

Théorie ergodique

Tommy Giroux

Reçu le 2024-09-20 et accepté le 2026-03-20

RÉSUMÉ Le but de cet article est de familiariser le lecteur avec la théorie ergodique. Pour ce faire, nous présenterons certaines notions relatives à la théorie de la mesure et à l'analyse fonctionnelle. Au cours de l'article, nous donnerons un exemple de système ergodique et énoncerons certains résultats fondamentaux de la théorie ergodique. Nous présenterons notamment, et démontrerons, le célèbre théorème ergodique de Birkhoff.

1 Introduction

En 1871, Ludwig Boltzmann, un physicien autrichien, formule pour la première fois l'hypothèse ergodique. Selon cette hypothèse fondamentale de la physique statistique, à l'équilibre, la valeur moyenne d'une grandeur sur l'espace étudié est égale à la moyenne d'un très grand nombre de mesures effectuées dans le temps. De l'analyse mathématique de cette hypothèse est née la théorie ergodique. Dans ce nouveau cadre, l'ergodicité exprime l'idée qu'un point en mouvement dans un système visitera toutes les régions de celui-ci de manière aléatoire. On retrouve aussi cette idée dans la théorie du chaos, où l'ergodicité est vue comme une propriété caractéristique des systèmes chaotiques. De plus, en théorie des probabilités, la théorie ergodique permet également de généraliser la loi des grands nombres à des familles de variables aléatoires dépendantes (voir [Kle13, chapitre 20]).

Dans notre exploration de la théorie ergodique, nous aborderons les espaces métriques [Mor20], la théorie de la mesure [Gal09, Ros06] et l'analyse fonctionnelle [Mus14]. À l'aide des définitions et des résultats rassemblés au fil de l'article, nous serons alors en mesure de présenter les notions et les résultats élémentaires de la théorie ergodique [Gre15, EW11].

Pour illustrer l'idée au cœur de la théorie ergodique, nous démontrons un résultat concernant les orbites des rotations irrationnelles du cercle. Ce dernier peut être représenté par l'intervalle $[0,1]$ dont on identifie les extrémités, autrement dit par l'espace quotient \mathbb{R}/\mathbb{Z} . En identifiant \mathbb{R}/\mathbb{Z} à l'intervalle $[0,1)$, nous munissons \mathbb{R}/\mathbb{Z} de l'ordre induit. Ainsi, pour $a, b \in \mathbb{R}/\mathbb{Z}$, l'expression $a < b$ signifie que les représentants choisis dans $[0,1)$ vérifient $a < b$.

Je tiens à remercier le Département de mathématiques de l'Université de Sherbrooke ainsi que le Conseil de recherches en sciences naturelles et en génie du Canada pour leur soutien financier. J'exprime également ma gratitude à Jean-Philippe Burelle pour son aide dans le cadre de mes travaux de recherche.

Lemme 1.1 ([Gre15], Théorème 1.1). *Soit $T : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ une rotation du cercle définie par $Tx = x + \alpha \pmod{1}$, où $0 < \alpha < 1$. Alors, l'orbite $(T^n x)_{n=0}^\infty$ (où T^n désigne la composition de l'application T avec elle-même n fois) est dense dans \mathbb{R}/\mathbb{Z} si et seulement si α est irrationnel.*

Démonstration. Pour la nécessité, nous allons procéder par contraposée. Supposons que α est rationnel. Alors, il existe $p, q \in \mathbb{N}$ tels que $\text{pgcd}(p, q) = 1$ et $\alpha = \frac{p}{q}$, ce qui implique qu'il existe $m \in \mathbb{N}$ tel que $\frac{mp}{q} = p$. En particulier, $m = q$, et comme $T^n x = x + n\alpha$, l'orbite consiste en q points distincts, séparés par un écart de $\frac{1}{q}$. Ainsi, l'orbite $(T^n x)_{n=0}^\infty$ n'est pas dense dans \mathbb{R}/\mathbb{Z} .

Pour la suffisance, soit α irrationnel et

$$d: \mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z} \rightarrow [0, \infty)$$

$$(x, y) \mapsto d(x, y),$$

la métrique sur \mathbb{R}/\mathbb{Z} définie par $d(x, y) = \min(|x - y|, 1 - |x - y|)$. Pour tous $x, y \in \mathbb{R}/\mathbb{Z}$, on remarque que comme α est irrationnel, il existe un entier naturel $n > 1$ tel que $x + (n - 1)\alpha \pmod{1} \leq y$ et $x + n\alpha \pmod{1} \geq y$, ce qui implique que $d(y, x + (n - 1)\alpha \pmod{1}) \leq \frac{\alpha}{2}$ ou que $d(y, x + n\alpha \pmod{1}) \leq \frac{\alpha}{2}$. Posons

$$n_1 = \begin{cases} n - 1 & \text{si } d(y, x + (n - 1)\alpha \pmod{1}) \leq \frac{\alpha}{2} \\ n & \text{si } d(y, x + n\alpha \pmod{1}) \leq \frac{\alpha}{2} \end{cases}.$$

Ce n_1 induit une nouvelle rotation :

$$T^{n_1}: \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$$

$$x \mapsto x + n_1\alpha \pmod{1}$$

Par récurrence, il est possible d'obtenir une suite $\{n_k\}_{k \in \mathbb{N}}$ d'entiers naturels telle que :

$$d(y, T^{n_1} x) \leq \frac{\alpha}{2}$$

$$d(y, T^{n_2+n_1} x) \leq \frac{\alpha}{2^2}$$

$$d(y, T^{n_3+n_2+n_1} x) \leq \frac{\alpha}{2^3}$$

$$\vdots$$

On a donc que $d(y, T^{n_k+\dots+n_2+n_1} x) \rightarrow 0$ lorsque $k \rightarrow \infty$, ce qui implique que pour tout $\varepsilon > 0$ et tout $x, y \in \mathbb{R}/\mathbb{Z}$, il existe $n \in \mathbb{N}$ tel que $d(y, T^n x) < \varepsilon$. Ainsi, l'orbite d'une rotation irrationnelle est dense dans \mathbb{R}/\mathbb{Z} . \square

Ce lemme jouera un rôle central dans la suite, puisqu'il servira de point de départ à la construction d'un système ergodique.

2 Définitions et premiers résultats

Chaque section commencera par un court rappel de certaines notions indispensables à la compréhension du sujet traité. Un lecteur expérimenté pourra, s'il le souhaite, passer cette sous-section.

2.1 Espace métrique et théorie de la mesure

Dans ce rappel, il sera question des espaces métriques et de la théorie de la mesure. En ce qui concerne les espaces métriques, la référence principale est [Mor20], tandis que pour la théorie de la mesure, nous nous appuyerons sur [Gal09] et [Ros06]. Pour plus de détails sur les espaces métriques, le lecteur pourra également consulter [Mus14]. Dans un premier temps, nous passerons en revue quelques définitions et résultats importants concernant les espaces métriques.

Définition 2.1. Soit X un ensemble non vide et $d : X \times X \rightarrow [0, \infty)$ une application telle que, pour tous $a, b, c \in X$, on a les propriétés suivantes :

- i) $d(a, b) \geq 0$ et $d(a, b) = 0$ si et seulement si $a = b$,
- ii) $d(a, b) = d(b, a)$, et
- iii) $d(a, c) \leq d(a, b) + d(b, c)$.

Alors, on appelle d une **métrique** sur X et on dit que (X, d) est un **espace métrique**.

En d'autres termes, un espace métrique n'est rien d'autre qu'un ensemble sur lequel on a défini une notion de distance. L'exemple le plus simple d'un tel espace est sans doute l'espace euclidien de dimension 1, c'est-à-dire le couple (\mathbb{R}, d) , où d désigne la distance euclidienne définie, pour tout $x, y \in \mathbb{R}$, par $d(x, y) = |x - y|$.

Dans un premier cours d'analyse, on se familiarise avec le concept de convergence dans le cas réel. Ici, nous utilisons les espaces métriques afin de généraliser ce concept. La définition suivante n'est rien d'autre que le prolongement naturel de la définition usuelle de la convergence au cadre des espaces métriques.

Définition 2.2. Soit (X, d) un espace métrique et $\{x_n\}_{n \in \mathbb{N}}$ une suite de points dans X . On dit que la suite **converge** vers $x \in X$, si pour chaque $\varepsilon > 0$, il existe un entier $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, $d(x, x_n) < \varepsilon$. On écrit $x_n \rightarrow x$. De plus, une suite $\{y_n\}_{n \in \mathbb{N}}$ de points de X est dite **convergente** dans (X, d) , s'il existe un point $y \in X$ tel que $y_n \rightarrow y$.

On note également que, dans le cas des espaces métriques, si la limite d'une suite existe, alors cette limite est unique. On termine ce rappel sur les espaces métriques par la définition des boules. Soient (X, d) un espace métrique, $p \in X$ et $r \in \mathbb{R}^{>0}$. On appelle **boule ouverte centrée en p** l'ensemble $B(p, r) := \{x \in X \mid d(p, x) < r\}$.

Il est maintenant temps de regarder du côté de la théorie de la mesure. Cette dernière nous permet de fournir un cadre rigoureux à la théorie ergodique.

Définition 2.3. Soit X un ensemble. On appelle **tribu** (ou **σ -algèbre**) sur X une famille \mathcal{A} de parties de X possédant les propriétés suivantes :

- i) $X \in \mathcal{A}$,
- ii) si $A \in \mathcal{A}$, alors $A^c \in \mathcal{A}$, et
- iii) si $A_n \in \mathcal{A}$, pour tout $n \in \mathbb{N}$, alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Les éléments de \mathcal{A} sont appelés les **ensembles mesurables** (ou **parties mesurables**) de (X, \mathcal{A}) et on dit que (X, \mathcal{A}) est un **espace mesurable**.

Les tribus nous permettent de définir de façon rigoureuse la notion d'ensemble mesurable, un concept fondamental en théorie de la mesure. Dans le cas où X serait un ensemble contenant un nombre dénombrable d'éléments, on pourrait tout simplement prendre $M = \mathcal{P}(X)$, l'ensemble des parties de X . La définition précise d'une tribu n'est pas essentielle pour une bonne compréhension de la théorie ergodique. Une tribu sur un ensemble X est simplement un sous-ensemble de l'ensemble des parties de X , possédant certaines propriétés. La nécessité des tribus vient du fait qu'il n'est généralement pas possible de définir une mesure possédant certaines propriétés intéressantes sur l'ensemble des parties d'un ensemble donné. Les définitions suivantes nous permettront de nous familiariser avec deux types de tribus bien connus.

Définition 2.4. Soit X un ensemble et $M \subset \mathcal{P}(X)$. La **tribu engendrée** par M sur X est l'intersection de toutes les tribus sur X contenant les éléments de M .

On peut caractériser la tribu engendrée par M sur X comme étant la plus petite tribu sur X contenant les éléments de M . Notons aussi que la tribu engendrée par M sur X existe, puisque $\mathcal{P}(X)$ est une tribu sur X contenant M ; l'intersection considérée est donc non vide.

Définition 2.5. Soit (X, d) un espace métrique. Notons \mathcal{O} l'ensemble des **ouverts** de (X, d) (c'est-à-dire les boules ouvertes, leurs unions et les intersections finies de boules ouvertes). On appelle **tribu borélienne** (ou **tribu de Borel**) de X , et on la note $\mathcal{B}(X)$ (ou simplement \mathcal{B} lorsqu'il n'y a pas de risque de confusion), la tribu engendrée par \mathcal{O} .

Après la fin de cette sous-section, lorsqu'il sera question de tribu, il s'agira de la tribu de Borel. Ceci dit, nous pouvons désormais discuter de mesure et d'espace mesuré.

Définition 2.6. Soient X un ensemble et \mathcal{A} une tribu sur X . On dit qu'une application $\mu : \mathcal{A} \rightarrow [0, +\infty]$ est une **mesure** sur (X, \mathcal{A}) , si elle vérifie les conditions suivantes :

- i) La mesure de la partie vide est nulle, c'est-à-dire, $\mu(\emptyset) = 0$.

- ii) Si $(A_n)_{n \in \mathbb{N}}$ est une famille dénombrable d'éléments de \mathcal{A} deux à deux disjoints (c'est-à-dire $A_i \cap A_j = \emptyset$ pour tous $i \neq j$), alors :

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

On appelle cette propriété la σ -**additivité**.

On dit que (X, \mathcal{A}, μ) est un **espace mesuré**.

Dans cet article, lorsque nous utiliserons le terme mesure, il s'agira en fait de **mesure de probabilité**. Plus précisément, considérons (X, \mathcal{A}, μ) un espace mesuré. Alors, μ est une mesure de probabilité si $\mu(X) = 1$. Le lecteur remarquera qu'il n'est pas plus général de considérer les mesures finies (c'est-à-dire les mesures μ telles que $\mu(X) \in \mathbb{R}^{>0}$) que de considérer les mesures de probabilité. En effet, toute mesure finie peut être normalisée pour en faire une mesure de probabilité. La prochaine proposition nous donne quelques propriétés très utiles sur les mesures.

Proposition 2.7. [Propriétés élémentaires d'une mesure] *Soit μ une mesure sur l'espace mesurable (X, \mathcal{A}) , alors μ possède les propriétés suivantes :*

- i) *Monotonie* : Si $A, B \in \mathcal{A}$ et $A \subset B$, alors $\mu(A) \leq \mu(B)$.
- ii) *Sous-additivité* : Si $A_n \in \mathcal{A}$, $\forall n \in \mathbb{N}$, alors

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n).$$

- iii) *Si $A_n \in \mathcal{A}$, $\forall n \in \mathbb{N}$ et si $A_n \subset A_{n+1}$, $\forall n \in \mathbb{N}$, alors*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

- iv) *Si $A_n \in \mathcal{A}$, $\forall n \in \mathbb{N}$ et si $A_{n+1} \subset A_n$, $\forall n \in \mathbb{N}$ avec $\mu(A_0) < \infty$, alors*

$$\mu\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Dans la théorie de la mesure, on rencontre fréquemment des ensembles de mesure nulle. On dit que ces ensembles sont **négligeables**.

Définition 2.8. Soit (X, \mathcal{A}, μ) un espace mesuré. On dit qu'une propriété \mathcal{P} est vraie **presque partout** (ou **p.p.**) si l'ensemble $\{x \in X \mid \mathcal{P}(x) \text{ est fausse}\}$ est de mesure nulle ou contenu dans un ensemble de mesure nulle.

La notion de propriété vraie presque partout est omniprésente dans tout l'article et elle est un concept fondateur de la théorie de la mesure. On poursuit avec un autre concept très important pour la suite de l'article.

Définition 2.9. Soient (X_1, \mathcal{A}_1) et (X_2, \mathcal{A}_2) deux espaces mesurables. On dit d'une application $T : X_1 \rightarrow X_2$ qu'elle est **mesurable** si $T^{-1}A \in \mathcal{A}_1, \forall A \in \mathcal{A}_2$ (où $T^{-1}A := \{x \in X \mid Tx \in A\}$ désigne l'image réciproque de l'ensemble A par T).

Par exemple, soit (X, \mathcal{B}) , un espace mesurable, toute application constante de X dans \mathbb{R} muni de sa tribu de Borel est mesurable, car l'image réciproque de tout ensemble mesurable de \mathbb{R} est X ou l'ensemble vide. Le lecteur prendra note que les applications continues de X dans \mathbb{R} muni de sa tribu de Borel sont mesurables (l'image réciproque d'un ouvert par une application continue est ouvert). De plus, il est important de remarquer que le concept d'application mesurable est indépendant de la mesure choisie. On termine cette sous-section par un exemple de mesure.

Définition 2.10. Soit l'espace mesuré $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, où $d \in \mathbb{N}^*$ et $\mathcal{B}(\mathbb{R}^d)$ est la tribu de Borel de \mathbb{R}^d . On appelle **mesure de Lebesgue** la mesure λ telle que pour tout pavé $P = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d] \subset \mathbb{R}^d$,

$$\lambda(P) = \prod_{i=1}^d (b_i - a_i).$$

Le lecteur notera que la mesure de Lebesgue est invariante par translation et que l'existence d'une telle mesure constitue un fait non trivial (voir, par exemple, [Ros06, chapitre 2]). La propriété suivante est, elle aussi, un peu moins évidente.

Proposition 2.11. *La mesure de Lebesgue est extérieurement régulière. C'est-à-dire, pour tout $E \in \mathcal{B}(\mathbb{R}^d)$, $\lambda(E) = \inf\{\lambda(V) \mid V \text{ ouvert}, V \supset E\}$.*

2.2 Système préservant la mesure et ergodicité

On commence notre initiation à la théorie ergodique par quelques définitions et résultats élémentaires. L'approche utilisée dans cet article pour présenter la théorie ergodique est largement inspirée de celle que l'on retrouve dans [Gre15]. En particulier, cette section s'inspire des chapitres 2 et 3 de ces notes de cours. Pour les étudiants qui aimeraient poursuivre leur apprentissage en théorie ergodique, [EW11] serait une très bonne référence pour débiter, on l'utilise aussi comme référence pour certains résultats.

Définition 2.12. Soient $(X_1, \mathcal{B}_1, \mu_1)$ et $(X_2, \mathcal{B}_2, \mu_2)$ deux espaces mesurés et $T : X_1 \rightarrow X_2$ une application. On dit que T est une **application préservant la mesure** si T est mesurable et si $\mu_1(T^{-1}A) = \mu_2(A)$, pour tout $A \in \mathcal{B}$. Dans le cas où (X, \mathcal{B}, μ) est un espace mesurable et $T : X \rightarrow X$ est une application préservant la mesure, on appelle (X, \mathcal{B}, μ, T) un **système préservant la mesure**.

Le prochain résultat nous fournit un outil pour montrer qu'une application préserve la mesure. De plus, puisque nous travaillons seulement avec les tribus de Borel, ce résultat s'applique à tout l'article.

Lemme 2.13. Soient $(X_1, \mathcal{B}_1, \mu_1)$ et $(X_2, \mathcal{B}_2, \mu_2)$ deux espaces mesurés et $T : X_1 \rightarrow X_2$ une application mesurable. Supposons que \mathcal{B}_2 soit engendrée par une collection $M \subset \mathcal{P}(X_2)$. Supposons que pour tout $A \in M$, on a $\mu_1(T^{-1}A) = \mu_2(A)$, alors T est une application préservant la mesure.

Voir la démonstration du lemme 2.2 de [Gre15], pour la démonstration. Considérons une nouvelle fois les rotations irrationnelles du cercle et montrons qu'il s'agit d'un système préservant la mesure pour la mesure de Lebesgue.

Remarque . Ici, et pour la suite de l'article, on identifie \mathbb{R}/\mathbb{Z} à l'intervalle $[0,1]$ où on identifie les extrémités.

Exemple 2.14. Soient $X = \mathbb{R}/\mathbb{Z}$ et $T : X \rightarrow X$ la rotation définie par $Tx = x + \alpha \pmod{1}$, où $0 < \alpha < 1$ est irrationnel. Alors, $(X, \mathcal{B}, \lambda, T)$ est un système préservant la mesure, pour \mathcal{B} la tribu de Borel de X et λ la mesure de Lebesgue.

Démonstration. On peut démontrer que $\mathcal{B}(X)$ est engendrée par M , la collection d'ensembles qui contient les ensembles de la forme (a,b) et ceux de la forme $[0,a) \cup (b,1)$ où $a,b \in X$. Comme la mesure de Lebesgue est invariante par translation, on a $\lambda(T^{-1}A) = \lambda(A)$ pour tout $A \in M$. Ainsi, par le lemme précédent, $(X, \mathcal{B}, \lambda, T)$ est un système préservant la mesure. \square

Maintenant que nous avons un peu de vocabulaire, on peut énoncer et démontrer le premier résultat en lien direct avec la théorie ergodique.

Théorème 2.15. [Théorème de récurrence de Poincaré][[Gre15], Théorème 2.1] Soit (X, \mathcal{B}, μ, T) un système préservant la mesure. Soient $A \in \mathcal{B}$ et A' l'ensemble des $x \in A$ tels que $T^n x \in A$ pour une infinité de $n \in \mathbb{N}$. Alors, $\mu(A \setminus A') = 0$, c'est-à-dire, il existe une infinité de $n \in \mathbb{N}$ tels que $T^n x \in A$ pour presque tout $x \in A$.

Démonstration. Posons $U_N := \bigcup_{n \geq N} T^{-n}A$ (où $T^{-n}A$ désigne l'image réciproque de l'ensemble A par T^n), l'ensemble des $x \in X$ tels que $T^n x \in A$ pour un certain $n \geq N$, et $U := \bigcap_{N \in \mathbb{N}} U_N$, l'ensemble des $x \in X$ tels que $T^n x \in A$ pour une infinité d'entiers n . On remarque que U est mesurable, car U est une intersection dénombrable d'unions dénombrables d'ensembles mesurables. De plus, par la définition de A' , on a $A' = A \cap U$. Maintenant, considérons la suite $(U_N)_{N \in \mathbb{N}}$, on sait que l'on a l'imbrication d'ensembles mesurables $U_0 \supseteq U_1 \supseteq U_2 \supseteq \dots$. De plus, $\mu(U_{N+1}) = \mu(T^{-1}U_N) = \mu(U_N)$, car U_N est mesurable et

$$T^{-1}U_N = T^{-1}\left(\bigcup_{n \geq N} T^{-n}A\right) = \bigcup_{n \geq N} T^{-(n+1)}A = \bigcup_{n \geq N+1} T^{-n}A = U_{N+1}.$$

Par l'égalité que nous avons, le fait que $\mu(U_0) < \infty$ (μ est une mesure de probabilité) et les propriétés élémentaires d'une mesure, on a $\mu(U) = \lim_{N \rightarrow \infty} \mu(U_N) = \lim_{N \rightarrow \infty} \mu(U_0) = \mu(U_0)$. Donc $\mu(U_0 \setminus U) = 0$, ce qui implique que $\mu(A \setminus U) = 0$, car $A \subset U_0$. Ainsi, $\mu(A \setminus A') = 0$, car $\mu(A \setminus A') = \mu(A \setminus (A \cap U)) = \mu(A \setminus U) = 0$. \square

On présente par la suite un résultat qui découle immédiatement du théorème de récurrence de Poincaré. Plus précisément, si on considère la démonstration du théorème de récurrence de Poincaré, le prochain résultat découle du fait que $U \subset U_1$.

Corollaire 2.16. [Conséquence du théorème de récurrence de Poincaré] *Soit (X, \mathcal{B}, μ, T) un système préservant la mesure, où μ est une mesure finie. Alors, pour tout $A \in \mathcal{B}$, $\mu\left(A \setminus \bigcup_{n=1}^{\infty} T^{-n}A\right) = 0$.*

Soient X un ensemble, E un sous-ensemble de X , et $T : X \rightarrow X$ une application. On dit que E est un ensemble **T -invariant** si $T^{-1}E = E$. De plus, considérons (X, \mathcal{B}, μ, T) un système préservant la mesure, on dit que E est un ensemble **presque T -invariant** si $\mu(E \Delta T^{-1}E) = 0$, où $E \Delta T^{-1}E := (E \setminus T^{-1}E) \cup (T^{-1}E \setminus E)$ désigne la différence symétrique des ensembles E et $T^{-1}E$. On peut maintenant entrer dans le vif du sujet et donner la définition de l'ergodicité.

Définition 2.17. Soit (X, \mathcal{B}, μ, T) un système préservant la mesure. On dit que T est **ergodique** si tous les ensembles presque T -invariants de (X, \mathcal{B}, μ, T) sont de mesure nulle ou 1.

En fait, ce n'est pas la transformation T qui est ergodique, mais le système au complet qui est ergodique. Dans une définition plus standard de l'ergodicité, on remplacerait "tous les ensembles presque T -invariants" par "tous les ensembles T -invariants". On pourrait être porté à penser que les deux définitions ne sont pas équivalentes, puisque qu'une condition sur tous les ensembles presque T -invariants est plus restrictive qu'une condition sur tous les ensembles T -invariants. En fait, les deux définitions sont équivalentes. On retrouve le même résultat (proposition 2.14) dans [EW11].

Proposition 2.18. *Soit (X, \mathcal{B}, μ, T) un système préservant la mesure. Alors, tous les ensembles presque T -invariants sont de mesure nulle ou 1 si et seulement si tous les ensembles T -invariants sont de mesure nulle ou 1.*

Démonstration. Supposons que tous les ensembles presque T -invariants sont de mesure nulle ou 1. Alors, tous les ensembles T -invariants sont de mesure nulle ou 1, car les ensembles T -invariants sont en particulier des ensembles presque T -invariants.

Au contraire, supposons que tous les ensembles T -invariants sont de mesure nulle ou 1. Soit E un ensemble presque T -invariant, alors $\mu(E \Delta T^{-1}E) = 0$. Posons $E' := \bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} T^{-n}E$. Alors, E' est un ensemble T -invariant, car $T^{-1}E' = \bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} T^{-(n+1)}E = \bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} T^{-n}E$. On peut également montrer, de manière assez directe, que

$$E \Delta E' \subset \bigcup_{n \geq 1} (E \Delta T^{-n}E).$$

De plus, comme pour tous ensembles A, B, C , on a $A\Delta B \subset (A\Delta C) \cup (B\Delta C)$, un simple calcul montre que

$$\mu(E\Delta T^{-n}(E)) \leq \sum_{i=1}^n \mu(T^{-(i-1)}E\Delta T^{-i}E) = n\mu(E\Delta T^{-1}E) = 0.$$

Donc $\mu(E\Delta E') = 0$, ce qui implique que $\mu(E) = \mu(E')$. Ainsi, tous les ensembles presque T -invariants sont de mesure nulle ou 1. \square

À l'aide du théorème de récurrence de Poincaré, on peut reformuler notre définition de l'ergodicité et obtenir une définition équivalente.

Lemme 2.19. *Un système préservant la mesure (X, \mathcal{B}, μ, T) est ergodique si et seulement si pour tout $A \in \mathcal{B}$ avec $\mu(A) > 0$, il existe $n \in \mathbb{N}$ tel que $T^n x \in A$, pour presque tout $x \in X$.*

Démonstration. Pour la suffisance. Supposons que (X, \mathcal{B}, μ, T) est ergodique. On prétend que l'ensemble $A_+ := \bigcup_{n=1}^{\infty} T^{-n}A$ des $x \in X$ pour lesquels il existe un certain $n \in \mathbb{N}$ tel que $T^n x \in A$ est presque T -invariant. On sait que $T^{-1}A_+ \subset A_+$ et que la différence symétrique de ces deux ensembles est $A_+ \setminus T^{-1}A_+$. De plus,

$$A_+ \setminus T^{-1}A_+ = T^{-1}A \setminus (T^{-2}A \cup T^{-3}A \cup T^{-4}A \cup \dots),$$

qui est de mesure nulle par le corollaire du théorème de récurrence de Poincaré et donc A_+ est presque T -invariant. Puisque T est ergodique, on a $\mu(A_+) = 0$ ou 1. Cependant, $\mu(A_+) > 0$, car $\mu(A) > 0$ et $\mu(A \setminus A_+) = 0$ (corollaire du théorème de récurrence de Poincaré). Ainsi, $\mu(A_+) = 1$, autrement dit presque tout les points de X sont dans A_+ .

Pour la nécessité, soit $E \in \mathcal{B}$ un ensemble presque T -invariant et E_+ défini comme A_+ . Par le théorème de récurrence de Poincaré $\mu(E) = \mu(E_+ \cap E)$. Sans perte de généralité, supposons que $\mu(E) > 0$. Alors, par hypothèse, $\mu(E_+) = 1$. Ainsi, $\mu(E) \in \{0, 1\}$, ce qui implique que le système est ergodique. \square

3 Théorème ergodique

Les sections de ce chapitre concernant directement la théorie ergodique sont inspirées des chapitres 3,4 et 5 de [Gre15].

3.1 Un peu plus de théorie de la mesure et les espaces de Hilbert

Afin d'approfondir notre étude de la théorie ergodique, il nous faut plus de connaissances sur la théorie de la mesure et les espaces de Hilbert. Ces deux sujets nécessitent une certaine connaissance des espaces métriques complets [Mor20]. Commençons par un rappel sur ces espaces et leurs propriétés.

Définition 3.1. Une suite $\{x_n\}_{n \in \mathbb{N}}$ de points dans un espace métrique (X, d) est appelée une **suite de Cauchy** si pour tout $\varepsilon > 0$, il existe un entier positif n_0 tel que pour tout $m \geq n_0$ et $n \geq n_0$, $d(x_m, x_n) < \varepsilon$.

Toute suite convergente est une suite de Cauchy, mais la réciproque n'est généralement pas vraie, ce qui nous mène à la prochaine définition.

Définition 3.2. Un espace métrique (X,d) est dit **complet** si toute suite de Cauchy dans (X,d) converge.

Considérons $(X, \|\cdot\|_X)$ un espace normé, alors (X, d_X) est un espace métrique pour $d_X(x,y) = \|x - y\|_X$ pour tout $x, y \in X$. On termine ce rappel sur les espaces métriques complets avec la définition d'un sous-espace fermé et une de ses propriétés en lien avec les espaces métriques complets.

Définition 3.3. Soit (X,d) un espace métrique. On dit de Y un sous-ensemble de X qu'il est un **sous-espace fermé** de (X,d) si et seulement si toute suite de points dans Y convergente dans X converge vers un point de Y .

Le lecteur prendra note qu'il ne s'agit pas d'une définition standard des sous-espaces fermés, mais plutôt d'une définition équivalente.

Lemme 3.4. Soient (X,d_1) un espace métrique complet, Y un sous-ensemble de X et d_2 la métrique induite sur Y par d_1 . Alors, (Y,d_2) est un sous-espace fermé de (X,d_1) si et seulement si (Y,d_2) est un espace métrique complet.

Maintenant que nous nous sommes dotés de ces définitions et résultats sur les espaces métriques complets, nous pouvons passer au complément sur la théorie de la mesure [Gal09]. On dit de f , une fonction, qu'elle est **simple**, si $f = \sum_{i=1}^n c_i 1_{E_i}$, où les c_i sont des constantes et les E_i sont des ensembles mesurables. Le lecteur prendra soin de noter que la somme est finie. On commence par définir l'intégrale d'une fonction simple et positive. Soient (X, \mathcal{B}, μ) , un espace mesuré et $f : X \rightarrow [0, \infty)$, une fonction simple et positive. Supposons que $f = \sum_{i=1}^n c_i 1_{E_i}$, où les c_i sont des constantes et les E_i sont des ensembles mesurables. Alors, on appelle intégrale de f (par rapport à la mesure μ), la somme $\sum_{i=1}^n c_i \mu(E_i)$ que l'on note $\int_X f d\mu$. Par la suite, on étend la définition d'intégrale aux fonctions mesurables et positives. Soit $f : X \rightarrow [0, \infty)$, une fonction mesurable et positive. On définit l'intégrale de f sur X par :

$$\int_X f d\mu := \sup \left\{ \int_X \Phi d\mu \mid \Phi \in \varepsilon_+(X) \text{ et } \Phi \leq f \text{ p.p.} \right\},$$

où $\varepsilon_+(X)$ est l'ensemble des fonctions simples et positives. De plus, on dit que f est intégrable si $\int_X f d\mu < \infty$. Il suffit maintenant de généraliser le concept aux fonctions quelconques. Soit $f : X \rightarrow \mathbb{R}$, une fonction mesurable. Alors, il existe $f^+ := \max(0, f)$ et $f^- := \min(0, f)$, deux fonctions positives et mesurables telles que $f = f^+ - f^-$ et on dit que f est intégrable si f^+ et f^- le sont. Dans ce cas, l'intégrale de f est $\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu$. La généralisation aux fonctions complexes se fait de façon analogue à la définition précédente. On considère f une fonction complexe, on a $f = \Re(f) + i\Im(f)$ et on dit que f est intégrable si $\Re(f)$ et $\Im(f)$ le sont. Nous ne présenterons pas toutes les propriétés de l'intégrale d'une fonction par rapport à une mesure, mais il est intéressant de

noter que cette intégrale partage certaines propriétés (par exemple, la linéarité) avec l'intégrale de Riemann.

On poursuit, en donnant la définition des espaces \mathcal{L}^p et de leurs normes p.

Définition 3.5. Soient $p \in [1, \infty)$, (X, \mathcal{B}, μ) un espace mesuré et \mathcal{M} l'ensemble des classes d'applications mesurables de X dans \mathbb{R} (respectivement \mathbb{C}) pour la relation d'égalité p.p.. On définit $L^p(X, \mathcal{B}, \mu)$ comme l'ensemble des fonctions $f \in \mathcal{M}$ vérifiant $\int_X |f|^p d\mu < +\infty$ (lorsqu'il n'y a pas de risque de confusion, on note $L^p(X)$). De plus, on note $L^\infty(X, \mathcal{B}, \mu)$, l'ensemble des classes de fonctions $f \in \mathcal{M}$ essentiellement bornées, c'est-à-dire, les fonctions $f \in \mathcal{M}$ telle qu'il existe $M > 0$ avec $|f| \leq M$ p.p. (lorsqu'il n'y a pas de risque de confusion, on note $L^\infty(X)$). Pour $f \in L^p(X)$, on appelle **la norme** \mathcal{L}^p de f et on note $\|f\|_p$, le nombre réel $\|f\|_p := (\int_X |f|^p d\mu)^{\frac{1}{p}}$. De façon analogue, pour $f \in \mathcal{L}^\infty$, on appelle **la norme** \mathcal{L}^∞ de f et on note $\|f\|_\infty$, le nombre réel $\|f\|_\infty := \inf\{M \in [0, \infty) \mid |f| \leq M \text{ p.p.}\}$. On définit pour tout $p \in \mathbb{N}^* \cup \infty$, **l'espace** $\mathcal{L}^p(X, \mathcal{B}, \mu)$ comme l'ensemble $L^p(X, \mathcal{B}, \mu)$ munie de la norme $\|\cdot\|_p$ (lorsqu'il n'y a pas de risque de confusion, on note $\mathcal{L}^p(X)$).

Pour conclure cette partie sur la théorie de la mesure nous présenterons trois résultats sur les espaces \mathcal{L}^p .

Théorème 3.6. *Les fonctions simples sont denses dans \mathcal{L}^1 .*

Théorème 3.7. [Théorème de Riesz-Fischer] *Pour tout $p \in [1, \infty]$, les espaces \mathcal{L}^p munis de leur norme $\|\cdot\|_p$ sont des espaces vectoriels complets.*

Le résultat nous dit que toute suite de Cauchy dans \mathcal{L}^p converge en norme p. Il faut interpréter ce résultat de la bonne façon. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions dans \mathcal{L}^p (où $p \neq \infty$), supposons que cette suite converge vers f en norme p. Le fait que la suite converge en norme p n'implique pas que f_n converge simplement p.p. vers f . En effet, il est très important de comprendre qu'il s'agit de deux convergences différentes et que l'une n'implique pas nécessairement l'autre.

Remarque 3.8. Un résultat de la théorie de la mesure nous dit que toute suite qui converge en norme \mathcal{L}^p admet une sous-suite convergente p.p..

Théorème 3.9. *Soit (X, \mathcal{B}, μ) un espace mesuré, de mesure finie. Alors, pour tout $p, q \in [1, \infty]$ tels que $p < q$, $\mathcal{L}^q \subset \mathcal{L}^p$.*

Puisque dans le cadre de cet article nous travaillons seulement avec des mesures de type finie, le dernier théorème est donc toujours valide.

Par la suite, nous allons avoir besoin de quelques notions sur les espaces de Hilbert. Encore une fois, pour plus de détails concernant les espaces de Hilbert, le lecteur est invité à consulter [Mus14], c'est aussi notre référence pour ce qui concerne les espaces de Hilbert. Tout d'abord, de façon informelle, les espaces de Hilbert sont des espaces vectoriels réels (respectivement complexes), dans lesquels nous pouvons appliquer nos techniques d'analyse. La première étape pour définir les espaces de Hilbert, était de définir les espaces métriques complets et maintenant, il nous reste juste à étendre le concept de produit scalaire aux espaces vectoriels complexes.

Définition 3.10. Soit E un espace vectoriel complexe muni d'une application $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{C}$ telle que pour $x, y, z \in E$ et $a, b \in \mathbb{C}$, on a :

- i) $\langle ax + y, z \rangle = \bar{a}\langle x, z \rangle + \langle y, z \rangle$,
- ii) $\langle x, by + z \rangle = b\langle x, y \rangle + \langle x, z \rangle$,
- iii) pour tout $x, y \in E$, $\langle x, y \rangle = \overline{\langle y, x \rangle}$,
- iv) pour tout $x \in E$, $\langle x, x \rangle \geq 0$, et
- v) pour tout $x \in E$, $\langle x, x \rangle = 0$ si et seulement si $x = 0$.

Alors, on dit que $\langle \cdot, \cdot \rangle$ est un **produit scalaire hermitien**.

Définition 3.11. Un **espace de Hilbert** H est un espace vectoriel complexe (respectivement réel) muni d'un produit scalaire hermitien (respectivement euclidien) $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{C}$ (respectivement \mathbb{R}) avec la propriété que H est complet par rapport à la norme $\|x\| = \sqrt{\langle x, x \rangle}$ induite par ce produit scalaire hermitien (respectivement euclidien).

Notons que, d'après le lemme 3.4, un sous-ensemble d'un espace de Hilbert est un espace de Hilbert si et seulement s'il est fermé. Le premier exemple d'un tel espace est l'espace \mathcal{L}^2 . Soit (X, \mathcal{B}, μ) un espace mesuré, alors $\mathcal{L}^2(X)$ est un espace de Hilbert pour le produit scalaire hermitien $\langle \cdot, \cdot \rangle : \mathcal{L}^2 \times \mathcal{L}^2 \rightarrow \mathbb{C}$ défini par $\langle f, g \rangle = \int_X \bar{f}g d\mu$ pour tout $f, g \in \mathcal{L}^2(X)$. Pour terminer cette sous-section, nous allons regarder quelques définitions et propriétés concernant les opérateurs et les fonctionnelles.

Définition 3.12. Soit H un espace de Hilbert. Un **opérateur linéaire continu** sur H est une transformation linéaire continue $T : H \rightarrow H$. C'est-à-dire, une application de $T : H \rightarrow H$ qui respecte les conditions suivantes :

- i) Pour tout $x, y \in H$, $T(x + y) = Tx + Ty$.
- ii) Pour tout $x \in H$ et $\lambda \in \mathbb{F}$, $T(\lambda x) = \lambda Tx$
- iii) Pour toute suite de Cauchy $(x_n)_{n \in \mathbb{N}}$ de H , $T(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} Tx_n$.

Une **fonctionnelle** sur H est une application linéaire continue $\phi : H \rightarrow \mathbb{F}$. L'ensemble des fonctionnelles est dénoté H^* et on dit que H^* est l'**espace dual** de H .

Définition 3.13. Soit H un espace de Hilbert, et notons $\|\cdot\|_H$ la norme induite par son produit scalaire. Un **opérateur linéaire borné** sur H est une application linéaire $\phi : H \rightarrow H$ telle qu'il existe une constante $C > 0$ vérifiant $\|\phi x\|_H \leq C\|x\|_H$ pour tout $x \in H$. L'infimum de toutes les constantes satisfaisant cette inégalité est appelé la **norme de l'opérateur** ϕ , et se note $\|\phi\|_{\text{op}}$.

En fait, dans les espaces normés, il n'y a pas de différence entre les applications linéaires bornées et les applications linéaires continues. Comme les espaces de Hilbert sont normés, les deux applications sont équivalentes. La prochaine proposition découle de la complétude des espaces de Hilbert et des propriétés des sous-espaces vectoriels fermés.

Proposition 3.14. *Soient H un espace de Hilbert muni d'un produit scalaire hermitien $\langle \cdot, \cdot \rangle_H$ et $V \subset H$ un sous-espace vectoriel fermé. On définit le **complément orthogonal** de V par $V^\perp := \{x \in H \mid \langle x, v \rangle_H = 0, \forall v \in V\}$. Alors, H se décompose comme $H = V \oplus V^\perp$, et il existe un unique opérateur linéaire $\pi : H \rightarrow V$ vérifiant :*

- i) $\pi|_V$ est l'identité sur V ;
- ii) $\pi|_{V^\perp}$ est l'application nulle ;
- iii) pour tout $x \in H$, $x - \pi x \in V^\perp$.

Le prochain résultat qui nous intéresse est le théorème de représentation de Riesz, ce dernier est sans doute l'un des théorèmes les plus importants concernant les espaces Hilbert. En effet, il établit un isomorphisme entre un espace de Hilbert H et son espace dual H^* muni de la norme $\|x^*\|_{H^*} = \|x\|_H$.

Théorème 3.15. [Théorème de représentation de Riesz] *Soit H un espace de Hilbert. Pour toute fonctionnelle $\phi \in H^*$, il existe un unique $x_\phi \in H$ tel que $\phi(x) = \langle x_\phi, x \rangle$ pour tout $x \in H$ et $\|\phi\|_{\text{op}} = \|x_\phi\|_H$.*

En fait, dans le cadre de cet article, le théorème de représentation de Riesz est surtout important pour justifier l'existence et l'unicité de l'adjoint d'un opérateur linéaire borné.

Définition 3.16. Soient H un espace de Hilbert et $\phi : H \rightarrow H$ un opérateur linéaire borné. On dit de ϕ^* qu'il est l'**adjoint** de ϕ , si $\phi^* : H \rightarrow H$ est un opérateur linéaire borné tel que $\langle \phi x, y \rangle = \langle x, \phi^* y \rangle$ pour tout $x, y \in H$. Dans ce cas, on a $\|\phi\|_{\text{op}} = \|\phi^*\|_{\text{op}}$.

3.2 Exemple de système ergodique

Nous n'avons toujours pas donné d'exemple de système ergodique, la raison principale étant le fait qu'il est assez fastidieux de montrer qu'un système est ergodique. Cependant, nous sommes à quelques résultats près de pouvoir montrer que le système de l'exemple 2.14 est ergodique. Le lecteur notera toutefois que les résultats présentés dans cette sous-section restent pertinents pour le reste de l'article. On commence par donner deux résultats sur les fonctions de \mathcal{L}^1 .

Lemme 3.17. *Supposons que (X, \mathcal{B}, μ, T) est un système préservant la mesure et que f est une fonction de $\mathcal{L}^1(X)$. Alors, $\int_X f \, d\mu = \int_X f \circ T \, d\mu$. (C'est-à-dire, $\int_X f(x) \, d\mu = \int_X f(Tx) \, d\mu$).*

Démonstration. On commence par montrer que le résultat est vrai pour les fonctions simples. Puisque l'intégrale est linéaire, on peut supposer sans perte de généralité que $f = 1_E$, où E est mesurable. Alors, $\int_X f \, d\mu = \int_X 1_E \, d\mu = \mu(E)$ et

$$\int_X f \circ T \, d\mu = \int_X 1_E \circ T \, d\mu = \int_X 1_{T^{-1}E} \, d\mu = \mu(T^{-1}E).$$

Donc comme T est une application préservant la mesure, on a $\int_X f \, d\mu = \int_X (f \circ T) \, d\mu$.

Considérons maintenant $f \in \mathcal{L}^1(X)$ quelconque et fixons $\varepsilon > 0$. Comme l'espace des fonctions simples est dense dans $\mathcal{L}^1(X)$, il existe une suite $(g_n)_{n \in \mathbb{N}}$ de fonctions simples telle que $\lim_{n \rightarrow \infty} \|f - g_n\|_1 = 0$. De plus, d'après la remarque 3.8, il existe une sous-suite $(g_{n_k})_{k \in \mathbb{N}}$ qui converge presque partout vers f . Comme T est mesurable, on en déduit que $(g_{n_k} \circ T)_{k \in \mathbb{N}}$ converge presque partout vers $f \circ T$. Par ailleurs, pour tous $n, m \in \mathbb{N}$, on a $\|g_n \circ T - g_m \circ T\|_1 = \|g_n - g_m\|_1$, puisque ce résultat a déjà été démontré pour les fonctions simples. Alors, $(g_n \circ T)_{n \in \mathbb{N}}$ est une suite de Cauchy dans $\mathcal{L}^1(X)$ et, comme $\mathcal{L}^1(X)$ est complet, elle converge en norme L^1 vers une certaine limite $h \in \mathcal{L}^1(X)$. Cependant, la sous-suite $(g_{n_k} \circ T)$ converge presque partout vers $f \circ T$. Par l'unicité de la limite presque partout, on a donc $h = f \circ T$ (dans \mathcal{L}^1). Donc la suite $(g_n \circ T)$ converge en norme \mathcal{L}^1 vers $f \circ T$. Par conséquent, il existe $n \in \mathbb{N}$ tel que $\|f - g_n\|_1 \leq \frac{\varepsilon}{2}$ et $\|f \circ T - g_n \circ T\|_1 \leq \frac{\varepsilon}{2}$. Cependant, par le fait que le résultat est vrai pour les fonctions simples, on a :

$$\begin{aligned} \left| \int_X f \, d\mu - \int_X f \circ T \, d\mu \right| &\leq \left| \int_X f - g_n \, d\mu \right| + \left| \int_X f \circ T - g_n \circ T \, d\mu \right| \\ &\leq \|f - g_n\|_1 + \|f \circ T - g_n \circ T\|_1 \\ &\leq \varepsilon \end{aligned}$$

Ainsi, comme ε était arbitraire, on a bien que $\int_X f \, d\mu = \int_X f \circ T \, d\mu$. \square

Par la suite, dans le but de donner notre premier exemple de système ergodique, on doit montrer que toutes les fonctions indicatrices d'un ensemble mesurable de $\mathcal{B}(\mathbb{R}/\mathbb{Z})$ peuvent être approximées par des fonctions continues de \mathbb{R}/\mathbb{Z} dans $[0,1]$ avec une précision arbitraire en norme \mathcal{L}^1 .

Lemme 3.18. *Soit $E \in \mathcal{B}(\mathbb{R}/\mathbb{Z})$ un ensemble mesurable. Alors, pour tout $\varepsilon > 0$, il existe une fonction continue $f : \mathbb{R}/\mathbb{Z} \rightarrow [0,1]$ telle que $\|f - 1_E\|_1 \leq \varepsilon$.*

Démonstration. Soient $E \in \mathcal{B}(\mathbb{R}/\mathbb{Z})$, un ensemble mesurable et $\varepsilon > 0$. Par le fait que la mesure de Lebesgue est extérieurement régulière, il existe U un ouvert de \mathbb{R}/\mathbb{Z} telle que $E \subset U$ et $\lambda(U \Delta E) \leq \frac{\varepsilon}{3}$. De plus,

$$\lambda(U \Delta E) = \lambda(E \setminus U) + \lambda(U \setminus E) = \lambda(U \setminus E).$$

On peut écrire U comme une union dénombrable d'intervalles ouvertes et disjoints de \mathbb{R}/\mathbb{Z} . Posons $U = \bigcup_{i=1}^{\infty} B_i$, où la collection $\{B_1, B_2, B_3, \dots\}$ est une

collection d'intervalles ouvertes et disjointes. Soit $U_J = \bigcup_{i=1}^J B_i$, une suite d'ensembles emboîtés telle que $U_1 \subset U_2 \subset U_3 \subset \dots$. Par les propriétés élémentaires des mesures on a :

$$\lambda\left(\bigcup_{J=1}^{\infty} U_J\right) = \lim_{J \rightarrow \infty} \lambda(U_J) = \lambda(U).$$

On a donc pour J assez grand, $\lambda(U \setminus U_J) \leq \frac{\varepsilon}{3}$. Cependant,

$$\begin{aligned} \|1_E - 1_{U_J}\|_1 &= \int_{\mathbb{R}/\mathbb{Z}} |1_E - 1_{U_J}| d\lambda \\ &= \int_{\mathbb{R}/\mathbb{Z}} |1_{E \setminus U_J} - 1_{U_J \setminus E}| d\lambda \\ &= \int_{\mathbb{R}/\mathbb{Z}} |1_{E \setminus U_J \cup U_J \setminus E}| d\lambda \\ &= \lambda(E \setminus U_J \cup U_J \setminus E) \\ &= \lambda(E \setminus U_J) + \lambda(U_J \setminus E) \\ &\leq \frac{2\varepsilon}{3}. \end{aligned}$$

Sans perte de généralité supposons que $B_i = (a_i, b_i)$, où $a_i, b_i \in \mathbb{R}/\mathbb{Z}$ et $a_i < b_i$. Posons,

$$X_{B_i, \delta}(x) = \begin{cases} \frac{x-a_i}{\delta} & \text{si } x \in (a_i, a_i + \delta) \\ 1 & \text{si } x \in [a_i + \delta, b_i - \delta] \\ \frac{b_i-x}{\delta} & \text{si } x \in (b_i - \delta, b_i) \\ 0 & \text{sinon} \end{cases},$$

pour δ petit. On remarque que $X_{B_i, \delta}$ est une fonction continue et que $\sum_{i=1}^J X_{B_i, \delta}$ est aussi une fonction continue, car les intervalles ouverts B_1, B_2, B_3, \dots sont deux à deux disjoints. De plus, un simple calcul montre que $\|X_{B_i, \varepsilon} - 1_{B_i}\|_1 \leq \varepsilon$. En prenant, $f = \sum_{i=1}^J X_{B_i, \frac{\varepsilon}{3J}}$, il est possible, en utilisant l'inégalité du triangle, de montrer que $\|f - 1_{U_J}\|_1 \leq \frac{\varepsilon}{3}$. Ainsi, comme $\|f - 1_E\|_1 \leq \|f - 1_{U_J}\|_1 + \|1_{U_J} - 1_E\|_1 \leq \varepsilon$, la démonstration se termine. \square

Reprenons l'exemple des rotations irrationnelles du cercle et montrons qu'il s'agit d'un système ergodique pour la mesure de Lebesgue. Nous avons déjà établi qu'il s'agit d'un système préservant la mesure; il reste donc à démontrer que tout ensemble presque T -invariant est de mesure nulle ou de mesure 1.

Exemple 3.19 ([Gre15], Proposition 3.1). Soient $X = \mathbb{R}/\mathbb{Z}$, \mathcal{B} la tribu de Borel de X , $T : X \rightarrow X$ la rotation $Tx = x + \alpha \pmod{1}$ où α est irrationnel et λ la mesure de Lebesgue. Alors, $(X, \mathcal{B}, \lambda, T)$ est un système ergodique.

Démonstration. Supposons que E est un ensemble presque T -invariant, notons que pour tout $n \in \mathbb{N}^*$, $1_E(T^n) = 1_E$ p.p.. Soit $\varepsilon > 0$. Le lemme 3.18 nous dit qu'il est possible de trouver une fonction continue f telle que $\|f - 1_E\|_1 \leq \varepsilon$. Du lemme 3.17 et du fait que T est une application préservant la mesure, on a

$\|(f - 1_E) \circ T^n\|_1 \leq \varepsilon$. De plus, comme $1_E = 1_E \circ T^n$ p.p., $\|1_E - f \circ T^n\|_1 \leq \varepsilon$. Par l'inégalité du triangle, il suit que $\|f - f \circ T^n\|_1 \leq 2\varepsilon$. De façon explicite, nous avons que

$$\int_0^1 |f(x) - f(x + n\alpha)| d\lambda(x) \leq 2\varepsilon.$$

Dans l'introduction, nous avons montré que les orbites des rotations irrationnelles sur le cercle étaient denses. Donc pour tout $t \in \mathbb{R}/\mathbb{Z}$, il existe une suite (n_i) d'entiers telle que $n_i\alpha \pmod{1} \rightarrow t$ lorsque $i \rightarrow \infty$. Il suit donc de la continuité de f que

$$\int_0^1 |f(x) - f(x + t)| d\lambda(x) \leq 2\varepsilon$$

pour tout $t \in \mathbb{R}/\mathbb{Z}$. Considérons l'application mesurable $\int_0^1 f : X \rightarrow \mathbb{R}$ définie par $\int_0^1 f(x) := \int_0^1 f(x + t) d\lambda(t)$. Alors,

$$\begin{aligned} \|f - \int_0^1 f\|_1 &= \int_0^1 |f(x) - \int_0^1 f(x + t) d\lambda(t)| d\lambda(x) \\ &\leq \int_0^1 \int_0^1 |f(x) - f(x + t)| d\lambda(t) d\lambda(x) \\ &\leq 2\varepsilon, \end{aligned}$$

où la dernière inégalité vient du théorème de Fubini adapté aux espaces mesurés. Donc par l'inégalité du triangle, $\|1_E - \int_0^1 f\|_1 \leq 3\varepsilon$. Ceci implique que

$$|\lambda(E) - \int_0^1 f| \leq \int_0^1 |1_E - \int_0^1 f| d\lambda \leq 3\varepsilon$$

et donc encore une fois par l'inégalité du triangle $\|1_E - \lambda(E)\|_1 \leq 6\varepsilon$. Cependant, $\varepsilon > 0$ est arbitraire. Donc $\|1_E - \lambda(E)\|_1 = 0$, ce qui implique que $1_E = \lambda(E)$ p.p.. Ainsi, il suit que $\lambda(E) = 0$ ou 1 . \square

3.3 Théorème ergodique moyen de von Neumann

Dans cette sous-section, nous allons caractériser l'espace I_T des fonctions presque T -invariantes et établir notre premier résultat ergodique. Considérons un système préservant la mesure (X, \mathcal{B}, μ, T) . Pour toute fonction $f \in \mathcal{L}^2(X)$, nous noterons $U_T f(x) := f(Tx)$. D'après le lemme 3.17 et le théorème 3.9, l'application $U_T : \mathcal{L}^2(X) \rightarrow \mathcal{L}^2(X)$ est une isométrie, c'est-à-dire que $\|U_T f\|_2 = \|f\|_2$ pour tout $f \in \mathcal{L}^2(X)$. En particulier, U_T est un opérateur linéaire borné sur $\mathcal{L}^2(X)$. L'espace des fonctions T -invariantes du système (X, \mathcal{B}, μ, T) est le sous-espace $I_T := \{f \in \mathcal{L}^2(X) \mid f = U_T f\}$. Le lecteur notera que I_T est un sous-espace vectoriel fermé de $\mathcal{L}^2(X)$.

Lemme 3.20. *Une transformation T est ergodique si et seulement si l'espace I_T des fonctions p.p. T -invariantes ne contient que des fonctions p.p. constantes.*

Démonstration. Pour la suffisance, supposons que I_T ne contient des fonctions constantes p.p.. Soit $E \in \mathcal{B}$, un ensemble mesurable et presque T -invariant, alors $1_E(x) = 1_E(Tx) = U_T 1_E(x)$, ce qui implique que $1_E \in I_T$ et par hypothèse 1_E est p.p. constante. Ainsi, on a bien que $\mu(E) \in \{0,1\}$.

Pour la nécessité, supposons que T est ergodique. Soit $f \in I_T$, c'est-à-dire $f \in \mathcal{L}^2(X)$ tel que $f(x) = f(Tx)$ p.p.. Notons que f est mesurable, puisqu'elle appartient à $\mathcal{L}^2(X)$. Supposons de plus que f est une fonction de X dans \mathbb{R} . Comme f est mesurable et presque T -invariante, pour tout ensemble mesurable $A \subset \mathbb{R}$, l'ensemble $\{x \in X \mid f(x) \in A\}$ est presque T -invariant et mesurable. De plus, comme T est ergodique, on a $\mu(\{x \in X \mid f(x) \in A\}) \in \{0,1\}$.

Posons

$$c := \inf\{b \in \mathbb{R} \mid \mu(\{x \mid f(x) \geq b\}) = 0\}.$$

Supposons que $a < c$. Alors, par définition de c , on a $\mu(\{x \in X \mid f(x) > a\}) > 0$. Comme l'ensemble $\{x \in X \mid f(x) > a\}$ est presque T -invariant et mesurable, l'ergodicité de T implique que $\mu(\{x \in X \mid f(x) > a\}^c) = \mu(\{x \in X \mid f(x) \leq a\}) = 0$. Ainsi, $f > a$ p.p. pour tout $a < c$, ce qui entraîne que $f \geq c$ p.p.

D'un autre côté, supposons que $a > c$. Alors, par définition de c , on a $\mu(\{x \in X \mid f(x) > a\}) = 0$. Ainsi, $f \leq a$ p.p., ce qui implique que $f \leq c$ p.p., puisque $a > c$ est arbitraire. On en déduit que $f = c$ p.p.

Montrons à présent que le résultat reste valable lorsque f est une fonction complexe. Écrivons $f = \Re(f) + i\Im(f)$. Si $f(x) = f(Tx)$, alors $\Re(f)(x) = \Re(f)(Tx)$ et $\Im(f)(x) = \Im(f)(Tx)$. On peut donc appliquer l'argument précédent séparément à $\Re(f)$ et à $\Im(f)$. \square

On sait que \mathcal{L}^2 est un espace de Hilbert, que I_T est un sous-espace fermé et que U_T est un opérateur linéaire borné. Le lemme suivant découle alors de la proposition 3.14 ainsi que de l'existence et de l'unicité de l'adjoint d'un opérateur linéaire borné.

Lemme 3.21. *Soit (X, \mathcal{B}, μ, T) un système préservant la mesure et $U_T : \mathcal{L}^2(X) \rightarrow \mathcal{L}^2(X)$ l'opérateur défini par $U_T f = f \circ T$. Alors,*

i) si $f \in I_T$, on a $U_T^ f = f$, où U_T^* désigne l'adjoint de U_T ;*

ii) si T est ergodique, pour tout $f \in \mathcal{L}^2(X)$,

$$\pi_T(f) = \int_X f \, d\mu,$$

où $\pi_T : \mathcal{L}^2(X) \rightarrow I_T$ est la projection orthogonale sur I_T donnée par la proposition 3.14.

Démonstration. Pour *i)* on sait que U_T est un opérateur linéaire borné et donc par le théorème de représentation de Riesz, il existe un unique opérateur U_T^* (l'adjoint de U_T) tel que pour tout $f, g \in \mathcal{L}^2(X)$, $\langle U_T f, g \rangle = \langle f, U_T^* g \rangle$. Cependant,

pour tout $f, g \in \mathcal{L}^2(X)$, on a $\langle U_T g, U_T f \rangle = \langle g, f \rangle$, car U_T est une isométrie. Considérons $f \in I_T$. Alors, $f = U_T f$ et pour tout $g \in \mathcal{L}^2(X)$ on a :

$$\begin{aligned} \langle g, f - U_T^* f \rangle &= \langle g, f \rangle - \langle g, U_T^* f \rangle \\ &= \langle U_T g, U_T f \rangle - \langle U_T g, f \rangle \\ &= \langle U_T g, U_T f - f \rangle \\ &= 0. \end{aligned}$$

En particulier, si $g = f - U_T^* f$, on a $\|f - U_T^* f\|_2 = 0$, ce qui implique que $U_T^* f = f$.

Pour la partie *ii*), on sait que I_T est un sous-espace vectoriel fermé de $\mathcal{L}^2(X)$. On a donc la décomposition orthogonale $\mathcal{L}^2(X) = I_T \oplus I_T^\perp$. Cependant,

$$I_T^\perp = \{f \in \mathcal{L}^2(X) \mid \langle f, g \rangle = 0 \ \forall g \in I_T\} = \{f \in \mathcal{L}^2(X) \mid g \langle f, 1 \rangle = 0 \ \forall g \in I_T\},$$

puisque, d'après le lemme 3.21 et le fait que T est ergodique, l'espace I_T ne contient que des fonctions p.p. constantes. De plus,

$$\{f \in H \mid g \langle f, 1 \rangle = 0 \ \forall g \in I_T\} = \{f \in H \mid \int_X f d\mu = 0\},$$

car tout $g \in I_T$ est une fonction constante presque partout, et en particulier il existe $g \neq 0$ p.p. dans I_T . Ainsi, si $f \in I_T$, alors $\pi_T(f) = \langle f, 1 \rangle = \int_X f d\mu$. \square

Pour la suite de l'article, on définit l'application $S_N : \mathcal{L}^p(X) \rightarrow \mathcal{L}^p(X)$ par $S_N f(x) = \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$, pour tout $f \in \mathcal{L}^p(X)$ et tout $x \in X$.

Lemme 3.22. *S_N est une application contractante dans $\mathcal{L}^p(X)$. C'est-à-dire, si $f \in \mathcal{L}^p(X)$, alors $\|S_N f\|_p \leq \|f\|_p$.*

Démonstration. Le résultat découle simplement de l'inégalité du triangle et du lemme 3.17. \square

Reprenons l'hypothèse ergodique évoquée dans l'introduction : dans un système à l'équilibre, la valeur moyenne dans l'espace d'une grandeur est égale à la moyenne d'un très grand nombre de mesures effectuées dans le temps. Il existe une analogie évidente entre cette hypothèse et les résultats qui suivent. Considérons comme grandeur physique les fonctions de l'espace \mathcal{L}^2 (ou \mathcal{L}^1 dans le cas du théorème de Birkhoff), et interprétons le temps comme le nombre d'itérations de la transformation T . Soit $f \in \mathcal{L}^2$ (ou \mathcal{L}^1). La valeur moyenne dans l'espace correspond alors à l'intégrale de f par rapport à la mesure du système, tandis que la moyenne d'un grand nombre de mesures prises dans le temps est donnée par $S_N f(x)$, pour presque tout $x \in X$ et pour N grand.

Le résultat suivant est le premier résultat *ergodique* de l'article, au sens où un résultat est dit ergodique lorsqu'il concerne la convergence de $S_N f$ vers $\int_X f d\mu$.

Théorème 3.23. *Soit (X, \mathcal{B}, μ, T) un système préservant la mesure. Alors, pour chaque $f \in \mathcal{L}^2(X)$, on a $\|S_N f - \pi_T(f)\|_2 \rightarrow 0$ quand $N \rightarrow \infty$.*

Démonstration. L'idée clé de la démonstration consiste à identifier le complément orthogonal I_T^\perp de I_T comme étant le sous-espace fermé engendré par les *cocycles*. Pour $g \in \mathcal{L}^2(X)$, on pose $\partial g = g - U_T g$ et on appelle ∂g le **cocycle** associé à g .

Soit M le sous-espace vectoriel fermé de $\mathcal{L}^2(X)$ engendré par les cocycles de $\mathcal{L}^2(X)$. Il est clair que $I_T \subset M^\perp$, car si $f \in I_T$, alors, d'après la partie *i*) du lemme 3.21, on a

$$\langle \partial g, f \rangle = \langle g - U_T g, f \rangle = \langle g, f \rangle - \langle U_T g, f \rangle = \langle g, f \rangle - \langle g, U_T^* f \rangle = 0.$$

Pour montrer que $M^\perp \subset I_T$, considérons $f \in M^\perp$. En particulier, $\langle \partial f, f \rangle = 0$, ce qui implique que

$$\begin{aligned} \|f - U_T f\|_2^2 &= \langle f, f - U_T f \rangle - \langle U_T f, f - U_T f \rangle \\ &= \|U_T f\|_2 - \langle U_T f, f \rangle \\ &= 0. \end{aligned}$$

Donc $f = U_T f$ et $M^\perp \subset I_T$.

Nous savons que I_T est un sous-espace vectoriel fermé et que l'on a la décomposition orthogonale $\mathcal{L}^2(X) = I_T \oplus I_T^\perp$, où I_T^\perp est le sous-espace fermé engendré par les cocycles. En particulier, on peut montrer que pour tout $f \in \mathcal{L}^2(X)$ et tout $\varepsilon > 0$, il existe $g, h \in \mathcal{L}^2(X)$ tels que $\|h\|_2 \leq \varepsilon$ et $f = \pi_T(f) + \partial g + h$. De plus, remarquons que pour tout $f \in \mathcal{L}^2(X)$, $S_N \pi_T(f) = \pi_T(f)$, car $\pi_T(f) \in I_T$.

Considérons $f \in \mathcal{L}^2(X)$ et $\varepsilon > 0$. Alors, il existe g et h comme précédemment et

$$\begin{aligned} \|S_N f - \pi_T(f)\|_2 &= \|S_N f - S_N \pi_T(f)\|_2 \\ &= \|S_N(f - \pi_T f)\|_2 \\ &= \|S_N(\partial g + h)\|_2 \\ &= \|S_N \partial g + S_N h\|_2 \\ &\leq \|S_N \partial g\|_2 + \|S_N h\|_2. \end{aligned}$$

Cependant, puisque S_N est une contraction, on a $\|S_N h\|_2 \leq \|h\|_2 \leq \varepsilon$. Il nous reste donc à vérifier que $\|S_N \partial g\|_2 \leq \frac{2}{N} \|g\|_2$:

$$\begin{aligned} \|S_N \partial g\|_2 &= \left\| \frac{1}{N} \sum_{n=0}^{N-1} (g \circ T^n - g \circ T^{n+1}) \right\|_2 \\ &= \left\| \frac{1}{N} (g - g \circ T + g \circ T - g \circ T^2 + \dots - g \circ T^N) \right\|_2 \\ &= \left\| \frac{1}{N} (g - U_T^N g) \right\|_2 \\ &\leq \frac{1}{N} (\|g\|_2 + \|U_T^N g\|_2) \\ &= \frac{2}{N} \|g\|_2. \end{aligned}$$

Alors, $\|S_N f - \pi_T(f)\|_2 \leq \frac{2}{N}\|g\|_2 + \varepsilon$. Ainsi, comme ε était arbitraire, $\|S_N f - \pi_T(f)\|_2 \rightarrow 0$ lorsque $N \rightarrow \infty$. \square

Le théorème précédent constitue une simple reformulation d'un résultat beaucoup plus connu.

Corollaire 3.24. [Théorème ergodique moyen de von Neumann][[Gre15], Théorème 4.2] *Supposons que (X, \mathcal{B}, μ, T) est un système ergodique. Alors, pour chaque $f \in \mathcal{L}^2(X)$, on a $S_N f \rightarrow \int_X f d\mu$ dans $\mathcal{L}^2(X)$ lorsque $N \rightarrow \infty$.*

Démonstration. Le résultat découle directement de la partie *ii*) du lemme 3.21 et du théorème précédent. \square

3.4 Théorème maximal ergodique

Cette sous-section est dédiée au théorème maximal ergodique. Le lecteur notera qu'il ne s'agit pas vraiment d'un théorème ergodique au sens strict, puisqu'il ne traite pas de la convergence de $S_N f$ vers $\int_X f d\mu$. En effet, dans ce théorème, on suppose seulement que le système préserve la mesure. Ce résultat reste cependant très important en théorie ergodique, notamment pour la démonstration du théorème ergodique de Birkhoff. Le lemme suivant est un cas particulier du lemme de recouvrement de Vitali.

Lemme 3.25. *Soit $\{I_m\}_{m \in S}$ une collection d'intervalles de la forme $I_m = [m, m + l(m)) \subset \mathbb{R}$ où S est un sous-ensemble fini de \mathbb{Z} et $l : \mathbb{Z} \rightarrow \mathbb{N}$ est une application. Alors il existe une sous-collection disjointe $\{I_m\}_{m \in S'}$ telle que l'union des intervalles de la sous-collection a une longueur d'au moins $\frac{1}{2}|S|$.*

Démonstration. Soit $\{I_m \mid m \in S_*\}$ la sous-collection minimale ayant la propriété que $\bigcup_{m \in S_*} I_m = \bigcup_{m \in S} I_m$. On remarque qu'un point y ne peut pas appartenir à trois intervalles de la sous-collection minimale, car si c'était le cas, un des trois intervalles serait contenu dans les deux autres et la sous-collection minimale ne serait pas minimale. En écrivant $S_* = \{m_1 < m_2 < \dots < m_k\}$, on peut construire deux collections $C_1 = I_{m_1} \cup I_{m_3} \cup \dots$ et $C_2 = I_{m_2} \cup I_{m_4} \cup \dots$, qui sont constituées d'intervalles disjoints. Cependant, l'union de ces deux collections est égale à $\bigcup_{m \in S} I_m$ et donc contient S . Ainsi, au moins une de ces deux collections a une longueur d'au moins $\frac{1}{2}|S|$. \square

Théorème 3.26. [Théorème ergodique maximal][[Gre15], Proposition 5.1] *Soit (X, \mathcal{B}, μ, T) un système préservant la mesure. Supposons que $f \in \mathcal{L}^1(X)$ est telle que $\|f\|_1 \leq \varepsilon$. Soit $E \subset X$, l'ensemble des $x \in X$ pour lesquels il existe $N \in \mathbb{N}$ tel que $|S_N f(x)| \geq \delta$. Alors E est mesurable et $\mu(E) \leq \frac{2\varepsilon}{\delta}$.*

Démonstration. Soient $\varepsilon > 0$, $\delta > 0$ et $f \in \mathcal{L}^1(X)$ telle que $\|f\|_1 \leq \varepsilon$. Posons $E(N_0) := \{x \in X \mid |S_N f(x)| \geq \delta \text{ pour un certain } N \leq N_0\}$, l'ensemble des $x \in X$ pour lesquels il existe $N \leq N_0$ tel que $|S_N f(x)| \geq \delta$. Le lecteur notera que la somme de fonctions mesurables est mesurable et que le module d'une fonction mesurable est également mesurable, ce qui implique que les ensembles

de la forme $\{x \in X \mid |f(x)| \geq c\}$ sont mesurables pour tout $c \in \mathbb{R}$ et pour toute fonction mesurable $f : X \rightarrow \mathbb{C}$. Par définition, $E(1) \subseteq E(2) \subseteq \dots \subseteq E$ et $\bigcup_{n \in \mathbb{N}^*} E(n) = E$. Soit $x \in E(N_0)$. Il existe alors $N(x) \leq N_0$ maximal tel que $|S_{N(x)} f(x)| \geq \delta$. En effectuant le changement d'indice $r = n + m$ pour m positif, on obtient :

$$\frac{1}{N(x)} \left| \sum_{r=m}^{N(x)+m-1} f(T^{r-m}x) \right| \geq \delta.$$

Ce qui implique que

$$\frac{1}{N(T^m x)} \left| \sum_{n=m}^{N(T^m x)+m-1} f(T^n x) \right| \geq \delta,$$

pour tout $x \in T^{-m}E(N_0)$.

Considérons $x \in X$ et notons $A(x)$ l'ensemble des $m \in \{0, \dots, N_1 - 1\}$, pour $N_1 \in \mathbb{N}$, tel que $x \in T^{-m}E(N_0)$. Donc pour $x \in X$ et $m \in A(x)$, on a :

$$\frac{1}{N(T^m x)} \left| \sum_{n=0}^{N(T^m x)-1} f(T^n(T^m x)) \right| \geq \delta.$$

Par construction, on sait que $N(T^m x) \in \{0, \dots, N_0\}$.

Maintenant, on a :

$$\int_X |A(x)| d\mu = \int_X \sum_{n=0}^{N_1-1} 1_{\{T^{-n}E(N_0)\}}(x) d\mu = \sum_{n=0}^{N_1-1} \mu(T^{-n}E(N_0)),$$

car $|A(x)| : X \rightarrow \mathbb{N}$ est l'application qui prend $x \in X$ et qui retourne la cardinalité de l'ensemble $A(x)$. De plus,

$$\sum_{n=0}^{N_1-1} \mu(T^{-n}E(N_0)) = \sum_{n=0}^{N_1-1} \mu(E(N_0)),$$

car T est une application préservant la mesure. En particulier, $\int_X |A(x)| d\mu = N_1 \mu(E(N_0))$.

Considérons $x \in X$ et la collection d'intervalles $\{I_m\}_{m \in A(x)}$, où $I_m = [m, m + N(T^m x))$. Par le lemme 3.25, on sait qu'il existe une sous-collection d'intervalles disjoints $\{I_m\}_{m \in A'(x)}$ telle que $\sum_{m \in A'(x)} |N(T^m x)| \geq \frac{1}{2} |A(x)|$, comme tous ces

intervalles sont contenus dans $[0, N_0 + N_1]$, on a :

$$\begin{aligned}
\left| \sum_{n=0}^{N_0+N_1} |f(T^n x)| \right| &\geq \left| \sum_{m \in A'(x)} \sum_{n=m}^{N(T^m x)+m-1} |f(T^n x)| \right| \\
&= \left| \sum_{m \in A'(x)} \frac{N(T^m x)}{N(T^m x)} \sum_{n=m}^{N(T^m x)+m-1} |f(T^n x)| \right| \\
&\geq \left| \sum_{m \in A'(x)} N(T^m x) \delta \right| \\
&\geq \frac{1}{2} \delta |A(x)|.
\end{aligned}$$

Donc

$$\int_X \left| \sum_{n=0}^{N_0+N_1} |f(T^n x)| \right| d\mu \geq \int_X \frac{1}{2} \delta |A(x)| d\mu = \frac{1}{2} \delta N_1 \mu(E(N_0)).$$

Cependant, d'après le lemme 3.17 et du fait que T préserve la mesure, on a :

$$\begin{aligned}
\int_X \left| \sum_{n=0}^{N_0+N_1} |f(T^n x)| \right| d\mu &= \int_X \sum_{n=0}^{N_0+N_1} |f(T^n x)| d\mu \\
&= (N_0 + N_1) \int_X |f| d\mu.
\end{aligned}$$

Alors,

$$(N_0 + N_1) \|f\|_1 \geq \frac{1}{2} N_1 \delta \mu(E(N_0)),$$

ce qui implique que

$$\mu(E(N_0)) \leq \frac{2\varepsilon}{\delta} \frac{N_0 + N_1}{N_1}.$$

Ainsi, comme $\frac{2\varepsilon}{\delta} \frac{N_0+N_1}{N_1} \rightarrow \frac{2\varepsilon}{\delta}$ lorsque $N_1 \rightarrow \infty$ et, par les propriétés élémentaires des mesures, $\mu(E(N_0)) \rightarrow \mu(E)$ lorsque $N_0 \rightarrow \infty$, on obtient bien $\mu(E) \leq \frac{2\varepsilon}{\delta}$. \square

3.5 Théorème ergodique de Birkhoff

Nous concluons cette introduction à la théorie ergodique par la présentation et la démonstration du théorème ergodique de Birkhoff. Ce théorème peut être vu comme une extension du théorème ergodique moyen de von Neumann. En effet, si l'on reprend la formulation de ce dernier en remplaçant $\mathcal{L}^2(X)$ par $\mathcal{L}^1(X)$ et la convergence en norme \mathcal{L}^2 de $S_N f$ vers $\int_X f d\mu$ par la convergence ponctuelle p.p. de $S_N f$ vers $\int_X f d\mu$, on obtient précisément le théorème ergodique de Birkhoff. Le lemme suivant nous sera utile pour contourner le fait que la convergence en norme \mathcal{L}^p n'implique pas, en général, la convergence ponctuelle p.p..

Lemme 3.27. Soient (X, \mathcal{B}, μ, T) est un système ergodique et $(f_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{L}^1(X)$ telle que pour tout $n \in \mathbb{N}$, on a $\lim_{N \rightarrow \infty} S_N f_n(x) = \int_X f_n d\mu$ p.p.. De plus, supposons que la suite $(f_n)_{n \in \mathbb{N}^*}$ converge en norme \mathcal{L}^1 vers f (c'est-à-dire $\lim_{n \rightarrow \infty} \|f_n - f\|_1 = 0$). Alors, $\lim_{N \rightarrow \infty} S_N f(x) = \int_X f d\mu$ p.p..

Démonstration. Soient $\varepsilon > 0$ et $0 < \delta < 4$. Choisissons n tel que $\|f_n - f\|_1 \leq \frac{1}{12}\varepsilon\delta$. Alors,

$$\|f_n - f\|_1 = \int_X |f_n - f| d\mu \geq \left| \int_X (f_n - f) d\mu \right| = \left| \int_X f_n d\mu - \int_X f d\mu \right|.$$

Donc $\left| \int_X f_n d\mu - \int_X f d\mu \right| \leq \frac{1}{12}\varepsilon\delta$. Soit A_M l'ensemble des $x \in X$ pour lesquels $|S_N f_n(x) - \int_X f_n d\mu| > \frac{1}{3}\varepsilon$ pour un certain $N \geq M$. Alors, $\bigcap_{M=1}^{\infty} A_M$ est contenu dans l'ensemble des $x \in X$ pour lesquels $S_N f_n(x) \not\rightarrow \int_X f_n d\mu$ et il est donc de mesure nulle. Notons que nous avons que $A_1 \supset A_2 \supset \dots$, alors $\mu(A_M) \rightarrow 0$ lorsque $M \rightarrow \infty$, ce qui implique qu'il existe un certain M^* tel que $\mu(A_{M^*}) \leq \frac{1}{2}\delta$. Alors, $|S_N f_n(x) - \int_X f_n(x) d\mu| \leq \frac{1}{3}\varepsilon$ pour $N \geq M^*$ et $x \notin A_{M^*}$. Maintenant, considérons B , l'ensemble des $x \in X$ pour lesquels il existe N tel que $|S_N f_n(x) - S_N f(x)| \geq \frac{1}{3}\varepsilon$. Alors, par le théorème ergodique maximal et un simple calcul, $\mu(B) \leq \frac{1}{2}\delta$. Donc pour $x \notin A_{M^*} \cup B$, ensemble de mesure inférieure à δ ,

$$\begin{aligned} \left| S_N f(x) - \int_X f d\mu \right| &\leq |S_N f(x) - S_N f_n(x)| + \left| S_N f_n(x) - \int_X f_n d\mu \right| \\ &\quad + \left| \int_X f_n d\mu - \int_X f d\mu \right| \\ &\leq \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon + \frac{1}{12}\delta\varepsilon \leq \varepsilon, \end{aligned}$$

pour tout $N \geq M^*$. Le résultat est vrai pour tout x en dehors d'un ensemble de mesure δ . Comme δ est arbitraire, le résultat est vrai pour presque tout $x \in X$.

Pour finir, il reste à montrer que $\lim_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu|$ existe et qu'elle est nulle. On sait que $\limsup_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu|$ existe et donc par ce qui précède

$$\limsup_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu| \leq \varepsilon$$

p.p.. Par la suite, considérons une suite $(\varepsilon_i)_{i \in \mathbb{N}}$ telle que $\varepsilon_i < \varepsilon_j$ pour $j < i$ et $\varepsilon_i \rightarrow 0$ lorsque $i \rightarrow \infty$. Pour tout i et presque tout x ,

$$\limsup_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu| \leq \varepsilon_i.$$

Puisque l'union dénombrable d'ensembles de mesure nulle est de mesure nulle, l'inégalité est vraie p.p. ce qui est vrai pour tout i simultanément. Donc pour presque tout $x \in X$,

$$\limsup_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu| = 0.$$

Cependant, on sait que $\liminf_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu| \geq 0$, car $|S_N f(x) - \int_X f d\mu| \geq 0$ pour tout N . Donc

$$\liminf_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu| = \limsup_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu|,$$

ce qui implique que $\lim_{N \rightarrow \infty} |S_N f(x) - \int_X f d\mu|$ existe et est nulle. Ainsi, on a $\lim_{N \rightarrow \infty} S_N f(x) = \int_X f d\mu$ p.p.. \square

Il est important de garder à l'esprit que nous travaillons avec des mesures type finie et donc que les conditions du théorème 3.9 sont vérifiées. En particulier, on a $\mathcal{L}^2(X) \subset \mathcal{L}^1(X)$.

Lemme 3.28. *Soit (X, \mathcal{B}, μ, T) est un système ergodique. Alors l'ensemble des fonctions qui s'écrivent comme une fonction constante plus un cocycle borné est dense dans $\mathcal{L}^1(X)$.*

Démonstration. Soient $\varepsilon > 0$ et $f \in \mathcal{L}^1(X)$. On sait que l'espace des fonctions simples est dense dans $\mathcal{L}^1(X)$, on peut donc supposer sans perte de généralité que $f = \sum_{i=0}^n c_i 1_{E_i}$, où $c_i \in \mathbb{C}$ pour tout $i \in I$ et E_i est un ensemble mesurable pour tout $i \in I$. Alors,

$$\begin{aligned} \int_X \left| \sum_{i=0}^n c_i 1_{E_i} \right|^2 d\mu &\leq \sum_{i=0}^n \int_X |c_i|^2 1_{E_i} d\mu \\ &= \sum_{i=0}^n |c_i|^2 \mu(E_i) \\ &\leq \infty. \end{aligned}$$

Donc $f \in \mathcal{L}^2(X)$; ce qui implique, par la démonstration du théorème 3.23, qu'il existe $g \in \mathcal{L}^2(X)$ telle que $\|f - \int_X f d\mu - \partial g\|_2 \leq \frac{1}{9}\varepsilon^2$. Par l'inégalité de Hölder, on a $\|f - \int_X f d\mu - \partial g\|_1 \leq \frac{1}{3}\varepsilon$. On a donc que les fonctions qui s'écrivent comme une fonction constante plus un cocycle sont denses dans $\mathcal{L}^1(X)$, car les fonctions simples sont denses dans $\mathcal{L}^1(X)$ et ε était arbitraire.

Maintenant, on aimerait montrer que le résultat est toujours vrai si on remplace le cocycle par un cocycle borné. Soit $g \in \mathcal{L}^2(X)$, alors $g \in \mathcal{L}^1(X)$ et on peut l'approximer par une fonction simple avec une précision arbitraire en norme \mathcal{L}^1 . De plus, avec une démonstration similaire à celle faite plus haut, on peut montrer que les fonctions simples sont bornées. Prenons $\bar{g} \in \mathcal{L}^\infty(X)$ tel que $\|g - \bar{g}\|_1 < \varepsilon$, alors par l'inégalité du triangle et le lemme 3.17, $\|\partial g - \partial \bar{g}\|_1 < \frac{2}{3}\varepsilon$. Donc par l'inégalité du triangle, on obtient que $\|f - \int_X f d\mu - \partial \bar{g}\|_1 \leq \varepsilon$. Ainsi, puisque ε était arbitraire, le résultat suit. \square

Théorème 3.29. [Théorème ergodique de Birkhoff] [[Gre15], Théorème 5.1] *Soient (X, \mathcal{B}, μ, T) un système ergodique et $f \in \mathcal{L}^1(X)$. Alors, $S_N f$ converge ponctuellement vers $\int_X f d\mu$ p.p..*

Démonstration. Supposons que $f = c + \partial g$, où c est une constante, $g \in \mathcal{L}^\infty(X)$ et $\partial g = g - U_T g$. Alors,

$$\begin{aligned} S_N f(x) &= \frac{1}{N} \sum_{n=0}^{N-1} (c + \partial g)(T^n x) \\ &= \frac{1}{N} \left(\sum_{n=0}^{N-1} c + \sum_{n=0}^{N-1} \partial g(x) \right) \\ &= c + \frac{1}{N} (g(x) - g(T^N x)). \end{aligned}$$

Donc $S_N f(x) \rightarrow c$ lorsque $N \rightarrow \infty$. De plus,

$$\begin{aligned} \int_X f d\mu &= \int_X c + \partial g d\mu \\ &= \int_X c d\mu + \int_X g - g \circ T^n d\mu \\ &= c + \int_X g d\mu - \int_X g \circ T^n d\mu \\ &= c, \end{aligned}$$

car T préserve la mesure. Donc $S_N f(x) \rightarrow \int_X f d\mu$ p.p. lorsque $N \rightarrow \infty$. Ainsi, la démonstration se termine en invoquant le lemme 3.28 et le lemme 3.27. \square

Références

- [EW11] Manfred EINSIEDLER et Thomas WARD : *Ergodic Theory with a view towards Number Theory*. Springer, 2011.
- [Gal09] Thierry GALLAY : *Théorie de la mesure et de l'intégration*, 2009.
- [Gre15] Ben GREEN : *Ergodique theory*, 2015.
- [Kle13] A. KLENKE : *Probability Theory : A Comprehensive Course*. Universitext. Springer London, 2013.
- [Mor20] Sidney Allen MORRIS : *Topology without tears*, 2020.
- [Mus14] Joseph MUSCAT : *Functional Analysis An Introduction to Metric Spaces, Hilbert Spaces, and Banach Algebras*. Springer, 2014.
- [Ros06] J.S. ROSENTHAL : *First Look At Rigorous Probability Theory, A (2nd Edition)*. World Scientific Publishing Company, 2006.

TOMMY GIROUX
 DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE
 Courriel: Tommy.Giroux@USherbrooke.ca

