\mu(E) = \sup\{\mu(K) \ \mid \ K \text{ compact de } X, \ K \subset E\} \quad \text{et}$$

$$\mu(E) = \inf\{\mu(U) \ \mid \ U \text{ ouvert de } X, \ U \supset E\}$$

3. Une mesure borélienne signée est dite *régulière* lorsque sa variation totale $|\mu|$ l'est.

On note $\mathcal{C}_c(X)$ l'ensemble des fonctions continues à support compact de X dans $\mathbb{R}$, muni de la norme $\| \ \|_\infty$. Le Théorème de représentation de Riesz, aussi appelé Théorème de représentation de Riesz-Markov, fait le lien entre le dual topologique de $\mathcal{C}_c(X)$ et l'ensemble des mesures boréliennes signées $\mathcal{S}(X, \mathcal{B})$. Il en existe de nombreuses versions, celle présentée dans [Rud75] est adaptée au cadre de notre étude.

Théorème C.3.1 (Théorème de représentation de Riesz-Markov)**.** *Pour toute forme linéaire continue Φ sur $\mathcal{C}_c(X)$, où X est un espace séparé localement compact, il existe une unique mesure borélienne, signée et régulière μ, telle que*

$$\forall f \in \mathcal{C}_c(X), \quad \Phi(f) = \int_X f d\mu$$

De plus, cette correspondance est une isométrie :

$$\|\Phi\| = \|\mu\|_{\mathcal{S}} = |\mu|(X)$$

On a donc une isométrie entre $\mathcal{C}_c(X)^*$ et l'espace $\mathcal{S}_r(X)$ des mesures boréliennes, signées et régulières sur X.

Dans le cas où X est métrique compact, l'utilisation de ce théorème est plus confortable. Ce qui suit est tiré de [Wal00].

Théorème C.3.2. *Soit X un espace métrique compact. Alors toute mesure borélienne signée sur X est régulière.*

Démonstration. Par définition, il suffit de démontrer le théorème dans le cas d'une mesure μ positive et finie sur X. Remarquons que dans le cas d'un espace compact, les fermés et les compacts coïncident.

On considère l'ensemble de parties

$$\mathcal{R} = \{A \in \mathcal{B} \ \mid \ \forall \varepsilon > 0,\ \exists U_\varepsilon \text{ ouvert},\ \exists K_\varepsilon \text{ compact},\ K_\varepsilon \subset A \subset U_\varepsilon \text{ et } \mu\left(U_\varepsilon \setminus K_\varepsilon\right) < \varepsilon\}$$

Nous allons montrer que $\mathcal{R}$ est une tribu.

Il est évident que $\mathcal{R}$ contient X et il est facile de vérifier que $\mathcal{R}$ est stable par passage au complémentaire.

Soit maintenant $(A_n)_{n\geq 1}$ une famille dénombrable d'éléments de $\mathcal{R}$ et $A = \bigcup_{n\geq 1} A_n$. Soit $\varepsilon > 0$.

Pour tout n, il existe un ouvert $U_{n,\varepsilon}$ et un compact $K_{n,\varepsilon}$ tels que

$$K_{n,\varepsilon} \subset A_n \subset U_{n,\varepsilon} \quad \text{et} \quad \mu\left(U_{n,\varepsilon} \setminus K_{n,\varepsilon}\right) < 3^{-n}\varepsilon$$

Soient $U_\varepsilon = \bigcup_{n\geq 1} U_{n,\varepsilon}$ et $\tilde{K_\varepsilon} = \bigcup_{n\geq 1} K_{n,\varepsilon}$.

Comme $\tilde{K_\varepsilon} = \bigcup_{k\geq 1} \bigcup_{n=1}^{k} K_{n,\varepsilon}$, on a

$$\mu\left(\tilde{K_\varepsilon}\right) = \lim_{k\to+\infty} \mu\left(\bigcup_{n=1}^{k} K_{n,\varepsilon}\right)$$

Donc, comme μ est finie, il existe un entier k_0 tel que

$$\mu\left(\tilde{K_\varepsilon} \setminus \bigcup_{n=1}^{k_0} K_{n,\varepsilon}\right) < \frac{\varepsilon}{2}$$

On pose $K_\varepsilon = \bigcup_{n=1}^{k_0} K_{n,\varepsilon}$. Alors, U_ε est ouvert et K_ε est compact. De plus,

$$K_\varepsilon \subset \tilde{K_\varepsilon} \subset A \subset U_\varepsilon$$

Enfin,

$$\begin{array}{rcl}
\mu\left(U_\varepsilon \setminus K_\varepsilon\right) & \leq & \mu\left(U_\varepsilon \setminus \tilde{K_\varepsilon}\right) + \mu\left(\tilde{K_\varepsilon} \setminus K_\varepsilon\right) \\
& \leq & \displaystyle\sum_{n\geq 1} \mu\left(U_{n,\varepsilon} \setminus K_{n,\varepsilon}\right) + \frac{\varepsilon}{2} \\
& \leq & \displaystyle\sum_{n\geq 1} 3^{-n}\varepsilon + \frac{\varepsilon}{2} = \varepsilon
\end{array}$$

Pour conclure, nous montrons que $\mathcal{R}$ contient tous les fermés de X.

Soit K un fermé de X et, pour $n \geq 1$,

$$U_n = \left\{x \in X \ \mid \ d\left(x, K\right) < \frac{1}{n}\right\}$$

Alors U_n est ouvert, par continuité de l'application $x \mapsto d(x, K)$ et comme K est fermé,

$$\bigcap_{n \geq 1} U_n = \{x \in X \mid d(x, K) = 0\} = K$$

La suite (U_n) étant décroissante, et la mesure μ finie, on a

$$\lim_{n \to +\infty} \mu(U_n) = \mu(K)$$

donc pour tout $\varepsilon > 0$, il existe $n_0 \geq 1$ tel que $\mu(U_{n_0} \setminus K) < \varepsilon$. Il suffit alors de poser $U_\varepsilon = U_{n_0}$ et $K_\varepsilon = K$ pour terminer cette démonstration. $\square$

Regrets et perspectives

Le présent mémoire n'a pas pu aborder un certain nombre de sujets très intéressants.

- À la fin de la section 4, il serait possible de démontrer le très important Théorème de Shannon-McMillan, en étudiant en profondeur le logarithme de certains jacobiens.
- Nous aurions aimé terminer ce mémoire en montrant le lien fondamental qu'il y a entre le principe variationnel et le Théorème de Perron-Frobenius-Ruelle : lorsque $\varphi \in \mathcal{B}_0$, alors la mesure $\nu = h\mu$ construite dans ce théorème est l'unique mesure telle que
$$\mathrm{P}(\varphi) = h_\nu(T) + \int_X \varphi d\nu$$
Elle est ergodique.
- Avec davantage de temps et beaucoup plus d'expérience dans ce domaine, il serait bon d'ajouter plus d'exemples.
- Les liens avec la physique n'ont pas été explorés par l'auteur, faute de compétence dans cette science. Elles apparaissent pourtant dès la génèse des systèmes dynamiques, puis elles fournissent aux mathématiciens des outils tels que l'entropie, la pression et les équilibres de Gibbs. L'apprentissage de ces notions d'un point de vue physique serait passionnant et sans doute fructueux.
- La motivation de cette étude dans l'ouvrage de Michel Zinsmeister est l'utilisation de tous ces outils pour le calcul de certaines dimensions de Hausdorff. Pour étudier cela, nous devrons nous plonger de nouveau dans un domaine inconnu.
- Le Théorème de Birkhoff est un résultat fondamental qui a de très nombreux corollaires. Nous aurions aimé, en l'utilisant, démontrer la loi forte des grands nombres ou étudier les chaînes de Markov à espace d'états infinis. Cela aurait nécessité d'autres théorèmes d'extension que celui de Carathéodory, afin de construire les bons espaces et les bonnes mesures.

Références

[Bir31] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 17(12) :656, 1931.

[BP00] Marc Briane and Gilles Pages. *Théorie de l'intégration : licence de mathématiques ; cours et exercices*. Vuibert, 2000.

[ELW11] Manfred Einsiedler, Elon Lindenstrauss, and Thomas Ward. Entropy in Dynamics. http://maths.dur.ac.uk/~tpcc68/entropy/welcome.html, 2011. Projet de livre en cours de finalisation.

[Gar65] AM Garcia. A simple proof of E. hopf's maximal ergodic theorem. *J. Math. Mech*, 14 :381–382, 1965.

[Laf12] Jacques Lafontaine. *Introduction aux variétés différentielles*. EDP sciences, 2012.

[Mau01] Bernard Maurey. Analyse fonctionnelle et théorie spectrale. https://www.imj-prg.fr/~bernard.maurey/ts012/poly/lths.pdf, 2001. Polycopié de cours de maîtrise de l'université Paris 7.

[RA95] Walter Rudin and Abdellatif Abouhazim. *Analyse fonctionnelle*. Ediscience international, 1995.

[Rud75] Walter Rudin. *Analyse réelle et complexe*. Masson, 1975.

[Sch95] Laurent Schwartz. *Analyse. 1. Théorie des ensembles et topologie*. Hermann, 1995.

[Wal00] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer, 2000.

[Wil91] David Williams. *Probability with martingales*. Cambridge university press, 1991.

[YK+39] Kôsaku Yosida, Shizuo Kakutani, et al. Birkhoff's ergodic theorem and the maximal ergodic theorem. *Proceedings of the Imperial Academy*, 15(6) :165–168, 1939.

[Zin96] Michel Zinsmeister. Publications de la smf. *Panoramas et Synthèses*, 4 :102, 1996.