INTRODUCTION À LA THÉORIE ERGODIQUE

1 Définitions

1.1 Cas d'une transformation

Définition 1 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : X \to X$ une application mesurable.*
1. On dit que T préserve les ensembles de mesure nulle (ou quasi invariante) si pour tout $A \in \mathcal{A}$, $\mu(A) = 0 \Rightarrow \mu(T^{-1}(A)) = 0$.
2. On dit que μ est T-invariante (ou que T conserve μ ou que T est un endomorphisme de $(X, \mathcal{A}, \mu)$) si pour tout $A \in \mathcal{A}$, $\mu(T^{-1}(A)) = \mu(A)$.
3. On dit qu'une partie mesurable A est invariante si $T^{-1}(A) = A$.
4. On dit qu'une partie mesurable A est invariante modulo μ si $\mu(T^{-1}(A)\Delta A) = 0$.
5. On dit qu'une fonction $f : X \to Y$ est invariante si $f \circ T = f$.
6. On dit qu'une fonction mesurable $f : X \to Y$ est invariante modulo μ si $f \circ T = f$ μ-presque partout.

Remarques. 1 signifie que l'image de μ par T est absolument continue par rapport à μ et 2 signifie que l'image de μ par T est égale à μ.
Lorsque T préserve les ensembles de mesure nulle, T agit sur les classes de fonctions mesurables égales μ-p.p. ; en effet si $\{f \neq g\}$ est un ensemble de mesure nulle alors $\{f \circ T \neq g \circ T\} = T^{-1}(\{f \neq g\})$ est de mesure nulle.
2. Si T préserve les ensembles de mesure nulle alors pour tous A appartenant à la tribu complétée, T^A appartient à la tribu complétée.
3. Si Y est une partie mesurable de X telle que $X \setminus Y$ soit de mesure nulle et si $T : Y \to X$ est une application telle que pour tout $A \in \mathcal{A}$, $\mu(A) = 0 \Rightarrow \mu(T^{-1}(A)) = 0$, alors il existe une partie mesurable $Z \subset Y$ telle que $T(Z) \subset Z$ et $X \setminus Z$ soit de mesure nulle (**exercice**).

Exemples-Exercices :
1. $X =$ un intervalle de $\mathbb{R}$ muni de la mesure de Lebesgue et $f : X \to X$ de classe $\mathcal{C}^1$ dont la dérivée ne s'annule qu'un nombre fini de fois. f préserve les ensembles de mesure nulle.
2. Soit U un ouvert de $\mathbb{R}^n$ muni de la mesure de Lebesgue. Un difféomorphisme $\varphi : U \to U$ préserve les ensembles de mesure nulle.
3. $X = \mathbb{R}/\mathbb{Z}$ muni de la mesure de Lebesgue, $\theta \in X$, $T : x \in X \to x + \theta$. T conserve la mesure de Lebesgue.
4. $X = \mathbb{R}/\mathbb{Z}$ muni de la mesure de Lebesgue, a un entier ≥ 1, $T : x \in X \to ax$. Idem.
5. $X =]0, 1[\setminus\mathbb{Q}$, $T : x \in X \to \{\frac{1}{x}\}$, $\mu = \frac{1}{1+x}\,dx$. μ est T-invariante.
6. Soit $A \in M_d(\mathbb{Z})$. A agit sur $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Montrer que si $\det A \neq 0$ alors A conserve la mesure de Lebesgue.
7. $b \in \mathbb{N}^*$. $\mathcal{A} = \{0, ..., b-1\}$ ensemble fini, $X_b = \mathcal{A}^{\mathbb{N}}$, $\mathcal{T} =$ tribu engendré par les cylindres $([a_0, ..., a_{k-1}] = \{x \in X_b : x_0 = a_0, ..., x_{k-1} = a_{k-1}\})$. $1 = p_1 + ... + p_b$. Probabilité μ qui est l'unique mesure sur $\mathcal{T}$ vérifiant $\mu([a_0, ..., a_{k-1}]) = p_{a_0}...p_{a_{k-1}}$. $S : X_b \to X_b$ défini par $S((a_i)_{n\geq 0}) = (a'_i)_{i\geq 0}$ avec $a'_i = a_{i+1}$ pour tout i. La mesure μ est S-invariante.

Exercice. On peut définir une addition dans X_b : $(a_0, ..., a_n, ...) + (b_0, ..., b_n, ...) = (c_0, ..., c_n, ...)$ où les c_i sont définis par récurrence grâce à la suite des retenues $r_0, ..., r_n, ...$ valant 0 ou 1 :
- si $a_0 + b_0 < b$ alors $c_0 = a_0 + b_0$ et $r_0 = 0$
- si $a_0 + b_0 \geq b$ alors $c_0 = a_0 + b_0 - b$ et $r_0 = 1$,

Lorsque $c_0, ..., c_n$ et $r_0, ..., r_n$ sont définis, c_{n+1} et r_{n+1} sont définis par :
- si $a_{n+1} + b_{n+1} + r_n < b$ alors $c_{n+1} = a_{n+1} + b_{n+1} + r_n$ et $r_{n+1} = 0$
- si $a_{n+1} + b_{n+1} + r_n \geq b$ alors $c_{n+1} = a_{n+1} + b_{n+1} + r_n - b$ et $r_{n+1} = 1$.

1. Justifier que $(X_b, +)$ est un groupe abélien.
2. Montrer que la mesure produit définie par $p_0 = ... = p_{b-1} = \frac{1}{b}$ est invariante par translation (les applications du type $\tau(x) = a + x$ où a est un élément fixé de X_b).

Proposition 1 *Soit $T : X \to X$ une application préservant les ensembles de mesure nulle et A une partie mesurable T-invariante modulo μ de X. Alors il existe A' mesurable tel que $T^{-1}(A') = A'$ et $\mu(A'\Delta A) = 0$.*

Démonstration. Si B est une partie mesurable telle que $\mu(A\Delta B) = 0$ alors B est invariante modulo μ. En effet, pour presque tout x on a

$$1_B(x) = 1_A(x) = 1_A \circ T(x) = 1_B \circ T(x).$$

Considérons l'ensemble A' des points de X dont la trajectoire passe par A (trajectoire (positive) de $x = \{x, T(x), ..., T^n(x), ...\}$). A' est mesurable car $A' = \cup_{n\geq 0} T^{-n}(A)$. On a évidement $A \subset A'$ et $T^{-1}(A') \subset A'$. Montrons que $\mu(A\Delta A') = 0$. Par hypothèse $T^{-1}(A) \setminus A$ est de mesure nulle et comme T préserve les ensembles de mesure nulle $T^{-n}(T^{-1}(A) \setminus A)$ est de mesure nulle pour tout $n \geq 0$. Si $x \in A' \setminus A$ alors il existe $n_x \geq 1$ minimal tel que $T^{n_x}(x) \in A$. On a $y = T^{n_x - 1}(x) \notin A$ et $T(y) \in A$ donc $x \in T^{-(n_x-1)}(T^{-1}(A) \setminus A)$ par conséquent $A' \setminus A \subset \cup_{n\geq 0} T^{-n}(T^{-1}(A) \setminus A)$ et $A' \setminus A$ est de mesure nulle. A' est donc un ensemble invariant modulo μ qui diffère de A d'un ensemble de mesure nulle et tel que $T^{-1}(A') \subset A'$.
Introduisons l'ensemble A'' des points de A' dont la trajectoire ne sort pas de A'. On a $A'' \subset A'$. Montrons que A'' est invariant et que A'' diffère de A' d'un ensemble de mesure nulle.
Il est clair que si $x \in A''$ alors $Tx \in A''$ et donc $A'' \subset T^{-1}(A'')$. Inversement, si $x \in T^{-1}(A'')$ alors $y = T(x) \in A''$ qui est inclus dans A'. Par conséquent $x \in T^{-1}A'$ et donc $x \in A'$. Comme la trajectoire de x est la réunion de $\{x\}$ et de la trajectoire de y, elle est incluse dans A' et donc par définition $x \in A''$.
Si $x \in A' \setminus A''$ alors la trajectoire de x sort une première fois de A' en un temps $m_x \geq 1$ donc $y = T^{m_x-1}(x) \in A'$, $T(y) \notin A'$ et $y \in A' \setminus T^{-1}(A')$. On en déduit que $A' \setminus A'' \subset \cup_{n\geq 0} T^{-n}(A' \setminus T^{-1}(A'))$ et que $A' \setminus A''$ est de mesure nulle. $\square$

Remarque A'' est l'ensemble des x de X dont la trajectoire passe une infinité de fois par A. On peut vérifier directement sur cette définition que A'' est invariant.

Exercices. 1. Montrer que l'ensemble $\mathcal{A}_T$ des parties mesurables invariantes est une tribu et que $f : X \to \mathbb{R}$ est mesurable invariante ssi f est mesurable par rapport à la tribu $\mathcal{A}_T$.
2. Soit $T : (X, \mathcal{A}, \mu) \to X$ une application préservant les ensembles de mesure nulle et $f : X \to \mathbb{R}$ une application mesurable telle que $f \circ T = f$ μ-presque partout. Montrer qu'il existe $f' : X \to \mathbb{R}$ mesurable, presque partout égale à f telle que $f' \circ T = f'$ partout.

Proposition 2 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : X \to X$ une application mesurable conservant μ. Pour tout $f : X \to \mathbb{C}$ mesurable et tout $p > 0$ on a $\|f \circ T\|_p = \|f\|_p$.*

Démonstration. Soit μ_T l'image de μ par T. Par hypothèse $\mu = \mu_T$ donc

$$\int_X |f \circ T|^p \, d\mu = \int_X |f|^p \, d\mu_T = \int_X |f|^p \, d\mu. \ \square$$

Notation Lorsque que T préserve la mesure, pour $f \in \mathcal{L}^p$ ou L^p on pose $Tf = f \circ T$.

1.2 Action d'un groupe

Définition 2 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et G un groupe topologique muni de sa tribu borélienne. Une action (à gauche) mesurable de G sur X est une application mesurable*

$$\begin{aligned} G \times X &\to X \\ (g, x) &\to gx \end{aligned}$$

telle que
i. $g_1(g_2x) = (g_1g_2)x$ pour tout $g_1, g_2 \in G$ et tout $x \in X$ et
ii. pour tout $g \in G$, l'application $T_g : x \in X \to gx \in X$ préserve les ensembles de mesure nulle.
Lorsque $G = \mathbb{R}$, et lorsque i et ii sont vérifiées, on dit que l'on a un flot.

On dit aussi que μ est quasi-invariante par G. Comme dans la section précedente, on définit les notions d'invariances par G des mesures, des parties et des fonctions et d'invariance modulo μ des parties et des fonctions. Par exemple $f : X \to \mathbb{R}$ est invariante modulo μ si pour tout $g \in G$ l'ensemble des x de X tels que $f(gx) \neq f(x)$ est de mesure nulle.

Remarque. Les applications T_g sont bijectives et $\mu(A) = 0 \Leftrightarrow \mu(T_g^{-1}A) = 0$.

Exemples.

1. Flots associés aux équations differentielles définies par un champ de vecteur $\mathcal{C}^1$ et dont les solutions sont toutes définies sur $\mathbb{R}$.
2. L'action naturelle de $GL(n, \mathbb{R})$ sur $\mathbb{R}^n$ muni de la mesure de Lebesgue.
3. L'action naturelle de $SL(n, \mathbb{R})$ sur $\mathbb{R}^n$ preserve la mesure de Lebesgue.

La relation entre l'invariance et l'invariance modulo μ subsiste mais nécessite quelques hypothèses.

Proposition 3 *Soit G un groupe locallement compact à base dénombrable agissant mesurablement sur un espace mesuré σ-fini $(X, \mathcal{A}, \mu)$. Si $f : X \to \mathbb{R}$ est une fonction mesurable invariante par G modulo μ alors il existe une fonction invariante $f' : X \to \mathbb{R}$ équivalente à f modulo μ.*

Démonstration. Comme G est à base dénombrable, il existe une probabilité λ sur G ayant une densité strictement positive par rapport à la mesure de Haar (à droite) de G. Notons que λ et la mesure de Haar ont les mêmes ensembles de mesure nulle. On peut appliquer le théorème de Fubini à la mesure produit $\lambda \times \mu$ car G et X sont σ-finis. On peut aussi supposer que f est à valeurs dans l'intervalle $]0, 1[$: il suffit de composer f avec une bijection bi-mesurable de $\mathbb{R}$ dans $]0, 1[$. Montrons que la fonction f est presque partout égale à la fonction $F : X \to \mathbb{R}$ définie par

$$F(x) = \int_G f(gx)\ d\lambda(g).$$

L'ensemble $\{x \in X : f(x) \neq F(x)\}$ est inclus dans l'ensemble Z des $x \in X$ tels que

$$\mathcal{N}_x = \{g \in G : f(gx) \neq f(x)\}$$

soit de mesure non nulle. Or le théorème de Fubini et l'invariance modulo μ de f entraîne que

$$\int_X \lambda(\mathcal{N}_x)\ d\mu(x) = \int_G \mu(\{x \in X : f(gx) \neq f(x)\}\ d\lambda(g) = 0,$$

donc $\mu(Z) = 0$ et $f = F$ sur $Y = X \setminus Z$.

Soit X' l'ensemble des $x \in X$ tels qu'il existe une constante c_x avec $f(gx) = c_x$ pour presque tout g de G. Clairement X' contient Y. Comme pour $h \in G$ et $c \in \mathbb{R}$,

$$\{g \in G : f(ghx) \neq c\} = \{g' \in G : f(g'x) \neq c\}h^{-1},$$

$x \in X'$ ssi $hx \in X'$ et la constante c_x ne dépend pas de x mais uniquement de l'orbite de x. Par conséquent X' est invariant par G et $F(hx) = f(x)$ pour tout $x \in X'$ et $h \in G$. On obtient une fonction invariante presque partout égale à f en prenant F sur X' et une constante arbitraire sur $X \setminus X'$. $\square$

2 Inégalité maximale de Hopf

Théorème 1 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : L^1(\mu) \to L^1(\mu)$ une application linéaire positive telle que $\forall f \in L^1$, $\|Tf\|_1 \leq \|f\|_1$. Pour $f \in L^1(\mu)$ et $n \geq 1$, posons*

$$S_n(f) = \sum_{k=0}^{n-1} T^k f,\ A_n(f) = \frac{1}{n} S_n f,\ M_n(f) = \max(A_1 f, ..., A_n f)$$

et

$$E_n(f) = \{x \in X : M_n(f)(x) \geq 0\},\ E_\infty(f) = \cup_{n \geq 1} E_n(f).$$

Alors pour tout $n \geq 1$,

$$\int_{E_n(f)} f\, d\mu \geq 0 \text{ et } \int_{E_\infty(f)} f\, d\mu \geq 0.$$

Démonstration. Posons $M_n^S(f) = \max(S_1 f, ..., S_n f)$. Pour tout $k = 1, ..., n$ $(M_n^S f)^+ \geq S_k f$ donc $f + T((M_n^S f)^+) \geq f + TS_k f = S_{k+1} f$. Ainsi $f \geq S_j f - T((M_n^S f)^+)$ pour $j = 2, ..., n+1$ mais c'est aussi vrai pour $j = 1$, donc en passant au maximum sur $j = 1, ..., n$, on obtient

$$f \geq M_n^S f - T((M_n^S f)^+).$$

En intégrant sur $E_n(f)$ on obtient

$$\int_{E_n(f)} f \, d\mu \geq \int_{E_n(f)} (M_n^S f - T((M_n^S f)^+) \, d\mu = \int_{E_n(f)} ((M_n^S f)^+ - T((M_n^S f)^+) \, d\mu$$

car $M_n^S f$ est positive sur $E_n(f)$. Par conséquent

$$\int_{E_n(f)} f \, d\mu \geq \int_{E_n(f)} (M_n^S f)^+ d\mu - \int_{E_n(f)} T((M_n^S f)^+) \, d\mu \geq 0$$

car T contracte la norme L^1. $\square$

Corollaire 1 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : X \to X$ une application mesurable conservant μ. Alors pour tout $f \in \mathcal{L}_1(\mu)$ et tout $\alpha > 0$ on a*

$$\mu(\{M_n f \geq \alpha\}) \leq \frac{1}{\alpha} \|f\|_1 .$$

Démonstration. On considère l'application $T : L^1(\mu) \to L^1(\mu)$ définie par $Tf = f \circ T$. Soit $A = \{M_n f \geq \alpha\}$. A est mesurable et comme f est dans L^1, A est de mesure finie, donc 1_A appartient à L^1. Considérons la fonction $g = f - \alpha 1_A$. On a $\int_{E_n(g)} g \, d\mu \geq 0$ d'après l'inégalité maximale de Hopf. Par définition de A, pour tout $x \in A$, il existe $k \leq n$ tel que $A_k f(x) \geq \alpha$, donc $A_k g(x) = A_k(f - \alpha 1_A)(x) \geq A_k(f - \alpha 1_X)(x) \geq 0$, d'où $A \subset E_n(g)$ et

$$\|f\|_1 \geq \int_{E_n(g)} f \, d\mu \geq \int_{E_n(g)} \alpha 1_A \, d\mu = \alpha \, \mu(A). \square$$

Remarque. En modifiant légèrement la démonstration du corollaire, on peut prouver que le corollaire est valable plus généralement pour une application linéaire $T : L^1(\mu) \to L^1(\mu)$ telle que $\|Tf\|_1 \leq \|f\|_1$ pour tout $f \in L^1(\mu)$ (T est une contraction), $Tf \geq 0$ pour tout $f \geq 0$ (T est positive) et $\|Tf\|_\infty \leq \|f\|_\infty$ pour tout $f \in L^1(\mu) \cap L^\infty(\mu)$.

3 Théorèmes de Birkhoff et von Neumann

Théorème 2 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, $T : X \to X$ une application mesurable conservant μ et $f \in \mathcal{L}^1(\mu)$. Alors la suite $(A_n f)_n$ converge μ-presque partout vers une fonction T-invariante $\overline{f}$ vérifiant $\left\|\overline{f}\right\|_1 \leq \|f\|_1$ et*

$$\forall A \in \mathcal{A}, \; T^{-1}(A) = A \textit{ et } \mu(A) < \infty \Rightarrow \int_A f \, d\mu = \int_A \overline{f} \, d\mu.$$

Démonstration. Supposons d'abord que f soit à valeurs réelles. Comme

$$A_{n+1} f = \frac{1}{n+1} f + \frac{n}{n+1} A_n f \circ T,$$

on a

$$\begin{aligned}\liminf A_n f = \liminf A_{n+1} f &= \liminf\left(\frac{1}{n+1} f + \frac{n}{n+1} A_n f \circ T\right) \\ &= \liminf (A_n f \circ T) = (\liminf A_n f) \circ T\end{aligned}$$

les fonctions $g = \liminf A_n f$ et $h = \limsup A_n f$ (à valeurs dans $\mathbb{R} \cup \{\pm\infty\}$), sont donc T-invariantes. Soit $\alpha > 0$. L'ensemble $D_h = \{h > \alpha\}$ est inclus dans $\cup_{n \geq 1} \{M_n f > \alpha\}$. D'après le corollaire de l'inégalité maximale on a

$$\mu(\{M_n f \geq \alpha\}) \leq \frac{1}{\alpha} \|f\|_1$$

donc, comme la suite $\{M_n f \geq \alpha\}$ est croissante, $\mu(D_h) \leq \frac{1}{\alpha}\|f\|_1$ et $\mu(h=\infty)=0$. De même en considérant $-f$ on voit que $\mu(\{g \leq \beta\}) \leq \frac{1}{|\beta|}\|f\|_1$ pour tout $\beta < 0$.
Pour montrer que la suite $(A_n f)_{n\geq 1}$ converge presque partout il suffit de prouver que si $a < b$ sont deux rationnels alors $B = \{g < a < b < h\}$ est de mesure nulle. Si $b > 0$ alors $\{h > b\}$ est de mesure finie (prendre $\alpha = b$ dans ce qui précède), de même si $a < 0$ alors $\{g < a\}$ est de mesure finie donc B est de mesure finie. Comme g et h sont invariantes, B est invariant. Considérons la fonction $f' = (f-b)1_B$. Si $x \in B$ alors il existe n tel que $A_n f(x) > b$ donc

$$\begin{aligned} 0 < \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k(x) - b &= \frac{1}{n}\sum_{k=0}^{n-1}(f\circ T^k(x) - b) \\ &= \frac{1}{n}\sum_{k=0}^{n-1}(f\circ T^k(x)-b)1_B(x) \\ &= \frac{1}{n}\sum_{k=0}^{n-1}(f-b)\circ T^k(x)1_B\circ T^k(x) = A_n f'(x). \end{aligned}$$

On en déduit que $B \subset E_\infty(f') = \cup_{n\geq 0}\{A_n f' \geq 0\}$. Donc

$$\int_B f'\,d\mu = \int_{E_\infty(f')} f' 1_B\,d\mu = \int_{E_\infty(f')} f'\,d\mu$$

qui est ≥ 0 d'après le théorème de Hopf, d'où $\int_B f\,d\mu \geq b\mu(B)$.
Le même raisonnement appliquer à la fonction $f'' = (a-f)1_B$ montre que $\int_B f\,d\mu \leq a\mu(B)$. On en déduit que $\mu(B) = 0$ et que la $(A_n f)_n$ converge presque partout.
On a $\overline{f} = \lim A_n f = g = h$ presque partout donc $\overline{f}$ est invariante.
Montrons que $\|\overline{f}\|_1 \leq \|f\|_1$. En écrivant $f = f^+ - f^-$ on se ramène au cas $f \geq 0$. Si f est positive, d'après le lemme de Fatou on a

$$\|\overline{f}\|_1 = \int_X \lim A_n f\,d\mu = \int_X \liminf A_n f\,d\mu \leq \liminf \int_X A_n f\,d\mu = \|f\|_1.$$

Il reste à prouver que $\forall A \in \mathcal{A},\ T^{-1}(A) = A$ et $\mu(A) < +\infty \Rightarrow \int_A f\,d\mu = \int_A \overline{f}\,d\mu$. On peut supposer $f \geq 0$. Pour $m \geq 0$, posons $f_m = \inf(f,m)$ et $g = 1_A f_m$. Comme A est invariant, $A_n g = 1_A A_n f_m \to 1_A \overline{f}_m$ μ-p.p.. De plus $\mu(A) < \infty$, donc d'après le théorème de convergence dominée

$$\int_X g\,d\mu = \int_X A_n g\,d\mu \to \int_X 1_A \overline{f}_m d\mu \leq \int_X 1_A \overline{f} d\mu.$$

Comme l'inégalité est vrai pour tout m,

$$\int_X f 1_A \leq \int_A 1_A \overline{f}\,d\mu.$$

L'inégalité inverse découle du lemme de Fatou.
Finalement si f est à valeurs complexes on applique les résultats précédents à $\operatorname{Re} f$ et $\operatorname{Im} f$ et on observe que $|\overline{f}| \leq \lim A_n |f|$. $\square$

Exemple $X = [0,1[$, $Tx = 10x \bmod 1$, $f = \sum_{k=0}^{9} k 1_{[\frac{k}{10},\frac{k+1}{10}[}$. $A_n f(x) =$ moyennes des n premières décimales de x. Pour presque tout x, la suite $(A_n f(x))_{n\geq 1}$ converge vers une fonction invariante $\overline{f}(x)$.

Corollaire 2 *Soit $(X, \mathcal{A}, T)$ un espace mesuré et $T : X \to X$ une application mesurable conservant μ. Alors pour tout $f \in \mathcal{L}^1(\mu)$, $\frac{1}{n} f\circ T^n(x)$ tend vers 0 pour presque tout x de X.*

Démonstration. C'est une conséquence du théorème de Birkhoff grâce à la relation $\frac{1}{n} f\circ T^n = S_n f - \frac{n-1}{n} S_{n-1} f$. $\square$

Corollaire 3 *Soit $(X, \mathcal{A}, T)$ un espace probabilisé et $T : X \to X$ une application mesurable conservant μ. Appelons $\mathcal{A}_T$ la sous tribu des ensembles T-invariants. Alors $\overline{f} = E(f|\mathcal{A}_T)$ μ-presque partout.*

Exercice. Soit $(X, \mathcal{A}, \mu)$ un espace probabilisé et $T : (X, \mathcal{A}, \mu) \to X$ une application bijective bimesurable conservant la mesure. Montrer que pour tout f de L^1 $\lim \frac{1}{n} \sum_{k=0}^{n} f \circ T^k = \lim \frac{1}{n} \sum_{k=0}^{n} f \circ T^{-k}$ presque partout.

Corollaire 4 ***Théorème de von Neumann.***
Soit $p \in [1, +\infty[$ et T une application d'un espace probabilisé $(X, \mathcal{A}, \mu)$ préservant la mesure. Pour tout $f \in L^p$, la suite $(A_n f)_n$ converge dans L^p vers une fonction invariante $\overline{f}$.

Démonstration. Si f est bornée alors $\|A_n f\|_\infty \le \|f\|_\infty$. D'après le théorème de Birkhoff, la suite $A_n f$ converge presque partout vers une fonction $\overline{f} = E(f|\mathcal{A}_T)$ où $\mathcal{A}_T$ est la tribu des ensembles invariants, le théorème de Lebesgue montre alors que $\int \left|A_n f - \overline{f}\right|^p \, d\mu \to 0$; on en déduit que $A_n f$ converge vers f dans L^p. Comme les A_n sont des applications linéaires de normes ≤ 1 et comme L^∞ est dense dans L^p, $A_n f$ converge vers $E(f|\mathcal{A}_T)$ dans L^p. $\square$

Théorème 3 ***Théorème de von Neumann dans un espace de Hilbert.***
Si T est une contraction d'un espace de Hilbert H, alors pour tout $u \in H$, la suite $(A_n u)_{n\ge1}$ converge dans H vers un élément $\overline{u}$ qui est la projection orthogonale de u sur $\ker(I - T)$. De plus $\ker(I - T)^\perp = \overline{\mathrm{Im}(I - T)}$.

Démonstration-exercice :
1. Montrer que si $u \in \overline{\mathrm{Im}(I - T)}$ alors $A_n u \to 0$ dans H.
2. Montrer que si $u \in H$ alors la suite $(A_n u)_n$ admet une valeur d'adhérence faible $v \in \ker(I - T)$.
3. Montrer que $u - v \in \overline{\mathrm{Im}(I - T)}$.
4. Montrer que pour tout $u \in H$, la suite $(A_n u)_{n\ge1}$ converge vers $v \in \ker(I - T)$.
5. Montrer que $\ker(I - T) = \ker(I - T^*)$. En déduire que $\ker(I - T)^\perp = \overline{\mathrm{Im}(I - T)}$. $\square$

Corollaire 5 ***Théorème de von Neumann dans*** L^2.
Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, $T : X \to X$ une application mesurable conservant μ et $f \in L^2(\mu)$. Alors la suite $(A_n f)_n$ converge dans $L^2(\mu)$ vers une fonction invariante $\overline{f}$ qui est la projection orthogonale de f sur le sous-espace des fonctions invariantes de $L^2(\mu)$.

3.1 Cas des flots

Théorème 4 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini, $(\phi_t)_{t\in\mathbb{R}} : X \to X$ un flot agissant sur X conservant μ et $f \in \mathcal{L}^1(\mu)$. Il existe un ensemble de mesure nulle $\mathcal{N}$ tel que pour tout $T > 0$, la fonction*

$$A_T f(x) = \frac{1}{T} \int_0^T f \circ \phi_t(x) \, dt$$

est définie sur $X \setminus \mathcal{N}$, appartient à $\mathcal{L}^1(\mu)$ et lorsque T tend vers l'infini, $A_T f$ converge presque partout vers une fonction $\overline{f}$ invariante par le flot, vérifiant $\left\|\overline{f}\right\|_1 \le \|f\|_1$ et

$$\forall A \in \mathcal{A},\ T^{-1}(A) = A \text{ et } \mu(A) < \infty \Rightarrow \int_A f \, d\mu = \int_A \overline{f} \, d\mu.$$

Démonstration. Pour chaque $T > 0$, considérons l'ensemble X_T des $x \in X$ tels que $\int_0^T |f \circ \phi_t(x)| \, dt < +\infty$. D'après le théorème de Fubini $\int_X \int_0^T |f \circ \phi_t(x)| \, dt d\mu(x) = T \|f\|_1 < \infty$ donc $\mathcal{N}_T = X \setminus X_T$ est de mesure nulle. Par conséquent l'ensemble $\mathcal{N} = \cup_{n\in\mathbb{N}^*} \mathcal{N}_n$ est de mesure nulle et, pour tout $x \notin \mathcal{N}$ et tout $T > 0$, $A_T f(x) = \frac{1}{T} \int_0^T f \circ \phi_t(x) \, dt$ est définie. Comme $A_T f(x) = A_T f^+(x) - A_T f^-(x)$ sur $X \setminus \mathcal{N}$ on peut supposer que f est positive.

Pour $s > 0$ fixé, considérons la fonction $g_s : X \to \mathbb{R}_+ \cup \{+\infty\}$ définie par $g_s(x) = \frac{1}{s} \int_0^s f \circ \phi_t(x) \, dt$. Comme g_s est intégrable, on peut appliquer le théorème de Birkhoff à l'application préservant la mesure ϕ_s et la fonction g_s :

$$\frac{1}{ns} \int_0^{ns} f \circ \phi_t(x) dt = \frac{1}{ns} \sum_{k=0}^{n-1} \int_{ks}^{(k+1)s} f \circ \phi_t(x) \, dt = \frac{1}{n} \sum_{k=0}^{n-1} g_s \circ \phi_s^k(x)$$

converge pour presque tout x vers une fonction $\overline{f_s} \in \mathcal{L}^1(\mu)$ ϕ_s-invariante. D'après le corollaire du théorème de Birkhoff nous savons aussi que $\frac{1}{n} g_s \circ \phi_s^n(x)$ tend vers 0 pour presque tout x. Or pour $ns \leq T \leq (n+1)s$, nous avons

$$\left| \int_0^T f \circ \phi_t(x)\ dt - \int_0^{ns} f \circ \phi_t(x)\ dt \right| \leq \int_{ns}^T f \circ \phi_t(x)\ dt \leq s g_s \circ \phi_s^n(x),$$

donc

$$\begin{aligned}
\left| \frac{1}{T} \int_0^T f \circ \phi_t(x)\ dt - \frac{1}{ns} \int_0^{ns} f \circ \phi_t(x)\ dt \right| &\leq \frac{1}{T} \left| \int_0^T f \circ \phi_t(x)\ dt - \int_0^{ns} f \circ \phi_t(x)\ dt \right| \\
&+ \frac{T - ns}{Tns} \int_0^{ns} f \circ \phi_t(x)\ dt \\
&\leq \frac{s}{T} g_s \circ \phi_s^n(x) + \frac{1}{n^2 s} \sum_{k=0}^{n-1} s g_s \circ \phi_s^k(x) \\
&\leq \frac{1}{n} g_s \circ \phi_s^n(x) + \frac{1}{n} \left(\frac{1}{n} \sum_{k=0}^{n-1} g_s \circ \phi_s^k(x) \right)
\end{aligned}$$

tend presque partout vers 0. Par conséquent, $\frac{1}{T} \int_0^T f \circ \phi_t(x)\ dt$ converge presque partout vers la fonction $\overline{f_s}$. Par unicité de la limite presque-sûre, les fonctions $\overline{f}_s$ sont égales presque partout ce qui montre que la limite est invariante par le flot. □

Remarque. Si f et g sont deux fonctions de $\mathcal{L}^1(\mu)$ égales presque partout, le théorème de Fubini montre que pour tout $T > 0$, les fonctions $A_T f$ et $A_T g$ sont égales presque partout.

Avec une démonstration analogue on prouve le théorème de von Neumann dans le cas des flots.

Théorème 5 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini, $(\phi_t)_{t\in\mathbb{R}} : X \to X$ un flot agissant sur X conservant μ et $f \in \mathcal{L}^2(\mu)$. Il existe un ensemble de mesure nulle $\mathcal{N}$ tel que pour tout $T > 0$, la fonction*

$$A_T f(x) = \frac{1}{T} \int_0^T f \circ \phi_t(x)\ dt$$

est définie sur $X \setminus \mathcal{N}$, appartient à $\mathcal{L}^2(\mu)$ et lorsque T tend vers l'infini, $A_T f$ converge dans $L^2(\mu)$ vers la projection orthogonale $\overline{f}$ de f sur le sous espace de $L^2(\mu)$ des fonctions invariante par le flot.

Démonstration : exercice.

4 Ergodicité

Définition 3 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : X \to X$ une application mesurable préservant les ensembles de mesure nulle. On dit T est ergodique si tous les ensembles A mesurables et invariants sont tels que $\mu(A)$ ou $\mu(X \setminus A) = 0$.*

Exemples. 1. Soit $X = \mathbb{T}^1$ et $T(x) = x + \theta$. T est ergodique ssi θ est irrationnel . Montrons que $\theta \notin \mathbb{Q} \Rightarrow \mathbb{T}$ ergodique. Soit A invariant. On a $1_A = \sum_{m\in\mathbb{Z}} c_m e^{imx}$ la convergence ayant lieu dans L^2. On a

$$1_A = 1_{T^{-1}(A)} = 1_A \circ T.$$

Comme la composition par T est continue sur L^2, on a

$$1_A(x) = 1_A \circ T(x) = \sum_{m\in\mathbb{Z}} c_m e^{i2\pi m(x+\theta)} = \sum_{m\in\mathbb{Z}} (c_m e^{i2\pi m\theta}) e^{i2\pi mx}$$

donc par unicité du développement en série de Fourier

$$c_m e^{i2\pi m\theta} = c_m \text{ pour tout } m$$

donc $c_m = 0$ pour tout $m \neq 0$ et $1_A = c_0$.$\square$
2. $T : x \in \mathbb{T}^1 \to 2x \in \mathbb{T}^1$. On utilise Fourier. Soit A invariant. On a

$$1_A \circ T(x) = \sum_{m \in \mathbb{Z}} c_m e^{i2\pi m 2x} = \sum_{m \text{ pair}} c_{m/2} e^{i2\pi m x}.$$

Par unicité des coefficients de Fourier, on obtient $c_m = 0$ si m est impair et $c_m = c_{m/2}$ si m est pair. Par conséquent $c_m = 0$ pour tous les $m \neq 0$.

Exercices. (voir exemples-exercices 7 page 1)
1. Considérons $(X_b, \mathcal{T})$ muni d'une probabilité produit quelconque μ et du décalage S.
On rappelle le résultat suivant : si deux mesures de même masse totale finie définies sur une même tribu coïncident sur une partie stable par intersection qui engendre la tribu alors elles coïncident sur toute la tribu.
a. Soit $A \in S^{-n}(\mathcal{T})$ et $C = [a_0, ..., a_{n-1}]$ un cylindre. Montrer que $\mu(A \cap C) = \mu(A)\mu(C)$.
b. Soit $A \in \mathcal{T}_\infty = \cap_{n \geq 0} S^{-n}(\mathcal{T})$. Montrer que $\mu(A) = 0$ ou 1. En déduire que S est ergodique.
2. On veut montrer que l'addition de 100... dans $(X_b, +)$ muni de la probabilité équilibrée produit est ergodique. Soit $\tau : x \in X_b \to x + 100....$
a. Soit $A \in \mathcal{T}$ une partie τ-invariante. Montrer que si C et C' sont deux cylindres de même longueur, alors il existe $n \in \mathbb{N}$ tel que $\tau^n(A \cap C) = A \cap C'$.
b. Soit $f : X_b \to \mathbb{R} \in \mathcal{L}^1(X_b)$. Pour chaque n, on définit la fonction $f_n : X_b \to \mathbb{R}$ par f_n est constante sur chaque cylindre C de longueur n de valeur $\frac{1}{\mu(C)} \int_C f d\mu$. Montrer que si f est continue pour la topologie produit alors la suite (f_n) converge uniformément vers f. En déduire que pour tout $f \in \mathcal{L}^1(X_b)$, (f_n) tend vers f dans $\mathcal{L}^1(X_b)$.
c. En déduire que τ est ergodique.

Proposition 4 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $T : X \to X$ une application mesurable préservant les ensembles de mesure nulle. T est ergodique ssi les seules fonctions mesurables invariantes sont les fonctions constantes presque sûrement.*

Démonstration. Si les seules fonctions invariantes sont presque sûrement constantes alors la fonction indicatrice d'une partie invariantes est presque surement constante ce qui implique T est ergodique. Réciproquement, supposons T est ergodique. Soit $f : X \to \mathbb{R}$ une fonction mesurable invariante. Pour tout $r \in \mathbb{Q}$, l'ensemble $X_r = \{x \in X : f(x) \leq r\}$ est invariant donc soit $\mu(X_r) = 0$ soit $\mu(X_r^C) = 0$. Soit $c = \sup\{r \in \mathbb{Q} : \mu(X_r) = 0\}$. Si $\mu(X) = 0$, il n'y a rien à prouver, sinon c est fini car $\mu(X_r) \to \mu(X)$ quand $r \to \infty$. Pour tout $n > 0$, on a donc $\mu(\{f < c - \frac{1}{n}\}) = \mu(\{f > c + \frac{1}{n}\}) = 0$ ce qui entraîne $f = c$ presque sûrement. $\square$

Proposition 5 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré de mesure finie et $T : X \to X$ une application mesurable préservant la mesure. Si T est ergodique alors pour tout $f \in L^1$ on a*

$$A_n f \to \frac{1}{\mu(X)} \int_X f \, d\mu \ \mu\text{-presque partout .}$$

Démonstration. La fonction $\overline{f} = \lim A_n f$ est invariante. D'après la proposition précédente $\overline{f}$ est presque surement égale à une constante c. De plus, $\int_X \overline{f} \, d\mu = \int_X f \, d\mu$ d'où $c = \frac{1}{\mu(X)} \int_X f \, d\mu$.
$\square$

Exercice. Montrer que pour presque tout $x \in [0, 1[$, la suite des moyennes des n premières décimales de x tend vers $4, 5$.

Proposition 6 *Soit $(X, \mathcal{A}, \mu)$ un espace probabilisé et $T : X \to X$ une application mesurable préservant la mesure. Les propriétés suivantes sont équivalentes :*
1. T est ergodique.
2. Pour tout $A, B \in \mathcal{A}$, $\mu(A) > 0$ et $\mu(B) > 0 \Rightarrow \exists k \in \mathbb{N}$, $\mu(T^{-k}(A) \cap B) > 0$.
3. Pour tout $A, B \in \mathcal{A}$, $\mu(A) > 0 \Rightarrow \mu(X \setminus \cup_{k \geq 0} T^{-k}(A)) = 0$.
4. Pour tout $A, B \in \mathcal{A}$, $\lim \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$.

Démonstration. Clairement, 2 et 3 sont équivalents. Supposons T ergodique. Soit A et B telles que $\mu(A) > 0$ et $\mu(B) > 0$. Posons $C = \cup_{k \geq 0} T^{-k}(A)$. On a $T^{-1}(C) \subset C$ et comme

$\mu(T^{-1}(C)) = \mu(C)$, on a $T^{-1}(C) = C \mod 0$ d'où $C = X \mod 0$ et 3 en découle. Pour la réciproque il suffit d'appliquer 2 à une partie invariante A et à sont complémentaire $B = A^C$.
Soit $A, B \in \mathcal{A}$. D'après le théorème de Von Neumann dans L^2, la suite

$$\langle A_n 1_A | 1_B \rangle = \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B)$$

est convergente. Si T est ergodique alors cette limite vaut $\int_X \mu(A) 1_B \, d\mu = \mu(A)\mu(B)$. Réciproquement si A est invariant, alors $\mu(A \cap B) = \lim \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$ pour tout $B \in \mathcal{A}$. Donc $0 = \mu(A \cap A^C) = \mu(A)\mu(A^C)$ et $\mu(A)\mu(A^C) = 0$. □

4.1 Cas des groupes

Définition 4 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré muni d'une action mesurable par un groupe topologique G. On dit que l'action de G est ergodique si tous les ensembles mesurables invariants A sont tels que $\mu(A)$ ou $\mu(X \setminus A) = 0$.*

Remarques. 1. Lorsque X est σ-fini et G est localement compact à base dénombrable, on peut remplacer invariant par invariant modulo μ.
2. L'ergodicité de l'action de G n'entraîne pas l'ergodicité des applications $T_g(x) = gx$. Prenons $\mathbb{R}$ agissant sur lui même par $(t, x) \to t + x$. Cette action est ergodique alors que pour tout t l'application T_t n'est pas ergodique.

Proposition 7 *Soit G un groupe locallement compact à base dénombrable agissant mesurablement sur un espace mesuré σ-fini $(X, \mathcal{A}, \mu)$. L'action de G est ergodique ssi les seules fonctions $f : X \to \mathbb{R}$ mesurables invariantes sont presque sûrement constantes.*

La démonstration est identique à celle faite dans les cas des applications préservant les ensembles de mesures nulles.

Exemple-exercice. On fait agir le groupe $G = SL(2, \mathbb{Z})$ sur le tore $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ muni de la mesure de Lebesgue : $gx =$ la classe du produit de la matrice g par un représentant de x. Identifions le tore $\mathbb{T}^2$ avec le carré semi-ouvert $[0, 1[^2$. Soit A une partie mesurable de $\mathbb{T}^2$ de mesure > 0 et invariante par l'action de G. Soit

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$s \in [0, 1[$ et $A_s = [0, 1[\times \{s\}$.
1. Montrer que pour presque tout s la mesure de Lebesgue de A_s est soit 0, soit 1. Notons I l'ensemble des s tels que A_s soit de mesure 1. On a donc $A = [0, 1[\times I$ modulo 0.
2. Conclure.

5 Mélanges

Définition 5 *Soit T une application préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$.*
1. On dit que T est faiblement mélangeante si $\forall A, B \in \mathcal{A}$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \mu(T^{-k}(A) \cap B) - \mu(A)\mu(B) \right| = 0.$$

2. On dit que T est mélangeante si $\forall A, B \in \mathcal{A}$,

$$\lim_{n \to \infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B).$$

Comme

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \mu(T^{-k}(A) \cap B) - \mu(A)\mu(B) \right| = 0 \Rightarrow \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$$

on a :

Proposition 8 *Si T est faiblement mélangeante alors T est ergodique.*

Pour caractériser le mélange et le mélange faible nous utilserons de trois lemmes.

Lemme 1 *Soit $(E, \|.\|)$ un espace vectoriel normé et $(B_n)_n$ une suite de forme bilinéaires sur E telles que $\forall x, y \in E$, $\forall n \in \mathbb{N}$, $|B_n(x,y)| \leq C\,\|x\|\,\|y\|$ où C est une constante.*
Supposons que pour tous les x et tous les y d'une partie F de E on ait
soit i. $\lim_{n\to\infty} B_n(x,y) = B(x,y)$,
soit ii. $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |B_k(x,y) - B(x,y)| = 0$,
soit iii. $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} B_k(x,y) = B(x,y)$.
Alors dans chacun des cas l'égalité s'étend aux vecteurs x et y de l'adhérence de l'espace vectoriel engendré par F.

Démonstration. Appelons G l'espace vectoriel engendré par F. Il est claire que i et ii s'étendent à G. Pour ii, soit $x = \sum_{i=1}^{p} \lambda_i x_i$ et $y = \sum_{j=1}^{q} \mu_j y_j$ deux combinaisons linéaires d'éléments de F. On a

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1} |B_n(x,y) - B(x,y)| &= \frac{1}{n}\sum_{k=0}^{n-1} \left| \sum_{i=1,j=1}^{p,q} \lambda_i\mu_j (B_n(x_i,y_j) - B(x_i,y_j)) \right| \\
&\leq \frac{1}{n}\sum_{k=0}^{n-1} \sum_{i=1,j=1}^{p,q} |\lambda_i\mu_j|\,|B_n(x_i,y_j) - B(x_i,y_j)| \\
&= \sum_{i=1,j=1}^{p,q} |\lambda_i\mu_j| \frac{1}{n}\sum_{k=0}^{n-1} |B_n(x_i,y_j) - B(x_i,y_j)|,
\end{aligned}$$

donc ii s'étend à G. Les extensions à $\overline{G}$ se prouvent toutes les trois de manière analogue. Soit x et $y \in \overline{G}$ et soit $\varepsilon > 0$. Il existe des vecteurs x_ε et y_ε de G tels que $\|x - x_\varepsilon\| \leq \varepsilon$ et $\|y - y_\varepsilon\| \leq \varepsilon$. On conclut grâce à la relation

$$\begin{aligned}
B_k(x,y) - B(x,y) &= B_k(x,y) - B_k(x_\varepsilon,y) + B_k(x_\varepsilon,y) - B_k(x_\varepsilon,y_\varepsilon) \\
&\quad + B_k(x_\varepsilon,y_\varepsilon) - B(x_\varepsilon,y_\varepsilon) + B(x_\varepsilon,y_\varepsilon) - B(x,y_\varepsilon) \\
&\quad + B(x,y_\varepsilon) - B(x,y),
\end{aligned}$$

combinée avec l'inégalité triangulaire et les inégalités $|B_k(a,b) - B_k(a',b)| \leq C\,\|a - a'\|\,\|b\|$. $\square$

Lemme 2 *Soit (a_n) et (b_n) deux suites réelles positives et bornées. Si les suites $|a_n - a|$ et $|b_n - b|$ tendent vers 0 au sens de Cesaro alors $|a_nb_n - ab|$ tend vers 0 au sens de Cesaro.*

Preuve du lemme. On écrit simplement

$$\begin{aligned}
\frac{1}{n}\sum_{k<n} |a_nb_n - ab| &= \frac{1}{n}\sum_{k<n} |a_n(b_n - b) + b(a_n - a)| \\
&\leq \frac{1}{n}\sum_{k<n} |a_n(b_n - b)| + \frac{1}{n}\sum_{k<n} |b(a_n - a)| \\
&\leq M \times (\frac{1}{n}\sum_{k<n} |b_n - b| + \frac{1}{n}\sum_{k<n} |a_n - a|). \ \square
\end{aligned}$$

Lemme 3 *Soit (a_n) une suite réelle positive et bornée. Si la suite (a_n^2) tend vers 0 au sens de Cesaro alors (a_n) tend vers 0 au sens de Cesaro.*

Preuve du lemme. Appelons M un majorant de cette suite. Soit $\varepsilon > 0$. Posons $A_n = \{k < n : a_n > \varepsilon\}$. Il existe N tel que pour tout $n \geq N$ on ait $\frac{1}{n}\sum_{k<n} a_n^2 \leq \varepsilon^3$, donc pour tout $n \geq N$, on a $\operatorname{card} A_n \leq \varepsilon n$. Par conséquent pour $n \geq N$,

$$\begin{aligned}
\frac{1}{n}\sum_{k<n} a_n &\leq \frac{1}{n}(\varepsilon \times (\text{le nb de } k \text{ tq } a_k \leq \varepsilon) + M \times (\text{le nb de } k \text{ tq } a_k > \varepsilon) \\
&\leq \frac{1}{n}(\varepsilon \times n + M \times \varepsilon n) = (M+1)\varepsilon. \ \square
\end{aligned}$$

On peut reformuler les notions de mélange et mélange faible.

Théorème 6 *Soit T une application préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$. On a les équivalences :*
1. T est mélangeante.
2. Pour toutes fonctions $f, g \in L^2(\mu)$, $\lim_{n\to\infty} \int_X f \circ T^n g \ d\mu = \int_X f \ d\mu \int_X g \ d\mu$.
3. Pour toutes fonctions $f, g \in L^2(\mu)$ d'intégrales nulles, $\lim_{n\to\infty} \int_X f \circ T^n g \ d\mu = 0$.

Théorème 7 *Soit T une application préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$. On a les équivalences :*
1. T est faiblement mélangeante.
2. Pour toutes fonctions $f, g \in L^2(\mu)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \int_X f \circ T^k \ g \ d\mu - \int_X f \ d\mu \int_X g \ d\mu \right| = 0.$$

3. Pour toute fonction $f \in L^2(\mu)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \int_X f \circ T^k \ f \ d\mu - (\int_X f \ d\mu)^2 \right| = 0.$$

Les deux théorèmes se prouvent en appliquant le lemme de bilinéarité à l'ensemble des fonctions indicatrices de parties mesurables qui forme une partie totale de L^2.

Le résultat suivant montre que la vérification du caractère ergodique, mélangeant où faiblement mélangeant d'une application peut se faire sur une classe de parties plus petite que l'ensembles des parties mesurables.

Théorème 8 *Soit T une application préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$ et $\mathcal{S}$ une semi-algèbre engendrant $\mathcal{A}$.*
1. Si $\forall A, B \in \mathcal{S}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$$

alors T est ergodique.
2. Si $\forall A, B \in \mathcal{S}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left|\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)\right| = 0$$

alors T est faiblement mélangeante.
3. Si $\forall A, B \in \mathcal{S}$,

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B)$$

alors T est mélangeante.

Rappels :
Une partie $\mathcal{S}$ de $\mathcal{P}(X)$ est une semi-algèbre si $\emptyset \in \mathcal{S}$, $\mathcal{S}$ est stable par intersection finie et si le complémentaire d'un élément de $\mathcal{S}$ est une réunion finie d'éléments disjoints de $\mathcal{S}$.

Théorème 9 *Soit $(X, \mathcal{A}, \mu)$ un espace probabilisé. Si $\mathcal{S}$ est une semi-algèbre engendrant $\mathcal{A}$ alors*

$$\forall \varepsilon > 0, \ \forall A \in \mathcal{A}, \ \exists B_1, ..., B_n \in \mathcal{S}, \ \textit{deux à deux disjoints}, \ \mu(A \Delta (B_1 \cup ... \cup B_n)) \leq \varepsilon.$$

Démonstration.
Le théorème sur les semi algèbre montre que l'espace vectoriel engendré par les fonctions indicatrices de parties appartenant à $\mathcal{S}$, contient les fonctions indicatrices de parties mesurables dans son adhérence. On conclut avec le lemme de bilinéarité. □

Exemples-Exercices.
1. $x \to 2x$ est mélangeante sur $\mathbb{T}^1$. On prend pour $\mathcal{S}$ les intervalles du type $I_{n,k} = [\frac{k}{2^n}, \frac{k+1}{2^n}[$. On a $T^{-m} I_{n,k} = \cup_{j \in E} I_{n+m,j}$ où $E = \{j \in \{0, ..., 2^{n+m} - 1\} : j = k + l2^n\} = \{j = k + l2^n : l < 2^m\}$. Ainsi pour $m \geq p$ on a

$$T^{-m}(I_{n,k}) \cap I_{p,q} = \cup_{j \in F} I_{n+m,j}$$

où $F = \{j = k + l2^n : l < 2^m \text{ et } (k + l2^n)2^{-n-m} \in I_{p,q}\}$.
Un compte précis donne card $F = 2^{m-p}$, d'où $\mu(T^{-m}(I_{n,k}) \cap I_{p,q}) = 2^{m-p} \times 2^{-n-m} = \mu(I_{n,p})\mu(I_{p,q})$.
2. Montrer à l'aide de Fourier que si $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ alors la application du tore $\mathbb{T}^2$, $x \to Ax$, est mélangeante.
3. Montrer que les translations du tore $\mathbb{T}^1$ ne sont pas mélangeantes.

Théorème 10 *Soit T une application préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$. On a les équivalences :*
1. T est faiblement mélangeante.
2. $T \times T$ est ergodique.
3. $T \times T$ est faiblement mélangeante.

Démonstration. Montrons que $1 \Rightarrow 3$. D'après le théorème précédent, il suffit de considérer des parties A et B du type $A = A_1 \times A_2$ et $B = B_1 \times B_2$. On a

$$\begin{aligned} \mu \times \mu((T \times T)^{-n}(A) \cap B) &= \mu \times \mu((T^{-n}(A_1) \cap B_1 \times T^{-n}(A_2) \cap B_2) \\ &= \mu(T^{-n}(A_1) \cap B_1)\mu(T^{-n}(A_2) \cap B_2). \end{aligned}$$

On utilise le premier lemme sur les moyennes de Césaro avec $|a_n - a| = |\mu(T^{-n}(A_1) \cap B_1) - \mu(A_1)\mu(B_1)|$ et $|b_n - b| = |\mu(T^{-n}(A_2) \cap B_2) - \mu(A_2)\mu(B_2)|$ et on obtient la première implication.
On a évidement $3 \Rightarrow 2$. Montrons $2 \Rightarrow 1$. Nous avons

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu \times \mu((T \times T)^{-k}(A \times X) \cap B \times X) = \mu(A) \times \mu(B),$$

de même

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B)^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu \times \mu((T \times T)^{-k}(A \times A) \cap B \times B) = \mu(A)^2 \times \mu(B)^2.$$

Par conséquent

$$\begin{aligned} &\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B))^2 \\ &= \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mu(T^{-k}(A) \cap B)^2 + \mu(A)^2\mu(B)^2 - 2\mu(T^{-k}(A) \cap B)\mu(A)\mu(B)) \\ &= 0. \end{aligned}$$

On conclut avec le second lemme sur les moyennes de Césaro. $\square$

Théorème 11 *Soit T une bijection bimesuarble et préservant la mesure d'un espace probabilisé $(X, \mathcal{A}, \mu)$. On a les équivalences :*
1. T est faiblement mélangeante.
2. Les seules fonctions propres de l'opérateur

$$\begin{aligned} U_T &: L^2(\mu) \to L^2(\mu) \\ &: f \to f \circ T \end{aligned}$$

sont les fonctions constantes.

Démonstration. Supposons T faiblement mélangeante. Soit f une fonction propre et λ la valeur propre associée. Comme U_T préserve le produit scalaire, λ est de module 1 et on a

$$\langle \lambda f, 1\rangle = \langle U_T(f), U_T(1)\rangle = \langle f, 1\rangle.$$

Par conséquent, soit $\lambda \neq 1$ et f est orthogonale aux constantes, soit $\lambda = 1$ et f est invariante. Dans ce second cas, comme T est ergodique f est constante. Montrons que dans le premier cas, f doit être nulle. Comme T est faiblement mélangeante, on a

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\ U_T^n f - 0\right| = 0,$$

or $\frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X fU_T^n f\right| = \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\lambda^n f\right| = \int_X f^2$, donc $f=0$ (si f est complexe il faut utiliser f et sa conjuguée).

Réciproquement, supposons que les seules fonctions propres de U_T soit les constantes.

Nous avons besoin de résultats de théorie spectrale.

Lemme 4 *(Wiener) Soit ν une mesure positive sur le cercle S^1. La mesure μ n'a pas d'atome ssi*

$$\lim_{n\to\infty}\frac{1}{2n+1}\sum_{k=-n}^{n}|\widehat{\nu}(k)|^2 = 0$$

où les $\widehat{\nu}(k)$ désignent les coefficients de Fourier de ν.

Soit f une fonction de $L^2(\mu)$ orthogonale aux constantes. La suite $(\langle U_T^k f, f\rangle)_{k\in\mathbb{Z}}$ est définie positive, donc d'après le théorème de Herglotz, il existe une mesure positive ν sur le cercle unité S^1 de masse $\langle U_T^0 f, f\rangle = \|f\|_2^2$ telle que pour tout $k\in\mathbb{Z}$, $\widehat{\nu}(k) = \langle U_T^k f, f\rangle$.

Montrons que la mesure ν n'a pas d'atome. Si $\nu(\{z_0\}) > 0$ alors la fonction $1_{\{z_0\}}$ est une limite dans $L^2(S^1,\nu)$ d'une suite $(P_n(z))_{n\in\mathbb{N}}$ de polynômes trigonométriques (applications de la forme $\sum_{j=-n}^{n} a_j z^j$ + densité des P_n dans les fonctions continues sur S^1 pour la norme uniforme et densité des fonctions continues dans L^2). Pour un polynôme $P(z) = \sum_{j=-n}^{n} a_j z^j$ un calcul montre que $\left\|\sum_{j=-n}^{n} a_j U_T^j f\right\|_2^2 = \int_{S^1}|P(z)|^2\,d\nu(z)$. On en déduit que la suite $(P_n(U_T)f)_n$ est une suite de Cauchy de $L^2(\mu)$ qui converge vers une $h\in L^2(\mu)$ de norme $\left\|1_{\{z_0\}}\right\|_{L^2(S^1,\nu)} \neq 0$. Or les suites $(zP_n(z))_n$ et $(z_0P_n(z))_n$ convergent vers la fonction $z_0 1_{\{z_0\}}$ dans $L^2(S^1,\nu)$ donc

$$U_T(h) = \lim_{n\to\infty} U_T P_n(U_T) f = \lim_{n\to\infty} z_0 P_n(U_T) f = z_0 h.$$

Par conséquent h est un vecteur propre. Mais la fonction f est orthogonale aux constantes et donc les fonctions $U_T^n f$ le sont aussi ainsi que h, une contradiction. La mesure ν n'a donc pas d'atome.

D'aprés le lemme de Wiener on a

$$\lim_{n\to\infty}\frac{1}{2n+1}\sum_{k=-n}^{n}\left|\langle U_T^k f, f\rangle\right|^2 = \lim_{n\to\infty}\frac{1}{2n+1}\sum_{k=-n}^{n}|\widehat{\nu}(k)|^2 = 0$$

et donc

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\left|\langle U_T^k f, f\rangle\right|^2 = 0.$$

Avec le second lemme sur les moyennes de Cesaro on en conclut que T est faiblement mélangeante.
□

Exercice. Montrer que les translations du tore $\mathbb{T}^1$ ne sont pas faiblement mélangeantes.

5.1 Cas des groupes

On se contente de définir les notions de mélange et de mélange faible.

Définition 6 *Soit G un groupe localement compact agissant mesurablement sur un espace probabilisé $(X, \mathcal{A}, \mu)$ dont l'action préserve μ. On dit que l'action de G est mélangeante si pour tout $\varphi, \psi \in L^2(\mu)$*

$$\lim_{g\to\infty} \langle \varphi \circ T_g, \psi \rangle = 0,$$

c'est à dire, si pour tout $\varepsilon > 0$ il existe un compact K de G tel que $\forall g \in G \setminus K$, $|\langle \varphi \circ T_g, \psi \rangle| \leq \varepsilon$.

Définition 7 *Soit G un groupe localement compact agissant mesurablement sur un espace probabilisé $(X, \mathcal{A}, \mu)$ dont l'action préserve μ. On dit que l'action de G est faiblement mélangeante si le sous-espace de $L^2(\mu)$ orthogonal aux constantes ne contient aucun sous espace de dimension fini stable par G.*

Dans le cas où $G = \mathbb{Z}$, la définition précédente est équivalente au théorème sur le spectre des applications faiblement mélangeantes. En effet, si une application n'est pas faiblement mélageante alors il existe un vecteur propre dans l'orthogonal des constantes et si il existe un sous espace de dimension fini invariant par G, i-e par le genérateur de G, alors la restriction de ce générateur au sous-espace invariant admet un vecteur propre.

Lemme 5 *Soit G un groupe localement compact non compact agissant mesurablement sur un espace probabilisé $(X, \mathcal{A}, \mu)$ dont l'action préserve μ. Si l'action de G est mélangeante alors elle est faiblement mélangeante et si elle faiblement mélangeante alors elle est ergodique.*

Démonstration. Supposons que l'action de G ne soit pas faiblement mélageante. Alors il existe un sous espace E de l'orthogonal des constantes de dimension finie et invariant par G. Soit $(g_n)_{n\in\mathbb{N}}$ une suite tendant vers l'infini dans G. Soit f un élément non nul de E. Comme E est de dimension finie et comme la suite $g_n f$ est bornée dans L^2, on peut à extraction près supposer qu'elle converge vers une fonction h. Cette fonction n'est pas nulle car $\|h\|_2 = \lim_n \|g_n f\|_2 = \|f\|_2 \neq 0$. On a alors

$$\lim_{n\to\infty} \langle g_n f, h \rangle = \langle h, h \rangle \neq 0$$

et donc l'action de G n'est pas mélageante.

Si l'action de G n'est pas ergodique il existe une partie A de X invariante par G. le sous espace engendré par la fonction indicatrice de A centrée est invariant par G. Donc l'action de G n'est pas faiblement mélangeante. $\square$

5.2 Cas de flots

Les définitions du mélange et du mélange faible d'un flot sont des cas particuliers des définitions générales pour les groupes. Mais on peut aussi donner des définitions directes calquées sur le cas des applications conservant la mesure. Les mêmes raisonnements montrent que l'on obtient des définitions équivalentes.

Définition 8 *Soit $(X, \mathcal{A}, \mu)$ un espace probabilisé, $(\phi_t)_{t\in\mathbb{R}} : X \to X$ un flot agissant sur X conservant μ.*
1. On dit que le flot $(\phi_t)_{t\in\mathbb{R}}$ est faiblement mélangeant si $\forall A, B \in \mathcal{A}$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |\mu(\phi_{-t}(A) \cap B) - \mu(A)\mu(B)| = 0.$$

2. On dit que $(\phi_t)_{t\in\mathbb{R}}$ est mélangeant si $\forall A, B \in \mathcal{A}$,

$$\lim_{t\to\infty} \mu(\phi_{-t}(A) \cap B) = \mu(A)\mu(B).$$

Les caractérisations avec des fonctions de $L^2(\mu)$ sont également valables et se démontrent de la même manière.

6 Récurrence

Définition 9 *Soit $T : (X, \mathcal{A}, \mu) \to X$ une application préservant les ensembles de mesure nulle.*
1. On dit que $A \in \mathcal{A}$ est absorbant si $A \subset T^{-1}(A)$.
2. On dit que $A \in \mathcal{A}$ est errant si les ensembles $T^{-k}(A)$, $k \geq 0$, sont disjoints.
3. On dit que T est conservative si tous les ensembles errants sont de mesure nulle.

Théorème 12 *Soit $T : (X, \mathcal{A}, \mu) \to X$ une application préservant les ensembles de mesure nulle. T est conservative ssi $\forall A \in \mathcal{A}$, pour presque tout $x \in A$, la suite $(T^k(x))_{k \geq 1}$ passe une infinité de fois dans A.*

Démonstration. Supposons T conservative. Soit $A \in \mathcal{A}$ de mesure > 0. Posons $A_r = \cup_{k \geq 1} T^{-k}(A)$ et $B = A \setminus A_r$. Si $x \in B$ alors pour tout $k \geq 1$, $T^k(x) \notin A$ et donc $T^k(x) \notin B$. Par conséquent $B \cap (\cup_{k \geq 1} T^{-k}(B)) = \emptyset$. On en déduit que si $0 \leq p < q$ alors $T^{-p}(B) \cap T^{-q}(B) = T^{-p}(B \cap T^{-(q-p)} B) = \emptyset$. B est donc errant. Comme T est conservative B est de mesure nulle.
$A \cap A_r$ est l'ensemble des points de A qui retourne au moins une fois dans A. Si x est un point de A tel que la suite $(T^k(x))$ ne visite qu'un nombre fini de fois A, alors il existe un $k \geq 0$ tel que $T^k(x) \in B$, donc l'ensemble des x de A qui ne visite A qu'un nombre fini de fois est inclus dans $\cup_{k \leq 0} T^{-k}(B)$ qui est de mesure nulle.
Réciproquement si T n'est pas conservative il existe un ensemble errant A de mesure strictement positive. Par définition les points de A ne retournent jamais dans A. $\square$

Exercice. Soit $T : (X, \mathcal{A}, \mu) \to X$ une application préservant les ensembles de mesure nulle. Démontrer que T est conservative ssi pour toute partie mesurable absorbante A, $\mu(T^{-1}A) = \mu(A)$.

Corollaire 6 *(Théorème de récurrence de Poincaré). Soit $T : (X, \mathcal{A}, \mu) \to X$ une application préservant la mesure. Si $\mu(X)$ est fini alors pour toute partie $A \in \mathcal{A}$ de mesure > 0 presque tout point de A retourne une infinité de fois dans A.*

Démonstration. Il suffit de montrer que T est conservative. Or si $A \in \mathcal{A}$ est tel que les ensembles $T^{-k}(A)$ soient disjoints, on a $\mu(X) \geq \sum_k \mu(T^{-k}(A)) = \infty \times \mu(A)$ d'où $\mu(A) = 0$. $\square$

Exercice. Soit X un espace topologique ayant une base dénombrable d'ouverts, T une application continue de X dans X et μ une mesure finie définie sur la tribu borélienne de X. On suppose que μ est T-invariante. Montrer que pour presque tout $x \in X$, la suite $(T^k(x))_{k \geq 0}$ retourne une infinité de fois dans tout voisinage de x (Poincaré 1899).

Théorème 13 *de décomposition de Hopf.*
Soit $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini et $T : (X, \mathcal{A}, \mu) \to X$ une application préservant les ensembles de mesure nulle. Alors il existe une partition de X en deux parties mesurables disjointes C et D, la partie conservative et la partie dissipative, telle que
1. C est absorbante,
2. C ne contient pas d'ensemble errant de mesure non nulle,
3. D est une réunion dénombrable d'ensembles errants.
Si de plus T est inversible d'inverse mesurable et préservant les ensembles de mesure nulle alors C est invariante et il existe un ensemble errant W avec $D = \cup_{k \in \mathbb{Z}} T^k W$.

Démonstration. On peut supposer μ finie quitte à la multiplier par une fonction de L^1 (on utilise l'hypothèse μ σ-finie). On construit par récurrence une suite $A_1, ..., A_n, ...$ d'ensembles errants (mesurables) tels que

$$\mu(A_1) \geq \frac{1}{2} \sup\{\mu(A) : A \subset X \text{ et } A \text{ errant}\},$$
$$\mu(A_{n+1}) \geq \frac{1}{2} \sup\{\mu(A) : A \subset X \setminus (A_1 \cup ... \cup A_n) \text{ et } A \text{ errant}\}.$$

Posons $D = \cup_{n \geq 1} A_n$ et $C = X \setminus D$. La suite

$$\alpha_n = \sup\{\mu(A) : A \subset X \setminus (A_1 \cup ... \cup A_n)) \text{ et } A \text{ errant}\}$$

est décroissante. Si elle ne tend pas vers 0 alors la suite $\mu(A_n)$ est minorée par $\alpha > 0$ ce qui est impossible car $\mu(X) < \infty$. Par conséquent, si A est errant inclus dans C alors $\mu(A) = 0$.

Comme l'image réciproque par T d'un ensemble errant est errant, $T^{-1}(A_n)$ est errant. Comme $T^{-1}(A_n) \cap C \subset T^{-1}(A_n)$, $T^{-1}(A_n) \cap C$ est errant et donc de mesure nulle. Ainsi $T^{-1}(D) \cap C = \cup_{n\geq 1}(T^{-1}(A_n) \cap C)$ est de mesure nulle. En remplaçant C par $C' = C \setminus (\cup_{n\geq 0} T^{-n}(T^{-1}(D) \cap C)))$ et D par $D' = X \setminus C'$, on peut supposer que $C \subset T^{-1}(C)$ car

$$D' = D \cup (\cup_{n\geq 0} T^{-n}(T^{-1}(D) \cap C)) = D \cup (\cup_{n\geq 0, k\geq 1} T^{-n}(T^{-1}(A_k) \cap C))$$

est réunion dénombrable d'ensemble errants.
Dans le cas où T est inversible, A est errant ssi $T(A)$ est errant. Si $T^{-1}(C) \setminus C$ est de mesure strictement positive alors il existe un A_k tel que $B = T^{-1}(C) \setminus C \cap A_k$ soit de mesure strictement positive. On aurait alors $T(B)$ errant, inclus dans C et de mesure > 0, donc $T^{-1}(C) \setminus C$ est de mesure nulle. Par conséquent C et D sont invariants mod 0, en les modifiant on peut supposer qu'ils sont invariants. Reprenons les A_i du début et remplaçons les par $A_i \cap D$, on peut donc supposer qu'ils sont inclus dans D et que leur réunion est D à un ensemble de mesure nulle près. Posons $A_i^* = \cup_{k\in\mathbb{Z}} T^k A_i$. Il suffit maintenant de prendre

$$W = \cup_{i=1}^{\infty}(A_i \setminus \cup_{k=1}^{i-1} A_j^*)). \square$$

Exemples. 1. $X = \mathbb{R}$ et $T(x) = x + a$. $D = \mathbb{R}$ et $C = \emptyset$.
2. $\mu(X) < \infty$ et si T conserve μ alors $D = \emptyset$.

Exercices 1. *Théorème de Hopf.* Soit $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini et $T : (X, \mathcal{A}, \mu) \to X$ une application préservant μ. Soit $f \in \mathcal{L}^1(X)$, $f \geq 0$. Montrer que
a. $\{x \in X : \sum_{n=0}^{\infty} f \circ T^n(x) = \infty\}$ est inclus modulo μ dans la partie conservative C de T.
b. Si de plus $f > 0$ sur X alors $\{x \in X : \sum_{n=0}^{\infty} f \circ T^n(x) = \infty\} = C$ modulo μ.

2. *Théorème de Maharam.* Soit $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini et $T : (X, \mathcal{A}, \mu) \to X$ une application préservant μ. Montrer que si il existe une partie mesurable A de mesure finie tel que $X = \cup_{n\geq 0} T^{-n} A$ modulo μ alors T est conservative.

3. Soit $T : \mathbb{R} \to \mathbb{R}$ définie par $T(x) = x - 1/x$. Montrer que T conserve la mesure de Lebesgue puis que T est conservative.

6.1 Application induite

Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, $T : X \to X$ une application préservant les ensembles de mesure nulle et $A \in \mathcal{A}$. Lorsque T est conservative presque tout point de A retourne une infinité de fois dans A. Cela permet de définir presque partout sur A une application de premier retour. Pour $x \in X$, on pose

$$r_A(x) = \inf\{k \geq 1 : T^k(x) \in A\},$$

r_A est fini presque partout sur A et on définit T_A par $T_A(x) = T^{r_A(x)}(x)$ pour les x de A tels que $r_A(x) < \infty$.

Théorème 14 *(Kac 1947). Si T est conservative et préserve la mesure alors pour tout $A \in \mathcal{A}$,*

$$\int_A r_A \, d\mu = \mu(\cup_{n=0}^{\infty} T^{-n}(A)).$$

Si T est ergodique et $\mu(A) > 0$, on obtient $\int_A r_A \, d\mu = \mu(X)$.

Démonstration. On peut supposer $0 < \mu(A) < \infty$ (si $\mu(A) = \infty$, les deux membres de l'égalité sont infinis). Posons $A_0 = A$ et pour $k \geq 1$,

$$A_k = T^{-1}(A_{k-1}) \cap A^C \text{ et } R_k = T^{-1}(A_{k-1}) \cap A.$$

Montrons par récurrence que pour $k \geq 1$, $r_A(x) = k$ ssi $x \in R_k \cup A_k$. C'est évident si $k = 1$. Supposons $k \geq 2$ et l'hypothèse vraie au rang $k - 1$. Si $x \in A_k \cup R_k$ alors $T(x) \in A_{k-1}$. Comme $A_{k-1} \subset A^C$, $r_A(x) > 1$ et d'après l'hypothèse de récurrence $r_A(T(x)) = k - 1$, donc $r_A(x) = 1 + (k - 1)$. Si $x \notin A_k \cup R_k$ alors $T(x) \notin A_{k-1}$ et soit $T(x) \notin A$, soit $T(x) \in A$. Dans le premier cas, $T(x) \notin R_{k-1}$, donc $r_A(T(x)) \neq k - 1$, par conséquent, $r_A(x) \neq k$. Dans le deuxième cas, $r_A(x) = 1 \neq k$.

Comme T conserve la mesure, nous avons $\mu(A_n) = \mu(T^{-1}(A_n)) = \mu(A_{n+1}) + \mu(R_{n+1})$. Une récurrence immédiate montre que pour $m > n \geq 1$,

$$\mu(A_n) = \sum_{k=n+1}^{m} \mu(R_k) + \mu(A_m).$$

Mais $\mu(A_0) = \sum_{k\geq 1} \mu(R_k)$ car T est conservative , d'où $\mu(A_m)$ tend vers 0 et $\mu(A_n) = \sum_{k=n+1}^{\infty} \mu(R_k)$. Finalement

$$\begin{aligned}\mu(\cup_{n=0}^{\infty} T^{-n}(A)) &= \sum_{n=0}^{\infty} \mu(T^{-n}(A) \setminus \cup_{k<n} T^{-k}(A)) \\ &= \sum_{n=0}^{\infty} \mu(A_n) = \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} \mu(R_k) = \sum_{k=1}^{\infty} k\mu(R_k) = \int_A r_A \, d\mu. \qquad \square\end{aligned}$$

Théorème 15 *(Kakutani 1943). Si T est conservative et conserve la mesure alors l'application T_A induite sur $(A, \mathcal{A} \cap A, \mu_A)$ conserve μ_A où μ_A est la restriction de μ à $\mathcal{A} \cap A$.*

Démonstration. Remarquons d'abord que T_A envoie $A_\infty = \{x \in A : T^k(x)$ visite A une infinité de fois$\}$ dans lui même et que $A \setminus A_\infty$ est de mesure nulle. On peut donc supposer que $A = A_\infty$. Soit B mesurable inclus dans A. Posons $B_0 = B$ et

$$B_k = T^{-1}(B_{k-1}) \cap A^C \text{ et } B'_k = T^{-1}(B_{k-1}) \cap A, \ k \geq 1.$$

Avec les notations du théorème précédent on voit par récurrence que $B_k = A_k \cap T^{-k}(B)$ et $B'_k = R_k \cap T^{-k}(B)$. De même,

$$\mu(B_n) = \sum_{k=n+1}^{m} \mu(B'_k) + \mu(B_m)$$

et

$$\mu(B) = \sum_{n=1}^{\infty} \mu(B'_n).$$

Pour conclure il suffit de prouver que $T_A^{-1}(B) = \cup_{n\geq 1} B'_n$. Or

$$T_A^{-1}(B) = \cup_{k\geq 1}(T_A^{-1}(B) \cap R_k) = \cup_{k\geq 1}(T^{-k}(B) \cap R_k) = \cup_{k\geq 1} B'_k. \ \square$$

7 Références

J. Aaronson: An Introduction to Infnite Ergodic Theory. AMS 1997.
U. Krengel: Ergodic Theorems. de Gruyter 1984.
P. Walters: An Introduction to Ergodic Theory. Springer 1982.