Introduction à la théorie ergodique en mesure infinie

É. Janvresse, T. de la Rue

Master 2 MFA Rouen 2014-2015

Dans tout ce chapitre, $(X, \mathscr{A}, \mu, T)$ est un système dynamique, où T est une transformation inversible qui préserve la mesure σ-finie μ (mais $\mu(X)$ n'est pas nécessairement finie). Cette partie du cours est en grande partie inspirée du 1er chapitre du livre de Jon Aaronson *Introduction to Infinite Ergodic Theory*.

1 Exemples

1.1 Addition sur $\mathbb{R}$

$X = \mathbb{R}$, μ est la mesure de Lebesgue et $Tx = x + 1$ pour tout $x \in \mathbb{R}$.

1.2 Automorphisme spécial

Considérons un système dynamique $(\bar{X}, \bar{\mathscr{A}}, \bar{\mu}, \bar{T})$ avec $\bar{\mu}(\bar{X}) = 1$ et une fonction $f : \bar{X} \to \mathbb{N}^*$. Posons $X := \big\{(x, n) \in \bar{X} \times \mathbb{N}, 0 \leq n < f(x)\big\}$ et définissons la transformation T sur X par

$$T : (x, n) \in X \longmapsto \begin{cases} (x, n+1) & \text{si } n < f(x) - 1 \\ (\bar{T}x, 0) & \text{sinon.} \end{cases}$$

(Voir Figure 1.) Partitionnons X en ensembles X_n, $n \geq 0$, selon la valeur de la seconde coordonnée d'un point dans X. Ainsi $X_0 = \bar{X} \times \{0\}$, et plus généralement, $X_n = \{x \in \bar{X} : f(x) > n\} \times \{n\}$. Sur X, on peut définir une mesure μ préservée par T, en posant pour $B \subset X_n$, $\mu(B) := \bar{\mu}(T^{-n}B)$. On a alors

$$\mu(X) = \sum_{n \geq 0} \mu(X_n) = \sum_{n \geq 0} \bar{\mu}(f > n) = \int_{\bar{X}} f \, d\bar{\mu}.$$

Ainsi, si la fonction f choisie au départ n'est pas intégrable, la mesure totale de X est infinie.

1.3 Chaînes de Markov

Soit E un ensemble dénombrable, et $P = (P_{i,j})_{(i,j) \in E \times E}$ le noyau de transition d'une chaîne de Markov sur E, ayant une mesure invariante γ infinie. (Par exemple, $E = \mathbb{Z}$, P le noyau de transition de la marche aléatoire simple symétrique sur $\mathbb{Z}$, et γ la mesure de comptage sur $\mathbb{Z}$.) Pour chaque $n \in \mathbb{N}$, on peut considérer la mesure μ_n définie sur $E^{\{-n, \ldots, n\}}$ par

$$\forall (x_{-n}, \ldots, x_n) \in E^{\{-n, \ldots, n\}}, \quad \mu_n(x_{-n}, \ldots, x_n) := \gamma(x_{-n}) \prod_{j=-n}^{n-1} P_{x_j, x_{j+1}}.$$

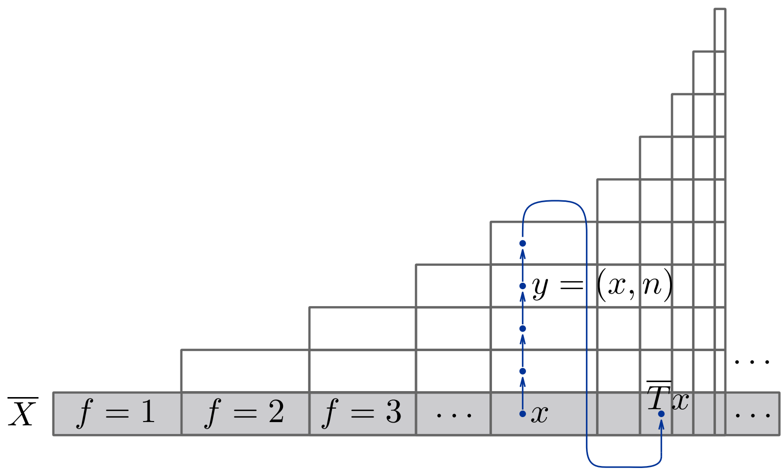

Figure 1 – Automorphisme spécial au-dessus de $(\bar{X}, \bar{\mathscr{A}}, \bar{\mu}, \bar{T})$.

Comme γ est invariante par P, la famille de mesures μ_n s'étend en une mesure infinie sur l'espace $X := E^{\mathbb{Z}}$ des trajectoires de la chaîne de Markov, et cette mesure est invariante par T, où T est le shift des coordonnées.

2 Conservativité et dissipativité

2.1 Union mesurable

Soit $(E, \mathscr{F}, m)$ un espace mesuré, A et B deux parties mesurables de E. On dit que A et B coïncident modulo m (et on note $A = B \mod m$) si $m(A \vartriangle B) = 0$. De même, on dit que A est inclus dans B modulo m (et on note $A \subset B \mod m$) si $m(A \setminus B) = 0$.

Définition 2.1. *Une collection $H \subset \mathscr{F}$ est dite* héréditaire *dans l'espace mesuré $(E, \mathscr{F}, m)$ si tout $A \in \mathscr{F}$ inclus dans un élément de H appartient aussi à H.*
Un ensemble $U \in \mathscr{F}$ recouvre *la collection héréditaire H si tout $A \in H$ est inclus $\mod m$ dans U.*
Un ensemble $U \in \mathscr{F}$ est saturé par *la collection héréditaire H si tout élément de $\mathscr{F}$ de mesure strictement positive inclus dans U contient un élément de H de mesure strictement positive :*

$$\forall A \in \mathscr{F}, m(A) > 0, A \subset U \Rightarrow \exists B \in H, m(B) > 0, B \subset A.$$

Si $U \in \mathscr{F}$ recouvre et est saturé par la collection héréditaire H, on dit que U est une union mesurable *de la collection H.*

Il y a unicité modulo m de l'union mesurable. En effet, supposons qu'il existe U et U' unions mesurables de H, telles que $m(U \setminus U') > 0$. Comme U est saturé par H, il existe $B \in H$, $m(B) > 0$ tel que $B \subset U \setminus U'$. D'autre part, comme U' recouvre H, on a $B \subset U' \mod m$, et on arrive à une contradiction. Donc $m(U \setminus U') = 0$ et par symétrie on arrive à $U = U' \mod m$.

Remarquons aussi que la notion d'union mesurable ne change pas si on remplace μ par une autre mesure qui lui est équivalente.

Lemme 2.2. *Soit $(E, \mathscr{F}, m)$ un espace mesuré où m est une mesure σ-finie. Soit $H \subset \mathscr{A}$ une collection héréditaire. Il existe une famille $(A_i)_{i\geq 1}$ dénombrable d'éléments de H disjoints telle que $\bigsqcup_{i\geq 1} A_i$ est l'union mesurable de H.*

Démonstration. On se ramène au cas où la mesure est finie en considérant une mesure finie ν équivalente à m. Soit $\varepsilon_1 := \sup\{\nu(A) : A \in H\}$, et fixons $A_1 \in H$ de mesure $\nu(A_1) \geq \varepsilon_1/2$. On considère alors $\varepsilon_2 := \sup\{\nu(A) : A \in H, A \cap A_1 = \emptyset\}$ et $A_2 \in H$, $A_2 \cap A_1 = \emptyset$ de mesure $\nu(A_2) \geq \varepsilon_2/2$. En continuant ainsi, on définit pour tout n

$$\varepsilon_n := \sup\{\nu(A) : A \in H, A \cap A_k = \emptyset, \forall k < n\},$$

et $A_n \in H$ de mesure $\nu(A_n) \geq \varepsilon_n/2$, tel que $A_n \cap A_k = \emptyset\, \forall k < n$.

Comme $\sum_n \varepsilon_n \leq 2\sum_n \nu(A_n) < \infty$, on a $\varepsilon_n \to 0$. Puisque les A_n sont tous dans H, $\bigcup_{n\geq 1} A_n$ est saturé par H. D'autre part, si $\bigcup_{n\geq 1} A_n$ ne recouvrait pas H, il existerait $A \in H$, $m(A) > 0$, tel que $A \cap A_n = \emptyset$ pour tout n. C'est impossible, car cela entraînerait $\nu(A) < \varepsilon_n$ pour tout n, ce qui impliquerait $\nu(A) = 0$ puis $m(A) = 0$ par l'équivalence de ν et m. □

2.2 Décomposition de Hopf

Définition 2.3. *On dit qu'un ensemble mesurable W est* errant *si les ensembles $(T^n W)_{n\in\mathbb{Z}}$ sont deux à deux disjoints (on notera $W \in \mathscr{W}$).*

Notons que $\mathscr{W}$ est une collection héréditaire dans $(X, \mathscr{A}, \mu)$. On peut donc considérer son union mesurable.

Définition 2.4. *La* partie dissipative *$D(T)$ est l'union mesurable des ensembles errants :*

- *Tout ensemble errant W est inclus $\bmod \mu$ dans $D(T)$: il existe $A \subset W$ tel que $A \subset D(T)$ et $\mu(W \setminus A) = 0$.*
- *Pour tout $A \subset D(T)$ tel que $\mu(A) > 0$, il existe un ensemble errant W inclus dans A, tel que $\mu(W) > 0$.*

La partie conservative *est $C(T) := X \setminus D(T)$ et on appelle* décomposition de Hopf *la partition de X en $\{D(T), C(T)\}$.*
Lorsque $X = D(T) \mod \mu$, on dit que T est totalement dissipative.
Lorsque $X = C(T) \mod \mu$, on dit que T est conservative *(tout ensemble errant est de mesure nulle).*

Exercice 2.1. *Vérifier que l'addition sur $\mathbb{R}$ définie en 1.1 est totalement dissipative, et que l'automorphisme spécial défini en 1.2 est toujours conservatif.*

Proposition 2.5. *Il existe $W \in \mathscr{W}$ tel que $D(T) = \bigsqcup_{n\in\mathbb{Z}} T^n W \mod \mu$.*

Démonstration. Fixons d'abord une mesure de probabilité ν équivalente à μ. Soit $\mathscr{I}$ la sous-tribu de $\mathscr{A}$ constituée des ensembles T-invariants :

$$\mathscr{I} := \{A \in \mathscr{A} : T^{-1}A = A\},$$

et remarquons que pour tout $W \in \mathscr{W}$, $\bigsqcup_{n\in\mathbb{Z}} T^n W \in \mathscr{I}$. Considérons la collection

$$\mathscr{C} := \left\{ \bigsqcup_{n\in\mathbb{Z}} T^n W : W \in \mathscr{W} \right\} \subset \mathscr{I}.$$

$\mathscr{C}$ est une collection héréditaire dans $(X, \mathscr{I}, \nu)$, et on peut donc considérer sa réunion mesurable, qui contient $D(T)$ modulo ν. De plus, d'après le lemme 2.2, il existe une famille dénombrable $(W_k)_k \subset \mathscr{W}$ telle que $\bigsqcup_k \bigsqcup_{n\in\mathbb{Z}} T^n W_k$ est l'union mesurable de $\mathscr{C}$. Posons alors $W := \bigsqcup_{k=1}^{\infty} W_k$. Alors $W \in \mathscr{W}$ et $\bigsqcup_{n\in\mathbb{Z}} T^n W$ est l'union mesurable de $\mathscr{C}$. On conclut donc que $D(T) = \bigsqcup_{n\in\mathbb{Z}} T^n W \mod \mu$. □

Théorème 2.6 (Théorème de récurrence d'Halmos)**.** *Soit $B \subset C(T)$, $\mu(B) > 0$. Alors*

$$\sum_{n\geq 1} \mathbb{1}_B(T^n x) = \infty \text{ pour } \mu\text{-presque tout } x \in B.$$

Démonstration. Soit $B \subset C(T)$, et notons r_B le temps de retour en B. Posons $W_B := B \setminus \bigcup_{n\geq 1} T^{-n}B = \{x \in B : r_B(x) = \infty\}$. Alors $W_B \in \mathscr{W}$ (voir le chapitre sur le temps de retour de Poincaré). Puisque $B \subset C(T)$, on en déduit que $\mu(W_B) = 0 = \mu(T^{-n}W_B)$ pour tout n. Posons ensuite $\overline{B} := B \setminus \bigcup_{n\geq 0} T^{-n}W_B$. On a clairement $B = \overline{B} \mod \mu$. De plus, par définition de $\overline{B}$, les points dans $\overline{B}$ reviennent dans B en dehors de W_B, donc reviennent une infinité de fois dans B. □

Théorème 2.7. *Soit $f \in L^1(\mu)$.*

$$f \geq 0 \Rightarrow \left[\sum_{n=1}^{+\infty} f \circ T^n = \infty\right] \subset C(T) \mod \mu.$$

$$f > 0 \Rightarrow \left[\sum_{n=1}^{+\infty} f \circ T^n = \infty\right] = C(T) \mod \mu.$$

Démonstration. Soit $f \geq 0$ dans $L^1(\mu)$. Montrons que $\sum_{n=1}^{+\infty} f \circ T^n < \infty$ μ-p.s. sur tout ensemble errant W. Pour tout $N \geq 1$ et tout W errant, on a

$$\begin{aligned}
\int_W \sum_{n=0}^{N} f \circ T^n d\mu &= \sum_{n=0}^{N} \int_X \mathbb{1}_W f \circ T^{N-n} d\mu = \sum_{n=0}^{N} \int_X \mathbb{1}_W \circ T^n f \circ T^N d\mu \\
&= \int_X \left(\sum_{n=0}^{N} \mathbb{1}_W \circ T^n\right) f \circ T^N d\mu \\
&\leq \int_X f \circ T^N d\mu = \int_X f d\mu.
\end{aligned}$$

Comme $f \in L^1(\mu)$, on a donc $\int_W \sum_{n=1}^{\infty} f \circ T^n d\mu < \infty$.

Supposons maintenant que $f > 0$ et montrons que $C(T) \subset \left[\sum_{n=1}^{+\infty} f \circ T^n = \infty\right]$ μ-p.s. Soit $A \subset C(T)$ avec $\mu(A) > 0$, et pour tout entier k, soit $A_k := \{x \in A : f(x) \geq 1/k\}$. D'après le théorème 2.6, puisque $A_k \subset A \subset C(T)$, on a alors

$$\sum_{n=1}^{+\infty} f \circ T^n \geq \frac{1}{k} \sum_{n=1}^{+\infty} \mathbb{1}_{A_k} \circ T^n = \infty \quad \mu\text{-p.s. sur } A_k,\ \forall k.$$

De plus, comme $f > 0$, $A = \cup_k A_k$ et donc, μ-p.s. sur A, $\sum_{n=1}^{+\infty} f \circ T^n = \infty$. □

2.3 Cas ergodique

La notion d'ergodicité se généralise en mesure infinie :

Définition 2.8. *Le système $(X, \mathscr{A}, \mu, T)$ est dit* ergodique *si toute partie A mesurable invariante par T vérifie $\mu(A) = 0$ ou $\mu(X \setminus A) = 0$.*

Théorème 2.9. *Si μ est une mesure continue et si $(X, \mathscr{A}, \mu, T)$ est ergodique, alors T est conservative.*

Démonstration. Supposons que T n'est pas conservative : alors il existe un ensemble errant W, avec $\mu(W) > 0$. Comme la mesure est continue, on peut trouver $W_1 \subset W$ avec $\mu(W_1) > 0$ et $\mu(W \setminus W_1) > 0$. Mais $A := \bigsqcup_{n\in\mathbb{Z}} T^n W_1$ vérifie $T^{-1}A = A$, $\mu(A) > 0$ et $\mu(X \setminus A) > 0$. Ceci contredit l'ergodicité du système. □

Exercice 2.2. *Soit $(X, \mathscr{A}, \mu, T)$ ergodique et conservatif, et soit $A \in \mathscr{A}$, avec $\mu(A) = 1$. Montrer que l'on peut représenter le système comme un automorphisme spécial au-dessus de (A, μ_A, T_A).*

Exercice 2.3. *Soit $(A, \mathscr{A}_A, \mu_A, T_A)$ un système dynamique ergodique où $\mu_A(A) = 1$. Prouver que pour toute fonction $f \geq 0$ mesurable, on a*

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n \xrightarrow[N\to\infty]{\mu_A\text{-p.s.}} \int_A f \, d\mu_A.$$

Autrement dit, la conclusion du théorème ergodique ponctuel de Birkhoff est valable pour toute fonction positive, même si elle n'est pas intégrable.

Théorème 2.10. *Soit $(X, \mathscr{A}, \mu, T)$ ergodique et conservatif, avec $\mu(X) = \infty$. Alors*

$$\forall h \in L^1(\mu), \quad \frac{1}{N}\sum_{n=0}^{N-1} h \circ T^n \xrightarrow[N\to\infty]{\mu\text{-p.s.}} 0.$$

Démonstration. Soit $A \in \mathscr{A}$ de mesure $\mu(A) = 1$. Comme $(X, \mathscr{A}, \mu, T)$ est ergodique et conservatif, on peut représenter le système comme un automorphisme spécial au-dessus de A. Soit $\bar{h}$ définie pour $x \in A$ par $\bar{h}(x) := h(x) + h(Tx) + \dots h(T^{r_A - 1}x)$. Notons que $\int_A \bar{h} d\mu_A = \int_X h d\mu < \infty$. On peut donc appliquer le théorème ergodique pour la transformation induite T_A sur A (car T ergodique implique T_A ergodique) :

$$\frac{1}{M}\sum_{m=0}^{M-1} \bar{h}(T_A^m x) \xrightarrow[M\to\infty]{\mu_A\text{-p.s.}} \int_A \bar{h} d\mu_A. \tag{1}$$

Soit $N(x, M) := r_A(x) + r_A(T_A x) + \dots + r_A(T_A^{M-1}x)$. Comme $\int_A r_A d\mu_A = \mu(X) = \infty$, en appliquant le résultat de l'exercice ci-dessus, on a

$$\frac{N(x,M)}{M} = \frac{1}{M}\sum_{m=0}^{M-1} r_A(T_A^m x) \xrightarrow[M\to\infty]{\mu_A\text{-p.s.}} \infty.$$

Ainsi,

$$\frac{1}{N(x,M)}\sum_{m=0}^{N(x,M)-1} h(T^m x) = \frac{M}{N(x,M)}\frac{1}{M}\sum_{m=0}^{M-1} \bar{h}(T_A^m x) \xrightarrow[M\to\infty]{\mu_A\text{-p.s.}} 0. \tag{2}$$

(observons que c'est encore vrai en remplaçant le coefficient $\frac{1}{N(x,M)}$ par $\frac{1}{N(x,M-1)}$). Regardons maintenant ce qui se passe lorsqu'on part d'un point qui n'est pas forcément dans A. Fixons $y \in X$. Soit $x \in A$ tel que $y = T^j x$ pour $0 \leq j < r_A(x)$. Soit $N > 1$ et considérons M tel que $N(x, M-1) \leq j + N - 1 < N(x, M)$. On a

$$\frac{1}{N+j} \sum_{n=0}^{N-1} h \circ T^n(y) \leq \frac{1}{N+j} \sum_{n=0}^{N+j-1} h \circ T^n(x) \leq \frac{1}{N(x, M-1)} \sum_{n=0}^{N(x,M)-1} h \circ T^n(x)$$

qui tend vers 0 pour μ-presque tout $x \in A$ d'après (2). l'ensemble des y correspondant à des $x \in A$ tels que la convergence n'a pas lieu étant négligeable, on en déduit que la moyenne des $h \circ T^n$ tend vers 0 μ-presque sûrement. $\square$

En utilisant (1) pour deux fonctions h et g, on montre de la même manière :

Théorème 2.11 (théorème ergodique quotient de Hopf)**.** *Soit $(X, \mathscr{A}, \mu, T)$ ergodique et conservatif, avec $\mu(X) = \infty$. Alors pour $g, h \in L^1(\mu)$ avec $\int_X g \, d\mu > 0$, on a*

$$\frac{\sum_{k=0}^{n} h \circ T^k}{\sum_{k=0}^{n} g \circ T^k} \to \frac{\int_X h d\mu}{\int_X g d\mu} \quad \mu\text{-p.s.}$$

La conclusion reste valable si $g \geq 0$ sans supposer l'intégrabilité de g.

En particulier, si g est la densité d'une mesure de probablité par rapport à la mesure μ, la convergence se fait vers $\int_X h d\mu$. Autrement dit, pour toute fonction dans $L^1(\mu)$, la convergence suivante a lieu pour une suite de fonctions à valeurs réelles $(a_n)_n$ convenable :

$$\lim_{n\to\infty} \frac{1}{a_n(x)} \sum_{k=0}^{n} h(T^k x) = \int_X h d\mu \quad \text{pour } \mu\text{-presque tout } x.$$

Mais les suites $(a_n(x))_n$ dépendent du point de départ x.

3 Type positif et type zéro

Définition 3.1. *Le système $(X, \mathscr{A}, \mu, T)$ est dit de* type positif *si, pour tout $A \in \mathscr{A}$ tel que $\mu(A) > 0$, on a $\limsup_n \mu(A \cap T^{-n} A) > 0$.*
Le système $(X, \mathscr{A}, \mu, T)$ est dit de type zéro *si, pour tout $A \in \mathscr{A}$ tel que $\mu(A) < \infty$, on a $\lim_n \mu(A \cap T^{-n} A) = 0$.*

Lemme 3.2. *Soit $A \in \mathscr{A}$ avec $0 < \mu(A) < \infty$ tel que $\lim_n \mu(A \cap T^{-n} A) = 0$. Alors pour tout $B \in \mathscr{A}$ de mesure finie, on a $\lim_n \mu(A \cap T^{-n} B) = 0$.*

Démonstration. Considérons

$$\mathscr{F} := \left\{ f \in L^2(\mu) : \int_X \mathbb{1}_A . \, f \circ T^n \, d\mu \xrightarrow[n\to\infty]{} 0 \right\}.$$

Par hypothèse, on a $\mathbb{1}_A \in \mathscr{F}$, et aussi $\mathbb{1}_B \in \mathscr{F}$ pour tout $B \subset A$. Comme T préserve la mesure μ, $\mathbb{1}_A \circ T^k \in \mathscr{F}$ pour tout k, et donc l'espace cyclique $S(\mathbb{1}_A)$ est contenu dans $\mathscr{F}$. De plus, si $g \perp S(\mathbb{1}_A)$, alors $\int_X \mathbb{1}_A \, g \circ T^n \, d\mu = \int_X \mathbb{1}_A \circ T^{-n} \, g \, d\mu = 0$. Donc $S(\mathbb{1}_A)^\perp \subset \mathscr{F}$. $\square$

Théorème 3.3. *Soit* $(X, \mathscr{A}, \mu, T)$ *ergodique et conservatif. S'il existe* $A \in \mathscr{A}$ *avec* $0 < \mu(A) < \infty$ *tel que* $\lim_n \mu(A \cap T^{-n}A) = 0$*, alors* $(X, \mathscr{A}, \mu, T)$ *est de type zéro.*

Démonstration. Considérons B_1 et B_2 disjoints tels que $\lim_n \mu(B_k \cap T^{-n} B_k) = 0$ pour $k = 1, 2$ et montrons que $B_1 \sqcup B_2$ vérifie la même propriété. En effet,

$$\mu\left((B_1 \sqcup B_2) \cap T^{-n}(B_1 \sqcup B_2)\right) = \sum_{k=1,2} \mu(B_k \cap T^{-n} B_k) + \mu(B_1 \cap T^{-n} B_2) + \mu(B_2 \cap T^{-n} B_1),$$

où les 2 premiers termes tendent vers 0 quand $n \to \infty$ par hypothèse et les 2 derniers tendent vers 0 par le lemme 3.2. Plus généralement, la propriété reste vraie par union dénombrable disjointe.

Soit $A \in \mathscr{A}$ donné dans l'énoncé du théorème. Comme $(X, \mathscr{A}, \mu, T)$ est ergodique et conservatif, on peut écrire $X \pmod{\mu}$ comme une union disjointe $\bigsqcup_k T^k \left(A \cap \{r_A > k\}\right)$. Donc tout B de mesure finie est une union disjointe d'ensembles de la forme $B_k := B \cap T^k \left(A \cap \{r_A > k\}\right) \subset T^k A$. Comme μ est invariante par T, pour tout k, $T^k A$ vérifie la même propriété que A. De plus la propriété reste vraie pour tout sous-ensemble, et donc pour B_k. Finalement $\lim_n \mu(B_k \cap T^{-n} B_)) = 0$. □

Corollaire 3.4. *Un système dynamique ergodique et conservatif est soit de type zéro, soit de type positif.*