# Absence de vitesse de convergence dans le théorème ergodique de Birkhoff.

**Rapport pour le séminaire de M2.**

Léo Daures

Pour Décembre 2021

## 1 Introduction et rappels de théorie ergodique

On considère un système dynamique, c'est-à-dire un espace probabilisé $(X, \mathcal{F}, \mu)$ muni d'une transformation $T : X \to X$ qui vérifie $\forall A \in \mathcal{F}\ \mu(T^{-1}A) = \mu(A)$. On dit qu'elle préserve la mesure. Cette transformation est dite ergodique si les seuls ensembles $A$ vérifiant $A = T^{-1}A$ sont des ensembles de mesure 0 ou 1. Un objectif de la théorie ergodique est d'étudier les trajectoires $(T^k(x))_k \in \mathbb{N}$ pour presque tout $x$, que $T$ soit ergodique ou non.

En ce sens, un résultat central en théorie ergodique est le théorème ergodique de Birkhoff, qui affirme que pour un système dynamique $(X, \mathcal{F}, \mu, T)$, la moyenne des $n$ premières évaluations d'une fonction en les points $T^k(x)$ converge $\mu$-presque partout. Si $T$ est une transformation ergodique, il se formule de la manière suivante :

**Théorème 1.1.** Soit $(X, \mathcal{F}, \mu, T)$ un système dynamique ergodique, et $f \in L^1(X, \mathcal{F}, \mu)$. Alors,

$$\frac{1}{n}\sum_{k=1}^{n} f \circ T^k \xrightarrow[n\to\infty]{\text{ps}} \int_X f(x)\mu(\mathrm{d}x).$$

On peut être frappé par la ressemblance de ce théorème avec la loi forte des grands nombres : une moyenne des $n$ premiers termes qui converge presque sûrement vers la moyenne de la fonction, ou autrement dit, vers son espérance (d'ailleurs, le théorème ergodique de Birkhoff implique la loi des grands nombres, en choisissant bien le système dynamique). Nous ne chercherons pas ici à démontrer ce résultat, mais au contraire à en observer les limites. En effet, il est naturel de se demander s'il peut être accompagné d'une vitesse de convergence, de la même manière que la loi des grands nombres est associée théorème centrale-limite. Le théorème central-limite donne une vitesse de convergence en loi, mais ici on s'intéressera à une vitesse de convergence presque sûre comme dans le théorème de Marcinkiewicz-Zygmund. On dira qu'une suite de réels $(v_n)$ qui croît vers $+\infty$ est une vitesse de convergence si pour tout $g \in L^1([0,1])$, il existe un certain $\ell(g)$ tel que :

$$v_n\left(\frac{1}{n}\sum_{k=1}^{n} g \circ T^k - \int_X g(x)\mu(\mathrm{d}x)\right) \xrightarrow[n\to\infty]{\text{ps}} \ell(g).$$

En fait, nous allons montrer qu'il n'existe pas de vitesse de convergence dans le théorème ergodique de Birkhoff.

Tout cet exposé s'appuie sur l'article [1] de del Junco et Rosenblatt, qui démontre le théorème 2.5 en tenant pour acquis les lemmes 4.5 et 3.5.

Dans la prochaine section, on énoncera le théorème central de cet exposé. Il semblera abscond au premier abord, mais permettra de trouver un contre-exemple à la tout fin de cet exposé. Afin de le démonter nous détaillerons un résultat de théorie de la mesure et un lemme intermédiaire dans les deux sections suivantes.

## 2 Le théorème à démontrer

**Définition 2.1.** Une suite double $(a_{k,n})_{k,n\geqslant 0}$ est appelée suite de sommation divergente si

1. pour tout $n$, la série $\sum_k a_{k,n}$ converge absolument

2. $\sup_n |\sum_{k=0}^{\infty} a_{k,n}| = +\infty$.

*Exemple* 2.2. Soit $\alpha_n$ une suite qui tend vers $+\infty$. Alors $a_{k,n} := \begin{cases} \frac{\alpha_n}{n} \text{ si } k \leqslant n \\ 0 \text{ sinon} \end{cases}$ définit une suite de sommation divergente.

Dans la suite $L^1_0(X)$ désigne l'ensemble des fonctions intégrables sur $X$, d'intégrale nulle. On dit de plus qu'un ensemble est un $\mathrm{G}_\delta$ si c'est une intersection dénombrable d'ouverts. Enfin rappelons qu'un ensemble de première catégorie est un ensemble contenu dans une réunion dénombrable de fermés d'intérieurs vide, et qu'il suffit d'être contenu dans le complémentaire d'un ensemble de première catégorie pour être un $\mathrm{G}_\delta$. On introduit de plus la distance suivante.

**Définition 2.3.** Soit $\mathcal{G}$ une tribu. Pour $A, B \in \mathcal{G}$, on définit $d_\triangle(A, B) = \mu(A \triangle B)$ la distance de la différence symétrique.

**Proposition 2.4.** $d_\triangle$ *est une pseudo-distance sur* $\mathcal{G}$*. D'autre part, si* $\mathcal{G}/\mathcal{N}$ *désigne la tribu* $\mathcal{G}$ *quotientée par l'ensemble des négligeables,* $d_\triangle$ *est une distance sur* $\mathcal{G}/\mathcal{N}$*.*

*Preuve.* La symétrie est claire, et l'inégalité triangulaire vient de $A \triangle B \subseteq (A \triangle C) \cup (B \triangle C)$. Enfin, la séparation découle du fait que dans $\mathcal{G}/\mathcal{N}$, seul $\varnothing$ est de mesure nulle. $\square$

On peut maintenant énoncer le théorème d'existence qui servira à créer un contre exemple en fin d'exposé.

**Théorème 2.5.** *Soit* $a_{k,n}$ *une suite de sommation divergente, et soit* $T : [0, 1] \to [0, 1]$ *une transformation préservant la mesure de Lebesgue* $\mu$*. Alors, il existe* $\mathcal{R}$ *une partie* $\mathrm{G}_\delta$ *dense de* $L^1_0([0, 1])$ *tel que*

$$\forall f \in \mathcal{R} \quad \sup_n \left| \sum_{k=} a_{k,n} f(T^k x) \right| = +\infty \qquad \mu\text{-pp.}$$

*De plus, il existe* $\mathcal{Q}$ *une partie* $\mathrm{G}_\delta$ *dense de* $\mathcal{B}([0, 1])$ *tel que*

$$\forall A \in \mathcal{Q} \quad \sup_n \left| \sum_{k=} a_{k,n} (\mathbb{1}_A(T^k x) - \mu(A)) \right| = +\infty \qquad \mu\text{-pp.}$$

Pour montrer le théorème 2.5, on va énoncer et démontrer deux théorèmes de théorie de la mesure et un lemme supplémentaire.

# 3 Théorie de la mesure

On considère $(X, \mathcal{F}, \mu)$ un espace probabilisé. On note $L^0(X)$ l'ensemble des fonctions mesurables $X \to \mathbb{R}$, quotienté par l'égalité $\mu$-presque sûre. Si $E$ est un espace topologique et $(T_n)_{n \in \mathbb{N}}$ est une suite d'applications $E \to L^0(X)$, on notera $T^*_N a := \max\{|T_0 a|, |T_1 a|, \ldots, |T_N a|\}$ et $T^* a := \sup\{|T_n a|, \quad n \in \mathbb{N}\}$. L'objectif de cette partie est de montrer les théorèmes 3.3 et 3.7 suivants.

**Définition 3.1.** Une application $T$ linéaire d'un espace vectoriel normé $E$ vers $L^0(X)$ est dite continue en mesure si pour tout $a_0$ dans $E$, on a $Ta \xrightarrow[a \to a_0]{\mu} Ta_0$ au sens de la convergence en mesure $\mu$.

**Proposition 3.2.** *Si* $T_1, T_2, \ldots, T_N$ *sont des applications continues en mesure, alors* $T^*_N$ *est aussi continue en mesure.*

*Preuve.* Remarquons que par propriété de max et inégalité triangulaire, la fonction $T^*_N$ est sous-additive, donc $T^*_N a - T^*_N b \leqslant T^*_N(a - b)$ et $T^*_N b - T^*_N a \leqslant T^*_N(a - b)$, d'où l'inégalité $|T^*_N a - T^*_N b| \leqslant T^*_N(a - b)$. Soit maintenant $(a_n)_{n \geqslant 0}$ une suite qui converge vers un certain $a$. Soit $\varepsilon > 0$. Alors,

$$\begin{aligned} \mu(|T^*_N a_n - T^*_N a| \geqslant \varepsilon) &\leqslant \mu(|T^*_N(a_n - a)| \geqslant \varepsilon) \\ &= \mu\left( \bigcup_{i=1}^{N} \{|T_i a_n - T_i a| \geqslant \varepsilon\} \right) \\ &\leqslant \sum_{i=1}^{N} \mu\left(|T_i a_n - T_i a| \geqslant \varepsilon\right) \xrightarrow[n \to \infty]{} 0. \end{aligned}$$

$\square$

**Théorème 3.3.** Soit $(B, \|\cdot\|)$ un espace de Banach et soit $T_n : B \to L(X)$ une suite d'applications linéaires et continues en mesure. Supposons qu'elle vérifie l'hypothèse (H1) : pour tous $\varepsilon > 0$ et $K > 0$, il existe $b \in B$ de norme $\|b\| \leqslant 1$ tel que $\mu(T^*b > K) \geqslant 1-\varepsilon$. Alors, il existe $\mathcal{R}$ un $\mathrm{G}_\delta$ dense de $B$ tel que $\forall b \in \mathcal{R},\ T^*b = +\infty$ $\mu$-presque sûrement.

*Remarque* 3.4. Avec $B = L^1([0,1])$, on a alors trouvé l'ensemble $\mathcal{R}$ du théorème 2.5, pour peu que la suite de transformations $T_n$ soit correctement choisie.

*Preuve.* Posons $\mathcal{R} := \{b \in B, \quad \mu(T^*b = +\infty) = 1\}$. On va montrer que $\mathcal{R}$ est un $\mathrm{G}_\delta$ dense en montrant que $\mathcal{E} := B \setminus \mathcal{R}$ est un ensemble de première catégorie. On aura alors bien exhibé un $\mathrm{G}_\delta$ vérifiant la bonne propriété. De plus, comme $B$ est complet, $\mathcal{R}$ sera bien dense dans $B$ en vertu du théorème de Baire. Montrons que $\mathcal{E}$ est de première catégorie en profitant du fait qu'une réunion dénombrable d'ensembles de première catégorie est encore de première catégorie. On a :

$$\begin{aligned}\mathcal{E} &= \{b \in B, \quad \mu(T^*b = +\infty) < 1\} \\ &= \{b \in B, \quad \exists k \in \mathbb{N}^* \quad \mu(T^*b = +\infty) < 1 - \frac{1}{k}\} \\ &= \bigcup_{k \in \mathbb{N}^*} \{b \in B, \quad \mu(T^*b = +\infty) < 1 - \frac{1}{k}\}.\end{aligned}$$

En posant $\mathcal{E}_k := \{b \in B, \quad \mu(T^*b = +\infty) < 1 - \frac{1}{k}\}$, on écrit donc $\mathcal{E}$ comme la réunion dénombrable de $\mathcal{E}_k$, et il suffit donc de montrer que les $\mathcal{E}_k$ sont de première catégorie. On va encore simplifier le problème en écrivant $\mathcal{E}_k$ comme réunion dénombrable des ensembles $\mathcal{E}_k(m) := \{b \in B, \quad \mu(T^*b > m) < 1 - \frac{1}{k}\}$ pour $m \in \mathbb{N}^*$ : on a

- d'une part, $\mathcal{E}_k(m) \subseteq \mathcal{E}_k$,

- d'autre part, pour $b \in \mathcal{E}_k$, la suite $(\mu(T^*b > m))_{m\in\mathbb{N}}$ décroît vers $\mu(T^*b = +\infty) < 1 - \frac{1}{k}$ donc à partir d'un certain $m_0$, on a $\mu(T^*b > m) < 1 - \frac{1}{k}$, d'où $\mathcal{E}_k \subseteq \bigcup_{m\in\mathbb{N}} \mathcal{E}_k(m)$, d'où finalement,

$$\mathcal{E}_k = \bigcup_{m\in\mathbb{N}} \mathcal{E}_k(m).$$

Il suffit donc de montrer que les $\mathcal{E}_k(m)$ sont inclus dans des réunions dénombrables de fermés d'intérieur vide. Pour cela, on pose $\mathcal{C}_k(m) := \{b \in B, \quad \mu(T^*b > m) \leqslant 1 - \frac{1}{k}\}$, qui contient $\mathcal{E}_k(m)$. Montrons que $\mathcal{C}_k(m)$ est un fermé d'intérieur vide.

- $\mathcal{C}_k(m)$ **est fermé.** Montrons d'abord que les $\mathcal{C}_k^N(m) := \{b \in B, \quad \mu(T_N^*b > m) \leqslant 1 - \frac{1}{k}\}$ sont fermés. Soit $(b_n) \in C_k^N(m)^{\mathbb{N}}$ une suite qui converge vers un certain $b \in B$. Soit $\varepsilon > 0$ Alors pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}\mu(T_N^*b \geqslant m+\varepsilon) &= \mu\big(T_N^*b \geqslant m+\varepsilon, T_N^*b_n \leqslant m\big) + \mu\big(T_N^*b \geqslant m+\varepsilon, T_N^*b_n > m\big) \\ &= \mu\big(T_N^*b - T_N^*b_n \geqslant (m - T_N^*b_n) + \varepsilon, T_N^*b_n \leqslant m\big) + \mu\big(T_N^*b \geqslant m+\varepsilon, T_N^*b_n > m\big) \\ &\leqslant \mu\big(T_N^*b - T_N^*b_n \geqslant 0 + \varepsilon, T_N^*b_n \leqslant m\big) + \mu\big(T_N^*b \geqslant m+\varepsilon, T_N^*b_n > m\big) \\ &\leqslant \mu\big(T_N^*b - T_N^*b_n \geqslant \varepsilon\big) + \mu\big(T_N^*b_n > m\big).\end{aligned}$$

  Or $T_N^*$ est continue en mesure par la proposition 3.2, donc le premier terme $\mu\big(T_N^*b - T_N^*b_n \geqslant \varepsilon\big)$ tend vers zéro (quand $n \to \infty$). D'autre part par hypothèse le deuxième terme est majoré par $1 - \frac{1}{k}$. En passant à la limite dans la dernière inégalité, on obtient

$$\mu(T_N^*b \geqslant m+\varepsilon) \leqslant 1 - \frac{1}{k}.$$

  C'est valable pour n'importe quel $\varepsilon > 0$, d'où

$$\mu(T_N^*b > m) = \mu\left(\bigcup_{j=1}^{\infty} (T_N^*b \geqslant m + \frac{1}{j})\right) = \lim_{j\to\infty} \mu(T_N^*b \geqslant m + \frac{1}{j}) \leqslant 1 - \frac{1}{k}.$$

  Cela achève de montrer que $b \in \mathcal{C}_k^N(m)$, qui est donc bien un fermé. Ensuite, on a $\mathcal{C}_k(m) = \bigcap_N \mathcal{C}_k^N(m)$ donc $\mathcal{C}_k(m)$ est un fermé.
  *Remarque : c'est pour cette dernière inégalité qu'il est nécessaire de considérer les $\mathcal{C}_k(m)$ plutôt que les $\mathcal{E}_k(m)$, puisque les inégalités strictes n'auraient pas résisté au passage à la limite.*

- $\mathcal{C}_k(m)$ **est d'intérieur vide.** Supposons par l'absurde qu'on peut trouver une boule $\mathfrak{B}$ de centre $c$ et de rayon $\delta$ contenue dans $\mathcal{C}_k(m)$. Soit $0 < \varepsilon < \gamma$, et $M \in \mathbb{N}$ tel que $M(\gamma - \varepsilon) > 1$. On utilise maintenant (enfin !) l'hypothèse (H1), grâce à laquelle il existe $b_0 \in B$ de norme 1 tel que $\mu(T^* b_0 > 2m\frac{M}{\delta}) \geqslant 1 - \varepsilon$. Posons $F_i = \{T^*(c - \frac{i\delta}{M} b_0) \leqslant m\}$ pour $i \in [\![0, M]\!]$. Alors les $F_i \cap \{T^* b_0 > 2m\frac{M}{\delta}\}$ sont deux à deux disjoints. En effet, comme $T^*$ est un sup de fonctions linéaires, elle est sous-additive, donc
$$T^*\left(\frac{i-j}{M}\delta b_0\right) \leqslant T^*\left(c - \frac{i}{M}\delta b_0\right) + T^*\left(c - \frac{j}{M}\delta b_0\right).$$
Donc si $i$ et $j$ sont distincts, sur $F_i \cap F_j \cap \{T^* b_0 > 2m\frac{M}{\delta}\}$, on a
$$2m \geqslant \left(\frac{|i-j|}{M}\delta\right) T^* b_0 > \left(\frac{|i-j|}{M}\delta\right) \times \left(2m\frac{M}{\delta}\right) = |i-j| \times 2m \geqslant 2m,$$
ce qui est absurde donc les ensembles sont bien disjoints. Mais d'autre part, la probabilité de $F_i \cap \{T^* b_0 > 2m\frac{M}{\delta}\}$ dépasse $\frac{1}{M}$. En effet, $c - \frac{i\delta}{M} b_0 \in \mathfrak{B} \subseteq \mathcal{C}_k(m)$ donc $\mu\left(T^*(c - \frac{i\delta}{M} b_0) \leqslant m\right) < \gamma$. Donc
$$\begin{aligned}\mu\left(F_i \cap \{T^* b_0 > 2m\frac{M}{\delta}\}\right) &= \mu(F_i) + \mu\left(T^* b_0 > 2m\frac{M}{\delta}\right) - \mu\left(F_i \cup \{T^* b_0 > 2m\frac{M}{\delta}\}\right)\\ &\geqslant (1-\varepsilon) + \gamma - 1\\ &= \gamma - \varepsilon > \frac{1}{M}.\end{aligned}$$
Comme ils sont disjoints, la probabilité de l'union des $F_i \cap \{T^* b_0 > 2m\frac{M}{\delta}\}$ s'écrit donc comme la somme $\sum_i \mu\left(F_i \cap \{T^* b_0 > 2m\frac{M}{\delta}\}\right) > 1$, ce qui est absurde. Donc $\mathcal{C}_k(m)$ est d'intérieur vide.

Finalement, on a montré que $\mathcal{E}$ s'écrit comme une réunion dénombrable de $\mathcal{E}_k(m)$, ceux-ci étant de première catégorie car chacun inclus dans le fermé d'intérieur vide $\mathcal{C}_k(m)$. Donc $\mathcal{E}$ est de première catégorie donc $\mathcal{R}$ est un $\mathrm{G}_\delta$ dense dans $B$. □

Comme on l'a déjà remarqué, ce théorème 3.3 permettra d'obtenir la première partie du théorème 2.5. Pour la deuxième partie, on considère un nouvel espace probabilisé $(Y, \mathcal{G}, \nu)$ et on suppose désormais que $(X, \mathcal{F}, \mu)$ est séparable (c'est le cas avec $X = Y = [0,1]$, $\mathcal{F} = \mathcal{G} = \mathcal{B}([0,1])$ et $\mu = \nu$ la mesure de Lebesgue, comme dans le théorème 2.5). On veut montrer un théorème similaire au théorème 3.3 en remplaçant l'espace de Banach $B$ par la tribu $\mathcal{G}$ munie de la topologie de $d_\triangle$. On veut effectuer un raisonnement similaire, et donc utiliser le théorème de Baire : on a donc besoin de complétude.

**Lemme 3.5.** L'espace métrique $\mathcal{G}/\mathcal{N}$ muni de $d_\triangle$ est complet.

*Preuve.* Considérons $(A_n)$ une suite de Cauchy de $\mathcal{G}/\mathcal{N}$, et montrons qu'elle converge. Pour ce faire, on va montrer que $\mathbb{1}_{A_n}$ converge $\nu$-presque sûrement (à sous-suite près) vers l'indicatrice d'un certain ensemble mesurable $A$. On utilise le résultat suivant : si $A$ et $B$ sont deux ensembles mesurables, $\nu(A \triangle B) = \mathbb{E}_\nu\left[|\mathbb{1}_A - \mathbb{1}_B|\right]$. La propriété de Cauchy de $(A_n)$ dans $(\mathcal{G}, d_\triangle)$ implique donc que $(\mathbb{1}_{A_n})$ est de Cauchy dans $L^1(Y)$. Mais $L^1(Y)$ est complet, donc la suite $(\mathbb{1}_{A_n})$ converge au sens de la norme 1 ! Elle admet donc une sous-suite $(\mathbb{1}_{A_{\varphi(n)}})$ qui converge $\nu$-presque sûrement. On veut identitifer sa limite. En fait, comme elle est à valeurs dans l'ensemble discret $\{0,1\}$, elle stationne presque sûrement en 0 ou en 1. Donc,
$$\exists N \in \mathcal{N}\ \forall x \in N^c\ (\exists n_0(x)\ \forall n \geqslant n_0(x),\ x \in A_{\varphi(n)}) \text{ ou } (\exists n_0(x)\ \forall n \geqslant n_0(x),\ x \notin A_{\varphi(n)}).$$
Donc, pour tout $x \in N^c$, la suite $(\mathbb{1}_{A_{\varphi(n)}}(x))$ converge vers 0 ou converge ver 1. Notons
$$A := \{x \in N^c, \quad \exists n_0(x)\ \forall n \geqslant n_0(x),\ x \in A_{\varphi(n)}\},$$
ensemble mesurable. Comme $N$ est négligeable, on a en fait montré que $(\mathbb{1}_{A_{\varphi(n)}})$ converge presque sûrement vers $\mathbb{1}_A$. Pour conclure, il ne reste plus qu'à montrer que $A_{\varphi(n)}$ converge vers $A$ au sens de la distance $d_\triangle$. C'est le cas par théorème de convergence dominée car
$$\mu(A_{\varphi(n)} \triangle A) = \mathbb{E}\left[|\mathbb{1}_{A_{\varphi(n)}} - \mathbb{1}_A|\right] \xrightarrow[n\to\infty]{} 0.$$
La suite de Cauchy $A_n$ admet donc une valeur d'adhérence, donc elle converge. □

**Définition 3.6.** On dit qu'une application $T : \mathcal{G} \to L^0(X)$ est linéaire si l'image par $T$ d'une réunion disjointe d'éléments de $\mathcal{G}$ est la somme des images par $T$ de ces éléments.

Remarquons que si $A \subseteq B$ et $T$ linéaire, on a $T(B \setminus A) = T(B) - T(A)$. On observe en fait que cette notion de linéarité est assez proche de la notion de linéarité entre deux espaces vectoriels. En particulier, la preuve de la proposition 3.2 est encore valable en remplaçant les signes moins par des signes d'exclusion. Sa conclusion reste donc la même, et si $T_1, \dots T_N$ sont linéaires continues, alors $T_N^*$ aussi.

**Théorème 3.7.** Soit $T_n : \mathcal{G} \to L^0(X)$ une suite d'applications linéaires continues en mesure. Supposons qu'elle vérifie l'hypothèse (H2) : pour tous $\varepsilon > 0$ et $K > 0$, il existe $A \in \mathcal{G}$ tel que $\nu(A) < \varepsilon$ et $\mu(T^*A > K) \geqslant 1 - \varepsilon$. Alors, il existe $\mathcal{Q}$ un $\mathrm{G}_\delta$ dense de $\mathcal{G}$ tel que $\forall A \in \mathcal{Q},\ T^*A = +\infty$ $\mu$-presque sûrement.

La preuve de ce théorème est très similaire à celle du théorème 3.3 : on montre qu'un certain ensemble est de première catégorie en l'écrivant comme union dénombrable d'ensembles qui sont de première catégorie car inclus dans un autre ensemble de première catégorie. On utilise des notations similaires à celles du théorème 3.3 pour faciliter la lecture.

*Schéma de preuve.* Comme $(X, \mathcal{F}, \mu)$ est séparable, on peut trouver une suite d'évènements $S_i \in \mathcal{F}$ de mesures strictement positives telle que pour tout $A \in \mathcal{F}$ et pour tout $\varepsilon > 0$, il existe un indice $i \in \mathbb{N}$ tel que $\mu(S_i \cap A) \geqslant (1-\varepsilon)\mu(S_i)$. On pose alors

$$\mathcal{E}_i := \{A \in \mathcal{G}, \quad \mu(S_i \cap (T^*A = \infty)) < \frac{1}{4}\mu(S_i)\} \quad \text{et} \quad \mathcal{E} := \bigcup_{i \in \mathbb{N}} \mathcal{E}_i.$$

On montre que $\mathcal{E}$ est de première catégorie en montrant que tous les $\mathcal{E}_i$ le sont. Le complémentaire de $\mathcal{E}$ contiendra donc un $\mathrm{G}_\delta$ noté $\mathcal{Q}$. On a pris soin de vérifier que $\mathcal{G}/\mathcal{N}$ est complet, donc $\mathcal{Q}$ sera même un $\mathrm{G}_\delta$ dense de $\mathcal{G}$ en vertu du théorème de Baire. Pour montrer que les $\mathcal{E}_i$ sont de première catégorie on introduit les $\mathcal{E}_i(m)$ définis par $\mathcal{E}_i(m) := \{A \in \mathcal{G}, \quad \mu(S_i \cap (T^*A > m)) < \frac{1}{4}\mu(S_i)\}$. On obtient alors par des calculs ensemblistes :

$$\mathcal{E}_i = \bigcup_{m \in \mathbb{N}} \mathcal{E}_i(m).$$

Il suffit donc de montrer que les $\mathcal{E}_i(m)$ sont bien de première catégorie. On se ramène à la méthode précédemment utilisée pour le théorème 3.3. Soit $\mathcal{C}_i(m) := \{A \in \mathcal{G}, \quad \mu(S_i \cap (T^*A > m)) \leqslant \frac{1}{4}\mu(S_i)\}$ qui contient $\mathcal{E}_i(m)$. On peut montrer que c'est un fermé d'intérieur vide, d'une manière similaire à celle utilisée précédemment.
*Remarque : c'est pour montrer que l'intérieur de $\mathcal{C}_i(m)$ est vide qu'on utilise depuis le début de la preuve le choix du facteur $\frac{1}{4}$.*

Une fois qu'on a montré que $\mathcal{C}_i(m)$ est un fermé d'intérieur vide, $\mathcal{E}_i(m)$ est de première catégorie, donc par réunion dénombrable on a montré que $\mathcal{E}$ est un ensemble de première catégorie de $\mathcal{G}$. Comme annoncé précédemment, et par le théorème de Baire, on obtient donc l'existence d'un $\mathrm{G}_\delta$ dense de $\mathcal{G}$ contenu dans le complémentaire de $\mathcal{E}$, et noté $\mathcal{Q}$. Pour compléter la preuve du théorème, il faut s'assurer que c'est bien celui qu'on cherche, *i.e.* qu'il vérifie $\forall A \in \mathcal{Q},\ T^*A = +\infty$ presque sûrement.

Soit $A \in \mathcal{Q}$. Supposons par l'absurde que $\mu(T^*A < +\infty) > 0$. Alors, par l'hypothèse de séparabilité, il existe $i \in \mathbb{N}$ tel que $\mu(S_i \cap (T^*A < +\infty)) > (1 - \frac{1}{4})\mu(S_i) = \frac{3}{4}\mu(S_i)$. Donc

$$\mu(S_i \cap (T^*A = \infty)) = \mu(S_i) - \mu(S_i \cap (T^*A < \infty)) < \frac{1}{4}\mu(S_i),$$

donc $A \in \mathcal{E}_i$ ce qui est absurde car $A$ est précisément un élément du complémentaire de $\mathcal{E}$. Donc on a montré que pour tout $A \in \mathcal{Q}$, $\mu(T^*A = \infty) = 1$. □

Armé de ces deux théorèmes 3.3 et 3.7, on a déjà les conclusions du théorème 2.5 à montrer. il suffit de trouver les bonnes suites $(T_n)$, et vérifier les hypothèses (H1) et (H2). Pour cela, on a besoin du lemme 4.7 de la section suivante.

# 4 Invariance approchée d'ensembles mesurables

On suppose dans cette section que l'espace probabilisé $(X, \mathcal{F}, \mu)$ est non-atomique, ce qui signifie que tout évènement non négligeable $A \in \mathcal{F}$ contient un évènement $B \in \mathcal{F}$ vérifiant $0 < \mu(B) < \mu(A)$. Dans la suite, on sera amené à se restreindre à des parties $Y \in \mathcal{F}$, qui seront systématiquement munies de leur tribu induite $\mathcal{F}_Y = \{Y \cap A, \quad A \in \mathcal{F}\}$, et donc elles aussi non-atomiques.

**Définition 4.1.** Soit $\varepsilon > 0$, et $T$ une transformation $X \to X$ préservant la mesure $\mu$. Un évènement $B \in \mathcal{F}$ est dit $\varepsilon$-invariant si $\mu(B \triangle T^{-1}B) \leqslant \varepsilon\mu(B)$.

Heuristiquement, cela signifie que la transformation $T$ préserve presque $B$ : $T^{-1}B$ est très proche de $B$. Remarquons que $B$ est $\varepsilon$-invariant $\Leftrightarrow \mu(B \cap T^{-1}B) \geqslant (1 - \frac{\varepsilon}{2})\mu(B)$ car

$$\mu(B \triangle T^{-1}B) = \mu(B) + \mu(T^{-1}B) - 2\mu(B \cap T^{-1}B) = 2\mu(B) - 2\mu(B \cap T^{-1}B),$$

d'où $\mu(B \cap T^{-1}B) = \mu(B) - \frac{1}{2}\mu(B \triangle T^{-1}B)$.

**Lemme 4.2.** Soit $A \in \mathcal{F}$ et $T$ une transformation $X \to X$ préservant $\mu$. Alors l'évènement

$$B_n := \bigcup_{i=0}^{n-1} T^{-i}A$$

est $\frac{2}{n}$-invariant pour $T$.

*Preuve.* Pour calculer la probabilité d'une réunion, on se ramène à une union disjointe. Soit $A_0 := A$ et $A_i := T^{-i}A \setminus B_i = (((T^{-i}A) \setminus T^{-(i-1)}A) \setminus \ldots) \setminus A$. On a alors la relation de récurrence $A_{i+1} = (T^{-1}A_i) \setminus A$ et on a par construction $B_0 = A_0$ et $B_{n+1} = B_n \cup T^{-n}A = B_n \cup (T^{-n}A \setminus B_n) = B_n \cup A_n$, donc pour tout $n \in \mathbb{N}$, on a

$$B_n = \bigsqcup_{i=0}^{n-1} A_i,$$

où l'union est disjointe. Donc on a

$$(T^{-1}B_n) \setminus B_n = (B_{n+1} \setminus A_0) \setminus B_n = B_{n+1} \setminus B_n \subseteq A_n.$$

D'autre part, comme $T$ préserve $\mu$, la suite $(\mu(A_i))_i$ est décroissante. En effet par la relation de récurrence sur les $A_i$,

$$\mu(A_{i+1}) = \mu((T^{-1}A_i) \setminus A) \leqslant \mu(T^{-1}A_i) = \mu(A_i).$$

Sachant cela, on a $\mu(B_n \triangle T^{-1}B_n) = \mu(B_n \cap (T^{-1}B_n)^c) + \mu(B_n^c \cap T^{-1}B_n)$. Mais comme $B_n \cap (T^{-1}B_n)^c$ et $B_n^c \cap T^{-1}B_n$ sont les complémentaires de $B_n \cap T^{-1}B_n$ respectivement dans $B_n$ et dans $T^{-1}B_n$, ils sont de mesure ils sont de mesure respectivement $\mu(B_n) - \mu(B_n \cap T^{-1}B_n)$ et $\mu(T^{-1}B_n) - \mu(B_n \cap T^{-1}B_n)$, donc comme $\mu$ est $T$-invariante, ils sont de même mesure. On a donc

$$\mu(B_n \triangle T^{-1}B_n) = 2\mu(B_n^c \cap T^{-1}B_n) \leqslant 2\mu(A_n) \leqslant 2\mu(A_i)$$

puisque $B_n^c \cap T^{-1}B_n \subseteq A_n$ et $(\mu(A_i))$ est décroissante. En sommant cette dernière inégalité, on obtient

$$n\mu(B_n \triangle T^{-1}B_n) \leqslant 2\sum_{i=0}^{n-1} \mu(A_i) = 2\mu(B_n),$$

d'où la $\frac{2}{n}$-invariance. □

L'objectif de la section est de trouver des ensembles $\varepsilon$-invariants de mesure donnée (voir proposition 4.7). Pour ce faire, on va utiliser un argument de connexité. Et pour amener cet argument, les trois lemmes suivants ont pour but de montrer une sorte de connexité pour $\mathcal{F}$.

**Lemme 4.3.** La tribu $\mathcal{F}$ possède des évènements de probabilité arbitrairement petite. Autrement dit, il existe une suite d'évènements non négligeables $(A_n)$ telle que $\mu(A_n) \to 0$.

*Preuve.* On construit $A_n$ de la manière suivante : $A_0 = X$, puis si $A_1, \ldots, A_n$ sont déjà construits, comme $\mathcal{F}$ est non-atomique, il existe $B \subseteq A_n$ tel que $0 < \mu(B) < \mu(A_n)$. Entre $B$ et $A_n \setminus B$, au moins l'un des deux est de mesure inférieure à $\frac{1}{2}\mu(A_n)$ : on le nomme $A_{n+1}$ et on itère le processus.
Avec cette construction, on obtient $0 < \mu(A_n) \leqslant 2^{-n}$, d'où la convergence. □

**Lemme 4.4.** La tribu $\mathcal{F}$ possède des évènements de probabilité arbitraire. Autrement dit, pour tout $t \in [0, 1]$, il existe un évènement $A \in \mathcal{F}$ tel que $\mu(A) = t$.

*Preuve.* On veut montrer qu'il existe des évènements de $\mathcal{F}$ de probabilité arbitraire (puisque $\mathcal{F}$ est non-atomique, on sait déjà qu'on peut y trouver des évènements de probabilité arbitrairement petite). Commençons par une étape intermédiaire :
**Étape 1 :** montrer que pour $Y \in \mathcal{F}$, pour tout $\delta > 0$, il existe un évènement $A \in \mathcal{F}_Y$ tel que $\frac{\delta}{2} \leqslant \mu(A) \leqslant \delta$.
Soit $\Gamma := \{A \in \mathcal{F}_Y, \quad \mu(A) \leqslant \delta\} \neq \varnothing$. Distinguons deux cas :

- Si $\Gamma$ n'est pas stable par réunion finie, on peut y trouver $A$ et $B$ tels que $A \cup B \notin \Gamma$. Alors $\delta \leqslant \mu(A \cup B) \leqslant \mu(A) + \mu(B)$ donc l'un des deux est de probabilité supérieure à $\frac{\delta}{2}$.
- Si $\Gamma$ est stable par réunion finie, on pose $\gamma := \sup\{\mu(A), \quad A \in \Gamma\}$, et on va montrer qu'il est atteint et vaut $\delta$. Pour tout $n \in \mathbb{N}^*$, $\gamma - \frac{1}{n} < \gamma$ donc par définition de la borne supérieure, il existe $A_n \in \Gamma$ tel que $\mu(A_n) \geqslant \gamma - \frac{1}{n}$. On définit une suite croissante d'évènements de $\Gamma$ par $B_n := A_1 \cup \ldots \cup A_n \in \Gamma$. Elle vérifie
$$\delta \geqslant \mu(B_n) \geqslant \mu(A_n) > \gamma - \frac{1}{n}.$$
On peut alors considérer $B := \bigcup_n B_n$, c'est un évènement et il vérifie $\mu(B) = \lim_n \mu(B_n) \leqslant \delta$ donc $B \in \Gamma$. D'autre part, il vérifie
$$\mu(B) = \lim_n \mu(B_n) \geqslant \lim_n \mu(A_n) \geqslant \lim_n (\gamma - \frac{1}{n}) = \gamma.$$
Donc $B$ est un élément de $\Gamma$ qui atteint le maximum $\gamma$.
Montrons maitenant que ce maximum est bien $\delta$. Si $\mu(B) = \gamma < \delta$, l'ensemble $B^c$ est non-atomique, donc il contient des parties de probabilités arbitrairement petites. En particulier, il contient un certain évènement $C$ de probabilité $\mu(C) < \delta - \gamma$. Comme $\Gamma$ est stable par réunion et $C \in \Gamma$, l'évènement $B \cup C$ est dans $\Gamma$ mais est de probabilité $\delta - \gamma + \gamma = \delta > \gamma$, ce qui est absurde. Donc $\gamma = \delta$, et en particulier on a trouvé un évènement $B$ de probabilité exactement $\delta$.

**Étape 2 :** soit $t \in [0, 1]$. Construire par récurrence un évènement de $\mathcal{F}$ de probabilité $t$.
En appliquant le résultat de l'étape 1 à $\delta = t$, on obtient un évènement $A_1$ tel que
$$\frac{t}{2} \leqslant \mu(A_1) \leqslant t.$$
Si $\mu(A_1) = t$, la construction est déjà terminée. Sinon, comme $A_1^c$ est encore non-atomique, on applique le résultat précédent à $\delta = t - \mu(A_1)$, et on obtient un certain $A_2 \subseteq A_1^c$ tel que $\frac{1}{2}(t - \mu(A_1)) \leqslant \mu(A_2) \leqslant t - \mu(A_1)$. Comme leur union est disjointe, on a donc
$$\frac{3}{4}t \leqslant \frac{1}{2}t + \frac{1}{2}\mu(A_1) \leqslant \mu(A_1) + \mu(A_2) \leqslant \mu(A_1 \cup A_2) \leqslant t.$$
Supposons qu'on ait construit $A_1, A_2, \ldots, A_n$ par récurrence tel que
$$\frac{2^n - 1}{2^n} t \leqslant \mu\left(\bigcup_{i=1}^n A_i\right) \leqslant t.$$
Si $\mu\left(\bigcup_{i=1}^n A_i\right) = t$, on peut arrêter la construction. Sinon on construit $A_{n+1}$ en utilisant le résultat précédent sur l'ensemble non-atomique $\left(\bigcup_{i=1}^n A_i\right)^c$ avec $\delta = t - \mu\left(\bigcup_{i=1}^n A_i\right) > 0$. Il existe donc un évènement $A_{n+1}$ disjoint de tous les $A_i$ déjà construits tel que
$$\frac{1}{2}\left(t - \mu\left(\bigcup_{i=1}^n A_i\right)\right) \leqslant \mu(A_{n+1}) \leqslant t - \mu\left(\bigcup_{i=1}^n A_i\right).$$
Comme l'union est disjointe, en ajoutant $\mu\left(\bigcup_{i=1}^n A_i\right)$ à chaque terme de cette inégalité, on obtient
$$\frac{2^{n+1} - 1}{2^{n+1}} t \leqslant \frac{1}{2}t + \frac{1}{2}\mu\left(\bigcup_{i=1}^n A_i\right) \leqslant \mu\left(\bigcup_{i=1}^n A_i\right) + \mu(A_{n+1}) \leqslant \mu\left(\bigcup_{i=1}^{n+1} A_i\right) \leqslant t.$$
Cette construction par récurrence se répète un nombre fini ou infini de fois : s'il n'y a qu'un nombre fini d'étapes on a trouvé un évènement de probabilité $t$, et sinon on considère l'évènement $A = \bigcup_{i=1}^\infty A_i$ (qui est bien mesurable comme union dénombrable). En passant à la limite dans l'encadrement de $\mu\left(\bigcup_{i=1}^n A_i\right)$, on obtient que $\mu(A) = t$. Dans les deux cas on a conclu qu'il existait un évènement de probabilité $t$. $\square$

**Lemme 4.5.** Si $\mathcal{N}$ désigne l'ensemble des évènements négligeables de $\mathcal{F}$, l'ensemble quotienté $\mathcal{F}/\mathcal{N}$ muni de la distance de la différence symétrique $d_\triangle$ est connexe par arcs.

*Preuve.* Soit$A \in \mathcal{F}$. On veut trouver un chemin (dans $\mathcal{F}/\mathcal{N}$) de $A$ vers $\varnothing$. On va commencer par exhiber une famille croissante d'évènements $(B_t)_{t\in[0,1]}$ telle que $\mu(B_t) = t$. On construit d'abord $B_s$ pour les $s$ rationnels dyadiques de la manière suivante :

- $B_0 = \varnothing$ et $B_1 = X$.
- Si les $\{B_{i/2^n},\ i \leqslant 2^n\}$ sont construits, on construit les $\{B_{i/2^{n+1}},\quad i \leqslant 2^{n+1},\ i \text{ impair}\}$ un par un. Pour construire $B_{(2j+1)/2^n}$, on se place dans l'espace non-atomique $B_{(j+1)/2^n} \setminus B_{j/2^n}$. par le lemme 4.4 précédent, il possède un évènement $A$ tel que $\mu(A) = 2^{-(n+1)}$. On peut alors poser $B_{(2j+1)/2^n} := B_{j/2^n} \cup A$, qui est bien de probabilité $\mu(B_{j/2^n}) + \mu(A) = \frac{j}{2^n} + \frac{1}{2^{n+1}} = \frac{2j+1}{2^{n+1}}$.

Pour compléter cette construction à tout $t \in [0,1]$, on pose

$$B_t = \bigcup_{n\in\mathbb{N}} B_{\lfloor 2^n t\rfloor/2^n}.$$

Comme $\lfloor 2^n t\rfloor/2^n$ croît vers $t$, on a $\mu(B_t) = \lim_n \mu(B_{\lfloor 2^n t\rfloor/2^n}) = \lim_n \lfloor 2^n t\rfloor/2^n = t$. De plus, si $t < s$, il existe un rationnel dyadique $\frac{i}{2^n}$ compris entre les deux, d'où par la construction $B_t \subseteq B_{i/2^n}$ et $B_{i/2^n} \subseteq B_s$, et la famille obtenue est bien comme annoncé. Alors, on considère le chemin $f : t \mapsto A \cap B_t$. Il vérifie $\mu(f(0)) = 0$, $\mu(f(1)) = \mu(A)$ et est bien continu pour $d_\triangle$. $\square$

**Proposition 4.6.** Soit $T_1, \ldots, T_k$ des transformations $X \to X$ qui commutent et qui préservent $\mu$. Soit $\varepsilon > 0$ et $\lambda \in [0,1]$. Alors il existe $B \in \mathcal{F}$ un ensemble $\varepsilon$-invariant pour $T_1, T_2, \ldots, T_k$ et de mesure $\mu(B) = \lambda$.

*Preuve.* Soit $N$ tel que $\varepsilon > \frac{2}{N}$. Avec $A$ un ensemble quelconque, on définit

$$\begin{aligned} A^N &:= \bigcup_{0\leqslant i_1,\ldots,i_k<N} T_1^{-i_1} T_2^{-i_2} \ldots T_k^{-i_k}(A) \\ &= \bigcup_{i_1=0}^{N-1} T_1^{-i_1} \left( \bigcup_{0\leqslant i_2,\ldots i_k<N} T_2^{-i_2} T_3^{-i_3} \ldots, T_k^{-i_k}(A) \right). \end{aligned}$$

Par le lemme 4.2, $A^N$ est un ensemble $\varepsilon$-invariant pour la transformation $T_1$. Comme les $T_i$ commutent, on peut écrire $A^N$ de la même manière en sortant chaque indice $1, \ldots, k$, et vérifier que $A^N$ est $\varepsilon$-invariant pour chacun des $T_i$. Mais $A^N$ n'est pas forcément de probabilité $\lambda$, il reste donc à trouver un $A$ tel que $\mu(A^N) = \lambda$. Pour cela, on va utiliser le théorème des valeurs intermédiaires grâce au lemme 4.5. En effet, $\mathcal{F}/\mathcal{N}$ est connexe par arcs et la fonction $f : \mathcal{F}/\mathcal{N} \to [0,1],\ A \mapsto \mu(A^N)$ est bien définie. Montrons qu'elle est continue, en montrant qu'elle est lipschitzienne. Soient $A_1, A_2, B_1, B_2$ des ensembles mesurables, on a :

$$\begin{aligned} (A_1 \cup A_2)\triangle(B_1 \cup B_2) &= \big((A_1 \cup A_2) \cap B_1^c \cap B_2^c\big) \ \cup\ \big(A_1^c \cap A_2^c \cap (B_1 \cup B_2)\big) \\ &= \big((A_1 \cap B_1^c \cap B_2^c) \cup (A_2 \cap B_1^c \cap B_2^c)\big) \ \cup\ \big((A_1^c \cap A_2^c \cap B_1) \cup (A_1^c \cap A_2^c \cap B_2)\big) \\ &\subseteq \big((A_1 \cap B_1^c) \cup (A_2 \cap B_2^c)\big) \ \cup\ \big((A_1^c \cap B_1) \cup (A_2^c \cap B_2)\big) \\ &= (A_1 \triangle B_1) \cup (A_2 \triangle B_2). \end{aligned}$$

En appliquant cette inclusion par récurrence, on montre que si $A_1, \ldots, A_n, B_1, \ldots, B_n$ sont des ensembles mesurables,

$$\bigcup_{i=1}^n A_i \ \triangle\ \bigcup_{i=1}^n B_i \subseteq \bigcup_{i=1}^n A_i \triangle B_i.$$

Soit maintenant deux ensembles mesurables $A$ et $B$. La propriété précédente implique que

$$A^N \triangle B^N \subseteq \bigcup_{0\leqslant i_1,\ldots,i_k<N} T_1^{-i_1} T_2^{-i_2} \ldots T_k^{-i_k}(A \triangle B),$$

donc en termes de mesure,

$$\mu(A^N \triangle B^N) \leqslant N^k \mu(A \triangle B).$$

On a donc enfin :

$$\begin{aligned}|f(A)-f(B)| = |\mu(A^N)-\mu(B^N)| &= |\mu(A^N\setminus(A^N\cap B^N)) - \mu(B^N\setminus(A^N\cap B^N))|\\ &\leqslant |\mu(A^N\setminus(A^N\cap B^N))| + |\mu(B^N\setminus(A^N\cap B^N))|\\ &= \mu(A^N\triangle B^N)\\ &\leqslant N^k\mu(A\triangle B) = N^k d_\triangle(A,B).\end{aligned}$$

Donc $f$ est $N^k$-lipschitzienne donc continue. Comme $f(\varnothing)=0$ et $f(X)=1$, par théorème des valeurs intermédiaires, il existe $A$ tel que $f(A)=\lambda$. Alors $A^N$ est l'ensemble $\varepsilon$-invariant recherché. □

Cette proposition implique immédiatement la proposition suivante, qu'on utilisera dans la section prochaine.

**Proposition 4.7.** Soit $T : X\to X$ une transformation qui préserve $\mu$. Soit $k\in\mathbb{N}$, $\varepsilon>0$ et $\lambda\in[0,1]$. Alors il existe $B\in\mathcal{F}$ un ensemble $\varepsilon$-invariant pour $T, T^2,\dots,T^k$ et de mesure $\mu(B)=\lambda$.

# 5 Preuve du théorème 2.5 et conclusion

On va montrer le théorème 2.5 en s'appuyant sur les théorèmes 3.3 et 3.7, et sur la proposition 4.7.

*Preuve du théorème 2.5.* On veut utiliser simultanément les théorèmes 3.3 et 3.7 sur $(X,\mathcal{F},\mu)=([0,1],\mathcal{B}([0,1]),\mu)$ avec $B=L^1_0([0,1])$ et $\mathcal{G}=\mathcal{B}([0,1])$. On définit deux suites d'applications linéaires par :

$$T_n : \begin{cases} L^1_0([0,1]) & \to L^1([0,1])\\ f & \mapsto \sum_{k=1}^\infty a_{k,n} f\circ T^k\end{cases} \quad \text{et} \quad \tilde{T}_n : \begin{cases}\mathcal{G}\to L^1([0,1])\\ A\mapsto \sum_{k=1}^\infty a_{k,n}(T^k 1\!\!1_A - \mu(A)).\end{cases}$$

Chacun des $T_n$ est bien continu en mesure car $\sum_k a_{k,n}$ converge absolument. En effet, si $(f_p)_p$ converge dans $L^1$ vers $f$,

$$\begin{aligned}\mu(\|T_n(f_p)-T_n(f)\|_1\geqslant\varepsilon) &= \mu\left(\left\|\sum_{k=1}^\infty a_{k,n}(f_p-f)\circ T^k\right\|_1\geqslant\varepsilon\right)\\ &\leqslant \mu\left(\sum_{k=1}^\infty |a_{k,n}|\|(f_p-f)\circ T^k\|_1\geqslant\varepsilon\right)\\ &= \mu\left(\|f_p-f\|_1\sum_{k=1}^\infty|a_{k,n}|\geqslant\varepsilon\right)\xrightarrow[p\to\infty]{} 0.\end{aligned}$$

Par un calcul similaire, les $\tilde{T}_n$ sont aussi continus en mesure. Il reste donc à vérifier l'hypothèse (H1) pour le théorème 3.3 et (H2) pour le théorème 3.7. Soit $\varepsilon>0$, et $K>0$. On va chercher une fonction $f$ vérifiant (H1) sous la forme $1\!\!1_A-\mu(A)$ grace à la proposition 4.7. En effet, si on choisit un $\lambda\in]0,\frac12[$ et $N\in\mathbb{N}$, il existe un ensemble $A$ de mesure $\mu(A)=\lambda$ qui est $\varepsilon$-invariant pour $T,T^2,\dots,T^N$. Alors, l'ensemble $B:=\cap_{i=0}^N T^{-i}A$ est de mesure supérieure à $\mu(A)-\varepsilon$. De plus, quitte à prendre $\varepsilon$ plus petit, la mesure de $C:=\cap_{i=0}^N T^{-i}A^c$ est supérieure à $\mu(A^c)-\varepsilon$. Montrons que $A$ vérifie $\mu(T^*1\!\!1_A>K)\geqslant\mu(|T_N1\!\!1_A|>K)\geqslant 1-2\varepsilon$ pour un bon choix de $N$. Comme $a_{k,n}$ est une suite de sommation divergente, il existe $n\in\mathbb{N}$ tel que $\lambda|\sum_{k=1}^\infty a_{k,n}|>K$. Par seconde inégalité triangulaire, la suite

$$\left(\lambda\left|\sum_{k=1}^N a_{k,n}\right| - \sum_{k=N+1}^\infty |a_{k,n}|\right)_{N\in\mathbb{N}}$$

est croissante et sa limite est strictement supérieure à $K$ par le choix de $n$. Alors, on peut donc considérer $N\in\mathbb{N}$ tel que $\lambda\left|\sum_{k=1}^N a_{k,n}\right|-\sum_{k=N+1}^\infty|a_{k,n}|>K$. Montrons que sur $B\cup C$, $|T_N1\!\!1_A|>K$.

- Soit $x \in B$.

$$\begin{aligned}
|T_N(\mathbb{1}_A)(x)| &= \left|\sum_{k=1}^{\infty} a_{k,n}(\mathbb{1}_A T^k(x) - \lambda)\right| \\
&= \left|\sum_{k=1}^{N} a_{k,n}(1-\lambda) + \sum_{k=N+1}^{\infty} a_{k,n}(\mathbb{1}_A T^k(x) - \lambda)\right| \\
&\geqslant (1-\lambda)\left|\sum_{k=1}^{N} a_{k,n}\right| - (1-\lambda)\sum_{k=N+1}^{\infty} |a_{k,n}| \qquad \text{(seconde inégalité triangulaire)} \\
&\geqslant \lambda\left|\sum_{k=1}^{N} a_{k,n}\right| - \sum_{k=N+1}^{\infty} |a_{k,n}| > K \qquad (\text{car } \lambda < \frac{1}{2} < 1-\lambda).
\end{aligned}$$

- Soit $x \in C$. On a par le même type de calculs,

$$\begin{aligned}
|T_N(\mathbb{1}_A)(x)| &= \left|\sum_{k=1}^{N} a_{k,n}(-\lambda) + \sum_{k=N+1}^{\infty} a_{k,n}(\mathbb{1}_A T^k(x) - \lambda)\right| \\
&\geqslant \lambda\left|\sum_{k=1}^{N} a_{k,n}\right| - \sum_{k=N+1}^{\infty} |a_{k,n}| > K.
\end{aligned}$$

Donc pour $x \in B \cup C$, on a bien $|T_N(\mathbb{1}_A)(x)| > K$. Donc

$$\mu(T^*\mathbb{1}_A > K) \geqslant \mu(|T_N \mathbb{1} A| > K) \geqslant \mu(B \cup C) \geqslant \mu(A) - \varepsilon + \mu(A^c) - \varepsilon = 1 - 2\varepsilon.$$

Au sens du théorème 3.3, $b = \mathbb{1}_A - \mu(A)$ est bien une fonction de la boule unité de $B = L^1_0([0,1])$ qui vérifie $\mu(T^*b > K) \geqslant 1 - 2\varepsilon$ ; et au sens du théorème 3.7, $A$ est bien un évènement de $\mathcal{G} = \mathcal{B}([0,1])$ de mesure $\lambda$ et qui vérifie $\mu(\tilde{T}^*A > K) \geqslant 1 - 2\varepsilon$. Les hypothèses des théorèmes 3.3 et 3.7 sont vérifiées, donc les $G_\delta$ denses $\mathcal{R}$ et $\mathcal{Q}$ existent. □

On peut maintenant conclure cet exposé.

**Théorème 5.1.** Il n'existe pas de vitesse de convergence presque sûre dans le théorème ergodique de Birkhoff.

*Preuve.* Supposons qu'il existe $(v_n)$ une suite de réels positifs croissant vers l'infini telle que pour toute fonction $g \in L^1([0,1])$, on ait

$$v_n\left(\frac{1}{n}\sum_{k=1}^{n} g \circ T^k - \int_X g\mu(\mathrm{d}x)\right) \xrightarrow[n\to+\infty]{\text{ps}} \ell(g),$$

pour une certaine limite finie $\ell(g)$. Considérons $a_{k,n}$ qui vaut 0 si $k > n$ et $\frac{v_n}{n}$ sinon. On a vu dans l'exemple 2.2 que $(a_{k,n})$ est une suite de sommation divergente. Mais alors, par le théorème 2.5 que nous venons de démontrer, il existe une (et même une infinité !) fonction $f$ intégrable et de moyenne nulle telle que

$$\sup_n \left|\frac{v_n}{n}\sum_{k=1}^{n} f \circ T^k\right| = +\infty \quad \text{presque sûrement.}$$

Mais par hypothèse sur $v$, la suite $(|\frac{v_n}{n}\sum_{k=1}^n f \circ T^k|)$ converge presque sûrement vers une limite finie, donc est bornée. C'est absurde donc $v$ n'existe pas. □

# Référence bibliographique


[1] Joseph Rosenblatt and Andres del Junco. “Counterexamples in ergodic theory and number theory”. In: *Mathematische Annalen* 245 (1979), pp. 185–197.