Notes de cours

Intoduction à la Théorie Ergodique

Master 2 de Probabilités et applications à l'UPMC

Janvier 2017

Table des matières

Chapitre 1. Systèmes dynamiques mesurés : Premières propriétés et exemples 5
1. Systèmes dynamiques 5
2. Récurrence de Poincarrré 6
3. Ergodicité 7
4. Mélange 9
5. Bernoulli 10
6. Exercices 10

Chapitre 2. Théorèmes ergodiques 13
1. Théorème ergodique en moyenne 13
2. Théorèmes ergodiques ponctuels 16
3. Preuve de l'inégalité maximale 20
4. Deux exemples d'applications 21
5. Exercices 22

Chapitre 3. Mesures invariantes d'un système dynamique topologique 25
1. Compacité 25
2. Structure convexe 26
3. Unique ergodicité 28

Chapitre 4. Théorie spectrale 33
1. Mélange faible 33
2. Spectre de l'opérateur de Koopman 36
3. Mesures spectrales et Théorème spectral 38
4. Exercices 41

Chapitre 5. Exemples 43
1. Sous décalage de type fini et mesures markoviennes 43
2. L'application de Gauss 44
3. Transformation de Chacon 46
4. Flot géodésique et horocyclique 47
5. Exercices 49

Chapitre 6. Entropie mesurée 51
1. Information 51
2. Entropie statique 52
3. KS-Entropie et théorème des générateurs 52
4. Formule de Katok, Newhouse et Shanon-Mcmillan-Breiman 54
5. Exemples 56
6. Exercices 57

Chapitre 7. Entropie topologique et principe variationnel 59
1. Complexité 59
2. Entropie topologique 60
3. Harmonicité de l'entropie 61
4. Entropie à la Bowen 61
5. Principe variationnel 62
6. Applications et exemples 64
7. Exercices 67

Chapitre 1

Systèmes dynamiques mesurés : Premières propriétés et exemples

Contents

1. Systèmes dynamiques **5**
1.1. Actions sur les mesures 5
1.2. Actions sur L^p 6
1.3. Isomorphismes 6
2. Récurrence de Poincarré **6**
3. Ergodicité **7**
4. Mélange **9**
5. Bernoulli **10**
6. Exercices **10**

On considérera toujours dans ce chapitre un espace mesurable $(X, \mathcal{B})$ supposé Borel standard, i.e. X est un espace métrique complet séparable et $\mathcal{B}$ sa tribu des Boréliens. Lorsque $f : X \to X$ est une application mesurable, on dira que le triplet $(X, \mathcal{B}, f)$ est un système dynamique mesurable.

1. Systèmes dynamiques

1.1. Actions sur les mesures. On note $\mathcal{M}(X)$ l'ensemble des mesures de proba sur $(X, \mathcal{B})$. Le système mesurable $(X, \mathcal{B}, f)$ induit une action f^* sur $\mathcal{M}(X)$ en définissant $f^*\mu$ comme étant l'élément de $\mathcal{M}(X)$ satisfaisant $f^*\mu(A) = \mu(f^{-1}A)$ pour tout $A \in \mathcal{B}$.

Définition 1.1. *Avec les notations précédentes, on dira qu'un mesure est f-invariante lorsque $f^*\mu = \mu$. Le quadruplet $(X, \mathcal{B}, f, \mu)$ est alors appelé un système dynamique mesuré.*

Exemples :

(1) Rotation : on considère la rotation f_α d'angle $\alpha \in \mathbb{R}$ sur le cercle $\mathbb{R}/\mathbb{Z}$ définie additivement par $f_\alpha(\overline{x}) = \overline{x + \alpha}$,

(2) Doublement de l'angle : on considère toujours sur le cercle $\mathbb{R}/\mathbb{Z}$ l'application f qui à $\overline{x}$ associe $\overline{2x}$,

(3) Soit $A \in GL_n(\mathbb{Z})$ une matrice carré d'ordre n inversible à coefficient entiers de déterminant 1 (et donc d'inverse à coefficients entiers), l'application $x \mapsto Ax$ descend sur le tore $\mathbb{R}^n/\mathbb{Z}^n$ en une application noté f_A.

On vérifie dans tous ces cas que la mesure de Lebesgue est invariante (utiliser par exemple le théorème de changement de variable).

1.2. Actions sur L^p.

Proposition 1.2. *Soit $(X, \mathcal{B}, f)$ un système dynamique mesurable et soit $\mu \in \mathcal{M}(X)$. Alors μ est f-invariante ssi pour tout $\phi \in L^1(\mu)$, on a*

$$\int \phi \circ f d\mu = \int \phi d\mu.$$

On en déduit que pour tout $p \in [1, +\infty[$ et tout système dynamique mesure $(X, \mathcal{B}, f, \mu)$ l' application $U_f : L^p(\mu) \to L^p(\mu)$ envoyant ϕ sur $\phi \circ f$ est une isométrie. On s'intéressera par la suite à l'opérateur U_f sur le Hilbert $L^2(\mu)$ (appelé l'opérateur de Koopman).

1.3. Isomorphismes. Deux systèmes dynamiques mesurés $(X, \mathcal{B}, f, \mu)$ et $(Y, \mathcal{C}, g, \nu)$ sont dit *isomorphes* (ou métriquement isomorphes) s'il existe $X' \in \mathcal{B}$ et $Y' \in \mathcal{C}$ respectivement f et g invariants avec $\mu(X') = \nu(Y') = 1$ et une bijection bimesurable $\psi : X' \to Y'$ tel que $\psi \circ f = g \circ \psi$ et $\psi^* \mu = \nu$.

Ces deux systèmes sont dits *spectralement isomorphes* si leur opérateurs de Koopman sont unitairement équivalents, i.e. s'il existe une isométrie bijective $V : L^2(\mu) \to L^2(\nu)$ telle que $VU_f = U_g V$.

Proposition 1.3. *Deux systèmes mesurés métriquement isomorphes sont spectralement isomorphes.*

On verra par la suite que l'inverse est faux. Une propriété stable par isomorphisme spectrale est dite spectrale.

2. Récurrence de Poincarré

Théorème 2.1. *Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique mesuré et $E \in \mathcal{B}$ avec $\mu(E) > 0$. Alors*

$$\sharp\{k \in \mathbb{N},\ f^k x \in E\} = +\infty \text{ pour } \mu\text{-p.t. } x \in E.$$

Démonstration. On note $E_0 := \{x \in E,\ f^k x \notin E \text{ pour tout } k \in \mathbb{N}^*\}$. Clairement $f^{-n}E_0 \cap f^{-m}E_0 = \emptyset$ pour $n \neq m \in \mathbb{N}$. Alors on a

$$+\infty > \mu\left(\bigcup_{n\in\mathbb{N}} f^{-n}E_0\right) = \sum_{n\in\mathbb{N}} \mu(f^{-n}E_0) = \sum_{n\in\mathbb{N}} \mu(E_0).$$

On a donc nécessairement $\mu(E_0) = 0$. Puis si on note $F := \{x \in E,\ \sharp\{k \in \mathbb{N},\ f^k x \in E\} < +\infty\}$, on doit montrer $\mu(F) = 0$. Mais $F = \bigcup_{n\in\mathbb{N}} f^{-n} E_0$... $\square$

DÉFINITION 2.2. *Soit $(X, \mathcal{B}, f)$ un système dynamique mesurable. Un point $x \in X$ est dit récurrent s'il existe une suite strictement croissante d'entiers $(n_k)_k$ telle que $f^{n_k} x \xrightarrow{k\to+\infty} x$.*

THÉORÈME 2.3. *Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique mesuré. Alors μ-p.t. $x \in X$ est récurrent.*

DÉMONSTRATION. Soit $\mathcal{V} = (V_n)_n$ une base dénombrable de voisinages de X. Soit F_n le sous-ensemble des points de V_n revenant finiment souvent dans V_n. D'après le théorème de récurrence de Poincarré, $\mu(F_n) = 0$. Alors on vérifie facilement que tout $x \notin \bigcup_{n\in\mathbb{N}} F_n$ est récurrent. $\square$

3. Ergodicité

Un système dynamique mesuré est ergodique s'il est irréductible dans le sens où on ne peut pas le décomposer en deux systèmes disjoints.

DÉFINITION 3.1. *Un système mesuré $(X, \mathcal{B}, f\mu)$ est dit ergodique lorsque tout ensemble f-invariant A (i.e. $f^{-1}A = A$), est de mesure nulle ou totale.*

PROPOSITION 3.2. *μ est ergodique ssi pour tout $\phi : X \to \mathbb{R}$ mesurable,*

$$[\phi \circ f = \phi\ \mu\ p.p.] \Rightarrow [\exists C \in \mathbb{R},\ \phi = C\ \mu\ p.p.]\,.$$

DÉMONSTRATION. L'implication $\Leftarrow$ suit directement en prenant $\phi = \chi_A$ avec $f^{-1}A = A$. Montrons maintenant $\Rightarrow$. On commence par considérer $\phi = \chi_E$ une indicatrice. On pose $E_0 := \{x \in X,\ f^k x \in E$ pour une infinité de $k \in \mathbb{N}\}$. L'ensemble E_0 est invariant, donc de mesure nulle ou totale. Il suffit pour conclure de voir que $\chi_E = \chi_{E_0}$ presque partout. D'après le théorème de récurence de Poincaré, $\mu(E \setminus E_0) = 0$. Puis on a $E_0 \setminus E \subset \left(\bigcup_{k\in\mathbb{N}} f^{-k-}E\right) \setminus E \subset \bigcup_{k\in\mathbb{N}} \left(f^{-k-1}E \setminus f^{-k}E\right)$ et $\mu(f^{-k-1}E \setminus f^{-k}E) = \mu(f^{-1}E \setminus E) = 0$ (car $\chi_E \circ f = \chi_E\ \mu$ p.p.). On suppose maintenant ϕ général. Pour tout réel t, l'ensemble $E_t = \{\phi > t\}$ est presque invariant, i.e. μ p.p. $\chi_{E_t} \circ f = \chi_{E_t}$. On déduit de la discussion précédente que $\mu(E_t) \in \{0, 1\}$. On conclut facilement que $\phi = c$ presque partout avec $c = \sup\{t,\ \mu(E_t) = 0\}$. $\square$

COROLLAIRE 3.3. *L'ergodicité est une propriété spectrale : μ est ergodique ssi l'espace propre de U_f associé à la valeur propre 1 est de dimension 1 (correspondant à l'ensemble des fonctions constantes presque partout).*

Exemples :

— la rotation d'angle α sur le cercle est ergodique (pour la mesure de Lebesgue) ssi α est irrationnel.
— un automorphime linéaire du tore f_A est ergodique ssi A n'a pas de valeurs propres racines de l'unité.

Pour montrer l'ergodicité de la mesure de Lebesgue dans ces deux exemples on peut par exemple utiliser le développement en série de Fourier pour établir qu'il n'existe pas de fonctions L^2 invariantes autre que les constantes. On conclut à l'aide du corollaire ci-dessus. Nous le démontrons pour les automorphismes du tore et laissons le cas plus facile de la rotation irrationelle. Sur le tore $\mathbb{R}^n/\mathbb{Z}^n$ les fonctions $(e_k)_{k\in\mathbb{Z}^n}$, $e_k(\overline{x}) = e^{2i\pi<k,x>}$ definissent une base hilbertienne de L^2. Soit $\phi \in L^2$ tel que $\phi \circ f_A = \phi$ et notons $\phi = \sum_k a_k(\phi)e_k$ son développement en série de Fourier. Alors pour tout $k \in \mathbb{Z}^n$,

$$\begin{aligned}
a_k(\phi \circ f_A) &= < \phi \circ f_A, e_k >, \\
&= \int_{\mathbb{R}^n/\mathbb{Z}^n} \phi(A\overline{x})e^{-2i\pi<k,x>}dLeb(\overline{x}), \\
&= \int_{\mathbb{R}^n/\mathbb{Z}^n} \phi(\overline{x})e^{-2i\pi<k,A^{-1}x>}dLeb(\overline{x}), \\
&= \int_{\mathbb{R}^n/\mathbb{Z}^n} \phi(\overline{x})e^{-2i\pi<(A^*)^{-1}k,x>}dLeb(\overline{x}), \\
&= a_{(A^*)^{-1}k}.
\end{aligned}$$

L'identité $\phi \circ f_A = \phi$ donne alors $a_{A^*l} = a_l$ pour tout $l \in \mathbb{Z}^n$. Si A n'a pas de racine de l'unité il en est de même de $A*$ (une matrice et son adjoint ont le même spectre). Donc pour $l \neq 0$ le sous ensemble de $\mathbb{Z}^n$ formé par les $(A^*)^k l$ est infini (pourquoi ?). La série des $(a_k)_k$ étant de carré sommable on a nécessairement $a_l = 0$ pour tout $l \neq 0$ et donc ϕ est constante. La mesure de Lebesgue est en fait dans ce cas totalement ergodique, i.e. elle est ergodique pour toutes les puissances de f_A. Enfin si A et donc A^* admet une valeur propre racine n^{me} de l'unité alors il existe $l \in \mathbb{Z}^n$ et $m \in \mathbb{N}^*$ tel que $(A^*)^m l = l$. La fonction L^2 $\sum_{0\leq p<m} e_{(A^*)^p l}$ est alors f_A-invariante.

Pour la rotation, au lieu des séries de Fourier on peut aussi utiliser "la distorsion nulle" des translations comme suit (voir aussi Lemme 1.1 Chapitre 5 sur l'application de Gauss). Soit A un ensemble invariant par f_α avec $Leb(A) \in]0,1[$ et établissons une contradiction. D'après le théorème de différentiation de Lebesgue, presque tout point $x \in A$ est de densité, i.e. $\frac{Leb(B(x,r)\cap A)}{Leb(B(x,r))} \xrightarrow{r\to 0} 1$. Soit $x_0 \in A$ un tel point de densité (par hypothèse $Leb(A) > 0$) et soit $r_0 > 0$ tel que pour tout $r < r_0$ on ait $\frac{Leb(B(x_0,r)\cap A)}{Leb(B(x_0,r))} > Leb(A)(< 1)$. Alors par densité de $\alpha\mathbb{Z}$ dans le cercle on peut recouvrir le cercle par une collection dénombrable de boules (qui sont ici juste des intervalles) fermées d'intérieur disjoint de la forme $f_\alpha^n B(x_0, r)$ avec $n \in \mathbb{Z}$ et $r < r_0$ (faites-le !). Enfin, on a pour chacune de ces boules, par invariance de A et de la mesure de Lebesgue, $Leb(f_\alpha^n [B(x_0, r)] \cap A) = Leb(f_\alpha^n [B(x_0, r) \cap A]) = Leb(B(x_0, r) \cap A) > Leb(A)Leb(B(x_0, r))$ d'où la contradiction en sommant l'inégalité précédente sur toutes les boules de la collection.

4. Mélange

Définition 4.1. *μ est dite mélangeante lorsque pour tout $A, B \in \mathcal{B}$, on a*

$$\mu\left(f^{-n}A \cap B\right) \xrightarrow{n\to+\infty} \mu(A)\mu(B).$$

Proposition 4.2. *μ est mélangeante ssi pour tout $\phi, \psi \in L^2$ on a*

$$\int \phi \circ f^n \psi d\mu \xrightarrow{n\to+\infty} (\int \phi d\mu)(\int \psi d\mu).$$

Autrement dit pour tout $\phi \in L^2$, la suite $(\phi \circ f^n)_n$ converge faiblement vers $\int \phi d\mu$. En particulier le mélange est une propriété spectrale.

Remarquez que pour montrer le mélange il suffit d'établir le critère précédent pour tout ϕ, ψ dans une partie dense de $L^2(I)$.

Proposition 4.3. *Le mélange entraine l'ergodicité.*

Démonstration. Soit μ mélangeante et C invariant, $f^{-1}C = C$. Alors en prenant $A = B = C$ dans la definition du mélange, on obtient $\mu(A) = \mu(A)^2$ et donc $\mu(A) \in \{0, 1\}$. $\square$

Exemples :

— la rotation sur le cercle n'est jamais mélangeante,

— les automorphismes linéaire du n-tore ergodiques sont mélangeant.

Montrons le deuxième point. Il suffit de montrer la Proposition 4.2 pour $\phi, \psi \in \{e_k,\ k \in \mathbb{Z}^n\}$ (pourquoi ?). Remarquons que U_{f_A} laisse invariant cet ensemble. Puis si $k, l \in \mathbb{Z}^n$, il existe au plus un entier n tel que $e_k \circ f_A^n = e_l$ (sinon e_k serait invariant par une certaine puissance de f_A, ce qui est exclu puisque A est totalement ergodique). Donc pour $m > n$ on a $< e_k \circ f_A^m, e_l >= 0$ ce qui entraine bien le mélange.

5. Bernoulli

Soit S l'ensemble fini $S = \{1, ..., s\}$. On note μ_p la mesure de Bernoulli de paramètre $p = (p_1, ..., p_s)$ (avec $p_i \geq 0$ et $\sum_i p_i = 1$) sur S. On considère le décalage σ sur le produit infini $S^{\mathbb{Z}}$, i.e. $\sigma((x_n)_n) = (x_{n+1})_n$. On note $\mu_p^{\mathbb{Z}}$ l'unique measure de proba sur $S^{\mathbb{Z}}$ (obtenu par le théorème de Carathéodory ou loi de consistence de Kolmogorov) satisfaisant pour tout cylindre $[x_k, ..., x_l] := \{(y_n)_n \in S^{\mathbb{Z}},\ y_i = x_i \text{ for } i = k, ..., l\}$ l'égalité suivante

$$\mu_p^{\mathbb{Z}}([x_k, ..., x_l]) = \prod_{i=k}^{l} p_{x_i}.$$

Clairement $\mu_p^{\mathbb{Z}}$ est σ-invariante.

PROPOSITION 5.1. *Tout schéma de Bernoulli* $(S^{\mathbb{Z}}, \mathcal{B}, \sigma, \mu_p^{\mathbb{Z}})$ *est mélangeant.*

DÉMONSTRATION. Il suffit de montrer le mélange pour ϕ et ψ des indicatrices de cylindres (pourquoi ?), disons respectivement $[x_k, ..., x_l]$ et $[y_m, ..., y_n]$. Alors $\phi \circ \sigma^n$ est l'indicatrice du cylindre $[x_{k-n}, .., x_{l-n}]$ de sorte que pour $l - n < m$ la fonction $\phi \circ \sigma^n \psi$ est l'indicatrice sur l'union des cylindres de la forme $[x_{k-n}, ..., x_{l-n}, z_{l-n+1}, ..., z_{m-1}, y_m, ..., y_n]$ pour tous les $z_{l-n+1}, ..., z_{m-1} \in S$. On calcule facilement que la mesure de cette union coincide avec le produit des mesures des deux cylindres initiaux $[x_k, ..., x_l]$ et $[y_m, ..., y_n]$. □

On peut montrer (faites le !) que le doublement de l'angle f_2 est isomorphe à un schéma de Bernoulli de paramètre $(1/2, 1/2)$. En particulier, il est mélangeant aussi.

6. Exercices

Exercice 1. *Formule de Kac.*
Soit $(X, \mathcal{B}, \mu, f)$ un système mesuré et $A \in \mathcal{B}$ avec $\mu(A) > 0$. D'après le théorème de récurrence de Poincarré il existe $A_0 \subset A$ avec $\mu(A_0) = \mu(A)$ tel que pour $x \in A_0$ on revient dans A. On note $\tau_A : A_0 \to A$ l'application de premier retour dans A, i.e. $\tau_A(x) = \inf\{n > 0,\ f^n(x) \in A\}$.

— montrez par récurrence que pour tout $\phi \in L^\infty(X)$ et pour tout entier N,

$$\int \phi d\mu = \int_A \sum_{j=0}^{N-1} \phi \circ f^j . 1_{\tau_A > j} d\mu + \int \phi \circ f^N . 1_{\bigcap_{j=1}^N f^{-j} A^c} d\mu,$$

— En déduire que $\int \phi d\mu = \int_A \sum_{k=0}^{\tau_A - 1} \phi \circ f^k d\mu$,

— Montrez que $\int_A \tau_A = \frac{1}{\mu(A)}$.

Exercice 2.
On se propose de montrer que si $([0,1], f, \mathcal{B}_{[0,1]}, \mu)$ est un système mesuré sur l'intervalle alors pour μ-p.t. x, on a

$$\liminf_n n|f^n(x) - x| \leq 1$$

— En raisonnant par l'absurde montrer qu'il existe $c > 1$ et N tel que $E = \{x,\ n|f^n(x) - x| > c \text{ pour tout } n > N\}$ est de mesure positive,

— Pour un point de densité a de E on considère $E_r = E \cap B(a, r)$ pour r petit. Donner une borne inférieure pour le temps de retour d'un point de E_r,

— Obtenez une contradiction en utilisant la formule de Kac.

Exercice 3. *Lemme de Rohlin.*
Soit $(X, f, \mathcal{B}, \mu)$ un système ergodique inversible apériodique. On se propose de montrer que pour tout $\epsilon > 0$ et tout $n > 0$ il existe un $A' \in \mathcal{B}$ tel que $A', fA', ..., f^{n-1}A'$ soient deux à deux disjoints et $\mu(\bigcup_{0 \leq k < n} f^k A') > 1 - \epsilon$.

— Montrer que pour tout $A \in \mathcal{B}$ avec $\mu(A) > 0$, on a $\mu(\bigcup_{n \in \mathbb{N}} f^n A) = 1$,

— Pour $A \in \mathcal{B}$ on note $A_k = A \cap \{\tau_A = k\}$ avec τ_A le premier temps de retour dans A. Montrer que les ensembles de la forme $B = f^{mn} A_k$ avec $mn < k$ vérifie que $B, fB, ..., f^{n-1}B$ sont deux à deux disjoints,

— On considère l'union A' de tous ces ensembles B. Montrez que A', fA',..., $f^{n-1}A'$ sont deux à deux disjoints puis que

$$\mu\left(X \setminus \left(\bigcup_{0 \leq k < n} T^k A'\right)\right) \leq (n-1)\mu(A),$$

— Conclure.

Chapitre 2

Théorèmes ergodiques

Contents

1.	**Théorème ergodique en moyenne**	**13**
2.	**Théorèmes ergodiques ponctuels**	**16**
3.	**Preuve de l'inégalité maximale**	**20**
4.	**Deux exemples d'applications**	**21**
4.1.	Nombres normaux	21
4.2.	Exposant de Lyapunov maximal positif	22
5.	**Exercices**	**22**

On considère toujours dans ce chapitre un système dynamique mesuré $(X, \mathcal{B}, f, \mu)$. On notera $\mathcal{I}$ la tribu des invariants, i.e.

$$\mathcal{I} = \{A \in \mathcal{B},\ f^{-1}A = A\}.$$

On voit facilement que $L^2(\mathcal{I})$ coincide avec $\{\phi \in L^2(\mathcal{B}),\ \phi \circ f = \phi\}$. Pour $\phi \in L^2 = L^2(\mathcal{B})$ l'espérance conditionnellle de ϕ relativement à $\mathcal{I}$ est la projection orthogonale de ϕ sur $L^2(\mathcal{I})$. Lorsque la mesure est ergodique, la tribu $\mathcal{I}$ des invariants est triviale (à des ensembles de mesure nulle près) et donc $E[\phi|\mathcal{I}] = \int \phi d\mu$ pour tout $\phi \in L^1$.

1. Théorème ergodique en moyenne

Rappelons enfin que l'opérateur de Koopman sur L^2 qui à ϕ associe $\phi \circ f$ est noté U_f et définit une isométrie.

THÉORÈME 1.1 (Théorème ergodique en moyenne). *Avec les notations ci-dessus, on a pour tout $\phi \in L^2$,*

$$\frac{1}{n} \sum_{k=0}^{n-1} U_f^k \phi \xrightarrow{n \to +\infty} E[\phi|\mathcal{I}] \text{ dans } L^2$$

En fait ce résultat suit du théorème plus général suivant (prendre $H = L^2$, $U = U_f$, $Ker(Id - U) = L^2(\mathcal{I})$) :

THÉORÈME 1.2. *Soit H un espace de Hilbert et $U : H \to H$ une contraction linéaire. Alors on a pour tout $x \in H$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k x \xrightarrow{n\to+\infty} p(x),$$

où $p(x)$ est le projeté orthogonal de x sur $Ker(Id-U)$.

On commence par rappeler quelques faits classiques d'analyse hilbertienne.

LEMME 1.3. *Soit H un hilbert. On a*

— *si F est un sev de H alors*

$$H = \overline{F} \oplus^{\perp} F^{\perp},$$

— *si $U : H \to H$ est une application linéaire continue, alors*

$$Im(U)^{\perp} = Ker(U^*),$$

— *si $U : H \to H$ est une contraction linéaire (i.e. $\|Ux\| \leq \|x\|$ pour tout x), alors U^* est aussi une contraction linéaire et*

$$Ker(Id-U) = Ker(Id-U^*).$$

DÉMONSTRATION. On montre juste le dernier point les deux autres étant classiques. On a par l'inégalité de Cauchy-Schwartz pour tout x, y :

$$< y, U^*x >=< Uy, x >\leq \|x\|\|y\|.$$

Donc U^* est aussi une contraction linéaire. Puis x est dans le noyau de $Id-U$ ssi on est dans le cas d'égalité du Cauchy-Schwartz ci-dessus avec $x=y$, qui entraine que x est aussi dans le noyau de $Id-U^*$. □

On montre maintenant le théorème 1.2. D'après le lemme précédent appliqué avec $F = Im(Id-U)$, on a

$$\begin{aligned} H &= \overline{Im(Id-U)} \oplus^{\perp} Im(Id-U)^{\perp}, \\ &= \overline{Im(Id-U)} \oplus^{\perp} Ker(Id-U^*), \\ &= \overline{Im(Id-U)} \oplus^{\perp} Ker(Id-U). \end{aligned}$$

On a

— si $x \in Ker(Id-U)$, alors $\frac{1}{n}\sum_{k=0}^{n-1} U^k x = x = p(x)$,

— si $x \in Im(Id-U)$, il existe $y \in X$ tel que $x = y - Uy$ et donc $\frac{1}{n}\sum_{k=0}^{n-1} U^k x = \frac{1}{n}(y - U^n y) \xrightarrow{n} 0 = p(x)$ puisque $\|U^n y\| \leq \|y\|$.

Par linéarité il s'ensuit que le théorème est vrai pour tout $x \in Im(Id-U) \oplus^{\perp} Ker(Id-U)$.

LEMME 1.4. *Soit $(U_n)_n$ une suite de contractions linéaires, alors*

$$G = \{x \in H,\ (U_n x)_n \text{ converge}\}$$

est un sev fermé de H et la limite $U : G \to H$ est linéaire continue.

Ce lemme conclut la preuve du Théorème 1.2 avec $U_n = \frac{1}{n}\sum_{k=0}^{n-1} U^k$ pour tout n. Puis G étant dense dans H dans notre cas, on a la convergence en tout point et la limite est continue. Puisqu'elle coincide avec p (continue sur H) sur G elle est égale à p.

DÉMONSTRATION DU LEMME 1.4. On note pour tout x le reste de Cauchy

$$\Delta(x) = \lim_N \sup_{q,q'>N} \|U_q x - U_{q'} x\|,$$

de sorte que $(U_n x)_n$ converge ssi $\Delta(x) = 0$. Montrons que G est fermé. Soit $(x_p)_p$ une suite de H telle que $\Delta(x_p) = 0$ pour tout p et qui converge vers x. On a alors

$$\begin{aligned}
\Delta(x) &\leq \Delta(x - x_p) + \Delta(x_p), \\
&\leq \Delta(x - x_p), \\
&\leq 2\sup_n \|U_n(x - x_p)\|, \\
&\leq \|x - x_p\| \xrightarrow{p} 0.
\end{aligned}$$

On a clairement par linéarité des $(U_n)_n$ que G est un espace vectoriel et que la limite U est linéaire. Les applications $(U_n)_n$ étant des contractions il en est de même de U. $\square$

COROLLAIRE 1.5. *μ est ergodique ssi $\forall \phi, \psi \in L^2$, on a $\frac{1}{n}\sum_{k=0}^{n-1} \int \phi \circ f^n \psi d\mu \xrightarrow{n} \int \phi d\mu \int \psi d\mu$.*

THÉORÈME 1.6 (Théorème ergodique en moyenne L^1). *Avec les notations ci-dessus, on a pour tout $\phi \in L^1$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} U_f^k \phi \xrightarrow{n\to+\infty} E[\phi|\mathcal{I}] \text{ dans } L^1.$$

DÉMONSTRATION. Dans la preuve précédente nous avons vu que pour la restriction U de U_f à L^2, le sous espace vectoriel F de L^2 défini par $F = Im(Id-U)+Ker(Id-U)$ était dense dans $(L^2, \|\|_2)$. L'espace L^2 étant dense dans $(L^1, \|\|_1)$ et l'injection canonique de $(L^2, \|\|_2)$ dans $(L^1, \|\|_1)$ étant continue, l'espace vectoriel G est aussi dense dans $(L^1, \|\|_1)$. De plus pour $\phi \in F$ on a la convergence L^2 et donc L^1 de $(U_n\phi)_n$ vers $E[\phi|\mathcal{I}]$. Le Lemme 1.4 permet de conclure la convergence L^1 de $(U_n\phi)_n$ pour tout $\phi \in L^1$. De plus la limite définit un opérateur linéaire continue. Celui-ci est égal sur F à l'espérance conditionnelle relativement à la tribu des invariants, qui est aussi un opérateur linéaire continue de $(L^1, \|\|_1)$. Par densité de F dans $(L^1, \|\|_1)$ ils coincident sur tout L^1. □

2. Théorèmes ergodiques ponctuels

Théorème 2.1 (Théorème ergodique ponctuel). *Soit* $(X, \mathcal{B}, f, \mu)$ *un système dynamique mesuré et* $\mathcal{I}$ *sa tribu des invariants. Alors pour tout* $\phi \in L^1(\mu)$,

$$\frac{1}{n}\sum_{k=0}^{n-1} \phi(f^k x) \xrightarrow{n} E[\phi|\mathcal{I}] \quad \textit{pour } \mu \textit{ p.t. } x.$$

La convergence a aussi lieu dans L^1.

Une conséquence directe du théorème ergodique est donnée par le lemme suivant, dont nous donnons une preuve directe indépendante.

Lemme 2.2. *Avec les notations précédentes,*

$$\frac{1}{n}\phi(f^n x) \xrightarrow{n} 0 \quad \textit{pour } \mu \textit{ p.t. } x.$$

Démonstration. Pour $\epsilon > 0$ on a

$$\begin{aligned} \sum_{n\geq 1} \mu(|\phi \circ f^n| > n\epsilon) &= \sum_{n \geq 1} \mu\left(\frac{|\phi|}{\epsilon} > n\right), \\ &\leq \int_{|\phi|>\epsilon} \frac{\phi}{\epsilon}\, d\mu. \end{aligned}$$

On conclut à l'aide du lemme de Borel-Cantelli. □

Théorème 2.3 (Théorème ergodique sous-additif). *Soit* $(X, \mathcal{B}, f, \mu)$ *un système dynamique mesuré et* $(\phi_n)_n \in L^1(\mu)^{\mathbb{N}}$ *une suite sous-additive (pour* f*), i.e.* $\phi_{n+m}(x) \leq \phi_n(x) + \phi_m(f^n x)$ *pour tout entiers* $m, n \geq 0$ *et tout* $x \in X$*. Alors il existe une fonction* $\overline{\phi} \in L^1(\mu)$ f*-invariante telle que*

$$\frac{\phi_n(f^n x)}{n} \xrightarrow{n} \overline{\phi} \quad \textit{pour } \mu \textit{ p.t. } x.$$

La convergence a aussi lieu dans L^1 *et* $\int \overline{\phi} = \inf_n \frac{\int \phi_n}{n}$.

Nous allons montrer le théorème ergodique sous-additif. Celui-ci entraine le théorème ergodique ponctuel. En effet la limite presque sûre $\overline{\phi}$ coincide nécessairement dans ce cas avec la limite L^1 qui est l'espérance conditionnelle relativement aux invariants d'après le Théorème 1.6 (on peut extraire de tout suite convergente dans L^1 une sous-suite qui converge presque sûrement vers la limite L^1).

L'outil fondamental pour montrer la convergence presque sûr d'une famille de fonctions est une inégalité dite maximale (voir aussi le théorème de différentiation de Lebesgue ou le théorème de Doobs de convergences des martingales).

LEMME 2.4 (Inégalité Maximale). *Soit* $(\phi_n)_n \in L^1(\mu)^{\mathbb{N}}$ *suradditive, i.e.* $(-\phi_n)$ *sousadditive, et* $\phi_n \geq 0$ *pour tout* n*. Alors pour tout* $\alpha > 0$

$$\alpha\mu\left(\sup_n \frac{\phi_n}{n} > \alpha\right) \leq \sup_n \frac{\int \phi_n d\mu}{n}.$$

Pour $\phi \in L^1(\mu)$ et $n \in \mathbb{N}^*$ on note $U_n(\phi)$ la suite (sous, sur)-additive donne par

$$U_n(\phi) = \sum_{k=0}^{n-1} \phi \circ f^k.$$

Clairement lorsque ϕ est un cobord borné (i.e. il existe $\psi \in L^\infty(\mu)$ tel que $\phi = \psi \circ f - \psi$) ou ϕ une fonction invariante (i.e $\phi \circ f = \phi$) ou encore une combinaison linéaires de telles fonctions, les suites (sous) additives $(U_n(\phi))_n$ verifient le théorème ergodique sous-additif. On va montrer :

— toute suite sous-additive est limite dans un certain sens d'une telle suite,

— la convergence p.p. et L^1 passe à la limite au moyen de l'inégalité maximale.

On va commencer par établir le premier point mais précisons tout d'abord la notion de convergence utilisée. Soit (ψ_n) et (ϕ_n^1), (ϕ_n^2),... des suites de fonctions L^1. On dira que "$(\phi_n)^p$ converge vers (ψ_n) quand p tend vers l'infini" s'il existe des suites $(\theta_n)^p$ dans $L^1(\mu)$, tels que

— $(\theta_n)^p$ est positive et suradditive pour f^p, i.e. $\theta_{(m+n)p}(x) \geq \theta_{mp}(x) + \theta_{np}(f^{pm}x)$ pour tout $m, n \in \mathbb{N}$ et tout $x \in X$,

— $\limsup_n \frac{\theta_{np}^p}{np} = \limsup_n \frac{\theta_n^p}{n}$ presque partout et dans L^1,

— $|\psi_n - \phi_n^p| \leq \theta_n^p$ pour tout n, p,

— $\sup_n \frac{\int \theta_{np}^p d\mu}{np} \xrightarrow{p} 0$.

On déduit facilement de la preuve du théorème ergodique en moyenne le lemme suivant.

Lemme 2.5. *L'espace vectoriel E_p de $L^\infty \cap L^1(\mu)$ engendré par $\{\phi \ : \ \exists \psi \in L^\infty(\mu),\ \phi = \psi \circ f^p - \psi\}$ et $\{\phi \ : \ \phi \circ f^p = \phi\} = L^1_{f^p}(\mathcal{I})$ est dense dans $L^1(\mu)$.*

Démonstration. On peut supposer $p = 1$. On a vu que l'espace vectoriel engendré par $\{\phi \ : \ \exists \psi \in L^2,\ \phi = \psi \circ f - \psi\}$ et $\{\phi \in L^2 \ : \ \phi \circ f = \phi\}$ est dense dans L^2. L'espace L^∞ étant dense dans L^2 l'espace E_1 est aussi dense dans L^2. Or L^2 est un sous ensemble dense de L^1 et l'injection de L^2 dans L^1 est continue. Il s'en suit que E_1 est aussi dense dans L^1.

$\square$

Pour tout p et toute suite $\phi = (\phi_n) \in L^1(\mu)^{\mathbb{N}}$ (avec $\phi_0 = 0$) on note $(U^p_n \phi)_n$ la suite définie pour tout x et tout entiers n, p par

$$U^p_n \phi(x) = \phi_{n-[n/p]p}(f^{[n/p]p}x) + \sum_{k=0}^{[n/p]} \phi_p(f^{kp}x).$$

Comme on l'a remarqué précédemment pour $\phi_p \in E_p$ et $\phi_1, ..., \phi_{p-1} \in L^\infty$ la suite $\left(\frac{U^p_n \phi(x)}{n}\right)_n$ converge pour presque tout x.

Lemme 2.6. *Soit $(\phi_n) \in L^1(\mu)^{\mathbb{N}}$ une suite sous-additive et $(\tilde{\phi}_n) \in \prod_n E_n$ avec $\|\phi_n - \tilde{\phi}_n\|_1 < 1$ pour tout n, alors $(U^p_n \tilde{\phi})^p$ "converge vers" (ϕ_n).*

Démonstration. On a $|U^p_n \phi - U^p_n \tilde{\phi}| \leq U^p_n |\phi - \tilde{\phi}|$ et ainsi $|U^p_n \tilde{\phi} - \phi_n| \leq U^p_n |\phi - \tilde{\phi}| + U^p_n \phi - \phi_n$. Le terme de droite θ^p_n définit bien une suite suradditive pour f^p et positive. De plus on vérifie $\limsup_n \frac{\theta^p_{np}}{np} = \limsup_n \frac{\theta^p_n}{n}$. En effet si on pose $V^p_n = (U^p_n \phi - \phi_n)_n$ alors on a pour $n = (k+1)p - s$ et $r = p - s$ avec $0 \leq s < p$ on a

$$\begin{aligned}
V_{(k+1)p}(x) &\geq U^p_{(k+1)p}(x) - \phi_n(x) - \phi_s(f^n x), \\
&\geq U^p_{kp}(x) + \phi_p(f^{kp}x) - \phi_n(x) - \phi_s(f^n x), \\
&\geq V_n(x) - \phi_r(f^{kp}x) + \phi_p(f^{kp}x) - \phi_s(f^n x).
\end{aligned}$$

Les trois termes de droites quotientés par n tendent vers 0 d'après le Lemme 2.2. De meme on a

$$U^p_{(k+1)p}|\phi - \tilde{\phi}|(x) \geq U^p_n |\phi - \tilde{\phi}|(x) + |\phi_p - \tilde{\phi}_p|(f^{kp}x) - |\phi_r - \tilde{\phi}_r|(f^{kp}x).$$

On conclut alors que $\limsup_n \frac{\theta^p_{np}}{np} = \limsup_n \frac{\theta^p_n}{n}$.

Enfin on a pour tout entier $n > 0$

$$\begin{aligned}
\frac{\int \theta_{np}^p d\mu}{np} &= \frac{1}{np}\int \left(U_{np}^p|\phi - \tilde{\phi}| + U_{np}^p\phi - \phi_{np}\right) d\mu \\
&\leq \frac{\|\phi_p - \tilde{\phi}_p\|_1}{p} + \sup_n \int \left(\frac{\phi_p}{p} - \frac{\phi_{np}}{np}\right) d\mu,
\end{aligned}$$

puis d'après le Lemme 3.2 énoncé plus loin

$$\sup_{n>0} \frac{\int \theta_{np}^p d\mu}{np} \leq 1/p + \int \frac{\phi_p}{p} d\mu - \lim_m \int \frac{\phi_m}{m} d\mu \xrightarrow{p\to+\infty} 0.$$

□

On montre maintenant le deuxième point.

LEMME 2.7. *Avec les notations precedentes on suppose que $(\phi^p)^p = (\phi_n^p)^p$ converge vers $\psi = (\psi_n)$.*

1. *si pour tout p la suite $(\frac{\phi_n^p(x)}{n})$ converge pour μ p.t. x, alors la suite $(\frac{\psi_n(x)}{n})$ converge aussi pour μ p.t. x,*
2. *si pour tout p la suite $(\frac{\phi_n^p}{n})$ converge dans L^1, alors la suite $(\frac{\psi_n}{n})$ converge aussi dans L^1.*

DÉMONSTRATION. Pour $x \in X$ et une suite de fonction $\phi = (\phi_n)_n$ on note $\Delta\phi(x)$ le reste de Cauchy en x de la suite $\frac{\phi_n(x)}{n}$, i.e. $\Delta(x) := \lim_N \sup_{q,q'\geq N} |\frac{\phi_q(x)}{q} - \frac{\phi_{q'}(x)}{q'}|$. On a pour tout x et tout p

$$\begin{aligned}
\Delta\psi(x) &\leq \Delta|\psi - \phi^p|(x) + \Delta(\phi_p), \\
&\leq \Delta|\psi - \phi^p|(x).
\end{aligned}$$

puis pour tout $\alpha > 0$

$$\begin{aligned}
\mu(\Delta\psi > \alpha) &\leq \mu(\Delta|\psi - \phi^p| > \alpha), \\
&\leq \mu(\limsup_n \frac{|\psi_n - \phi_n^p|}{n} > \alpha), \\
&\leq \mu(\limsup_n \frac{\theta_n^p}{n} > \alpha).
\end{aligned}$$

Par hypothèse $\limsup_n \frac{\theta_{np}^p}{np} = \limsup_n \frac{\theta_n^p}{n}$ et donc en utilisant ensuite l'inegalite maximale :

$$\begin{aligned}
\mu(\Delta\psi > \alpha) &\leq \mu(\limsup_n \frac{\theta_n^p}{n} > \alpha), \\
&\leq \mu(\limsup_n \frac{\theta_{np}^p}{np} > \alpha), \\
&\leq \frac{1}{\alpha} \sup_n \frac{\int \theta_{np}^p d\mu}{np}.
\end{aligned}$$

Comme ce dernier terme tends vers 0 quand p tends vers l'infini on en deduit que $\Delta\psi(x) = 0$ pour μ pp x.

Montrons enfin le second point qui n'utilise pas l'inégalité maximale. En notant comme dans le Lemme 1.4 la notation $\Delta(\Xi)$ pour désigner le reste de Cauchy d'une suite Ξ de $L^1(\mu)$ on a avec $\Psi = (\psi_n/n)_n$, $\Phi^p = (\phi_n^p/n)$, $\Theta^p = (\theta_n^p/n)_n$ et $\tilde{\Theta}^p = (\theta_{np}^p/np)_n$:

$$\begin{aligned}
\Delta(\Psi) &= \Delta(|\Psi - \Phi^p|), \\
&\leq \Delta(\Theta^p), \\
&\leq \Delta(\tilde{\Theta}^p) = \xrightarrow{p} 0.
\end{aligned}$$

□

3. Preuve de l'inégalité maximale

Il nous reste a montrer l'inégalité maximale.

INEGALITE MAXIMALE. On aura besoin du lemme élémentaire suivant dû à F. Riesz :

LEMME 3.1. *Soient* $u_1, ..., u_n \in \mathbb{R}$. *On note* $v_j = \max(0, u_{j+1}, u_{j+1} + u_{j+2}, ..., u_{j+1} + ... + u_{j+n})$ *pour* $j = 0, ..., n-1$ *et* $v_n = 0$. *Alors*

$$\sum_{j=0}^{n-1} u_{j+1} 1_{v_j > 0} \geq 0.$$

Le lemme se déduit facilement des observations suivantes. Pour tout $j \geq 0$, on a

$$v_j \leq v_{j+1} + u_{j+1} 1_{v_j > 0}$$

puis en sommant sur $j = 0, ..., n-1$ on obtient

$$0 \leq v_0 - v_n = v_0 \leq \sum_{j=0}^{n-1} u_{j+1} 1_{v_j > 0}.$$

On revient maintenant à la preuve de l'inégalité maximale. On applique le lemme de Riesz au n-uplet $u_j = \phi_j - \phi_{j-1} - \alpha$ pour $j = 1, ..., n$:

$$\sum_{j=0}^{n-1} (\phi_j - \phi_{j-1} - \alpha) 1_{v_j > 0} \geq 0$$

Puis $(\phi_n)_n$ est suradditive positive, donc croissante d'où pour tout j

$$\phi_j - \phi_{j-1} \geq (\phi_j - \phi_{j-1}) 1_{v_j > 0} \text{ et donc}$$

$$\phi_n \geq \alpha \sum_{j=0}^{n-1} 1_{v_j > 0}$$

puis en intégrant

$$\int \phi_n d\mu \geq \alpha \sum_j \mu(v_j > 0)$$

Mais par suradditivité on a $(v_j > 0) = (\max_{n \geq l > j}(\phi_l - \phi_j) > (l-j)\alpha) \supset (\max_{n-j \geq l-j > 0}(\phi_{l-j} \circ f^j) > (l-j)\alpha)$. Par invariance de μ ce dernier ensemble à même mesure que $(\max_{n-j \geq l-j > 0}(\phi_{l-j}) > (l-j)\alpha)$ et donc en changeant les variables, $k = n - j$ et $m = l - j$,

$$\sum_j \mu(v_j > 0) \quad \geq \quad \sum_{k=1,\dots,n} \mu(\max_{0 \leq m \leq k} \frac{\phi_m}{m} > \alpha).$$

et donc puisque $\mu(\max_{0 \leq m \leq k} \frac{\phi_m}{m} > \alpha) \xrightarrow{k \to +\infty} \mu\left(\sup_n \frac{\phi_n}{n} > \alpha\right)$ on a par Césaro

$$\liminf_n \frac{\int \phi_n d\mu}{n} \geq \alpha\mu\left(\sup_n \frac{\phi_n}{n} > \alpha\right).$$

Il suffit pour conclure de verifier que $\liminf_n \frac{\int \phi_n d\mu}{n} = \sup_n \frac{\int \phi_n d\mu}{n}$. Cela vient de la suradditivité de $(\int \phi_n d\mu)_n$ et du lemme classique suivant

Lemme 3.2. *Soit $(a_n)_n \in \mathbb{R}^{\mathbb{N}}$ une suite suradditive, i.e. $a_{n+m} \geq a_n + a_m$ pour tout entiers n, m. Alors la suite $(a_n/n)_n$ converge vers $\sup_n a_n/n$.*

□

Dans le théorème ergodique de Birkhof, si on suppose ϕ dans L^p pour $p \in [1, +\infty[$ alors les moyennes de Birkhof $\frac{1}{n}\sum_{k<n} \phi \circ f^k$ converge vers $E[f|\mathcal{I}]$ dans L^p. En effet si $\phi \in L^\infty$ il suffit d'appliquer le théorème ergodique ponctuel et théorème de convergence dominée. Puis on conclut avec le Lemme 1.4.

4. Deux exemples d'applications

4.1. Nombres normaux. Un nombre $x \in [0, 1)$ est dit normal s'il admet un unique developement en base 2 avec une fréquence 1/2 de 0 et de 1.

La mesure de Lebsegue sur le cercle étant invariante par $x \mapsto 2x \ mod\ 1$, il suit du théorème ergodique :

Corollaire 4.1. *Lebesgue presque tout point $x \in [0, 1)$ est normal.*

4.2. Exposant de Lyapunov maximal positif. Soit $(X, \mathcal{B}, f, \mu)$ un système mesuré ergodique et soit $A : X \to \mathcal{M}_d(\mathbb{R})$ une application mesurable. On considère le cocycle $\left(A^n : X \to \mathcal{M}_d(\mathbb{R})\right)_n$ associé à A définie pour $n > 1$ par recurrence :

$$\forall x \in X, \; A^n(x) = A(f^n x) \circ A^{n-1}(f^{n-1}x).$$

Pour une norme multiplicative sur $\mathcal{M}_d(\mathbb{R})$ la suite $\log \|A^n(x)\|$ est sous-additive. On déduit alors du théorème ergodique sous additif :

COROLLAIRE 4.2. *On suppose $x \mapsto \log^+ \|A(x)\| \in L^1(\mu)$. La suite $(\frac{1}{n} \log \|A^n(x)\|)_n$ converge pour p.t. x vers $\inf_n \frac{1}{n} \int \log \|A^n(x)\| d\mu(x)$. Cette limite (indépendante de la norme $\|$ sur $\mathcal{M}_d(\mathbb{R})$) est appelée l'exposant de Lyapunov (maximal) du cocycle A.*

5. Exercices

Exercice 4. *Récurrence de Kintchine.*
Un ensemble d'entiers est dit relativement dense s'il intersecte tout intervalle d'entiers de longueur $\geq L$ pour un certain L. Soit $(X, \mathcal{B}, \mu, f)$ un système mesuré, on se propose de montrer que pout tout $A \in \mathcal{B}$ et tout $\epsilon > 0$ il existe un ensemble relativement dense d'entiers $\mathcal{N}$, tel que pour tout $n \in \mathcal{N}$ on a :

$$\mu(A \cap f^{-n}A) > \mu(A)^2 - \epsilon$$

— En utilisant le théorème ergodique de Von Neuman montrer que pour n assez grand on a pour tout entier l,

$$\|\frac{1}{n} \sum_{k=0}^{n-1} (1_A \circ f^k - 1_A \circ f^{k+l}\|_2 < 2\epsilon,$$

— En déduire que $m(A)^2 < \epsilon + \frac{1}{N^2} \sum_{0 \leq n,m \leq N-1} \mu(f^{-m}A \cap f^{-n-l}A)$,
— Conclure.

Exercice 5. *Capacité orbitale.*
Soit X un espace compact métrisable et $T : X \to X$ une application continue. Pour tout borélien $A \subset X$ on note $ocap(A)$ sa capacité orbitale définie par :

$$ocap(A) := \lim_n \frac{1}{n} \sup_{x \in X} \sharp\{0 \leq k < n, \; T^k x \in A\}.$$

— Montrer que la limite est bein définie,
— Montrez que pour tout A et toute mesure de proba T-invariante on a

$$\mu(A) \leq ocap(A),$$

— Montrez que pour A fermé, on a

$$\sup_{\mu} \mu(A) = ocap(A),$$

— Donner un exemple de A où l'inégalité est stricte.

Exercice 6.
Soit $(X, \mathcal{A}, \mu, f)$ un système dynamique mesuré et soit $\phi \in L^1$.

— Pour tout $\epsilon > 0$ on pose $A_\epsilon := \{x,\ S_n\phi(x) > \epsilon\}$ et $B_\epsilon := \bigcup_{k\in\mathbb{N}} f^{-k}A_\epsilon$.
Montrer que si $S_n\phi(x) \xrightarrow{n} +\infty$ alors $x \in B_\epsilon$.

— On suppose que μ p.t. $x \in X$ $\sum_{n\in\mathbb{N}} \phi(f^n x) = +\infty$. Montrer que

$$\int \phi d\mu > 0.$$

Chapitre 3

Mesures invariantes d'un système dynamique topologique

Contents

1. Compacité **25**
2. Structure convexe **26**
3. Unique ergodicité **28**
3.1. Exercices 31

On considère dans tout ce chapitre un système dynamique topologique (X, f), i.e. X est un espace metrisable compact et $f : X \to X$ une application continue. On note $\mathcal{C}(X)$ l'ensemble des applications continues de X dans $\mathbb{C}$ muni de la topologie de la convergence uniforme et $\mathcal{M}(X)$ l'ensemble des mesures de probabilités boréliennes muni de la topologie faible-*. Rappelons que cette topologie est la plus petite toplogie rendant les applications $\mu \mapsto \int \phi d\mu$ pour $\phi \in \mathcal{C}(X)$. Muni de cette topologie $\mathcal{M}(X)$ est compacte et cette topologie est métrisable. Par exemple si $(f_n)_n$ est une partie dense dénombrable de $\mathcal{C}(X)$ alors la métrique d suivante induit la topologie faible-*

$$d(\mu, \nu) := \sum_n \frac{|\int \phi_n d\mu - \int \phi_n d\nu|}{2^n(1 + \|\phi_n\|_\infty)}$$

U_f agit sur $\mathcal{C}(X)$ par composition $U_f(\phi) = \phi \circ f$ et sur $\mathcal{M}(X)$ par dualité $f^*\mu = \mu(f^{-1}.)$:

$$\int \phi(f^*\mu) = \int \phi \circ f \; d\mu \text{ pour tout } (\mu, f\phi \in \mathcal{M}(X) \times \mathcal{C}(X).$$

1. Compacité

Lemme 1.1. *L'application $f^* : \mathcal{M}(X) \to \mathcal{M}(X)$ est continue.*

Lemme 1.2. *L'ensemble $\mathcal{M}(X, f) := \{\mu \in \mathcal{M}(X),\ \mu \text{ est } f\text{-invariante}\}$ est un compact convexe non vide de $\mathcal{M}(X)$.*

Démonstration. C'est un fermé comme ensemble des points fixes d'une application continue. Montrons qu'il est non vide. Pour $x \in X$ on note δ_x la mesure de Dirac en x. On note aussi $\mu_n^x := \frac{1}{n}\sum_{k=0}^{n-1}\delta_{f^k x}$. On vérfie facilement,

$$f^*\mu_n^x = \frac{1}{n}\sum_{k=1}^{n}\delta_{f^k x}, \tag{1.1}$$

$$= \mu_n^x + \frac{1}{n}\left(\delta_{f^n x} - \delta_x\right). \tag{1.2}$$

Si $\mu = \lim_k \mu_{n_k}$ est une limite faible dans $\mathcal{M}(X)$ de $(\mu_n)_n$ alors par continuité de f^*, on a $\lim_k f^*(\mu_{n_k}) = f^*(\mu)$ et donc en passant à la limite dans (1.1) on obtient $f^*\mu = \mu$, c'est à dire $\mu \in \mathcal{M}(X, f)$. □

En fait le même argument permet de montrer l'énoncé suivant (faites-le!) :

Lemme 1.3. *Avec les notations précédentes et* $U_n^f := \frac{1}{n}\sum_{k=0}^{n-1}(f^*)^k$, *la suite de compacts* $U_n^f\mathcal{M}(X)$ *converge pour la topologie de Hausdorf vers* $\mathcal{M}(X, f)$.

On notera par la suite $\mathcal{M}_e(X, f)$ le sous ensemble de $\mathcal{M}(X, f)$ formé par les mesures ergodiques.

Lemme 1.4. *On suppose* $\mu \in \mathcal{M}_e(X, f)$. *Alors pour* μ *p.t.* x, *on a*

$$\frac{1}{n}\sum_{k=0}^{n-1}\delta_{f^k x} \xrightarrow{n\to+\infty} \mu \text{ dans } \mathcal{M}(X).$$

Démonstration. Soit $(\phi_n)_n$ une partie dense dénombrable de $\mathcal{C}(X)$. En raisonnnant par densité (comme dans le lemme 1.4) il suffit de montrer qu'il existe un ensemble de mesure totale E telle que pour tout $x \in E$ on est $\frac{1}{n}\sum_{k=0}^{n-1}\phi_n(f^k x)$ converge vers $\int \phi_n d\mu$ quand n tends vers l'infini. Mais on sait d'après le théorème de Birkhof que pour chaque n il existe un ensemble E_n de mesure totale tel que tout $x \in E_n$ satisfait cette propriété de convergence. Il sufffit donc de prendre $x \in E := \bigcap_n E_n$. □

2. Structure convexe

Pour un ensemble convexe C l'ensemble extrémal, $ex(C)$, est l'ensemble des points de C qui ne sont pas dans des segments ouverts de C.

Lemme 2.1.

$$\mathcal{M}_e(X, T) = ex(\mathcal{M}(X, T))$$

En particulier d'après le théorème de Krein-Millman, il existe toujours des mesures ergodiques f*-invariantes.*

DÉMONSTRATION. — $\supset$ Soit E invariant alors $\mu|_E$ et $\mu|_{X\setminus E}$ sont aussi f-invariantes et donc

$$\mu = \mu(E)\mu|_E + (1-\mu(E))\mu|_{X\setminus E},$$

— $\subset$. On propose deux preuves. On suppose par l'absurde que $\mu \in \mathcal{M}_e(X,f)$ n'est pas extrémal, i.e. il existe $t \in]0,1[$ et $\mu_1 \neq \mu_2 \in \mathcal{M}(X,f)$ tels que $\mu = t\mu_1 + (1-t)\mu_2$.

1.Soit $E \in \mathcal{B}$. D' après le théorème ergodique ponctuelle, on a la convergence μ p.p.

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_E \circ f^n \to \mu(E), E_{\mu_i}[1_E|\mathcal{I}].$$

En particulier $\mu(E) = \mu_i(E)$ pour $i = 1, 2$, mais $\mu_1 \neq \mu_2$.

2. Puisque $\mu_i << \mu$, il existe d'après le théorème de Radon-Nykodin $\phi_i \in L^1(\mu)$ tels que $\mu_i = \phi_i d\mu$. L'invariance de μ_i et μ entraine celle de ϕ_i. Par ergodicité de μ les fonctions ϕ_i sont constantes d'intégrale 1 donc égales à 1.

□

LEMME 2.2. *Soient μ et ν deux mesures ergodiques distinctes. Alors $\mu \perp \nu$, i.e. il existe $E \in \mathcal{B}$ avec $\mu(E) = 0$ et $\nu(E) = 1$.*

DÉMONSTRATION. D'après le lemme précédent, la mesure $\chi = \frac{1}{2}(\mu + \nu)$ n'est pas ergodique. Soit donc E un ensemble f-invariant avec $\chi(E) \neq \{0,1\}$. Mais μ et ν étant ergodiques on a $(\mu(E), \nu(E)) \in \{0,1\}^2 \neq (0,0), (1,1)$.
□

Lorsque M est une mesure de proba sur $\mathcal{M}(X,T)$ on peut lui associer un barycentre, $bar(M)$. C'est l'unique élément de $\mathcal{M}(X,f)$ tel que pour tout fonction affine continue $\phi : \mathcal{M}(X,T) \to \mathbb{R}$

$$\int f(\nu)dM(\nu) = \phi(bar(M)).$$

Le convexe compact $\mathcal{M}(X,f)$ étant métrisable on peut montrer le théorème suivant (voir R. Phelps, "Lectures on Choquet's theorem") :

THÉORÈME 2.3. *Pour toute mesure $\mu \in \mathcal{M}(X,f)$ il existe une unique mesure M_μ (appelée la décomposition ergodique de μ,) supportée sur les mesures ergodiques telle que $bar(M_\mu) = \mu$.*

3. Unique ergodicité

Le système topologique (X, f) est dtit uniquement ergodique ssi il existe une unique mesure invariante (ergodique).

THÉORÈME 3.1. *(X, f) est uniquement ergodique ssi pour tout $\phi \in C(X)$ la suite de fonction $(U_n^f \phi)$ converge uniformément vers une fonction constante (égale à $\int \phi d\mu$ avec μ l'unique mesure invariante).*

DÉMONSTRATION. Puisque $\mathcal{M}(X, f)$ est limite pour la topologie de Hausodrf de $U_f^n \mathcal{M}(X)$ le diamètre de $U_f^n \mathcal{M}(X)$ tend vers 0 quand n tend vers l'infini. En particulier si ϕ est une fonction continue

$$\int \phi d(U_f^n \delta_x) = \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k x) \xrightarrow{n} \int \phi d\mu$$

uniformément en x. □

PROPOSITION 3.2. *Pour α irrationnel, $(\mathbb{S}^1, f_\alpha)$ est uniquement ergodique.*

Une preuve classique attribué à Weyl consiste à montrer la convergence uniformes des sommes de Birkhof associées aux monômes trigonométriques puis à conclure en utilisant la densité des polynômes trigonométriques dans $C(\mathbb{S}^1)$. On propose ici une autre preuve utilisant la "rigidité" de la mesure de Lebesgue.

DÉMONSTRATION. Soit μ une mesure invariante par f_α. On va montrer que μ est invariante par toute rotation f_β. Clairement μ est invariante par toutes les rotations f_β pour $\beta \in \alpha\mathbb{Z}$. Puisque $\alpha\mathbb{Z}$ est dense dans $\mathbb{R}$ il suffit donc de vérifier que $\gamma \mapsto f_\gamma^* \mu$ est continue pour la topologie faible*. Mais pour $\phi \in C(\mathbb{S}^1)$ on a $\gamma \mapsto \int \phi d(f_\gamma^* \mu) = \int \phi(x + \gamma) d\mu(x)$ est continue (intégrale à paramètre). □

On dit qu'une suite $(x_k) \in [0, 1]^{\mathbb{N}}$ est équidistribuée ssi pour tout $\phi \in C(X)$ on a

$$\frac{1}{n} \sum_{k=0}^{n-1} f(x_k) \xrightarrow{n} \int f dLeb.$$

PROPOSITION 3.3. *Si P est un polynôme de degré supérieur ou égal à 1 avec un coefficient dominant irrrationnel, alors $(P(n) \ mod \ 1)_n$ est equidistribuée.*

On se propose de montrer ce résultat pour $P(X) = X(X - 1)\alpha/2$, la preuve du cas général suit les mêmes idées. En utilisant les séries de Fourier comme pour montrer l'ergodicité de la mesure de Lebesgue précédemment, on montrer que $Leb \times Leb$ est une mesure ergodique pour le produit semi-direct, $S(x, y) = (x + \alpha, x + y)$ sur le tore $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$.

On utilise ensuite le lemme suivant :

Lemme 3.4. *Soit (X, T) un système uniquement ergodique soit $\phi : X \to \mathbb{R}/\mathbb{Z}$ une fonction continue. On suppose que $\mu \times Leb_{\mathbb{R}/\mathbb{Z}}$ est ergodique pour le produit semi-direct $T_\phi(x, y) = (Tx, y + \phi(x))$ avec μ l'unique mesure T-invariante. Alors $(X \times \mathbb{R}/\mathbb{Z}, T_\phi)$ est uniquement ergodique.*

Démonstration. Soit ν_0 une mesure ergodique T_ϕ invariante. Pour tout t on note g_t le flot $g_t(x, y) = (x, y + t)$. Il commute avec T_ϕ si bien que $\nu_t = g_t^* \nu_0$ est T_ϕ invariante et ergodique et donc aussi la mesure $\nu := \int \nu_t dLeb(t)$. On vérifie maintenant que $\nu = \mu \times Leb$ est invariante. Pour toute fonction continue $f \in C(X \times \mathbb{R}/\mathbb{Z})$ on a

$$\begin{aligned} \int f(x,y) d\nu(x,y) &= \int \left(\int f(x,y) d\nu_t(x,y) \right) dLeb(t), \\ &= \int \left(\int f \circ g_t(x,y) \right) dLeb(t), \\ &= \int \int f(x, y+t) d\nu_0(x,y) dLeb(T), \\ &= \int \left(\int f(x,z) dLeb(z) \right) d\nu_0(x,y). \end{aligned}$$

Mais si on note $\pi : X \times \mathbb{R}/\mathbb{Z} \to X$ la projection sur le premier facteur X, on a $\pi^* \nu \in \mathcal{M}(X, T) = \{\mu\}$. Donc

$$\int f(x,y) d\nu(x,y) = \int \left(\int f(x,z) dLeb(z) \right) d\pi^* \nu_0(x,y) = \int f(x,z) d\mu(x) dLeb(z).$$

On en conclut que $\nu = \mu \times Leb$. Cette mesure étant supposé ergodique, elle est extrémale dans l'ensemble des mesures invariantes et donc $\nu_t = \mu \times Leb$ pour Lebesgue presque tout t. Pour un tel t on a $g_{-t}^* \nu_t = \nu_0 = g_{-t}^*(\mu \times Leb) = \mu \times Leb$. □

En appliquant ce lemme on a donc que S est uniquement ergodique. En particulier pour toute fonctione ϕ continue sur le cercle

$$\frac{1}{n} \sum_{k=0}^{n-1} (1 \times \phi)(S^n(0,0)) \xrightarrow{n} \int \phi dLeb$$

Or on calcule facilement par récurrence : $S^n(0, 0) = (n\alpha, n(n-1)\alpha/2)$.

La Proposition 3.3 peut aussi être démontrée à l'aide du lemme technique suivant que l'on utilisera au chapitre suivant.

LEMME 3.5 (Van der Corput). *Soit H un espace préhilbertien et soit $(v_n)_n$ une suite d'éléments de H de norme inférieure ou égale à 1. Alors*

$$\|\frac{1}{N}\sum_{n=1}^{N} v_n\| \leq \left(\frac{2}{H}\sum_{h=0}^{H-1}|\frac{1}{N}\sum_{n=1}^{N} < v_{n+h}, v_n > |\right)^{\frac{1}{2}} + O(H/N).$$

DÉMONSTRATION. Pour tout $0 \leq h < H$ on a

$$\frac{1}{n}\sum_{n=1}^{N} v_n = \frac{1}{n}\sum_{n=1}^{N} v_{n+h} + O(H/N)$$

avec un grand O uniforme en h.

En sommant sur h puis en appliquant l'inégalité triangulaire on obtient

$$\|\frac{1}{N}\sum_{n=1}^{N} v_n\| \leq \frac{1}{N}\sum_{n=1}^{N}\|\frac{1}{H}\sum_{h=0}^{H-1} v_{n+h}\| + O(H/N),$$

puis par l'inégalité de Cauchy-Schwartz (dans $\mathbb{R}^N$)

$$\|\frac{1}{N}\sum_{n=1}^{N} v_n\| \leq \left(\frac{1}{N}\sum_{n=1}^{N}\|\frac{1}{H}\sum_{h=0}^{H-1} v_{n+h}\|^2\right)^{\frac{1}{2}} + O(H/N).$$

En développant la norme au carré, on a

$$\|\frac{1}{N}\sum_{n=1}^{N} v_n\| \leq \left(\frac{1}{H^2}\sum_{h=0}^{H-1}\sum_{k=0}^{H-1}\left|\frac{1}{N}\sum_{n=1}^{N} < v_{n+h}, v_{n+k} >\right|\right)^{\frac{1}{2}} + O(H/N).$$

On a encore

$$\left|\frac{1}{N}\sum_{n=1}^{N} < v_{n+h}, v_{n+k} >\right| = \left|\frac{1}{N}\sum_{n=1}^{N} < v_{n+|h-k|}, v_n >\right| + O(H/N),$$

puis en sommant suivant les diagonales $h = k + cst$ on obtient le résultat voulu.

□

COROLLAIRE 3.6. *Soit $(x_n)_n$ une suite telle que les suites $(x_{n+h} - x_n)_n$ soient équidistribuées modulo 1 pour tout entier $h > 0$, alors la suite $(x_n)_n$ est aussi équidistribuée modulo 1.*

DÉMONSTRATION. On suit la preuve classique de Weyl de la Proposition 3.2 en montrant pour tout entier k non nul la convergence vers 0 de $\frac{1}{N}\sum_{n=1}^{N} e^{2ik\pi x_n}$. On applique pour cela le lemme de Van der Corput à la suite $(v_n)_n = (e^{2ik\pi x_n})_n$ (observez que $v_{n+h}\overline{v_n} = e^{2ik\pi(x_{n+h}-x_n)}$).

□

3.1. Exercices.

Exercice 7.

Soit $(\mathbb{S}^1, f_\alpha)$ une rotation irrationnelle du cercle. On considère $\phi : \mathbb{S}^1 \to \mathbb{R}$ une fonction continue et $E \subset \mathbb{S}^1$ de mesure de Lebesgue totale. Montrer que pour tout $\epsilon > 0$ et tout $y \in \mathbb{S}^1$ il existe $x \in \mathbb{S}^1$ tel que

$$|U_n^{f_\alpha}\phi(y) - U_n^{f_\alpha}\phi(x)| < \epsilon.$$

En déduire que f_α est uniquement ergodique (supposant connue l'ergodicté de la mesure de Lebesgue).

Exercice 8.

Soient α et β deux réels rationnellement indépendants. Montrer $R_\alpha \times R_\beta$ est uniquement ergodique sur $\mathbb{S}^1 \times \mathbb{S}^1$.

Chapitre 4

Théorie spectrale

Contents

1. Mélange faible **33**
1.1. Définitions et premières caractérisations 33
1.2. Mélange faible d'ordre supérieur 34
1.3. Récurrence multiple et principe de correspondance de Furstenberg 35
2. Spectre de l'opérateur de Koopman **36**
2.1. Valeurs/vecteurs propres 36
2.2. Spectre purement ponctuel 37
2.3. Spectre de Lebesgue 37
3. Mesures spectrales et Théorème spectral **38**
3.1. Liens avec le mélange/spectre 39
4. Exercices **41**

1. Mélange faible

1.1. Définitions et premières caractérisations. Soit $(X, \mathcal{B}, \mu, f)$ un système dynamique mesuré. On dit que μ est faiblement mélangeante lorsque pour tout $A, B \in \mathcal{B}$ on a

$$\frac{1}{n}\sum_{k=0}^{n-1}\left|\mu(A\cap f^{-k}B)-\mu(A)\mu(B)\right| \xrightarrow{n} 0,$$

ou de façon equivalente pour tout $\phi, \psi \in L^2$

$$\frac{1}{n}\sum_{k=0}^{n-1}\left|\int \phi\circ f^n\psi d\mu-\int\phi d\mu\int\psi d\mu\right| \xrightarrow{n} 0.$$

D'après le Corollaire 1.5 toute transformation faiblement mélangeante est ergodique.

On rappelle qu'un sous ensemble E de $\mathbb{N}$ est de densité 1 lorsque $\frac{\sharp(E\cap[1,N])}{N} \xrightarrow{1}$.

Lemme 1.1. *Soit $(a_k)_k$ une suite bornée de réels positifs. Alors les propriétés suivantes sont equivalentes :*

— $\lim_n \frac{1}{n}\sum_{k=0}^{n-1} a_k = 0$,
— *il existe un sous ensemble* Z *de* $\mathbb{N}$ *de densité* 1, *tel que* $\lim_{n\to+\infty,\ n\in Z} a_n = 0$.

Démonstration. Supposons $\lim_n \frac{1}{n}\sum_{k=0}^{n-1} a_k = 0$ (l'autre implication suit immédiatement des définitions et du caractère borné de la suite). Pour tout p l'ensemble $Z_p := \{k \ : \ a_k \geq \frac{1}{p}\}$ est de densité nulle, en particulier il existe un entier n_p tel que pour tout $n \geq n_p$ on a

$$\frac{\sharp\,(Z_p \cap [1,n])}{n} < 1/p.$$

On peut supposer de plus que $(n_p)_p$ est croissante. On vérifie alors que le complémentaire de $Z := \bigcup_{p\geq 1}(Z_p \cap [n_p, \infty[)$ est de densité 1 : en effet pour un entier n, il existe un entier $p = p_n$ tel que $n_p \leq n < n_{p+1}$ et alors $Z \cap [1,n] = \bigcup_{q\leq p}(Z_q \cap [1,n]) = Z_p \cap [1,n]$ et $\frac{\sharp(Z\cap[1,n])}{n} < 1/p$. Aussi $p_n \xrightarrow{n} +\infty$. Enfin si n est dans le complémentaire de Z $a_n \leq \frac{1}{p_n}$. □

Théorème 1.2. *Les propriétés suivantes sont équivalentes :*

(1) μ *est faiblement mélangeante,*

(2) $\mu \times \mu$ *est faiblement mélangeante,*

(3) $\mu \times \mu$ *est ergodique.*

Démonstration. Comme on l'a mentionné ci-dessus, (2) entraine (3). Par définition de la mesure produit on vérifie facilement que (1) entraine (2). Montrons donc que (3) entraine (1) : d'après le lemme précédent et le corollaire 1.5, il suffit de vérifier que pour tout $A, B \in \mathcal{B}$ on a

$$\frac{1}{n}\sum_{k=0}^{n-1} |\mu(f^{-k}B \cap A) - \mu(A)\mu(B)|^2 \xrightarrow{n} 0.$$

Mais le terme de droite s'ecrit aussi comme

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu\times\mu((f\times f)^{-k}(B\times B)\cap(A\times A) - \mu(A)^2\mu(B)^2 - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu(f^{-k}B\cap A) - \mu(A)\mu(B)$$

et par ergodicté de $\mu \times \mu$ et donc de μ on a toujours par le Corollaire 1.5 que ce terme tends vers 0 quand n tends vers l'infini. □

1.2. Mélange faible d'ordre supérieur.

Théorème 1.3. *Soit* $(X, \mathcal{B}, f, \mu)$ *un système faiblement mélangeant alors pour tout* $\phi_1, ..., \phi_k \in L^\infty(\mu)$ *on a la convergence suivante dans* $L^2(\mu)$

$$\frac{1}{N}\sum_{n=1}^{N} U_f^n(\phi_1)...U_f^{kn}(\phi_k) \xrightarrow{N\to+\infty} \int \phi_1 d\mu ... \int \phi_2 d\mu.$$

DÉMONSTRATION. La preuve se fait par induction sur k. Le cas $k = 1$ correspond au théorème ergodique en moyenne de Von Neuman. Le pas de la récurrence se démontre à l'aide du théroème de Van der Corput comme suit. On peut supposer $\int \phi_k d\mu = 0$. On note alors $v_n = U_f^n(\phi_1)...U_f^{kn}(\phi_k)$ si bien que

$$< v_{n+h}, v_n >= \int \phi_1 U^h \phi_1 ... U^{(k-1)n} \phi_k U^{(k-1)n+kh} \phi_k d\mu.$$

Par l'hypothése de récurrence $\frac{1}{N}\sum_{n=0}^{N-1} < v_{n+h}, v_n >$ tends vers $s_h = \int \phi_1 U^h \phi_1 d\mu ... \int \phi_k U^{kh} \phi_k d\mu$ quand n tends vers l'infini. Puis on conclut avec le lemme de Van der Corput car $s_h \leq \|\phi_1\|_\infty^2 \|\phi_{k-1}\|_\infty^2 \times |\int \phi_k U^{kh} \phi_k d\mu|$ et la somme de Césaro de ce dernier facteur tends vers 0, le système (f^k, μ) étant faiblement mélangeant (pourquoi ?). □

COROLLAIRE 1.4. *Soit $(X, \mathcal{B}, f, \mu)$ un système faiblement mélangeant, alors pour tout entier $k \geq 2$ et pour tout $A_1, ..., A_k \in \mathcal{B}$,*

$$\frac{1}{N}\sum_{n=1}^{N} \left|\mu(A_1 \cap T^{-n}A_2 \cap ... \cap T^{-kn}A_k) - \mu(A_1)...\mu(A_k)\right| \xrightarrow{N\to+\infty} 0.$$

DÉMONSTRATION. D'après le lemme 1.1 il suffit de montrer que la série de Césoaro associé aux carrés de $\left(\mu(A_1 \cap T^{-n}A_2 \cap ... \cap T^{-kn}A_k) - \mu(A_1)...\mu(A_k)\right)_n$ converge vers 0. On développe alors le carré. La suite $\mu(A_1 \cap T^{-n}A_2 \cap ... \cap T^{-kn}A_k)^2 = \mu \times \mu \left((A_1 \times A_1) \cap ... \cap (T^{-kn}A_k \times T^{-kn}A_k)\right)$ converge vers $\mu(A_1)^2 ... \mu(A_k)^2$ au sens de Césaro d'après le théorème précédent, la mesure $\mu \times \mu$ étant faiblement mélangeante pour le produit $f \times f$ d'après le théorème 1.2. Les autres termes du développement convergeant au sens de Césaro vers $-\mu(A_1)^2 ... \mu(A_k)^2$, cela conclut la preuve. □

1.3. Récurrence multiple et principe de correspondance de Furstenberg.

THÉORÈME 1.5. *Récurrence multiple Soit $(X, \mathcal{B}, f, \mu)$ un système mesuré. Pour tout $A \in \mathcal{B}$ avec $\mu(A) > 0$ on a pour tout entier $k \in \mathbb{N}$*

$$\liminf_N \frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap ... \cap T^{-kn}A) > 0.$$

Le mélange faible d'ordre supérieure entraine facilement le théorème de récurrence multiple pour des systèmes faiblement mélangeants. Nous n'aborderons le cas général dans ces notes, mais nous expliquons maintenant comment ce résultat entraine le théorème suivant dû à Szemeredi.

THÉORÈME 1.6 (Szemeredi). *Soit $\mathbb{A}$ un sous ensemble de $\mathbb{Z}$ de densité supérieure de Banach positive, i.e. $\overline{d}(\mathbb{A}) := \limsup_N \frac{\sharp \mathbb{A}\cap[-N,N]}{2N+1} > 0$, alors $\mathbb{A}$ contient des suites arithmétiques de longueur arbitrairement grande.*

Ce théorème suit de la récurrence multiple du principe suivant. On associe à la suite $\mathbb{A}$ la suite $(x_n^{\mathbb{A}})_n$ de $\{0,1\}^{\mathbb{Z}}$ définie par $x_n^{\mathbb{A}} = 1$ ssi n appartient à A. Si $(n_j)_j$ est une suite d'entiers avec $\lim_j \frac{\sharp \mathbb{A}\cap[-n_j,n_j]}{2n_j+1} = \overline{d}(\mathbb{A})$ alors toute limite faible μ de $(\mu_j)_j := \left(\frac{1}{2n_j+1}\sum_{|l|\leq n_j}\delta_{\sigma^l(x^{\mathbb{A}})}\right)_j$ vérifie $\mu(A) = \overline{d}(\mathbb{A}) > 0$ avec $A = [1] := \{(x_n)_n,\ x_0 = 1\}$. Le théorème de Szemeredi suit alors du principe de récurrence appliqué au système mesuré $(\{0,1\}^{\mathbb{Z}}, \sigma, \mu)$. En effet il existe pour tout entier k un entier n tel que $\mu(A\cap\sigma^{-n}A\cap...\cap\sigma^{-kn}A) > 0$. L'ensemble $B = A\cap\sigma^{-n}A\cap...\cap\sigma^{-kn}A$ étant ouvert et fermé, on a $\mu(B) = \lim_j \mu_j(B)$. En particulier pour j assez grand on a $0 < \mu_j(B) = \frac{1}{2n_j+1}\sum_{|l|\leq n_j}\delta_{\sigma^l(x^{\mathbb{A}})}(A)$. Ainsi il existe l tel que $\sigma^l x^{\mathbb{A}}$ appartient à B, ce qui est équivalent à $l, l+k, ..., l+kn \in \mathbb{A}$.

2. Spectre de l'opérateur de Koopman

Pour un opérateur linéaire continue U d'un Banach dans lui-même, on appelle spectre de U, l'ensemble des complexes λ tel que $U - \lambda Id$ n'est pas inversible.

2.1. Valeurs/vecteurs propres. Le spectre contient toujours l'ensemble des valeurs propres, appelé aussi spectre ponctuel.

LEMME 2.1. *Le spectre ponctuel de U_f est un sous ensemble dénombrable de $\mathbb{S}^1$ et les fonctions propres associées à des valeurs propres distinctes sont orthogonales. Si μ est de plus ergodique alors le spectre ponctuel est un sous groupe de $\mathbb{S}^1$ et les fonctions propres sont de module constant.*

DÉMONSTRATION. Clairement U_f étant une isométrie toute valeur propre est dans le cercle unité. De même on vérifie facilement l'orthogonalité de fonctions propres associées à des valeurs propres distinctes. L'espace L^2 étant séparable, l'ensemble des valeurs propres est donc au plus dénombrable.

Si μ ergodique et ϕ une fonction propre non nulle. Alors $|\phi|$ est invariant et donc constant. Si $\phi\circ f = \alpha\phi$ et $\psi\circ f = \beta\psi$ alors $1/f$ et fg sont propres pour α^{-1} et $\alpha\beta$ respectivement. □

Rappelons qu'on a vu que μ était ergodique ssi l'espace propre associé à la valeur propre 1 était de dimension 1, réduit aux fonctions constantes.

THÉORÈME 2.2. *Si μ est faiblement mélangeante alors son spectre ponctuel réduit à $\{1\}$ et les constantes sont les seules fonctions propres.*

L'inverse est aussi vrai et sera montré un peu plus tard.

DÉMONSTRATION. Soit ϕ une fonction propre associé à une valeur propre $\lambda \neq 1$. Alors $\int \phi \, d\mu = 0$ et $\int \phi \circ f^k \overline{\phi} \, d\mu = \lambda^k \int |\phi|^2 \, d\mu$ pour tout entier k. Il s'en suit que

$$\frac{1}{n} \sum_{k=0}^{n-1} | \int \phi \circ f^j \overline{\phi} \, d\mu| = \int |\phi|^2 d\mu \neq 0,$$

ce qui contredit le mélange faible. □

2.2. Spectre purement ponctuel. Lorsque le spectre coincide avec le spectre ponctuel, i.e. quand les fonctions propres engendrent un espace vectoriel dense, on dit que le spectre est purement ponctuel.

Le spectre de la rotation f_α sur le tore est purement ponctuel : en effet $(e^{2i\pi k \cdot})_k$ forme une base Hilbertienne de L^2 et pour tout k la fonction $e^{2i\pi k \cdot}$ est une fonction propre pour $e^{2i\pi k\alpha}$.

THÉORÈME 2.3 (Halmos-Von Neuman). *Deux systèmes mesurés ergodiques à spectre purement ponctuel sont métriquement isomorphes ssi ils ont le même ensemble de valeurs propres. De plus un tel système est métriquement isomorphe à la mesure de Haar pour une rotation sur un groupe abélien compact. En particulier il n'est pas faiblement mélangeant.*

2.3. Spectre de Lebesgue. On suppose ici pour simplifier les systèmes inversibles. Un système mesuré inversible est dit de spectre de Lebesgue dénombrable lorsqu'il existe un sous espace vectoriel fermé G de L^2 de dimension infinie tel que $(U_f^k G)_{k\in\mathbb{Z}}$ sont deux à deux orthogonaux et

$$L^2 = 1 \oplus \left(\bigoplus_k U_f^k G \right).$$

THÉORÈME 2.4. *Deux systèmes mesurés avec un spectre de Lebesgue dénombrable sont spectralement isomorphes.*

DÉMONSTRATION. Il suffit de remarquer que deux espaces de Hilbert séparables de dimension infine G et G' sont isomorphes puis de prolonger cet isomorphisme ψ à tout L^2 en posant $\psi(U^k x) = U^k \psi(x)$ pour tout k et tout $x \in G$. □

THÉORÈME 2.5. *Tout schéma de Bernoulli a un spectre de Lebesgue dénombrable, en particulier tous les schémas de Bernoulli sont spectralement, isomorphes.*

Démonstration. Pour tout sous tribu $\mathcal{A}$ on note $L_0^2(\mathcal{A})$ le sous-espace de $L^2(\mathcal{A})$ orthogonal aux constantes, i.e. $L_0^2(\mathcal{A}) = \{f \in L^2(\mathcal{A}),\ \int f d\mu = 0\}$. On considère $\mathcal{A}$ la tribu engendré par les cylindres positifs $[C_0, ..., C_n]$ avec $n \in \mathbb{N}$ et on note $G = \{\phi \in L_0^2(\mathcal{A}),\ \phi \perp U_f(L_0^2(\mathcal{A}))\}$. Clairement les $U_f^k G$ sont deux à deux orthogonaux. De plus $L_0^2(\mathcal{A}) = \bigoplus_{k\geq 0} U_f^k G$ et l'espace vectoriel engendré par les $U_f^k G$ pour $k \leq 0$ est dense dans $L_0^2 = L_0^2(\mathcal{B})$. On en déduit que $H = 1 \oplus \left(\bigoplus_k U_f^k G\right)$. Il reste à voir que G est de dimension infinie. Pour cela on remarque que si ϕ est dans G et si $(A_k)_{k\in\mathbb{N}}$ est une famille infinie de Boréliens de $f{-}1(\mathcal{A})$ deux à deux disjoints alors $(1_{A_k}\phi)_k$ est une famille orthogonale de G. $\square$

Théorème 2.6. *Tout système inversible avec un spectre de Lebesgue dénombrable est mélangeant.*

Démonstration. Il suffit par linéarité et densité de montrer le mélange pour $\phi \in G$ et $\psi = 1$ ou ψ dans $U_f^n G$ pour un certain n. L'orthogonalité des $U_f^k G$ donne immédiatement le résultat voulu. $\square$

3. Mesures spectrales et Théorème spectral

On rappelle que toute mesure ν sur le cercle est déterminé par ses coefficients de Fourier $(\hat{\nu}(n))_{n\in\mathbb{Z}}$, définis pour tout n par

$$\hat{\nu}(n) = \int_{\mathbb{S}^1} z^n d\nu(z).$$

Proposition 3.1. *Pour tout $\phi \in L^2$ il existe une unique mesure de proba ν_ϕ telle que $\int U_f^n \phi\overline{\phi} d\mu = \hat{\nu_\phi}(n)$. De plus, sur la fermeture $Z(\phi)$ de l'espace cyclique engendré par $U_f^n\phi$ avec $n \in \mathbb{Z}$ l'opérateur U_f est spectralement isomorphe à l'opérateur de multiplication par z dans $L^2(\mathbb{S}^1, \nu_\phi)$. La mesure ν_ϕ est dite spectrale.*

Démonstration. On considère pout tout $z \in \mathbb{S}^1$ et ϕ avec $\|\phi\|_2 = 1$,

$$\nu_N(z) := \frac{1}{N} \| \sum_{0\leq n<N} z^n U_f^n \phi\|_2^2.$$

On a

$$\nu_N(z) = \frac{1}{N}\left(1 + \sum_{0\leq k\neq l\leq N} z^{l-k} \int U_f^{l-k}\phi\overline{\phi} d\mu\right),$$

de sorte que $\int \nu_N(z) dLeb(z) = 1$ et $\hat{\nu_N}(p) = \frac{N-p}{N}\int U_f^p \phi\overline{\phi} d\mu \xrightarrow{N} \int U_f^p\phi\overline{\phi} d\mu$. D'après le théorème de Levy sur le cercle, il existe une unique mesure ν_ϕ sur le cercle telle que $\nu_n(z)dLeb(z)$ converge vers ν_ϕ en topologie faible-*. De plus $\hat{\nu_\phi}(p) = \int U_f^p \phi\overline{\phi} d\mu$ pour tout p.

Maintenant on verifie facilement, que

$$\int |P(z)|^2 d\nu_\phi(z) = \int |P(U_f)\phi|^2 d\mu,$$

Par densité des polynomes trigonométriques dans L^2 on en déduit que $L^2(\nu_\phi, \mathbb{S}^1)$ est isométrique à $L^2(\mu, Z(\phi))$. De plus $(U_f, L^2(X,\mu))$ est clairement isomorphe à l'opérateur de multiplication par z sur $L^2(\mathbb{S}^1, \nu_\phi)$. $\square$

THÉORÈME 3.2. *Il existe une collection (au plus dénombrable) de mesure spectrales sur le cercle $\nu_1 >> \nu_2 >> ... >> \nu_r$ telle que $U_f : L^2_0 \circlearrowleft$ soit isométriquement conjugué à la multiplication par z dans $\bigoplus_{k=1}^r L^2(\mathbb{S}^1, \nu_k)$. De plus à equivalence près des mesures $(\nu_k)_k$ cette decomposition est unique. La mesure ν_1 est appelé la mesure spectrale maximale. Le spectre de U_f coïncide avec le support de la mesure spectrale maximale.*

REMARK 3.3. Si ν est une mesure absolument continue relativement à ν_1 il existe alors $\phi \in L^2$ tel que $\nu_\phi = \nu$. En effet si $\nu_1 = \nu_\psi$ et $g = \frac{d\nu}{\nu_1} \in L^1$ est le dérivée de Radon-Nykodin associée, alors la mesure spectrale associée à $g^{\frac{1}{2}}(U)\psi$ coincide avec ν.

3.1. Liens avec le mélange/spectre. On fait dans cette partie le lien entre les propriétés du mélange et celles du spectre. On commence tout d'abord par le lemme suivant :

LEMME 3.4.
- *μ est ergodique ssi pour tout $\phi \in L^2_0$, on a $\frac{1}{n}\sum_{k=0}^{n-1}\int \phi \circ f^k \overline{\phi} d\mu \xrightarrow{n} 0$,*
- *μ est faiblement mélangeante ssi pour tout $\phi \in L^2_0$, on a $\frac{1}{n}\sum_{k=0}^{n-1}|\int \phi \circ f^k \overline{\phi} d\mu| \xrightarrow{n} 0$,*
- *μ est mélangeante ssi pour tout $\phi \in L^2_0$, on a $\int \phi \circ f^n \overline{\phi} d\mu \xrightarrow{n} 0$.*

DÉMONSTRATION. On suppose pour simplifier que (f,μ) est inversible. Alors si ψ est dans (la fermeture de) l'espace cyclique engendré par ϕ, on peut remplacer $\overline{\phi}$ par $\overline{\psi}$ dans les limites des termes de droite. De même si ψ est dans son orthogonal. $\square$

THÉORÈME 3.5.
- *(f,μ) ergodique a un spectre purement ponctuel ssi toute mesure spectrale (resp. la mesure spectrale maximale) est atomique,*
- *(f,μ) est faiblement mélangeante ssi sur $L^2_0(\mu)$ toute mesure spectrale (resp. la mesure spectrale maximale) est non atomique,*
- *(f,μ) est de Lebesgue ssi les mesures spectrales associées à une famille dense de $L^2_0(\mu)$ sont toutes équivalentes à la mesure de Lebesgue. De*

plus toutes les mesures spectrales sont alors équivalentes à la mesure de Lebesgues.

REMARK 3.6. On peut aussi montrer que (f, μ) est de Lebesgue denombrable ssi $r = \infty$ et tous les ν_k sont équivalentes à la mesure de Lebesgue.

DÉMONSTRATION. Pour le premier point, il suffit de remarquer que si ϕ est une fonction propre de valeur propre λ, alors ν_ϕ a ses coefficient de Fourier satisfaisant $\hat{\nu_\phi}(n) = \lambda^n$ et donc ν_ϕ coincide avec la mesure de Dirac en λ. En effet si $\phi = \sum_n a_n \phi_n$ avec ϕ_n fonction propre de valeur propre (simple) λ_n et de norme 1 alors $\nu_\phi = \sum_n |a_n|^2 \delta_{\lambda_n}$.

Pour le deuxième, on a d'après le lemme 1.1 equivalence avec $\frac{1}{n}\sum_{k=0}^{n-1} | \int \phi \circ f^k \overline{\phi} d\mu|^2 \xrightarrow{n} 0$. Mais ce terme s'écrit aussi

$$\begin{aligned} \frac{1}{n}\sum_{k=0}^{n-1} | \int \phi \circ f^k \overline{\phi} d\mu|^2 &= \frac{1}{n}\sum_{k=0}^{n-1} | \int z^k d\nu_\phi(z)|^2, \\ &= \int\int \frac{1}{n}\sum_{k=0}^{n-1} (z'\overline{z})^k d\nu_\phi(z) d\nu_\phi(z'). \end{aligned}$$

On conclut par convergence dominée que ce dernier terme tends vers

$$\int\int 1_{z=z'} d\nu_\phi(z) d\nu_\phi(z') = \sum_{a\in\mathbb{S}^1} \nu_\phi^2(\{a\}).$$

En particulier le mélange faible est équivalent au caractère continu des mesures spectrales de L_0^2.

Pour le dernier point il suffit de voir (pourquoi ?) que si $\phi \in G$ de norme 1, alors ν_ϕ est la mesure de Lebesgue. Mais $\hat{\nu_\phi}(n) = \int U_f^n \phi \overline{\phi} = 0$ ssi $n \neq 0$ et vaut 1 si $n = 0$. Une mesure étant déterminée par ses coefficients de Fourier il s'en suit que ν_ϕ est la mesure de Lebesgue.

□

On retrouve aussi le fait que tout système de spectre de Lebesgue dénombrable est mélangeant. En effet il suffit d'après le lemme précédent de voir que $\int U_f^n \phi \overline{\phi} d\mu \xrightarrow{n} 0$ pour tout $\phi \in L_0^2$ mais $\int U_f^n \phi\overline{\phi} d\mu = \hat{\nu_\phi}(n)$. Puisque ν_ϕ est equivalent à la mesure de Lebesgue alors $\hat{\nu_\phi}(n) \xrightarrow{n} 0$ d'après le lemme de Rieman-Lebesgue.

On peut montrer que toute mesure de proba sur le cercle est la mesure spectrale maximale pour une certaine dynamique.

4. Exercices

Exercice 9.
A l'aide du théorème spectral, redémontrer le théorème ergodique en moyenne de Von Neuman dans le cas ergodique.

Exercice 10. *Théorème ergodique en moyenne pour les carrés.*
Soit $(X, \mathcal{B}, f, \mu)$ un système mesuré inversible. On se propose de montrer le théorème ergodique en moyenne pour les carrés, i.e. la convergence dans L^2 de $\left(\frac{1}{n}\sum_{k=0}^{n-1}\phi\circ f^{k^2}\right)_n$ pour tout $\phi \in L^2$.

- montrer qu'il suffit de montrer que pour tout $\phi \in L^2$ et tout $l \in \mathbb{Z}$ on a convergence de $\left(\frac{1}{n}\sum_{k=0}^{n-1}\int \phi\circ f^{k^2+l}\overline{\phi}d\mu\right)_n$,
- montrer que pour tout $z \in \mathbb{C}$ de module 1 la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} z^{k^2}\right)_n$ converge (distinguer le cas où z est racine de l'unité),
- Conclure. En travaillant encore un peu, on peut identifier la limite comme étant $\sum_{\lambda \text{ vp}} \lambda^2 P_\lambda$ avec P_λ le projecteur orthogonal sur $Ker(\lambda Id - U_f)$.

Exercice 11.
On considère l'application $f : \mathbb{R}/\mathbb{Z}\times\mathbb{R}/\mathbb{Z} \circlearrowleft$ qui à (x, y) associe $(x+\alpha, x+y)$. Déterminer les mesures spectrales ν_ϕ pour ϕ dans la base hilbertienne de Fourier $\{e_{k,l}\}_{k,l\in\mathbb{Z}}$.

Chapitre 5

Exemples

Contents

1.	**Sous décalage de type fini et mesures markoviennes**	**43**
2.	**L'application de Gauss**	**44**
3.	**Transformation de Chacon**	**46**
4.	**Flot géodésique et horocyclique**	**47**
5.	**Exercices**	**49**

1. Sous décalage de type fini et mesures markoviennes

Soit Y un sous ensemble fermé de $\{1,...,d\}^{\mathbb{Z}}$ invariant pas le décalage S. Le système dynamique (Y,S) est appelé un sous décalage à alphabet fini. Un sous décalage $Y \subset \{1,...,d\}^{\mathbb{Z}}$ est de type fini, lorsqu'il existe une matrice $A \in M_d(\{0,1\})$ (appelée matrice d'adjacence) telle que

$$Y := \{(x_n)_n \in \{1,...,d\}^{\mathbb{Z}},\ A_{x_n,x_{n+1}} = 1 \text{ pour tout } n\}.$$

Une matrice $P = (p_{ij}) \in M_d([0,1])$ est dite stochastique lorsque

$$\forall i,\ \sum_j p_{ij} = 1.$$

Une proba π sur $\{1,...,d\}$ est dite stationnaire lorsque $\pi P = \pi$. On appelle mesure markovienne la mesure invariante μ_π associée à une proba stationnaire π défini pour tout cylindre

$$\mu_\pi([k_i,...,k_j]) = \pi(k_i)\prod_{l=i}^{j-1} p_{k_l,k_{l+1}}.$$

On peut associer à toute matrice stochastique P le sous décalage de type fini dont la matrice d'adjacence $A = (a_{ij})$ est définie par $a_{ij} = 1$ ssi $p_{ij} > 0$.

Théorème 1.1. *(Perron-Froebenius) Il existe toujours des probas stationnaires π relativement à P. De plus celle-ci est unique lorsque A_P est irréductible et $\frac{1}{n}\sum_{k=0}^{n-1}\psi P^k \stackrel{n}{\to} \pi$ pour tout vecteur de proba ψ. Si A_P est aussi apériodique, on a alors $\psi P^n \stackrel{n}{\to} \pi$ pour tout vecteur de proba ψ.*

Démonstration. L'existence de mesure stationaire suit du théorème de point fixe de Brower appliquée à la fonction $\psi \mapsto \psi P$ du simplexe des mesure de proba sur $\{1, ..., d\}$ dans lui-même. De plus toute limite de suite de la forme $\left(\frac{1}{n}\sum_{k=0}^{n-1}\psi P^k\right)_n$ est stationaire. Lorsque A_P est irréductible on voit facilement que tout vecteur de proba stationaire a ses coordonnées strictement positives. L'unicité de la proba stationnaire dans le cas irréductible (resp. la convergence forte dans le cas apériodique) résulte du théorème de Perron-Froebenius, qui garantit pour une matrice positive irréductible que le rayon spectral est valeur propre et que l'espace propre associé est de dimension 1 (resp. les autres valeurs propres sont de module strictement inférieurs). □

On a alors les propriétés suivantes pour la mesure Markovienne associée.

Théorème 1.2. *Soit P une matrice stochastique et π une measure de proba stationnaire. Alors la mesure de Markov associé à (π, P) est :*
— *ergodique ssi A_P est irréductible,*

— *mélangeante (même Bernoulli, en fait) ssi A_P est apériodique.*

Démonstration. On montre tout d'abord le cas mélangeant. On raisonne comme dans la preuve du mélange de la mesure de Bernoulli : on montre pour deux cylindres quelconques $A := [k_0, ..., k_j]$ et $B := [l_0, ..., l_i]$ que $\mu_\pi(A \cap \sigma^{-n}B) \xrightarrow{n} \mu_\pi(A)\mu_\pi(B)$. On a pour $n > j$

$$\mu_\pi(A \cap \sigma^{-n}B) = \pi_{k_0}P_{k_0k_1}...P_{k_{j-1}k_j}(P^{n-j})_{k_jl_0}P_{l_0l_1}...P_{l_{i-1}l_i}.$$

D'après le théorème de Perron-Froebenius $(P^{n-j})_{k_jl_0}$ converge vers π_{l_0} quand n tend vers l'infini et donc le terme de droite tends vers $\mu_\pi(A)\mu_\pi(B)$. Réciproquement le mélange entraine facilement l'apériodicité. Pour le cas ergodique on suit le même raisonnement en utilisant le Théorème 1.1.

□

2. L'application de Gauss

On considère pour $x \in (0, 1]$ l'application de Gauss $x \mapsto 1/x - [1/x]$.

Lemme 2.1. *La mesure $\mu_G := \frac{1}{\log 2}\frac{dx}{1+x}$ est G-invariante.*

DÉMONSTRATION. Il suffit de montrer que $\mu_G(G^{-1}[0,s]) = \mu_G([0,s])$. On a

$$\begin{aligned}
G^{-1}([0,s]) &= \bigcup_{n\geq 1}[1/n+s, 1/n], \\
\mu_G(G^{-1}([0,s])) &= = \sum_{n\geq 1}\int_{1/s+n}^{1/n}\frac{dx}{\log 2(1+x)}, \\
&= \frac{1}{\log 2}\sum_{n\geq 1}\log\left(\frac{1+1/n}{1+1/s+n}\right), \\
&= \frac{1}{\log 2}\sum_{n\geq 1}\log(1+s/n)-\log(1+s/n+1), \\
&= \frac{\log(1+s)}{\log 2} = \mu_G([0,s]).
\end{aligned}$$

□

On montre désormais que $\mu = \mu_G$ est ergodique.

LEMME 2.2. *Pour tout $x \neq 0$ on a*
— $|G'(x)| \geq 1$,
— $|(G\circ G)'(x)| \geq 2$,
— $|\left(\ln G'\circ G^{-1}\right)'(x)| \leq 2$.

LEMME 2.3 (Distorsion bornée). *Il existe une constante C telle que pour tout x, y dans la même n-branche inverse de G, on a*

$$\left|\frac{(G^n)'(x)}{(G^n)'(y)}\right| \leq C.$$

DÉMONSTRATION. On a

$$\begin{aligned}
\log\frac{(G^n)'(x)}{(G^n)'(y)} &= \sum_{k=1}^{n}\log G'\circ G^{-1}(G^k x)-\log G'\circ G^{-1}(G^k y), \\
\left|\log\frac{(G^n)'(x)}{(G^n)'(y)}\right| &\leq 2\sum_{k=1}^{n}|G^k x-G^k y|, \\
&\leq 2\sum_{k=1}^{n}2^{-k/2}|x-y| \leq C.
\end{aligned}$$

□

COROLLAIRE 2.4. *Pour tout A, B dans une même n-branche inverse, on a*

$$\frac{\mu(G^n A)}{\mu(G^n B)} \leq C\frac{\mu(A)}{\mu(B)}.$$

On conclut maintenant l'ergodicité de μ en utilisant l'argument de point de densité déja évoqué précédement. Soit A un ensemble invariant de μ mesure mesure non nulle. On considère un point de densité de A pour $\mu << Leb$. On note I_n la n-branche monotone contenant a on a alors

$$\frac{\mu(A \cap I_n)}{\mu(I_n)} \xrightarrow{n} 1.$$

Mais d'après le corollaire précédent et par invariance de A on a

$$\begin{aligned} \frac{\mu\left(G^n(A^c \cap I_n)\right)}{\mu(G^n I_n)} &\leq C\frac{\mu(A^c \cap I_n)}{\mu(I_n)}, \\ \mu(A^c) &\leq C\left(1 - \frac{\mu(A \cap I_n)}{\mu(I_n)}\right) \xrightarrow{n} 0. \end{aligned}$$

Donc $\mu(A) = 1$ et ceci conclut l'ergodicité de μ. En travaillant encore un peu on peut montrer que μ est en fait mélangeante.

3. Transformation de Chacon

On présente ici un exemple de transformation faiblement mélangeante mais pas mélangeante. C'est un exemple d'une méthode de construction plus générale connu sous le nom de "cutting and stacking" (découpage et empilement). Ces transformations sont définies sur l'intervalle comme limite de transformations $(f_n)_n$. Pour tout n la transformation f_n est défine sur des piles P_n obtenus avec les intervalles découpés. Enfin en dehors de ces tours on garde un intervalle ("the spacer") que l'on redécoupe et qu'on pourra intercaler avec les piles à l'étapes n pour former les piles à l'étape $n+1$.

Dans le cas de la trasformation de Chacon on a à toute étape une seule pile. La $n+1^{eme}$ pile est obtenu en découpant verticalement la n^{eme} pile en trois sous-piles de même taille, puis en empilant dnas l'ordre suivant les deux premières sous-piles, un intervalle du "spacer" puis la dernière sous-pile. On vérifie facilement que la transformation f_n consistant à monter verticalement dans la n^{eme} pile converge simplement sur $[0,1)$ et que la limite f est inversible sur $(0,1)$ et préserve la mesure de Lebesgue.

Théorème 3.1. *La transformation de Chacon est faiblement mélangeante mais non mélangeante.*

Démonstration. 1. On montre tout d'abord l'ergodicité. Soit $\phi : (0,1) \to \mathbb{R}$ une application measurable telle que $\phi \circ f = \phi$. Soit $\epsilon > 0$. D'après le théorème de Lusin ϕ est continue sur un ensemble compact E avec $\mu(E) > 1 - \epsilon$. Puis par uniforme continuité de ϕ sur E il existe N tel que pour tout entier $n \geq N$ on ait $|\phi(x) - \phi(y)| < \epsilon$ pour $x, y \in E$ dans une même "assiette" de la n^{eme} pile P_n. De plus pour n assez grand le spacer est de longueur inférieur à *epsilon*de sorte que l'on peut choisir une telle assiette A_n de sorte que $\mu(A_n \cap E) \geq \frac{1-2\epsilon}{l_n}$ avec l_n la hauteur de la n^{eme} pile. Par invariance de ϕ on obtient un ensemble E' avec $\mu(E') > 1 - 2\epsilon$ tel que $|\phi(x) - \phi(y)| < \epsilon$ pour $x, y \in E'$. Ceci étant vrai pour tout ϵ on en déduit facilement que ϕ est égale à une constante presque partout.

2. On vérifie maintenant la propriété de mélange faible. Il suffit de voir que toute fonction propre est constante ou encore vu qu'on vient de montrer l'ergodicité que le spectre ponctuel est réduit à $\{1\}$. Soit $\phi \in L^2$ tel que $\phi \circ f = \lambda\phi$. Par ergodicité on peut suppposer $|\phi| = 1$. On choisi E comme en 1. On note l_n la hauteur et A_n la base de la n^{me} pile. Enfin on prend x_0 dans l'intersection $f^{-l_n}E \cap E \cap f^k A_{n+1}$ pour un $k < l_n$. C'est toujours possible car $Leb(f^{-l_n}E \cap E) > 1 - 2\epsilon$ et $Leb(\bigcup_{k<l_n} f^k A_{n+1}) = \frac{Leb(P_n)}{3}$. Maintenant par la procédure d'empilement on a $f^{l_n}x_0 \in f^k A_n$ et donc $|\phi(x_0) - \phi(f^{l_n}x_0)| < \epsilon$. Mais $\phi \circ f^{l_n} = \lambda^{l_n}\phi$ donc $|\lambda^{l_n} - 1| < \epsilon$. La même chose s'applique à $n+1$ de sorte que $|\lambda^{l_{n+1}} - 1| < \epsilon$. Mais $l_{n+1} = 3l_n + 1$ d'où $\lambda = 1$.

3. La transformation f n'est pas mélangeante. La base A_0 de la première pile est $[0, 2/9[$. L'empilement fait que $Leb(A_0 \cap T^{l_0=1}A_0) \geq \frac{Leb(A_0)}{3}$. Puis si on note A_0^i pour $i = 1, ..., 3$ les trois sous intervalles découpés de A_0 on a de même $Leb(A_0^i \cap T^{l_1}A_0^i) \geq \frac{Leb(A_0^i)}{3}$ pour tout i, et donc $Leb(A_0 \cap T^{l_1}A_0) \geq \frac{Leb(A_0)}{3}$. Par récurrence on montre de même que $Leb(A_0 \cap T^{l_n}A_0) \geq \frac{Leb(A_0)}{3}$ pour tout n. Mais si la transformation était mélangeante on aurait alors $Leb(A_0 \cap T^{l_n}A_0) \to_n Leb(A_0)^2 = (\frac{2}{9})^2$. Or $\frac{Leb(A_0)}{3} = \frac{2}{9} \times \frac{3}{9}...$

□

4. Flot géodésique et horocyclique

La mesure de Haar sur le groupe (unimodulaire) $SL2(\mathbb{R}) = \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \ \alpha\delta - \gamma\beta = 1 \right\}$ est donnée par $\mu = \frac{d\delta d\beta d\gamma}{|\delta|}$. Mais on a $\mu(SL^2(\mathbb{R})) = +\infty$. On s'intéresse au sous groupe discret $\Gamma = SL_2(\mathbb{Z})$ et au quotient $X = SL^2(\mathbb{Z}) \setminus SL^2(\mathbb{R})$ muni de la topologie quotient (X n'est pas compact). On peut montrer que $SL_2(\mathbb{R})$ admet une distance invariante à gauche D. On considère alors la distance quotient d sur X définit par $d(\Gamma g, \Gamma g') = \inf_{\gamma \in \Gamma} D(g, \gamma g')$. Le groupe $\Gamma = SL^2(\mathbb{Z})$ étant discret la mesure μ induit une mesure ν sur X encore invariante par translation à droite. On peut alors montrer que $\nu(X) < +\infty$.

Le flot géodésique est l'action associée au flot $(g_t)_{t\in\mathbb{R}}$ défini par $g_t(\Gamma h) = \Gamma h G_t$ avec

$$G_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}.$$

On définit aussi les flots $(h_s^-)_s$ et $(h_s)_s$ par $h_s(\Gamma h) = \Gamma h H_s$ et $h_s(\Gamma h) = \Gamma h H_s$ avec $H_s^- = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$ et $H_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$.

La mesure ν étant invariante par translation a droite elle est en particulier invariante par $(g_t)_t$, $(h_t)_t$ et $(h_t^-)_t$. On rappelle que $SL_2(\mathbb{R})$ est engendré par les matrices $H_s, H_{s'}^-$ pour tout $s, s' \in \mathbb{R}$.

THÉORÈME 4.1. *La mesure ν est ergodique pour le flot géodésique sur X.*

DÉMONSTRATION. On doit vérifier que si $f \in L^2$ est g_t-invariante alors f est constante. Pour cela il suffit montrer que f est aussi h_s et h_s^- invariante car ces éléments engendrent tout le groupe $SL_2(\mathbb{R})$.

Pour $g \in SL_2(\mathbb{R})$, on note $T_g : L^2(X) \to L^2(X)$ l'isométrie de $L^2(X)$ qui à $f \in L^2(X)$ associe $\Gamma h \mapsto f(\Gamma h g)$. L'application $g \mapsto T_g(f)$ est continue de $SL^2(\mathbb{R})$ dans $L^2(X)$ (on peut supposer f continue à support compact par densité, puis utiliser l'uniforme continuité de f).

Soient $h, g \in Sl_2(\mathbb{R})$ et $f \in L^2(X)$ invariante par g alors

$$\|T_h f - f\| = \|T_{g^n h g^{-n}} f - f\|$$

Si $g^n h g^{-n}$ tend vers l'identité quand n tend vers $+\infty$, alors on aura à la limite $T_h f = f$.

On conclut en remarquant que $G_t H_s G_t^{-1} = H_{se^t}$ et $G_t H_s G_t^{-1} = H_{se^{-t}}$, d'où

$$(G_t)^n H_s G_t^{-n} \xrightarrow{n\to+\infty} Id \text{ for } t < 0,$$

$$(G_t)^n H_s^- G_t^{-n} \xrightarrow{n\to+\infty} Id \text{ for } t > 0.$$

$\square$

THÉORÈME 4.2. *La mesure ν est ergodique pour le flot horocyclique $(h_s)_s$ sur X.*

DÉMONSTRATION. D'après la preuve précédente il suffit de montrer que toute fonction $f \in L^2(X)$ qui est h_s-invariante pour tout s est g_t-invariante pour un certain $t > 0$. Pour cela on raisonne comme précédemment en utilisant pour $s \neq 0$:

$$H_s^{-2n}\begin{pmatrix}1 & 0\\ -\frac{1}{2ns} & 1\end{pmatrix}H_s^n \xrightarrow{n\to+\infty} \begin{pmatrix}2 & 0\\ 0 & 1/2\end{pmatrix}$$

□

La mesure ν est en fait mélangeante pour les flots géodésique et horocyclique (Théorème de Howe-Moore, cf Exercices ci-dessous). Le flot géodésique est même Bernoulli, mais pas le flot horocyclique...

5. Exercices

Exercice 12.
Montrez que la mesure de Parry satisfait la propriété de Gibbs (Proposition 1.4).

Exercice 13. *Théorème de Howe-Moore.*
Avec les notations de la section 5.4 on se propose de montrer que pour tout $f \in L_0^2(\nu)$ $T_g f$ converge faiblement vers 0 quand g tends vers $+\infty$. Montrer que ceci entraine le mélange du flot horocyclique et géodésique pour ν. On passe maintenant à la preuve du théorème. Nous rappelons que tout $g \in SL2(\mathbb{R})$ s'écrit sous la forme $g = k_1 a k_2$ avec k_i orthogonales et a diagonale à coeffcient positifs.

— Soit u une limite faible de $(T_{g_n} f)_n$. On note (k_1, k_2) une limite de $k_1(g_n)$ et $k_2(g_n)$. Montrez que $T_{a_n k_2} f$ converge faiblement vers $T_{k_1^{-1}} u$.
— En déduire que $T_{k_1^{-1}} u$ est h_s-invariante.
— Conclure.

Exercice 14. *Transformation de Von Neuman-Kakutani.*
On considère la transformation définie par cutting and stacking comme suit. On découpe l'intervalle $[0,1]$ en deux et on empile la partie droite $[1/2,1)$ sur la gauche $[0,1/2)$. On effectue de nouveau l'opération sur la pile de deux segments ainsi obtenue. On répète cette trasformation comme dans la construction de la transformationde Chacon. La transformation limite f est appelée transformation de Von Neuman-Kakutani.

— Etabllir l'ergodicité.
— Montrer que cette transformation n'est pas faiblement mélangeante.

Chapitre 6

Entropie mesurée

Contents

1.	**Information**	**51**
2.	**Entropie statique**	**52**
3.	**KS-Entropie et théorème des générateurs**	**52**
4.	**Formule de Katok, Newhouse et Shanon-Mcmillan-Breiman**	**54**
5.	**Exemples**	**56**
6.	**Exercices**	**57**

1. Information

Pour α, β deux partitions finies de X on dit que α est plus fine que β et on note $\alpha > \beta$, lorsque tout élément de β est l'union d'éléments de α. Enfin la partition jointe $\alpha \vee \beta$ est définie conmme $\alpha \vee \beta := \{A \cap B,\ A \in \alpha \text{ et } B \in \beta\}$.

Définition 1.1. *L'information de α relativement à β est la fonction*

$$I(\alpha|\beta) := -\sum_{A\in\alpha} 1_A \log E[1_A|\beta] = -\sum_{A\in\alpha, B\in\beta} 1_{A\cap B} \log\left(\frac{\mu(A\cap B)}{\mu(B)}\right).$$

On notera aussi $\mu(A|\beta)$ pour $E[1_A|\beta]$. Remarquez que lorsque α et β sont indépendantes, i.e. $\mu(A\cap B) = \mu(A)\mu(B)$ pour tout $(A,B) \in \alpha \times \beta$ alors $E[1_A|\beta] = \mu(A)$ et donc $I(\alpha|\beta) = I(\alpha)$.

Lemme 1.2.

$$I(\alpha \vee \beta|\gamma) = I(\alpha|\beta \vee \gamma) + I(\beta|\gamma).$$

Démonstration. Il suffit d'écrire puis de sommer pour tout $A \in \alpha$, $B \in \beta$, $C \in \gamma$, l'égalité suivante :

$$\log\left(\frac{\mu(A\cap B\cap C)}{\mu(C)}\right) = \log\left(\frac{\mu(A\cap B\cap C)}{\mu(B\cap C)}\right) + \log\left(\frac{\mu(B\cap C)}{\mu(C)}\right).$$

□

2. Entropie statique

DÉFINITION 2.1. *L'entropie statique conditionelle de α sachant β est*

$$H_\mu(\alpha|\beta) = \int I(\alpha|\beta)d\mu.$$

Lorsque β est la partition triviale on note l'entropie statique de α (relativement à β)

$$H(\alpha) = -\sum_{A\in\alpha} \mu(A)\log\mu(A).$$

On a alors

$$H_\mu(\alpha|\beta) = \sum_{B\in\beta} \mu(B)H_{\mu_B}(\alpha).$$

LEMME 2.2. *Soient α, β, γ trois partitions finies,*

— $H_\mu(\alpha \vee \beta|\gamma) = H_\mu(\alpha|\beta\vee\gamma) + H_\mu(\beta|\gamma)$,

— *si $\beta > \gamma$, on a $H_\mu(\alpha|\beta) \leq H_\mu(\alpha|\gamma)$,*

— $H_\mu(\alpha) \leq \log\sharp\alpha$ *avec égalité ssi* $\mu(A) = 1/\sharp\alpha$ *pour tout* $A\in\alpha$.

DÉMONSTRATION. La première identité s'obtient en intégrant l'identité associée pour l'information. Pour la seconde inégalité remarquez que pour tout $C\in\gamma$ on a pour tout $A\in\alpha$ par convexité de $\phi(t) = -t\log t$

$$\sum_{B\in\beta,\ B\subset C} \mu(B)\left(\phi(\mu(A\cap B)/\mu(B))\right) \leq \mu(C)\phi(\mu(A\cap C)/\mu(C)).$$

Enfin pour la dernière égalité il suffit d'appliquer l'inégalité de Jensen (et son cas d'égalité) à la variable aléatoire equidistribuée sur l'ensemble $\{\mu(A),\ A\in\alpha\}$. □

Lorsque les partitions α et β sont indépendantes la première inégalité devient pour γ triviale

$$H_\mu(\alpha\vee\beta) = H_\mu(\alpha) + H_\mu(\beta).$$

3. KS-Entropie et théorème des générateurs

Pour α, β deux partition finies il suit du lemme ci-dessus que la suite $H_\mu(\alpha^n|\beta^n)$ est sous additive. La limite $h(f,\alpha) = \lim_n \frac{H_\mu(\alpha^n)}{n} = \inf_n \frac{H_\mu(\alpha^n)}{n}$ est bien définie. L'entropie de μ est alors définie comme

$$h(f,\mu) = \sup_{\alpha\ finie} h(f,\alpha).$$

LEMME 3.1. — $h(f,\alpha) - h(f,\beta) \leq H(\alpha|\beta)$,

— *pour* (f,μ) *inversible on a* $h(f,\bigvee_{k=-N}^{N}\alpha^k)=h(f,\alpha)$.

On dit qu'une partition finie P est génératrice pour un système inversible lorsque $\bigvee_{n\in\mathbb{Z}}P^k=\mathcal{B}$, i.e. la plus petite tribu contenant les éléments de P^k pour tout k est la tribu $\mathcal{B}$.

Théorème 3.2. *(Générateurs de Sinaï) Soit P une partition génératrice pour un système inversible, alors*

$$h(\mu)=h(\mu,P).$$

Démonstration. Il suffit de montrer que $H(\beta|\bigvee_{k=-N}^{N}\alpha^k)$ tends vers 0 quand N tend vers l'infini.

$$\begin{aligned}
H(\beta|\bigvee_{k=-N}^{N}\alpha^k) &= -\sum_{B\in\beta}\int 1_B\log\mu(B|\bigvee_{k=-N}^{N}\alpha^k),\\
&= -\sum_{B\in\beta}\int \mu(B|\bigvee_{k=-N}^{N}\alpha^k)\log\mu(B|\bigvee_{k=-N}^{N}\alpha^k),\\
&= \sum_{B\in\beta}\int\phi\left(\mu(B|\bigvee_{k=-N}^{N}\alpha^k)\right)d\mu,
\end{aligned}$$

avec $\phi(t)=-t\log t$ pour $t\in[0,1]$. Par le théorème de convergence des martingales $\mu(B|\bigvee_{k=-N}^{N}\alpha^k)$ converge ponctuellement vers 1_B. Puis par convergence dominée $\int\phi\left(\mu(B|\bigvee_{k=-N}^{N}\alpha^k)\right)d\mu\to_N 0$. □

On montre de même que si $(\alpha_k)_k$ est une suite croissante de partitions finies telle que $\bigvee_k\alpha_k=\mathcal{B}$ alors $h(\mu)=\lim_k h(\alpha_k)$.

Lemme 3.3. — $h(\nu\times\mu)=h(\nu)+h(\mu)$,
— *si* $\pi:(Y,\nu)\to(X,\mu)$ *est un facteur alors* $h(\nu)\le h(\mu)$,
— $h(f^k,\mu)=kh(f,\mu)$.

Démonstration. Par indépendance l'entropie pour la mesure produit d'un produit de partition est donnée par la somme des entropies relatives à ces partitions. La formule sur l'entropie du produit suit alors de la remarque précédente puisque les partitions produits engendrent la tribu produit. Pour le deuxième point on vérifie facilement que pour tout partition P on a

$$h(\mu,P)=h(\nu,\pi^{-1}P),$$

et donc $h(\mu)\le h(\nu)$ en prenant le supremum sur toutes les partitions finies. Enfin l'entropie des puissances suit de $h_{f^k}(\mu,P^k)=kh_f(\mu,P)$ et $H_{f^k}(\mu,P^n)\le H_f(\mu,P^{kn})$.

□

Il suit en particulier du deuxième point que l'entropie est invariante par isomorphisme métrique.

4. Formule de Katok, Newhouse et Shanon-Mcmillan-Breiman

On définit maintenant l'entropie de façons différentes pour un système ergodique :

$$s^{+/-} := \overline{\underline{\lim}}_n - \frac{\log \mu(A_x^n)}{n}$$

$$k_\lambda^{+/-} := \overline{\underline{\lim}}_n \frac{1}{n} \log \min \left\{ \sharp C_n,\ C_n \subset \alpha^n \text{ et } \mu(\bigcup_{A^n \in C_n} A^n) > \lambda \right\}$$

$$n_\lambda^{+/-} := \inf_{E,\ \mu(E)>\lambda} \overline{\underline{\lim}}_n \frac{1}{n} \log \min \left\{ \sharp C_n,\ C_n \subset \alpha^n \text{ et } E \subset \bigcup_{A^n \in C_n} A^n \right\}$$

On va montrer que pour un système ergodique ces quantités sont égales à l'entropie de Kolmogorov-Sinaï.

LEMME 4.1. *$s^{+/-}$ sont des constantes (presque partout).*

DÉMONSTRATION. Puisque $\int f d\mu = \int f \circ T d\mu$ pour tout $f \in L^1$ par invariance de la mesure, toute fonction mesurable intégrable vérifiant $f \circ T \leq f$ satisfait $f \circ T = f$ p.p. et donc f est constante p.p. par ergodicité. Mais $\log \mu(A_x^n) \leq \log \mu(A_{Tx}^{n-1})$ implique $s^{+/-}(Tx) \leq s^{+/-}(x)$. □

On va maintenant comparer ces différentes quantités entre elles. Par une inversion triviale d'infimum on a tout d'abord :

LEMME 4.2.

$$k_\lambda^{+/-} \leq n_\lambda^{+/-}$$

LEMME 4.3.

$$n_\lambda^{+} \leq s^{-}$$

DÉMONSTRATION. Soit $\epsilon > 0$. Pour μ p.p. x on note $n(x)$ le premier entier $n > 1/\epsilon$ tel que $\mu(A_x^n) \geq e^{-n(s^- + \epsilon)}$. Soit $G \subset X$ avec $\mu(G) > 1 - \epsilon$ tel que $n(x) < N_0$ pour tout $x \in G$. D'après le théorème ergodique de Birkhof, il existe $E \subset X$ avec $\mu(E) > \lambda$ et N_1 tels que pour $n > N_1$ et $x \in E$ on ait $\frac{\sharp\{0 \leq k < n,\ T^k x \in G\}}{n} \simeq 1 - \epsilon$. Pour x fixé dans E et $n > N_1$ on decoupe l'intervalle $[0, n]$ comme suit. On marque avec [le premier instant k_0 avec $f^{k_0} x \in G$ puis par] l'instant $k_0 + n(f^{k_0} x)$. Puis on marque de nouveau le premier instant k_1 dans G avec [avec $k_1 > k_0 + n(f^{k_0} x)$. On continue jusqu'à dépasser n.

On obtient ainsi une découpe de l'intervalle $[0,n]$ avec des symboles [et], qui marquent des entrées x_i dans G et les longueur $n(x_i)$ associées. Puisque $n(x) > 1/\epsilon$ pour tout x le nombre de configurations possibles est plus petit que $(C_n^{\epsilon n})^2$. Maintenant dans les intervalles délimités par [] le nombre de α-codages possibles est borné par $e^{n'(s^-+\epsilon))}$ où n' est la somme des longueurs de ces intervalles. Or en dehors de ces intervalles on est dans le complémentaire de E ou dans le dernier intervalle dépassant n (qui est de longueur au plus N_0), donc $n' \gtrsim (1-\epsilon)n - N_0$. Par conséquent

$$\sharp\{C_n \subset \alpha^n,\ E \subset C_n\} \leq (C_n^{\epsilon n})^2 \sharp\alpha^{\epsilon n + N_0} e^{n(s^-+\epsilon))}$$

On conclut en utiisant le fait qu'il existe ϕ avec $\lim_{\epsilon\to 0}\phi(\epsilon) = 0$ tel que $C_n^{\epsilon n} \leq e^{\phi(\epsilon)n}$ pour tout n. $\square$

Lemme 4.4.

$$n_\lambda^+ \geq s^+$$

Démonstration. Supposons par l'absurde que $n_\lambda^+ < s^+$. Soit $\epsilon > 0$ tel que $n_\lambda^+ + \epsilon < s^+ - \epsilon$ et soit E avec $\mu(E) > \lambda$ et $C_n \subset \alpha^n$ un recouvrement minimal de E tel que pour tout $n \in \mathbb{N}$ on ait $\sharp C_n \lesssim Ce^{(n_\lambda^+ + \epsilon)n}$ (pour un $C > 0$). Alors pour $n_\lambda^+ + \epsilon < b < s^+$ on a

$$\sum_{n\in F} \mu\left(\underbrace{\{x,\ \mu(A_x^n) < e^{-bn} \text{ et } A^n \in C_n\}}_{B_n}\right) \leq \sum_{n\in F} e^{-bn}\sharp C_n < +\infty.$$

Par le lemme de Borel-Cantelli, μ-presque tout x appartient seulement à un nombre fini de B_n pour $n \in \mathbb{N}$. Mais par définition de s^+ presque tout point de X verifie $\mu(A_x^n) < e^{-(s^+-\epsilon)n}$ pour une infinité de n. C'est en particulier le cas pour μ p.p. $x \in E \subset \bigcap_{n\in\mathbb{N}} C_n$. Contradiction. $\square$

Lemme 4.5.

$$k_\lambda^- \geq s^-$$

Démonstration. Pour tout ϵ, λ il existe un ensemble F de mesure plus grande que $1 - \lambda/2$ et un entier N tel que pour tout $n > N$ on a $\mu(A_x^n) \lesssim e^{-(s^- - \epsilon)n}$ pour $x \in F$. Si $C_n \subset \alpha^n$ tel que $\mu(C_n) > \lambda$ et donc $\mu(C_n \cap F) > \lambda/2$. Par conséquent $\sharp C_n \gtrsim \frac{2}{\lambda} e^{(s^- - \epsilon)n}$ et donc $k_\lambda^- \geq s^-$. $\square$

On compare désormais ces quantités à l'entropie métrique. Une application directe du lemme de Fatou donne

Lemme 4.6.

$$s^- \leq h$$

L'inégalité suivante permet de conclure :

LEMME 4.7.

$$h \leq \lim_{\lambda \to 1} k_\lambda^+$$

DÉMONSTRATION. Pour une infinité de n on dispose d'une collection $C_n \subset \alpha^n$ de mesure totale plus grande que α telle que $\frac{\log C_n}{n} \simeq k_\lambda^+$. On a

$$\begin{aligned} H(\alpha^n) &\leq H(\alpha^n|\{X \setminus C_n, C_n\}) + H(\{X \setminus C_n, C_n\}), \\ &\leq \mu(X \setminus C_n) H_{\mu_{X \setminus C_n}}(\alpha^n) + \mu(C_n) H_{\mu_{C_n}}(\alpha^n) + \log 2, \\ &\leq (1-\lambda) n \log \sharp\alpha + \lambda \log \sharp C_n + \log 2. \end{aligned}$$

En divisant par n et en prenant la limite supérieure en l'infini on obtient le resultat voulu.

□

5. Exemples

PROPOSITION 5.1. *(1) L'entropie des mesures périodiques est toujours nulle. L'entropie des rotations sur le cercle est nulle,*

(2) l'entropie d'un Bernoulli de paramètre $(p_1, ..., p_n)$ est $\sum_i -p_i \log p_i$,

(3) l'entropie d'une mesure de Markov avec proba stationnaire (π_i) et matrice de transition (p_{ij}) est $-\sum_i \pi_i p_{ij} \log p_{ij}$,

(4) l'entropie de la mesure de Gauss est $\int \log |G'| d\mu$,

(5) l'entropie de la mesure de Lebesgue pour un autmorphisme linéaire f_A du tore est $\sum_i |\lambda_i|$ où les λ_i sont les logarithmes des modules des valeurs propres de A de module > 1.

DÉMONSTRATION. 1. Il suffit de montrer que pour toute partition en intervalle α on a $h(\mu, \alpha) = 0$. En effet on peut trouver une suite croissante $(\alpha_k)_k$ de telles partitions telles que $(\alpha_k)_k$ engendre la tribu des boréliens. Il suffit en effet de considérer une suite de partitions dont le diamètre tend vers 0. On conclut à l'aide du théorème des générateurs. Maintenant si α est une partition en intervalle on voit facilement que $\sharp\alpha^n \leq n\sharp\alpha$ et donc $h(\mu, \alpha) = \lim_n \frac{H(\alpha^n)}{n} \leq \lim_n \frac{\log \sharp(\alpha^n)}{n} = 0$.

2. La partition P_0 en coordonnée zero est clairement génératrice. Il suffit donc de calculer $h(\mu, P_0)$. Clairement $H(P_0) = \sum_i -p_i \log p_i$. Mais par définition de la mesure de Bernoulli les partitions P_0 et $\bigvee_{k=1}^{n-1} P_0^{-k}$ sont indépendantes de sorte que $H_\mu(P_0^n) = H_\mu(P_0) + H_\mu(P_0^{n-1})$. Donc $\frac{H_\mu(P_0^n)}{n} = H_\mu(P_0)$ pour tout n...

3. La partition $P = \{[1/n, 1/n+1[,\ n \in \mathbb{N}\}$ en branches inversibles est génératrice. On a vu en effet que $diam(P^n) \leq C/2^n$. Maintenant on a $\mu(A_x^n) \simeq m(A_x^n) = \int_{0,1} \frac{1}{|(G^n)'|(y)} dy$. Mais par le lemme de distorsion, on a $C^{-1}\frac{1}{|(G^n)'|(x)} \leq \int_{0,1} \frac{1}{|(G^n)'|(y)} dy \leq C\frac{1}{|(G^n)'|(x)}$ pour une constante $C > 1$. On a donc pour n grand $-\frac{\log \mu(A_x^n)}{n} \simeq \frac{1}{n} \log |(G^n)'|(x) = \frac{1}{n}\sum_{k=0}^{n-1} \log |G'(G^k x)|$. D'après le théorème ergodique (rappelons que la mesure de Gauss est ergodique) cette dernière quantité tend vers $\int \log |G'| d\mu$. Enfin $h = s^+ = \int \log |G'| d\mu$.

4. On verra dans les chapitres suivants que $h(\mu) \leq \sum_i |\lambda_i|$ pour toute mesure f_A-invariante. Il suffit donc de montrer que $h(Leb, \beta) \geq \sum_i |\lambda_i|$ pour une partition β. On considère la partition $\beta = \{\overline{A^{-1}|_{[0,1[^d}\gamma}\}$ avec γ la partition de $\mathbb{R}^d$ en cubes de taille 1, i.e. $x, y \in [0,1[^d$ sont dans le même élément de β ssi Ax et Ay sont dans le même cube de γ . On peut décomposer $\mathbb{R}^d = E_u \oplus E_{cs}$ avec E_u, resp. E_{cs}, les sous espaces vectoriels réels invariants correspondant aux espaces propres associées aux valeurs propres de modules > 1, resp ≤ 1. Tout élément B^n de β^n est un polygone qui s'envoie par A^n dans un polygone de longueur bornée par une constante (indépendante de n) dans la direction E_u et tel que $Leb_{E_{cs}(x)}(A^n B^n \cap E_{cs}(x)) \leq Jac(A^n|E_{cs})$ pour tout x. En utilisant le théorème de Fubini on obtient donc

$$Leb(A^n B_x^n) \leq C \int_{E_u(x)} \left(\int_{E_{cs}(y)} 1_{A^n B_x^n} dz \right) dy \leq C Jac(A^n|E_{cs}).$$

Par changement de variable on a donc pour tout μ p.p. x

$$-\frac{\ln Leb(B_x^n)}{n} \geq \frac{\ln C Jac\,(f^n|_{E_u})}{n} \to_n \sum_i |\lambda_i|.$$

On conclut en appliquant la formule de Shanon-Mcmillan-Breiman.

□

On a vu que deux systèmes de Bernoulli sont toujours spectralement isomorphes (ils ont le spectre de Lebesgue dénombrable). Cependant ils n'ont en général pas la même entropie et ne sont donc pas métriquement isomorphes. Ornstein a montré que l'entropie classifiait complètement les Bernoulli, i.e. deux systèmes de Bernoulli avec la même entropie sont (métriquement) isomorphes.

6. Exercices

Exercice 15. *Formule d'Abramov.*
Soient $(X, \mathcal{B}, \mu, f)$ un système mesuré ergodique inversible et $A \in \mathcal{B}$ avec

$\mu(A) > 0$. On note $f_A = f^{\tau_A}$ l'application de premier retour induite sur A et $\mu_A = \mu(. \cap A)/\mu(A)$.

— Vérifiez que μ_A est f_A-invariante et ergodique,

— Montrez que
$$h_{f_A}(\mu_A) = \frac{h_f(\mu)}{\mu(A)}.$$

Chapitre 7

Entropie topologique et principe variationnel

Contents

1. Complexité **59**
2. Entropie topologique **60**
3. Harmonicité de l'entropie **61**
4. Entropie à la Bowen **61**
5. Principe variationnel **62**
6. Applications et exemples **64**
6.1. Inégalité de Ruelle en dimension 1 64
6.2. Entropie des applications de l'intervalle monotone par morceaux 65
6.3. Sous-décalage à alphabet fini 66
7. Exercices **67**

On considère dans cette partie un système dynamique topologique (X, T).

1. Complexité

Pour un recouvrement fini d'ouverts $\mathcal{U}$ de X, on note $\mathcal{U}^{[k,l]}$ le recouvrement itéré $\mathcal{U}^{[k,l]} := \bigcap_{j=k}^{l} T^{-j}\mathcal{U}$. On note aussi $\mathcal{U}_V$ le recouvrement induit sur $V \subset X$. Pour deux recouvrements d'ouverts $\mathcal{U}$ et $\mathcal{V}$ on dit que $\mathcal{V}$ est plus fin que $\mathcal{U}$, noté $\mathcal{V} > \mathcal{U}$ lorsque tout élément de $\mathcal{V}$ est inclus dans un élément de $\mathcal{U}$.

On note $H(\mathcal{U})$ la complexité de $\mathcal{U}$:

$$H(\mathcal{U}) = \log \min\{\sharp\mathcal{W},\ \mathcal{W} > \mathcal{U}\}$$

Si $\mathcal{U}$ et $\mathcal{V}$ sont deux recouvrements d'ouverts on définit la complexité conditionnelle :

$$H(\mathcal{U}|\mathcal{V}) = \sup_{V \in \mathcal{V}} \log \min\{\sharp\mathcal{W},\ \mathcal{W} \text{ recouvrement de } V \text{ avec } \mathcal{W} > \mathcal{U}_V\}$$

Remarquez que les minimas ci-dessus sont toujours atteints pour un sous-recouvrement $\mathcal{W} \subset \mathcal{U}$ de $\mathcal{U}$.

Lemme 1.1. — *$H(\mathcal{U} \vee \mathcal{V}|\mathcal{W}) \leq H(\mathcal{U}|\mathcal{V} \vee \mathcal{W}) + H(\mathcal{V}|\mathcal{W})$,*
— *$H(T^{-1}\mathcal{U}|T^{-1}\mathcal{V}) \leq H(\mathcal{U}|\mathcal{V})$,*
— *si $\mathcal{V} > \mathcal{W}$ alors $H(\mathcal{U}|\mathcal{V}) \leq H(\mathcal{U}|\mathcal{W})$,*

— $H(\mathcal{U}) \leq \log \sharp\mathcal{U}$.

2. Entropie topologique

Comme pour l'entropie mesurée, on montre facilement avec le lemme précédent que la suite $H(\mathcal{U}^n|\mathcal{V}^n)$ est sous-additive et il s'en suit que $h_{top}(f,\mathcal{U}) = \lim_n \frac{H(\mathcal{U}^n)}{n} = \inf_n \frac{H(\mathcal{U}^n)}{n}$ est bien définie (cette quantité est aussi notée $h(\mathcal{U})$ ou $h(f,\mathcal{U})$ par la suite). L'entropie topologique est alors définie comme

$$h_{top}(f) = \sup_{\mathcal{U}} h(f,\mathcal{U}).$$

LEMME 2.1. — *$h(\mathcal{U}) - h(\mathcal{V}) \leq H(\mathcal{U}|\mathcal{V})$,*

— *pour un homéomorphisme T on a $h(\bigvee_{k=-N}^{N} T^{-k}\mathcal{U}) = h(\mathcal{U})$.*

On dit qu'un homéomorphisme de X est expansif s'il existe un recouvrement ouvert $\mathcal{U}$ tel que le diamètre de $\mathcal{U}^n$ tendent vers 0 quand n tends vers l'infini ou de façon équivalente les éléments de $\mathcal{U}^\infty$ sont des singletons. On dira alors que $\mathcal{U}$ est un générateur topologique.

Par compacité de X pour tout recouvrement ouvert $\mathcal{U}$ de X il existe $\epsilon_{\mathcal{U}}$ tel que toute boule de diamètre inférieur à $\epsilon_{\mathcal{U}}$ est inclus dans un élément de $\mathcal{U}$. Le maximum des tels $\epsilon_{\mathcal{U}}$ est appelé la constante de Lebesgue de $\mathcal{U}$ et est notée $Leb(\mathcal{U})$. Par définition si α est une partition dont le diamètre est inférieur à la constante de Lebesgue d'un générateur topologique $\mathcal{U}$ alors α est un générateur (au sens mesuré) pour toutes les mesures T-invariantes.

THÉORÈME 2.2. *Soit $\mathcal{V}$ un générateur topologique, alors*

$$h_{top}(T) = h(\mathcal{V}).$$

DÉMONSTRATION. On commence comme pour le théorème des générateurs de Sinaï : pour tout recouvrement d'ouvert $\mathcal{U}$ on a

$$h(\mathcal{U}) - h(\mathcal{V}) \leq H(\mathcal{U}|\bigvee_{k=-N}^{N} T^{-k}\mathcal{V})$$

Maintenant si N est choisi assez grand de sorte que le diamètre de $\bigvee_{k=-N}^{N} T^{-k}\mathcal{V}$ soit plus petit que la constante de Lebesgue du recouvrement $\mathcal{U}$, alors

$$H(\mathcal{U}|\bigvee_{k=-N}^{N} T^{-k}\mathcal{V}) = 0$$

□

Comme dans le cas mesuré, on montre de la même façon que si $(\mathcal{U}_k)_k$ est une suite de recouvrements ouverts dont le diamètre tends vers 0 alors $h_{top}(T) = \lim_k h_{top}(T, \mathcal{U}_k)$. Le lemme suivant se montre comme dans le cas mesuré.

Lemme 2.3. — *$h_{top}(T \times S) = h_{top}(T) + h_{top}(S)$,*

— *si $\pi : (Y, S) \to (X, T)$ est un facteur topologique alors $h_{top}(T) \leq h_{top}(S)$,*

— *$h_{top}(T^k) = k h_{top}(T)$.*

3. Harmonicité de l'entropie

Théorème 3.1. *Soit (X, f) un système mesuré et soient μ, ν deux mesures f-invariantes, alors pour tout $\lambda \in [0,1]$*

$$h(\lambda\mu + (1-\lambda)\nu) = \lambda h(\mu) + (1-\lambda) h(\nu).$$

Démonstration. Tout d'abord par concavité on a $h(\lambda\mu + (1-\lambda)\nu) \geq \lambda h(\mu) + (1-\lambda) h(\nu)$. Montrons l'inégalité inverse. On considère $X' = X_1 \coprod X_2$ avec X_i des copies de X. Sur X_1 on considère la mesure ξ correspondant à $\lambda\mu$ et $(1-\lambda)\nu$ sur X_2. Pour α une partition de X on note α' la partition de X' donnée par les ensembles de la forme $A_1 \cup A_2$ avec A_1, A_2 la copie du même élément de $A \in \alpha$. Alors si β est la partition à deux éléments $\beta = \{X_1, X_2\}$ on a

$$\begin{aligned} H_\xi(\alpha'^n) &\leq H_\xi(\beta) + H_\xi(\alpha'^n|\beta), \\ &\leq \log 2 + \xi(X_1) H_{\xi_{X_1}}(\alpha'^n) + \xi(X_2) H_{\xi_{X_2}}(\alpha'^n), \\ &\leq \log 2 + \lambda H_\mu(\alpha^n) + (1-\lambda) H_\nu(\alpha^n). \end{aligned}$$

On vérifie aussi directement que $H_\xi(\alpha'^n) = H_{\lambda\mu+(1-\lambda)\nu}(\alpha^n)$. En divisant par n on conclut facilement en prenant la limite quand n tends vers l'infini. □

En travaillant encore un peu, on peut en fait montrer que h est harmonique, i.e. si $\mu = \int \nu dM_\mu(\nu)$ est la décomposition ergodique de μ alors

$$h(\mu) = \int h(\nu) dM_\mu(\nu).$$

4. Entropie à la Bowen

Pour $x \in X$, $n \in \mathbb{N}$ et $\epsilon > 0$ on note $B(x, n, \epsilon)$ la (ϵ, n)-boule dynamique

$$B(x, n, \epsilon) := \{y \in X,\ d(T^k x, T^k y) < \epsilon \ \text{ pour tout } 0 \leq k < n\}.$$

Clairement si $\mathcal{U}$ est un recouvrement d'ouvert de X avec $diam(\mathcal{U}) < \epsilon$ (resp. $Leb(\mathcal{U}) < \epsilon$) alors $U_x^n \subset B(x,n,\epsilon)$ (resp. $B(x,n,\epsilon) \subset U_n^x$) pour tout $U_x^n \in \mathcal{U}^n$ contenant x. En particulier on a

$$h_{top}(T) = \lim_{\epsilon \to 0} h_{top}(T,\epsilon)$$

$$\text{avec } h_{top}(T,\epsilon) := \limsup_n \frac{1}{n} \log \min\{\sharp C_n,\ \bigcup_{x \in C_n} B(x,n,\epsilon) = X\}.$$

On montre aussi facilement que

$$h_{top}(T) = \lim_{\epsilon \to 0} h'_{top}(T,\epsilon)$$

$$\text{avec } h'_{top}(T,\epsilon) := \limsup_n \frac{1}{n} \log \max\{\sharp C_n,\ y \notin B(x,n,\epsilon) \text{ pour } x \neq y \in C_n\}.$$

On vérifie en effet que

$$\min\{\sharp C_n,\ \bigcup_{x \in C_n} B(x,n,\epsilon) = X\} \leq \max\{\sharp C_n,\ y \notin B(x,n,\epsilon) \text{ pour } x \neq y \in C_n\}$$

et

$$\max\{\sharp C_n,\ y \notin B(x,n,\epsilon) \text{ pour } x \neq y \in C_n\} \leq \min\{\sharp C_n,\ \bigcup_{x \in C_n} B(x,n,\epsilon/2) = X\}.$$

LEMME 4.1. *On suppose μ ergodique. Soit α une partition avec $\mu(\partial\alpha) = 0$. Alors pour tout $\delta > 0$ il existe ϵ_0 tel que pour tout $\epsilon < \epsilon_0$ pour μ pp x*

$$\limsup_n \frac{1}{n} \log \sharp\{A^n \in \alpha^n,\ A^n \cap B(x,n,\epsilon) \neq \emptyset\} < \delta.$$

DÉMONSTRATION. Soit $\epsilon > 0$ tel que le ϵ-voisinage $\partial^\epsilon\alpha$ de $\partial\alpha$ soit de mesure plus petite que δ. D'après le théorème ergodique ponctuel, pour μ presque tout x on a $\frac{1}{n}\sharp\{0 \leq k < n,\ T^k x \in \partial^\epsilon\alpha\} \to_n \mu(\partial^\epsilon\alpha) < \delta$. Soit $\epsilon_0 = \max_{A \in \alpha} d\left(A, X \setminus (A \cup \partial^\epsilon\alpha)\right)$. Alors $B(x,n,\epsilon_0)$ rencontre au plus $(\sharp\alpha)^{a_n}$ éléments de α^n avec $a_n = \sharp\{0 \leq k < n,\ T^k x \in \partial^\epsilon\alpha\}$. $\square$

5. Principe variationnel

Le principe variationel relie l'entropie des mesures et l'entropie topologique.

THÉORÈME 5.1.

$$h_{top}(T) = \sup_{\mu \in \mathcal{M}_{(e)}(X,T)} h(\mu).$$

DÉMONSTRATION. Montrons tout d'abord que $h(\mu) \leq h_{top}(T)$ pour toute mesure T-invariante. Par harmonicité de l'entropie il suffit de le montrer pour des mesures ergodiques μ. Soit $\lambda \in (0,1)$ et soit E un ensemble avec $\mu(E) > \lambda$. On considère une partition α avec $\mu(\partial\alpha) = 0$ et $h(\mu,\alpha) \simeq$

$h(\mu)$. Pour δ fixé, d'après le lemme 4.1 il existe $\epsilon > 0$ et $E' \subset E$ avec $\mu(E') > \lambda$ tel que $\limsup_n \frac{1}{n} \sup_{x \in E'} \log \sharp\{A^n \in \alpha^n,\ A^n \cap B(x, n, \epsilon) \neq \emptyset\} < \delta$. On considère un recouvrement C_n de X par des boules dynamiques de rayon $\epsilon/2$, i.e. $\bigcup_{x \in C_n} B(x, n, \epsilon/2) = X$. En particulier il existe $C'_n \subset E$ avec $\sharp C'_n \leq \sharp C_n$ tel que $E \subset \bigcup_{x \in C'_n} B(x, n, \epsilon)$. Par conséquent E' est couvert par au plus $e^{n\delta} \sharp C_n$ éléments de α^n pour n assez grand. En particulier $h(\mu) = n_\lambda^+(\mu) \leq h_{top}(T) + \delta$.

On montre désormais qu'il existe pour tout $\epsilon > 0$ une mesure T-invariante μ avec $h(\mu) \geq h'_{top}(T) - \epsilon$. Soit $\epsilon > 0$ tel que $h'_{top}(T, \epsilon) \simeq h_{top}(T)$ et α une partition de diamètre $< \epsilon$. On considère pour tout n un ensemble C_n de cardinal maximal parmi les ensembles E_n vérifiant $y \notin B(x, n, \epsilon)$ pour $x \neq y \in E_n$. On pose $\mu_n := \frac{1}{\sharp C_n} \sum_{x \in C_n} \delta_x$. Par définition il y au plus un élément de E_n dans chaque $A^n \in \alpha^n$. Il s'en suit que

$$H_{\mu_n}(\alpha^n) = \log \sharp E_n.$$

On fixe maintenant un entier positif $p < n$ on a par le lemme 2.2 du chapitre précédent pour tout $0 \leq j < p$ avec $n - j = ([n/p] - 1)p + i_j$

$$\begin{aligned} H_{\mu_n}(\alpha^n) &= H_{\mu_n}(\alpha^j) + H_{T_*^{([n/p]-1)p+j}\mu_n}(\alpha^{i_j}) + \sum_{q=0}^{[n/p]-2} H_{T_*^{qp+j}\mu_n}(\alpha^p), \\ &\leq 3p \log \sharp\alpha + \sum_{q=0}^{[n/p]-2} H_{T_*^{qp+j}\mu_n}(\alpha^p), \end{aligned}$$

puis par concavité on obtient en sommant pour tous les $0 \leq j < p$

$$\begin{aligned} \frac{1}{p} H_{\nu_n}(\alpha^p) &\geq \frac{1}{p} \sum_{0 \leq j < p} \sum_{q=0}^{[n/p]-2} \frac{1}{p([n/p]-1)} H_{T_*^{qp+j}\mu_n}(\alpha^p), \\ &\geq \frac{1}{n} H_{\mu_n}(\alpha^n) - \frac{3p}{n} \log \sharp\alpha, \end{aligned}$$

où on a noté $\nu_n := \frac{1}{p([n/p]-1)} \sum_{k=0}^{p([n/p]-1)-1} T_*^k \mu_n$. On montre comme dans le lemme 1.2 que toute limite faible ν de $(\nu_n)_n$ est invariante (ν_n ne l'était pas a priori). Soit $(n_k)_k$ une sous suite telle que $\lim_k \frac{\log E_{n_k}}{n_k} = h'_{top}(T, \epsilon)$. Soit ν une limite faible de $(\nu_{n_k})_k$. On peut choisir la partition α avec $\nu(\partial\alpha) = 0$ de sorte que $\mu \mapsto H_\mu(\alpha^p)$ est une fonction réelle de $\mathcal{M}(X)$ continue en ν.

On a donc en passant à la limite en n (à p fixé) :

$$\frac{H_\nu(\alpha^p)}{p} \geq \lim h'_{top}(T, \epsilon).$$

On conclut en passant à la limite en p que

$$h(\mu) \geq h_{top}(T', \epsilon).$$

□

En général pour un système topologique il n'existe pas de mesure d'entropie maximale, i.e. de mesure réalisant le supremum dans le principe variationel.

THÉORÈME 5.2. *On suppose que $T : X \to X$ soit un homéomorphisme expansif. Alors il existe une mesure d'entropie maximale.*

DÉMONSTRATION. Soit $\mu_n \in \mathcal{M}(X,T)$ tel que $h(\mu_n) > h_{top}(T) - \frac{1}{n}$. Soit μ une limite faible de $(\mu_n)_n$. On considère une partition α de diamètre inférieur à la constante de Lebesgue d'un générateur topologique avec $\mu(\partial\alpha) = 0$. Comme nous l'avons déja remarqué on a $h(\mu,\alpha) = h(\mu)$ pour tout μ. Pour conclure il suffit de vérifier que $\limsup_n h(\mu_n,\alpha) \le h(\mu,\alpha)$. Soit n tel que $h(\mu,\alpha) \simeq \frac{1}{n} H_\mu(\alpha^n)$. Puisque $\mu(\partial\alpha) = 0$ la fonction $\mu \mapsto H_\mu(\alpha^n)$ est continue donc pour ν proche de μ on a $h(\nu) \le \frac{H_\nu(\alpha^n)}{n} \simeq \frac{H_\mu(\alpha^n)}{n} \simeq h(\mu,\alpha) = h(\mu)$.

□

6. Applications et exemples

6.1. Inégalité de Ruelle en dimension 1.

THÉORÈME 6.1. *Pour tout application f continue C^1 par morceaux et pour tout $\mu \in \mathcal{M}([0,1],f)$, on a*

$$h(f,\mu) \le \int \log^+ |f'| d\mu.$$

En particulier

$$h_{top}(f) \le \log^+ \|f'\|_\infty.$$

DÉMONSTRATION. Soit α une partition de $[0,1]$ en intervalles de longueur ϵ. On a

$$h(\mu,\alpha) = \lim_n \frac{1}{n} \sum_{k=0}^{n-1} H_\mu(\alpha^{k+1}|\alpha^k).$$

Puis pour tout k on a

$$\begin{aligned}
H_\mu(\alpha^{k+1}|\alpha^k) &\le \sum_{B\in\alpha^k} \mu(B) \log \sharp\{A \in \alpha^{k+1},\ A \subset B\}, \\
&\le \sum_{B=(B_0,\dots,B_{k-1})\in\alpha^k} \mu(B) \log \sharp\{A \in \alpha,\ f(B_{k-1}) \cap A \ne \emptyset\}, \\
&\le \sum_{B=(B_0,\dots,B_{k-1})\in\alpha^k} \mu(B) \log \left(\max(\sup_{x\in B_{k-1}} |f'(x)|, 1) + 2 \right), \\
&\le \log 3 + \int \log^+ |f'(f^{k-1}x)|_\alpha d\mu(x) = \log 3 + \int \log^+ |f'(x)|_\alpha d\mu(x),
\end{aligned}$$

avec $|f'(f^{k-1}x)|_\alpha := \sup_{y\in\alpha(x)}|f'(y)|$, où $\alpha(x)$ désigne l'élément de α contenant x. Quand le diamètre de α tend vers 0 cette dernière intégrale vers $2+\int \log^+|f'(x)|_\alpha d\mu(x)$. On obtient donc à la limite

$$h(f,\mu) \leq \log 3 + \int \log^+|f'|d\mu.$$

En considérant f^p pour p grand on a

$$h(f,\mu) = \frac{h(f^p,\mu)}{p} \leq \frac{\log 3}{p} + \frac{1}{p}\int \log^+|(f^p)'|d\mu \simeq \int \log^+|f'|d\mu.$$

□

6.2. Entropie des applications de l'intervalle monotone par morceaux.

THÉORÈME 6.2. *Soit f une application continue par morceaux et soit α la partition en branches monotones de α. Alors*

$$h_{top}(f) = \lim_n \frac{1}{n}\log \sharp\alpha^n = \lim_n \frac{1}{n}\log \sharp\{n\text{-}branches\ monotones\}.$$

DÉMONSTRATION. Nous montrons tout d'abord que l'on peut toujours recouvrir une n-branche monotone par ϵ/n (ϵ,n) boules dynamiques. Ceci entrainera que $h_{top}(f) \leq \lim_n \frac{1}{n}\log\sharp\{n\text{-branches monotones}\}$. En effet si I_n est une n-monotone branche, i.e. I_n est un intervalle tel que $f^k|_{I_n}$ est monotone pour tout $1\leq k\leq n$ alors on vérifie que $\bigcup_{x\in E} B(x,n,\epsilon)\supset I_n$ avec $E := \bigcup_{0\leq k<n} f^{-k}F_\epsilon$ et $F_\epsilon = \{k\epsilon,\ k\in\mathbb{N}\}\cap[0,1]$. En effet on vérifie facilement que si $x\in I_n$ est compris entre deux éléments consécutifs y_0 et y_1 de E alors par monotonie x est dans la boule $B(y_i,n,\epsilon)$.

Il suffit maintenant pour conclure que $h_{top}(f)\geq \lim_n \frac{1}{n}\log\sharp\alpha^n$ car chaque élément de α^n est inclus dans une n-branche monotone. Soit $\mathcal{U}$ un recouvrement d' intervalle ouverts de $[0,1]$ tel que tout élément de $\mathcal{U}$ rencontre au plus trois éléments de α. On a facilement

$$h_{top}(f) \geq h_{top}(f,\mathcal{U}) \geq \lim_n \frac{1}{n}\log\sharp\alpha^n - \log 3.$$

En appliquant ceci non pas à (f,α) mais à (f^p,α^p) pour un large p, on obtient $h_{top}(f)\geq \lim_n \frac{1}{n}\log\sharp\alpha^n$. □

On s'intéresse maintenant au cas particulier des applications tentes. Pour tout $0\leq a\leq 2$ on pose T_a l'application de l'intervalle definie par

$$T_a(x) = ax \text{ pour } x\leq 1/2,$$

$$T_a(x) = a(1-x) \text{ pour } x\geq 1/2.$$

THÉORÈME 6.3. *Pour tout a, on a*

$$h_{top}(T,a) = \log^+ a$$

DÉMONSTRATION. Par l'inégalité de Ruelle il suffit de montrer que $h_{top}(T,a) \geq \log a$ pour $a > 1$. D'après la formule pour l'entropie des applications monotones par morceaux, l'entropie de T_a est donnée par le taux de croissance exponentielle en n des n-branches monotones. Mais l'image par T_a^{n-1} d'une telle branche est de longueur $\leq 1/2$. Donc toutes n-branche monotone est de longueur $\leq 1/2a^{n-1}$. En particulier le nombre de n-branches est $\geq 2a^{n-1}$. □

6.3. Sous-décalage à alphabet fini. Soit (Y,S) un sous-décalage à alphabet fini, alors

$$h_{top}(Y,S) = \lim_n \frac{\log \sharp Y_n}{n}$$

avec Y_n le nombre de mots à n lettres que l'on peut trouver dans Y.

THÉORÈME 6.4. *Soit (Y,σ_A) un sous décalage de type fini avec A pour matrice d'adjacence, alors*

$$h_{top}(\sigma_A) = \log \rho(A).$$

DÉMONSTRATION. On suppose pour simplifier (faites le cas général) que pour tout $i,j \in \{1,...,K\}$ il existe un entier $p_{i,j}$ tel que $(A^{p_{i,j}})_{i,j} > 0$. Pour calculer l'entropie de σ il suffit d'après la remarque ci-dessus d'estimer le cardinal de Y_n. Pour tout entier n il existe $i,j \in \{1,...,K\}$ tel que $\sharp Y_n(i,j) \geq \frac{\sharp Y_n}{K^2}$ où $Y_n(i,j)$ est l'ensemble des mot à n lettres commençant par i et finissant par j. Puisque $(A^{p_{j,i}})_{j,i} > 0$ on a aussi

$$\sharp Y_{n+p_{j,i}}(i,i) \geq \sharp Y_n(i,j) \geq \frac{\sharp Y_n}{K^2}$$

et donc

$$\log \rho(A) = \lim_n \frac{\log Tr(A^n)}{n} \geq \lim_n \sharp \frac{\log \sharp Y_n}{n} = h_{top}(\sigma).$$

Comme $Tr(A^n)$ est plus petit que le cardinal de Y_n, on a en fait une égalité. □

Soit $A \in M_d(0,1)$ une matrice aperiodique et $(u_1,...,u_d)$ et $(v_1,...,v_d)$ les vecteurs propres à gauche et à droite associés à la valeur propre maximale $\lambda > 0$.

DÉFINITION-PROPOSITION 6.5. *La mesure de Parry du sous shift de type fini de matrice d'adjacence A est la mesure Markovienne associé à la* *matrice stochastique P et à la proba stationnaire π défine par*

$$P_{ij} = \frac{A_{ij}}{\lambda v_j}$$

$$\pi_i = \frac{u_i v_i}{\sum_j u_j v_j}.$$

La mesure de Parry μ satisfait la propriété suivante (dite de Gibbs).

PROPOSITION 6.6. *Il existe $0 < a < b$ tels que pour tout n-cylindre C_n on a*

$$\frac{a}{\lambda^n} \leq \mu(C_n) \leq \frac{b}{\lambda^n}$$

THÉORÈME 6.7. *La mesure de Parry μ est l'unique mesure d'entropie maximale de (Y, σ_A).*

DÉMONSTRATION. On raisonne par l'absurde. Soit m est une autre mesure d'entropie maximale. On peut supposer m ergodique (pourquoi ?). Les mesures ergodiques m et μ sont orthogonales : il existe un Borel E tel que $\mu(E) = 0$ et $m(E) = 1$. On considère une union E_n de $[-n, n]$-cylindres tels que $\mu(E_n) \simeq 0$ et $m(E_n) \simeq 1$. On a alors avec P_0 la partition en 0-coordonnées

$$\begin{aligned}
h(m) &= \log \lambda, \\
&= \lim_n \frac{1}{2n+1} H_m(P_0^{[-n,n]}), \\
&\leq \frac{1}{2n+1} \Big(H_m(\{E_n, Y \setminus E_n\}) + H_m(P_0^{[-n,n]} | \{E_n, Y \setminus E_n\}) \Big).
\end{aligned}$$

Puis la mesure μ étant de Gibbs on a :

$$\begin{aligned}
(2n+1)\log\lambda &\lesssim m(E_n)\left[(2n+1)\log\lambda + \log\mu(E_n)\right] \\
&\quad + (1 - m(E_n))\left[(2n+1)\log\lambda + \log(1-\mu(E_n))\right], \\
0 &\leq m(E_n)\log\mu(E_n) + (1-m(E_n))\log(1-\mu(E_n)) \xrightarrow{n} -\infty.
\end{aligned}$$

□

7. Exercices

Exercice 16.
Montrer que tout homéomorphisme du cercle est d'entropie topologique nulle.

Exercice 17.
Soit $f : M \to M$ un système dynamique C^1 sur une variété compacte lisse M de dimension d. On note $R(f) = \lim_n \frac{1}{n} \log \sup_x \|D_x f^n\|$. Vérifiez que la

limite est bien définie et ne dépend pas de la structure Riemanienne $\|\|$ sur M. Montrez enfin que
$$h_{top}(T) \leq dR(T).$$

Exercice 18.
Soient $(X_n, T_n)_n$ des systèmes topologiques. On compactifie l'union disjointe des X_n's par un point $*$ (i.e. une base de voisinages $(V(x))_x$ de l'espace toplogique X ainsi obtenue est donnée par $V(x)$ étant les ouverts de X_n lorsque $x \in X_n$ et $V(*)$ étant les complémentaires des compacts de $\bigcup_n X_n$. On considère la dynamique T sur X coincidant avec T_n sur X_n et avec $T* = *$.

— Montrer que $h_{top}(T) = \sup_n h_{top}(T_n)$,
— On suppose que $(h_{top}(T)_n)_n$ est une suite bornée strictement croissante. Montrez que (X, T) n'a pas de mesure d'entropie maximale.

David BURGUET
LPMA - CNRS UMR 7599
Universite Paris 6
75252 Paris Cedex 05
FRANCE david.burguet@upmc.fr