MAT551 :

Introduction aux systèmes dynamiques et à la théorie ergodique.

Notes de cours

D. Burguet

Octobre 2021

Table des matières

Chapitre 1

Introduction

Sommaire

1. Qu'est ce qu'un système dynamique ?

Un système dynamique est la donnée d'un espace des phases X, d'un semi-groupe $(T, \cdot)$ et d'une loi d'évolution $f^t : X \to X$ avec $f^t \circ f^s = f^{t \cdot s}$ pour $t, s \in T$ (autrement dit, il s'agit d'une action de T sur X). Lorsque T est un groupe le système dynamique est dit inversible. Ainsi pour $T = \mathbb{R}^+$ la phase $f^t(x)$ décrit l'état du système à l'instant t partant de x en $t = 0$. Le système $(X, (f^t)_t)$ est alors appelé un flot. Lorsque $T = \mathbb{N}$ la phase $f^n(x)$ représente l'état du système après n itérations, on parle alors de dynamiques discrètes. On se bornera dans ce cours à l'étude des actions de T avec $T = \mathbb{N}$ ou $\mathbb{Z}$.

L'objet de la théorie des systèmes dynamiques consiste à décrire qualitativement le comportement asymptotique (dans notre cadre, quand $t \to +\infty$) des orbites $\mathcal{O}_f(x) := (f^t(x))_{t \in T}$ pour $x \in X$. On s'intéresse par exemple :

- aux propriétés statistiques des orbites : à quelle fréquence une orbite visite-t-elle une région donnée de l'espace des phases ?
- à la complexité du système : combien d'orbites peut-on distinguer avec une certaine précision ?
- à la stabilité structurelle et aux bifurcations : deux dynamiques discrètes f et g proches sont-elles comparables ?

L'étude diffère suivant la structure de (X, f). On distingue essentiellement les deux classes suivantes de dynamiques discrètes :

— *les systèmes dynamiques probabilistes* $(X, \mathcal{B}, f, \mu)$. L'espace des phases X est muni d'une tribu $\mathcal{B}$. L'application $f : X \to X$ est mesurable et laisse invariante une mesure de probabilité μ sur $\mathcal{B}$.

— *les systèmes dynamiques topologiques* (X, f). L'espace des phases X est un espace métrisable compact et l'application $f : X \to X$ est une fonction continue.

Parmi les systèmes dynamiques topologiques on étudie plus particulièrement :

— *les systèmes dynamiques différentiables* (X, f). L'espace des phases X est une variété C^k compacte et l'application $f : X \to X$ est une fonction de classe C^k avec $k \in \mathbb{N}^* \cup \{\infty\}$.

— *les systèmes dynamiques holomorphes* (X, f). L'espace des phases X est une surface de Riemann et l'application $f : X \to X$ est une fonction holomorphe.

2. Classification et invariant

Pour deux systèmes discrets (X, f) et (Y, g) un **morphisme** de (Y, g) vers (X, f) est une application $\pi : Y \to X$ satisfaisant $\pi \circ g = f \circ \pi$. Lorsque π est surjective on dit alors que (Y, g) est une **extension** de (X, f) ou encore que (X, f) est un **facteur** de (Y, g). L'application π est alors appelée une **semi-conjugaison**. Dans ce cas la dynamique de (X, f) est complètement déterminée par celle de (Y, g) : pour tout $x \in X$ et $y \in \pi^{-1}\{x\}$ on a $f^n x = \pi(f^n y)$. Si π est inversible on dit que π est un **isomorphisme** ou encore une **conjugaison**.

Suivant la classe étudiée on demande de plus que le morphisme préserve la structure associée aux deux systèmes dynamiques. Ainsi un morphisme $\pi : Y \to X$ entre deux systèmes probabilistes $(Y, \mathcal{C}, g, \nu)$ et $(X, \mathcal{B}, f, \mu)$ est aussi mesurable (relativement aux tribus $\mathcal{C}$ et $\mathcal{B}$) et envoie la mesure ν sur μ, i.e. $\mu(A) = \nu(\pi^{-1}A)$ pour tout $A \in \mathcal{B}$. De plus la relation de commutativité $\pi \circ g = f \circ \pi$ ainsi que l'inversibilité de π dans le cas d'un isomorphisme est requise seulement à des ensembles de mesures nulles près. On parle d'**extension, facteur et conjugaison mesurée**. Pour des systèmes topologiques on requiert la continuité du morphisme. On parle alors d'**extension, facteur et de conjugaison topologique**.

L'ultime objectif du dynamicien est de classifier les systèmes dynamiques à conjugaison près, c'est à dire de déterminer les classes de conjugaison [1], de savoir identifier la classe d'un système donné et d'avoir un

1. les classes d'équivalence pour la relation de conjugaison.

modèle explicite dans chaque classe. Il existe une telle richesse de systèmes dynamiques que cet objectif est rarement atteint même en considérant des classes de dynamiques plus restreintes. Un **invariant** est une propriété (ou un objet) mathématique associée à tout système dynamique d'une certaine classe, qui est invariante par conjugaison. Pour simplifier le problème de classification à conjugaison près on peut alors chercher à classifier les systèmes relativement à un invariant donné.

EXEMPLE 1.1. *Pour un système dynamique discret (X, f) un point $x \in X$ est dit **périodique** lorsque $f^k x = x$ pour un certain entier strictement positif k. Un système dynamique est dit apériodique s'il n'admet pas de point périodique. Le nombre de points périodiques de période k pour un entier k donné ainsi que la propriété d'apériodicité sont des invariants.*

3. Quelques exemples

3.1. Rotations et Homéomorphismes du cercle. Les homéomorphismes du cercle représentent une classe de dynamique aujourd'hui bien comprise. Parmi eux on distingue les rotations que l'on peut facilement classifier.

PROPOSITION 1.2. *Deux rotations du cercle $\mathbb{R}/\mathbb{Z}$ d'angle α et β sont topologiquement conjuguées si et seulement si $\alpha = \beta$ ou $-\beta$ $[mod\ 1]$.*

Nous verrons au chapitre 5 que tout difféomorphisme apériodique de classe C^2 du cercle, qui préserve l'orientation, est topologiquement conjugué à une rotation. L'angle de cette rotation est appelé le nombre de rotation qui constitue dans ce cadre un invariant "complet" (i.e. deux tels systèmes avec le même nombre de rotation sont topologiquement conjugués).

D'un point de vue mesuré, on verra que la mesure de Lebesgue est l'unique mesure invariante des rotations irrationnelles.

3.2. Décalage sur $\{0,1\}^{\mathbb{Z}}$ et Fer à cheval de Smale. Le fer à cheval est un modèle dynamique hyperbolique important décrit sur la Figure 1 ci-dessous.

On s'intéresse à la dynamique de f restreint au compact invariant $K = \bigcap_{n \in \mathbb{Z}} f^n C$ formé des points dont l'orbite est contenue dans C. Chaque point x de K est codé par une suite de 0 et de 1, le n^{me} terme de la suite valant i lorsque $f^i(x)$ appartient au sous-rectangle C_i de C. Ainsi le code de $f(x)$ correspond au code de x que l'on a décalé d'une valeur vers la gauche.

Muni de la topologie produit l'ensemble $\{0,1\}^{\mathbb{Z}}$ des suites bilatères à valeurs dans $\{0,1\}$ est compact et le décalage σ qui à une suite $(x_n)_n \in \{0,1\}^{\mathbb{Z}}$ associe la suite $\sigma((x_n)_n) = (x_{n+1})_n$ est continue.

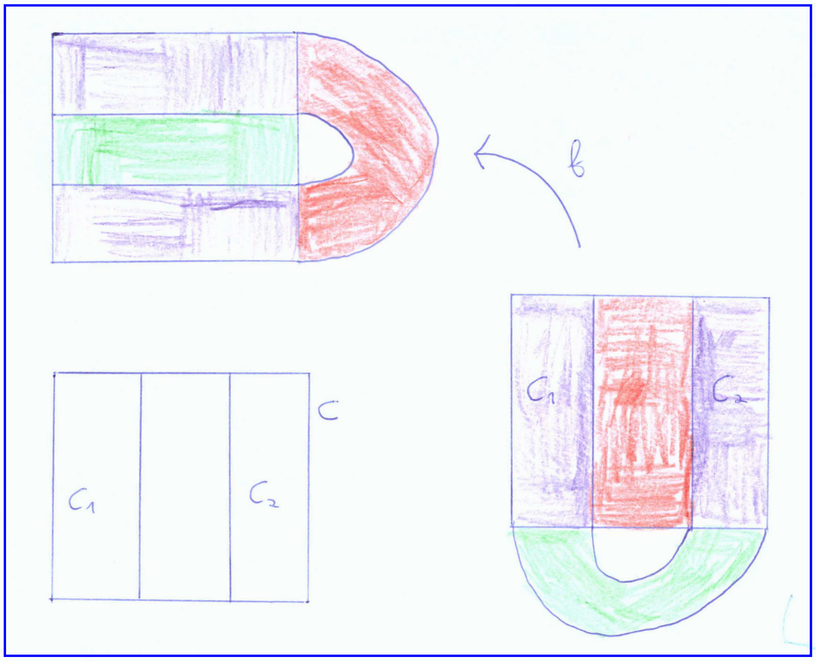


FIGURE 1: *Le fer à cheval de Smale.*

THÉORÈME 1.3. *Le système* (f, K) *est topologiquement conjugué à* $(\sigma, \{0,1\}^{\mathbb{Z}})$.

L'ensemble des mesures invariantes par $(\{0,1\}^{\mathbb{Z}}, \sigma)$ est extrêmement riche : tout système probabiliste *pas trop complexe* est isomorphe à $(\{0,1\}^{\mathbb{Z}}, \mathcal{B}, \sigma, \mu)$ pour une certaine mesure de probabilité σ-invariante μ.

3.3. Application de Gauss et développement en fractions continues. Notons $\mathcal{A}$ l'ensemble des suites $(a_n)_{n\in\mathbb{N}}$ finies ou dénombrables, avec $a_0 \in \mathbb{Z}$ et $a_k \in \mathbb{N}^*$ pour $k \geq 1$. Pour $(a_n)_n \in \mathcal{A}$ on définit par récurrence les fractions continues associées $([a_0, ..., a_n])_n$ par $[a_0] = a_0$ puis pour tout $n > 0$ par

$$[a_0, ..., a_n] = a_0 + \frac{1}{[a_1, ..., a_n]}$$

ce que l'on note sous la forme

$$[a_0, ..., a_n] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{... + \frac{1}{a_n}}}}.$$

PROPOSITION 1.4. *Pour tout $x \in \mathbb{R}$ il existe une unique suite $(a_n(x))_n \in \mathcal{A}$ telle que*[2]

$$x = \lim_n [a_0(x), ..., a_n(x)].$$

La suite $(a_n(x))_n$ est appelée le développement en fraction continue de x. Cette suite est finie si et seulement si x est rationnel.

Pour $x \in \mathbb{R}$ les fractions irréductibles $\frac{p_n}{q_n} := [a_0(x), a_1(x), ..., a_n(x)]$ sont les meilleures approximations rationnelles de x. On rappelle que la fraction irréductible $\frac{p}{q}$ est une meilleure approximation rationnelle de x si $|qx - p| = \|qx\| := \min_{r \in \mathbb{Z}} |qx - r| < \|q'x\|$ pour tout $q' < q$.

Le développement en fraction continue d'un réel $x \in [0, 1]$ est relié au système dynamique de l'intervalle $[0, 1]$, appelé application de Gauss, défini par $G(x) = \frac{1}{x} - [\frac{1}{x}]$ pour tout $x \in (0, 1]$ avec $G(0) = 0$. En effet pour x irrationnel on peut écrire pour tout $n > 0$ le nombre x sous la forme $x = [a_0(x), ..., a_n(x) + b_n(x)]$ avec $b_n(x) \in]0, 1[$. On vérifie alors facilement que

$$b_{n+1}(x) = G(b_n(x)) \text{ et } a_{n+1}(x) = \left[\frac{1}{b_n(x)}\right]$$

puis

$$a_n(x) = \left[\frac{1}{G^{n-1}(x - a_0(x))}\right] = a_1\left(G^{n-1}(x - a_0(x))\right).$$

PROPOSITION 1.5. *La mesure de probabilité de densité $\frac{1}{(1+x)\log 2}$ est G-invariante. Cette mesure est appelée la mesure de Gauss.*

Le système $([0, 1], G)$ admet plusieurs mesures de probabilité invariantes mais seule la mesure de Gauss est absolument continue par rapport à la mesure de Lebesgue.

2. Pour une suite finie $(u_n)_n = (u_0, u_1, ...u_k)$ on définit par convention la limite $\lim_n u_n$ de $(u_n)_n$ comme étant le dernier terme u_k de la suite.

Chapitre 2

Ergodicité et Mélange

Sommaire

1. Systèmes dynamiques

1.1. Actions sur les mesures. On considérera toujours dans ce chapitre un espace mesurable $(X, \mathcal{B})$, i.e. $\mathcal{B}$ est une tribu sur l'ensemble X. Lorsque $f : X \to X$ est une application mesurable, on dira que le triplet $(X, \mathcal{B}, f)$ est un système dynamique mesurable.

On note $\mathcal{M}(X)$ l'ensemble des mesures de probabilité 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 2.1. *Avec les notations précédentes, on dira qu'une mesure de probabilité sur $(X, \mathcal{B})$ est f-invariante lorsque $f^*\mu = \mu$. Le quadruplet $(X, \mathcal{B}, f, \mu)$ est alors appelé un système dynamique probabiliste.*

La *théorie ergodique* est l'étude des systèmes dynamiques probabilistes.

Remarque 2.2. *Pour montrer qu'une mesure est f-invariante, il est suffisant de vérifier que $f^*\mu(A) = \mu(A)$ pour tout A appartenant à un π-système générant la tribu $\mathcal{B}$.*

Exemples :

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

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

(3) Soit $A \in GL_n(\mathbb{Z})$ une matrice carré d'ordre n inversible à coefficients entiers de déterminant 1 ou -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 continue notée f_A.

Dans les trois exemples précédents (où on considère la tribu $\mathcal{B}$ des Boréliens sur X), la mesure de Lebesgue est invariante.

1.2. Actions sur L^p. Un système dynamique probabiliste induit une action naturelle sur les fonctions mesurables par composition : $\phi \mapsto \phi \circ f$. Pour $\mu \in \mathcal{M}(X)$ et $p \geq 1$, on note $L^p = L^p(\mu)$ l'espace vectoriel des fonctions mesurables $f : X \to \mathbb{C}$ avec $\int |f|^p \, d\mu < +\infty$ muni de la norme $\|f\|_p := \left(\int |f|^p \, d\mu\right)^{1/p}$.

Proposition 2.3. *Soit $(X, \mathcal{B}, f)$ un système dynamique mesurable et soit $\mu \in \mathcal{M}(X)$. Alors μ est f-invariante si et seulement si pour tout $\phi \in L^1(\mu)$, on a*

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

Démonstration. Lorsque ϕ est une fonction indicatrice $\phi = \chi_E$ on a

$$\int \phi \ d\mu = \mu(E) \text{ et } \mu(f^{-1}E) = \int \phi \circ f \ d\mu.$$

Ceci prouve le caractère suffisant.

Si μ est f-invariante on a par linéarité d'après les inégalités ci-dessus $\int \phi \circ f \ d\mu = \int \phi \ d\mu$ pour toute fonction étagée ϕ. Enfin tout fonction ϕ positive de $L^1(\mu)$ est la limite presque sûre d'une suite croissante de fonctions étagées. Le théorème de convergence monotone permet de conclure la proposition pour les fonctions positives intégrables puis pour toute fonction intégrable, de nouveau par linéarité. □

On en déduit que pour tout $p \in [1, +\infty[$ et tout système dynamique probabiliste $(X, \mathcal{B}, f, \mu)$ l' application $U_f : L^p(\mu) \to L^p(\mu)$ envoyant ϕ sur $\phi \circ f$ est une isométrie. L'opérateur U_f sur l'espace de Hilbert $L^2(\mu)$ est appelé l'opérateur de Koopman. Une branche importante de la théorie ergodique consiste à étudier les propriétés spectrales de cet opérateur.

2. Récurrence

On s'intéresse à la propriété de récurrence dans les systèmes probabilistes :

Etant donné un ensemble mesurable E, une orbite typique de E visite-t-elle l'ensemble E une infinité de fois ?

THÉORÈME 2.4 (Récurrence de Poincaré). *Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste et $E \in \mathcal{B}$ avec $\mu(E) > 0$. Alors*

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

En particulier, il existe un entier $n > 0$ tel que

$$\mu(E \cap f^{-n}E) > 0.$$

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

$$F := \left\{x \in E,\ \sharp\{k \in \mathbb{N},\ f^k x \in E\} < +\infty\right\},$$

on doit montrer que $\mu(F) = 0$. Mais $F = \bigcup_{n\in\mathbb{N}} f^{-n}E_0$... □

Le théorème suivant, dont la preuve beaucoup plus difficile ne sera pas abordée dans ce cours, permet de raffiner le résultat précédent : l'ensemble des temps de retour dans E est non seulement infini mais contient des suites arithmétiques de longueur arbitrairement grande.

THÉORÈME 2.5 (Récurrence multiple de Fürstenberg). *Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste. Pour tout $E \in \mathcal{B}$ avec $\mu(E) > 0$ on a pour tout entier $k \in \mathbb{N}$*

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

En particulier il existe un entier $n > 0$ tel que

$$\mu(E \cap f^{-n}E \cap ... \cap f^{-kn}E) > 0.$$

3. Ergodicité

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

DÉFINITION 2.6. *Un système probabiliste $(X, \mathcal{B}, f, \mu)$ est dit ergodique lorsque tout ensemble f-invariant $A \in \mathcal{B}$ (i.e. $f^{-1}A = A$), est de mesure nulle ou totale.*

La propriété d'ergodicité est un invariant : tout système probabiliste isomorphe à un système ergodique est lui-même ergodique.

Proposition 2.7. *Le système probabiliste* $(X, \mathcal{B}, f, \mu)$ *est ergodique si et seulement si 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. Le caractère suffisant suit directement de la définition en prenant $\phi = \chi_A$ avec $f^{-1}A = A \in \mathcal{B}$. Montrons maintenant l'implication réciproque. On commence par considérer $\phi = \chi_E$ une indicatrice d'un ensemble mesurable E. On pose $E_0 := \{x \in X, \ f^k x \in E \text{ pour une infinité de } k \in \mathbb{N}\}$. L'ensemble E_0 est mesurable 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 récurrence de Poincaré, on a $\mu(E \setminus E_0) = 0$. Puis

$$\begin{aligned} E_0 \setminus E \quad &\subset \quad \left(\bigcup_{k \in \mathbb{N}} f^{-k-1} E \right) \setminus E, \\ &\subset \quad \bigcup_{k \in \mathbb{N}} \left(f^{-k-1} E \setminus f^{-k} E \right) = \bigcup_{k \in \mathbb{N}} f^{-k} \left(f^{-1} E \setminus E \right). \end{aligned}$$

Or $\chi_E \circ f = \chi_{f^{-1}E} = \chi_E$ presque partout, c'est à dire qu'à un ensemble de mesure nulle près, un point appartient E si et seulement si il appartient à $f^{-1}(E)$, en particulier $\mu(f^{-1}E \setminus E) = 0$. Enfin par invariance de la mesure on a $\mu\left(f^{-k}\left(f^{-1}E \setminus E\right)\right) = \mu\left(f^{-1}E \setminus E\right) = 0$ pour tout entier $k \in \mathbb{N}$. Finalement, on a donc $\mu(E_0 \setminus E) = 0$ et $\chi_E = \chi_{E_0}$ presque partout. On suppose maintenant que ϕ est une fonction mesurable quelconque vérifiant $\phi \circ f = \phi$ presque partout. Pour tout réel t, l'ensemble $E_t = \{\phi > t\}$ vérifie $\chi_{E_t} \circ f = \chi_{E_t}$ presque partout. On déduit de la discussion précédente que $\mu(E_t) \in \{0, 1\}$, puis on conclut que $\phi = c$ presque partout avec $c = \inf\{t, \ \mu(E_t) = 0\}$. En effet c est fini (car $\mu(E_t) \xrightarrow{t \to -\infty} 1$) et pour tout $\epsilon > 0$ les ensembles $\{\phi > c - \epsilon\}$ et $\{\phi > c + \epsilon\}$ sont respectivement de mesure totale et nulle. $\square$

Corollaire 2.8. *Le système est ergodique si et seulement si l'espace propre de l'opérateur de Koopman associé à la valeur propre* 1 *est de dimension* 1 *(correspondant à l'ensemble des fonctions constantes presque partout).*

Exemple 2.9. *Dans les deux exemples suivants la mesure considérée est la mesure de Lebesgue.*

(i). la rotation d'angle α *sur le cercle est ergodique si et seulement si* α *est irrationnel.*

(ii). un automorphisme linéaire du tore f_A *est ergodique si et seulement si* A *n'a pas de valeurs propres racines de l'unité.*

Démonstration. (ii). Pour montrer l'ergodicité de la mesure de Lebesgue, notée Leb, dans ces deux exemples on peut utiliser le développement en série de Fourier afin d'établir qu'il n'existe pas de fonctions $L^2 = L^2(Leb)$ 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 irrationnelle. Sur le tore $\mathbb{R}^n/\mathbb{Z}^n$ les fonctions $(e_k)_{k\in\mathbb{Z}^n}$, $e_k(x) = e^{2i\pi<k,x>}$ définissent 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 $k \in \mathbb{Z}^n$,

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

et donc $e_k \circ f_A = e_{A^*k}$. 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 pour valeur propre, 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 $(A^*)^p l$, $p \in \mathbb{Z}$, 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 m^{eme} de l'unité. La matrice $(A^*)^m \in M_n(\mathbb{Q})$ admet un vecteur propre rationnel associé à la valeur propre associé à la valeur propre 1. En considérant un multiple il existe $l \in \mathbb{Z}^n$ et $m \in \mathbb{N}^*$ tel que $(A^*)^m l = l$. La fonction L^2 donnée par $\phi = \sum_{0\leq p<m} e_{(A^*)^p l}$ est alors f_A-invariante.

(i). Pour la rotation, au lieu des séries de Fourier on peut aussi utiliser "la distorsion nulle" des translations comme suit (voir aussi Exo 14 pour l'ergodicité de la mesure de Gauss). On raisonne par l'absurde en considérant un ensemble A invariant par f_α avec $Leb(A) \in]0,1[$. D'après le théorème de différentiation de Lebesgue (cf Rappels), presque tout point $x \in A$ est de densité, i.e. $\frac{Leb(]x-s,x+r[\cap A)}{Leb(]x-s,x+r[)} \xrightarrow{0<r,s\to 0} 1$. Soit $x_0 \in A$ un tel point de densité (par hypothèse $Leb(A) > 0$) et soit $r_0 > 0$ tel que pour tout $0 < r, s < r_0$ on ait $\frac{Leb(]x_0-s,x_0+r[\cap A)}{Leb(]x_0-s,x_0+r[)} > Leb(A)(< 1)$. Alors par densité de $\alpha\mathbb{Z}$ dans le cercle on peut recouvrir celui-ci, à un ensemble fini près, par une collection finie d'intervalles disjoints de la forme $f_\alpha^n\left(]x_0-s,x_0+r[\right)$ avec $n\in\mathbb{Z}$ et $0<r,s<r_0$ (faites-le !). Enfin, on a pour chacun de ces intervalles :

$$\begin{aligned}
Leb\left(f_\alpha^n\left(]x_0-s,x_0+r[\right)\cap A\right) &\overset{\text{invariance de } A}{=} Leb\left(f_\alpha^n\left(]x_0-s,x_0+r[\cap A\right)\right),\\
&\overset{\text{invariance de } Leb}{=} Leb(]x_0-s,x_0+r[\cap A),\\
&> Leb(A)Leb(]x_0-s,x_0+r[),\\
&\overset{\text{invariance de } Leb}{>} Leb(A)Leb\left(f_\alpha^n\left(]x_0-s,x_0+r[\right)\right).
\end{aligned}$$

On obtient une contradiction en sommant l'inégalité précédente sur tous les intervalles de la collection. □

4. Mélange

DÉFINITION 2.10. *Le système probabiliste $(X,\mathcal{B},f,\mu)$ est mélangeant si et seulement si pour tout $\phi,\psi\in L^2(\mu)$ on a*

$$\int(\phi\circ f^n)\times\psi\;d\mu\xrightarrow{n\to+\infty}\left(\int\phi\;d\mu\right)\left(\int\psi\;d\mu\right).\quad(*)$$

Autrement dit pour tout $\phi\in L^2$, la suite $(U_f^n\circ\phi)_n=(\phi\circ f^n)_n$ converge faiblement dans L^2 vers la fonction constante égale à $\int\phi\;d\mu$. En particulier 1 est l'unique valeur propre de l'opérateur de Koopman U_f et l'espace propre associé est réduit aux constantes. D'après le Corollaire 2.8 on a alors :

PROPOSITION 2.11. *Le mélange entraine l'ergodicité.*

Comme l'ergodicité, le mélange est invariant par isomorphisme. Remarquez aussi que pour montrer le mélange il suffit d'établir (*) pour tout ϕ,ψ dans une partie générant un sous-espace vectoriel dense de $L^2(\mu)$. En particulier en considérant les fonctions indicatrices, on observe que le système est mélangeant lorsque pour tout $A,B\in\mathcal{B}$, on a

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

L'ergodicité d'un système mélangeant se retrouve aussi facilement à partir de cette caractérisation : si μ est mélangeante et C invariant, $f^{-1}C=C$, alors en prenant $A=B=C$ dans la limite ci-dessus, on obtient $\mu(C)=\mu(C)^2$ et donc $\mu(C)\in\{0,1\}$.

EXEMPLE 2.12. *(i). la rotation sur le cercle n'est jamais mélangeante,*

(ii). les automorphismes linéaires du n-tore ergodiques sont mélangeants.

Montrons le point (ii). Il suffit de montrer la propriété (*) pour ϕ, ψ dans la base hilbertienne $\{e_k,\ k \in \mathbb{Z}^n\}$. Remarquons que U_{f_A} laisse invariant cet ensemble. Puis si $k, l \in \mathbb{Z}^n$ avec $k \neq 0$, il existe au plus un entier M tel que $e_k \circ f_A^M = e_l$ (sinon e_l serait invariant par une certaine puissance de f_A, ce qui est exclu puisque A et toutes ses puissances sont ergodiques). Donc pour $m > M$ on obtient $< e_k \circ f_A^m, e_l >= 0$, ce qui entraine bien le mélange (le cas $k = 0$ étant trivial).

5. Bernoulli

Soit S l'ensemble fini $S = \{1, ..., s\}$ muni de la topologie discrète. On note μ_p la mesure de Bernoulli sur S de paramètre $p = (p_1, ..., p_s)$ avec $p_i \geq 0$ et $\sum_i p_i = 1$. On considère le décalage σ sur le produit infini (compact) $S^{\mathbb{Z}}$, i.e. $\sigma((x_n)_n) = (x_{n+1})_n$. On note $\mu_p^{\mathbb{Z}}$ l'unique mesure de probabilité sur $S^{\mathbb{Z}}$, obtenue par les lois de consistance de Kolmogorov (cf Rappels Théorème 8.18), 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

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

La mesure $\mu_p^{\mathbb{Z}}$ est σ-invariante. En effet $\sigma^*\mu_p$ et μ_p coïncide sur le π-système donné par les unions finies de cylindres (qui engendrent la tribu $\mathcal{B}^{\mathbb{Z}}$ des boréliens sur $S^{\mathbb{Z}}$).

Proposition 2.13. *Tout schéma de Bernoulli $(S^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}}, \sigma, \mu_p^{\mathbb{Z}})$ est mélangeant.*

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

Remarque 2.14. *On peut aussi considérer le décalage unilatère sur $S^{\mathbb{N}}$ muni de $\mu_p^{\mathbb{N}}$. Le système probabiliste ainsi obtenu n'est plus inversible mais est toujours mélangeant.*

Le doublement de l'angle f sur le cercle muni de la mesure de Lebesgue est isomorphe à un schéma de Bernoulli de paramètre $(1/2, 1/2)$ sur $\{0, 1\}^{\mathbb{N}}$, via l'application qui à $x \in \mathbb{R}/\mathbb{Z}$ associe son développement dyadique (voir Proposition 3.14). En particulier, il est aussi mélangeant.

Remarque 2.15. *De façon plus générale on peut considérer le décalage* $(X^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}}, \mu^{\mathbb{Z}}, \sigma)$ *(encore appelé schéma de Bernoulli) pour n'importe quel espace probabilisé* $(X, \mathcal{B}, \mu)$*. Le mélange s'établit alors de la même façon.*

6. Exercices

Exercice 1. *Questions du cours.*
Etablir les points suivants vus en cours :

(1) le fer à cheval est topologiquement conjugué à $\{0,1\}^{\mathbb{Z}}$,
(2) la mesure de Gauss est invariante par la transformation de Gauss,
(3) la mesure de Lebesgue est invariante par les automorphismes linéaires du tore,
(4) les rotations irrationnelles du cercle ne sont pas mélangeantes (pour la mesure de Lebesgue).

Exercice 2.
Soit $(X, \mathcal{B}, f, \mu)$ un système ergodique et soit $\phi \in L^1(\mu)$ avec $\phi \circ f \geq \phi$. Montrer que ϕ est constante.

Exercice 3. *Mesure périodique.*
Soit $(X, \mathcal{B}, f)$ un système mesurable. Soit x un point n-périodique de f. Montrer que $\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k x}$ est une mesure invariante ergodique mais pas mélangeante.

Exercice 4.
Montrer que la rotation d'angle $\alpha = (\alpha_1, ..., \alpha_d)$ sur le tore $\mathbb{T}^d = \left(\frac{\mathbb{R}}{\mathbb{Z}}\right)^d$ est ergodique pour la mesure de Lebesgue si et seulement si les réels α_i, $i = 1, ..., d$, sont rationnellement indépendants.

Exercice 5.
Soit $\alpha \notin \mathbb{Q}$. Montrer que la mesure de Lebesgue est ergodique mais pas mélangeante pour la transformation du tore $(x, y) \mapsto (x + \alpha, x + y)$.

Exercice 6. *Spectre de l'opérateur de Koopman.*

(1) Montrer que le spectre et l'ensemble des valeurs propres des opérateurs de Koopman associés à deux systèmes probabilistes isomorphes sont égaux.
(2) On considère un système probabiliste ergodique inversible. Montrer que le spectre de l'opérateur de Koopmann est un sous groupe du cercle unité.

(3) On considère un système mélangeant. Déterminer les valeurs et les vecteurs propres de l'opérateur de Koopman.

Exercice 7. *Spectre purement ponctuel.*
Déterminer le spectre et les valeur propres de l'opérateur de Koopman pour les rotations du cercle. En déduire (indépendamment) les propriétés suivantes :

— Les systèmes probabilistes R_α et R_β donnés par deux rotations d'angle irrationnel sont isomorphes si et seulement si $\alpha = \beta$ ou $-\beta$ [1]. Qu'en est il pour le cas d'angle rationnel ?
— Les rotations ne sont pas mélangeantes.

Exercice 8. *Spectre de Lebesgue.* (1) Soit f_A un automorphisme linéaire ergodique du tore $\mathbb{T}^2$. Déterminer le spectre et les valeurs propres de l'opérateur de Koopman associés. On pourra considérer pour $\lambda \in \mathbb{C}$ avec $|\lambda| = 1$ la suite $(f_n)_n$ de $L^2(Leb)$ définie pour tout $n \in \mathbb{N}$ par $f_n := \frac{1}{2n+1} \sum_{k=-n}^{n} \lambda^{-k} U_A^k f$ avec $f = e_{1,0}$ et montrer que $U_A f_n - \lambda f_n \to 0$ quand n tends vers $+\infty$.

(2) De façon similaire montrer que le spectre des schémas de Bernoulli est aussi le cercle tout entier.

Exercice 9. *Autre définition du mélange.*
Montrer que $(X, \mathcal{B}, \mu, f)$ est mélangeant ssi $\mu(A \cap f^{-n}A) \xrightarrow{n} \mu(A)^2$ pour tout $A \in \mathcal{B}$.

Exercice 10. *Opérateur de transfert.*
On rappelle qu'une mesure signée sur un espace mesurable $(X, \mathcal{B})$ est la différence de deux mesures positives, dont au moins une est finie. Soient μ une mesure positive et ν une mesure signée sur $\mathcal{B}$, on dit que ν est absolument continue par rapport à μ si tout $A \in \mathcal{B}$ avec $\mu(A) = 0$ vérifie aussi $\nu(A) = 0$.

Soient $(X, \mathcal{B}, m)$ un espace probabilisé et $T : X \to X$ une application mesurable absolument continue par rapport à m, i.e. pour tout $A \in \mathcal{B}$ avec $m(A) = 0$ on a aussi $m(T^{-1}A) = 0$.

(1) Soit $f \in L^1(m)$. Vérifiez que la mesure signée finie $T^*(fdm)$, définie par $T^*(fdm)(A) = \int_{T^{-1}A} fdm$ pour tout $A \in \mathcal{B}$, est absolument continue par rapport à m. On notera $P_T f$ la dérivée de Radon-Nykodym associée à $T^*(fdm)$ relativement à m.

(2) Montrer que $P_T f$ est l'unique élément de $L^1(m)$ telle que pour tout $g \in L^\infty(m)$,

$$\int_X P_T f \cdot g \, dm = \int_X f \cdot (g \circ T) \, dm.$$

L'opérateur $P_T : L^1(m) \to L^1(m)$ est appelé l'opérateur de transfert associé à (T, m). Montrer que P_T est linéaire et positif (i.e. $P_T f \geq 0$ pour $f \geq 0$) puis que $|||P_T||| = 1$.

(3) Soit $f \geq 0$ avec $\int f dm = 1$. Montrer que $f dm$ est une mesure de probabilité T-invariante si et seulement si f est un point fixe de l'opérateur de transfert P_T.

(4) On considère une application $T : [0,1] \to [0,1]$ de classe C^2 avec un nombre fini de points critiques. Montrer que pour tout $f \in L^1([0,1])$, on a pour Lebesgue presque tout x,

$$P_T f(x) = \sum_{y,\ T(y)=x} \frac{f(y)}{|T'|(y)}.$$

(5) Montrer que $\frac{dx}{\pi\sqrt{x(1-x)}}$ est invariante par l'application logistique $T : [0,1] \to [0,1]$ donnée par $T(x) = 4x(1-x)$ pour tout $x \in [0,1]$.

Exercice 11. *Système induit et formule de Kac.*
Soit $(X, \mathcal{B}, \mu, f)$ un système probabiliste et $A \in \mathcal{B}$ avec $\mu(A) > 0$. D'après le théorème de récurrence de Poincaré 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. $\tau_A(x) = \inf\{n > 0,\ f^n(x) \in A\}$ et $f_A = f^{\tau_A}$.

— Montrez que la mesure de probabilité μ_A induite par μ sur A est f_A-invariante.
— Montrer que μ_A est f_A ergodique lorsque μ est f-ergodique.
— Montrez par récurrence que pour tout $\phi \in L^\infty(X)$ et pour tout entier N,

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

— En déduire pour μ ergodique, $\int \phi d\mu = \int_A \left(\sum_{k=0}^{\tau_A - 1} \phi \circ f^k \right) d\mu$, puis établir la formule de Kac $\int_A \tau_A \, d\mu_A = \frac{1}{\mu(A)}$.

Exercice 12. *Récurrence de Boschernitzan.*
On se propose de montrer que si $([0,1], f, \mathcal{B}_{[0,1]}, Leb)$ est un système probabiliste ergodique sur l'intervalle alors pour Leb-p.t. x, on a

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

— En raisonnant par l'absurde montrer qu'il existe $c > 1, s > 0$ et N tel que $E = \{x,\ n|f^n(x) - x| > c$ pour tout $n > N$ et $|f^n x - x| > 2s$ pour tout $n = 0, ..., N\}$ est de mesure positive,
— Pour un point de densité a de E on considère $E_r = E \cap B(a, r)$ pour $r < s$ petit. Montrez que $\tau_{E_r} \geq \frac{c}{2r}$,
— Obtenez une contradiction en utilisant la formule de Kac.

Exercice 13. *Lemme de Rohlin.*
Soit $(X, f, \mathcal{B}, \mu)$ un système ergodique inversible. On suppose que [1] $\inf_{A \in \mathcal{B},\ \mu(A) > 0} \mu(A) = 0$. 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 $\mu(\bigcup_{0 \leq k < n} f^k A') > 1 - \epsilon$.
— Vérifier que pour tout $C \in \mathcal{B}$ avec $\mu(C) > 0$, on a $\mu(\bigcup_{l \in \mathbb{N}} f^l C) = 1$,
— Soit $A \in \mathcal{B}$ avec $\mu(A) > 0$. On note τ_A le premier temps de retour dans A. En considérant pour tout k les ensembles $A_k = A \cap \{\tau_A = k\}$ pour tout $k \in \mathbb{N}^*$, montrer la formule de Kac pour un système ergodique inversible, i.e. $\int_A \tau_A \, d\mu = 1$.
— En déduire que pour tout entier K il existe $A \in \mathcal{B}$ avec $\mu(A) > 0$ et $\tau_A \geq K$.
— Montrer alors que les ensembles de la forme $B_k = f^{mn} A_k$ avec $0 \leq mn + n - 1 < k$ vérifie que $B_k, fB_k, ..., f^{n-1}B_k$ sont deux à deux disjoints,
— On considère l'union A' de tous ces ensembles B_k. Montrez que A', fA',..., $f^{n-1}A'$ sont deux à deux disjoints puis que

$$\mu\left(X \setminus (\bigcup_{0 \leq l < n} f^l A')\right) \leq \frac{n-1}{K},$$

— Conclure.

Exercice 14. *Ergodicité de la mesure de Gauss.*
1. Montrez que pour tout $x \neq 0$ on a
— $|G'(x)| \geq 1$,
— $|(G \circ G)'(x)| \geq 4$,

1. Cette propriété est par exemple vérifiée lorsque μ n'a pas d'atome et $\mathcal{B}$ est la tribu des boréliens associée à un espace métrique X : en effet on a alors $0 < \mu(B(x,r)) \xrightarrow{r \to 0} \mu(\{x\}) = 0$ pour tout x dans le support de μ.

— $\left|\left(\ln |G'| \circ G^{-1}\right)'(x)\right| \leq 2$.

2. Vérifiez que pour tout $x \notin \bigcup_{k\in\mathbb{N}} G^{-k}0$ et pour tout $n \in \mathbb{N}$ on a

$$\log |(G^n)'(x)| = \sum_{l=0}^{n-1} \log |G'(G^l(x))|.$$

3. On appelle branche monotone de G^n tout intervalle maximal sur lequel G^n est monotone. Montrez que $G^n : I_n \to (0,1)$ est un difféomorphisme pour toute branche monotone I_n de G^n. En déduire que la longueur d'une branche monotone de G^n est inférieure à $2^{-\frac{n}{2}+1}$.

4. Montrez qu'il existe une constante C telle que pour tout x, y dans la même branche monotone de G^n, on a

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

5. En déduire que pour tout A, B dans une même branche monotone de G^n, on a

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

6. Conclure l'ergodicité de μ en utilisant l'argument de point de densité déjà évoqué dans le cours pour les rotations du cercle.

Chapitre 3

Théorèmes ergodiques

Sommaire

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

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

Alors $L^2(\mathcal{I})$ coïncide avec l'ensemble des fonctions de $L^2(\mathcal{B})$ invariantes par f :

$$L^2(\mathcal{I}) := \{\phi \in L^2(\mathcal{B}),\ \phi \circ f = \phi\}.$$

En effet $\phi \in L^2(\mathcal{B})$ vérifie $\phi \circ f = f$ si et seulement si pour tout borélien C de $\mathbb{C}$ on a $f^{-1}(\{\phi \in C\}) = \{\phi \circ f \in C\} = \{\phi \in C\}$.

Pour $\phi \in L^2 = L^2(\mathcal{B})$ l'espérance conditionnelle (voir Rappels) 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

Pour tout $\phi : X \to \mathbb{C}$ dans $L^p(\mathcal{B})$ avec $p \geq 1$ on note $U_f(\phi)$ la fonction $\phi \circ f \in L^p(\mathcal{B})$. On rappelle que U_f définit une isométrie de $L^2 := L^2(\mathcal{B})$ dans lui-même. Nous montrons tout d'abord le théorème ergodique dans $L^2(\mathcal{B})$ par des méthodes hilbertiennes.

Théorème 3.1 (Théorème ergodique en moyenne dans L^2). *Avec les notations ci-dessus, on a pour tout $\phi \in L^2$,*

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

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

THÉORÈME 3.2. *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 3.3. *Soit H un espace de 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 :

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

Donc U^* est aussi une contraction linéaire. Puis en prenant $x = y$ dans l'inégalité (*), on a les équivalences suivantes

$$\begin{array}{cl} & x \text{ est dans le noyau de } Id - U, \\ \Leftrightarrow & \text{l'inégalité (*) est une égalité,} \\ \Leftrightarrow & x \text{ est aussi dans le noyau de } Id - U^*. \end{array}$$

□

On montre maintenant le théorème 3.2. 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 3.4. *Soit $(U_n)_{n\in\mathbb{N}}$ une suite de contractions linéaires de H, alors*

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

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

Ce lemme conclut la preuve du Théorème 3.2 avec $U_n = \frac{1}{n}\sum_{k=0}^{n-1} U^k$ pour tout entier $n > 0$ (pour $n = 0$ on pose $U_0 = 0$). En effet $Im(Id - U) \oplus Ker(Id - U)$ étant dense dans H et contenu dans G fermé, on obtient $G = H$. La suite $(U_n(x))_n$ converge donc en tout point $x \in H$ et la limite est continue. Puisqu'elle coïncide sur l'espace dense $Im(Id - U) \oplus Ker(Id - U)$ avec p, qui est continue sur H, elle est égale à p.

Démonstration du Lemme 3.4. 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 si et seulement si $\Delta(x) = 0$. 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. □

Corollaire 3.5. *Le système $(X, \mathcal{B}, f, \mu)$ est ergodique si et seulement si pour tout $\phi, \psi \in L^2$, on a*

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

Théorème 3.6 (Théorème ergodique en moyenne L^1). *Avec les notations ci-dessus, on a pour tout $\phi \in L^1$,*

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

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 3.4 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 coïncident sur tout L^1. □

2. Théorème ergodique ponctuel

Théorème 3.7 (Théorème ergodique ponctuel [1]). *Soit* $(X, \mathcal{B}, f, \mu)$ *un système dynamique probabiliste et* $\mathcal{I}$ *sa tribu des invariants. Alors pour tout* $\phi \in L^1(\mu)$,

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

Corollaire 3.8. *Avec les notations précédentes,*

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

Corollaire 3.9. *Le système* $(X, \mathcal{B}, f, \mu)$ *est ergodique si et seulement si pour tout* $\phi \in L^1(\mu)$, *on a*

$$\frac{1}{n}\sum_{k=0}^{n-1} \phi(f^k x) \xrightarrow{n} \int \phi \, d\mu \quad \textit{pour } \mu \textit{ p.t. } x.$$

Dans ce cas on a en particulier pour tout $A \in \mathcal{B}$

$$\frac{1}{n}\sharp\left\{0 \le k \le n-1,\ f^k x \in A\right\} \xrightarrow{n} \mu(A).$$

1. Dans l'énoncé ci-dessous nous utilisons la notation $\phi \circ f^k$. On préférera la notation $U_f^k(\phi)$ pour étudier la composition par f^k comme un opérateur de L^p dans lui-même.

L'outil fondamental pour montrer la convergence presque sûre d'une famille de fonctions est une inégalité dite maximale (par exemple, voir aussi le théorème de différentiation de Lebesgue ou le théorème de Doobs de convergence des martingales). On rappelle que pour tout $n > 0$ et tout $\phi \in L^1(\mu)$ on a $U_n(\phi) = \frac{1}{n}\sum_{l=0}^{n-1} \phi \circ f^l$ (par convention on pose $U_0 = 0$).

LEMME 3.10 (Inégalité Maximale). *Soit $\phi \in L^1$, alors on a*

$$\int_A \phi \, d\mu \geq 0 \quad \textit{avec } A = \{\sup_{k\in\mathbb{N}} U_k(\phi) > 0\}.$$

Pour $\alpha > 0$ on obtient en appliquant le lemme précédent à $|\phi| - \alpha$

$$\alpha\mu\left(\sup_{k\in\mathbb{N}} U_k(|\phi|) > \alpha\right) \quad \leq \quad \|\phi\|_1. \tag{2.1}$$

Clairement lorsque ϕ est un cobord borné (i.e. il existe $\psi \in L^\infty$ tel que $\phi = \psi \circ f - \psi$) ou $\phi \in L^1$ est une fonction invariante (i.e $\phi \circ f = \phi$) ou encore une combinaison linéaire de telles fonctions, les suites $(U_n(\phi))_n$ converge μ presque sûrement. On va montrer les deux points suivants :

— l'espace vectoriel engendré par les cobords bornés et les fonctions invariantes est dense dans $(L^1, \|\|_1)$,

— la convergence presque sûre de $(U_n(\phi))_n$ *passe à la limite* au moyen de l'inégalité maximale.

Une fois la convergence presque sûre établie, la limite presque sûre coïncide nécessairement avec la limite L^1 qui est l'espérance conditionnelle relativement aux invariants d'après le Théorème 3.6. En effet on peut extraire de tout suite convergente dans L^1 une sous-suite qui converge presque sûrement.

On commence par établir le premier point qui se déduit de la preuve du théorème ergodique en moyenne.

LEMME 3.11. *L'espace vectoriel F de L^1 engendré par $\{\phi \ : \ \exists\psi \in L^\infty,\ \phi = \psi \circ f - \psi\}$ et $\{\phi \in L^\infty \ : \ \phi \circ f = \phi\} = L^\infty(\mathcal{I})$ est dense dans L^1.*

DÉMONSTRATION. 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 $\{\phi \ : \ \exists\psi \in L^\infty,\ \phi = \psi\circ f - \psi\}$ est dense dans $\{\phi \ : \ \exists\psi \in L^2,\ \phi = \psi \circ f - \psi\}$. Par conséquent l'espace vectoriel engendré par $\{\phi \ : \ \exists\psi \in L^\infty,\ \phi = \psi \circ f - \psi\}$ et $\{\phi \in L^2 \ : \ \phi \circ f = \phi\}$ est 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 F est aussi dense dans L^1.

$\square$

On montre maintenant le deuxième point.

Lemme 3.12. *L'ensemble*

$$G = \{\phi \in L^1,\ (U_n(\phi))_n \text{ converge presque sûrement } \}$$

est un sous-espace vectoriel fermé $(L^1, \|\|_1)$.

Démonstration. Soit $(\phi_k)_k$ une suite de fonctions de L^1 convergeant vers ϕ dans L^1. On suppose que pour tout k la suite $(U_n(\phi_k))_n$ converge presque sûrement. Montrons alors que la suite $(U_n(\phi))_n$ converge presque sûrement. Pour $x \in X$ et une fonction $\psi : X \to \mathbb{R}$ on note $\Delta\psi(x)$ le reste de Cauchy en x de la suite $(U_n(\psi))_n$, i.e. $\Delta\psi(x) := \lim_N \sup_{q,q' \geq N} |U_q(\psi) - U_{q'}(\psi)|(x)$. On a pour tout x et tout k

$$\begin{aligned} \Delta\phi(x) &\leq \Delta(\phi - \phi_k)(x) + \Delta\phi_k(x), \\ &\leq \Delta(\phi - \phi_k)(x), \end{aligned}$$

puis pour tout $\alpha > 0$

$$\begin{aligned} \mu(\Delta\phi > \alpha) &\leq \mu(\Delta(\phi - \phi_k) > \alpha), \\ &\leq \mu\left(\sup_n U_n(|\phi - \phi_k|) > \alpha/2\right). \end{aligned}$$

En utilisant ensuite l'inégalité maximale (2.1) pour $\phi - \phi_k$ on obtient

$$\mu(\Delta\phi > \alpha) \leq \frac{2}{\alpha}\|\phi - \phi_k\|_1.$$

Comme ce dernier terme tends vers 0 quand k tends vers l'infini on en déduit que $\Delta\phi(x) = 0$ pour μ p.t. x. □

Il nous reste à montrer l'inégalité maximale.

Preuve de l'inégalité maximale. Pour tout entier $n \geq 0$ et tout $\phi \in L^1$, on notera :

$$S_n\phi = nU_n(\phi) \text{ et}$$
$$S_n^*\phi := \max_{0\leq k\leq n} S_k\phi.$$

Pour $n > 0$ on a $S_n\phi = \sum_{l=0}^{n-1} \phi\circ f^l$. Cette somme est appelée la somme de Birkhof de ϕ d'ordre n. Sur l'ensemble $A_n = \{\max_{0\leq k\leq n} U_k(\phi) > 0\} = \{S_n^*\phi > 0\}$ on a

$$\begin{aligned} S_n^*\phi \circ f + \phi &= \max_{1\leq k\leq n+1} S_k\phi, \\ &\geq \max_{1\leq k\leq n} S_k\phi, \\ &\geq \max_{0\leq k\leq n} S_k\phi = S_n^*\phi. \end{aligned}$$

En intégrant sur A_n, on obtient

$$\begin{aligned}\int_{A_n} (S_n^*\phi \circ f + \phi)\ d\mu &\geq \int_{A_n} S_n^*\phi\ d\mu,\\ \int S_n^*\phi \circ f\ d\mu + \int_{A_n} \phi\ d\mu &\geq \int S_n^*\phi\ d\mu.\end{aligned}$$

La mesure μ étant f-invariante on a $\int S_n^*\phi \circ f\ d\mu = \int S_n^*\phi\ d\mu$ et on obtient donc

$$\int_{A_n} \phi\ d\mu \geq 0.$$

Or on a $A = \bigcup_n^{\nearrow} A_n = \{\sup_{k\in\mathbb{N}} U_k(\phi) > 0\}$. Par convergence dominée on conclut donc :

$$\int_A \phi\ d\mu = \lim_n \int_{A_n} \phi\ d\mu \geq 0.$$

$\square$

Au lieu des sommes de Birkhoff on peut s'intéresser à la convergence de suites *sous-additives*. Le théorème ergodique ponctuel se généralise alors de la façon suivante.

Théorème 3.13 (Théorème ergodique sous-additif). *Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste et $(\phi_n)_n \in L^1(\mu)^{\mathbb{N}}$ une suite sous-additive, i.e. $\phi_{n+m}(x) \leq \phi_n(x) + \phi_m(f^n x)$ pour tout $m, n \in \mathbb{N}$ et presque tout $x \in X$. Alors il existe une fonction f-invariante $\overline{\phi} \in L^1(\mu)$ telle que*

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

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

3. Deux exemples d'applications

3.1. Nombres normaux. On note $\mathcal{B}$ l'ensemble des suites $(a_n)_{n\in\mathbb{N}}$ avec $a_0 \in \mathbb{Z}$ et $(a_n)_{n\in\mathbb{N}^*} \in \{0,1\}^{\mathbb{N}}$ qui ne sont pas constantes égales à 1 à partir d'un certain rang. Pour $(a_n)_{n\in\mathbb{N}} \in \mathcal{B}$ et pour $n \in \mathbb{N}$ on note $(a_0, ..., a_n)$ le dyadique $(a_0, ..., a_n) := \sum_{k=0}^n \frac{a_k}{2^k}$.

Proposition 3.14. *Pour tout $x \in \mathbb{R}$ il existe une unique suite $(a_n(x))_{n\in\mathbb{N}} \in \mathcal{B}$ avec*

$$x = \lim_n (a_0(x), ..., a_n(x)).$$

Cette suite est appelée le développement en base 2 de x. Le développement est nul à partir d'un certain rang si et seulement si x est dyadique.

Le développement en base 2 est relié au doublement de l'angle $f : [0,1] \circlearrowleft$ qui à $x \in [0,1]$ associe la partie fractionnaire de $2x$ notée $\{2x\}$. En effet pour tout $n > 0$ et tout $x \in \mathbb{R}$ on a $a_n(x) = [2f^{n-1}(x - a_0(x))] = a_1\left(f^{n-1}(x - a_0(x)\right)$.

Un nombre $x \in \mathbb{R}$ est dit normal si son développement en base 2 admet une fréquence $1/2$ de 0 et de 1, i.e.

$$\frac{1}{n}\sharp\{1 \leq k \leq n,\ a_k(x) = 0\} \xrightarrow{n\to+\infty} \frac{1}{2}.$$

La mesure de Lebesgue étant f-invariante et ergodique par $x \mapsto \{2x\}$, il suit du théorème ergodique appliqué à $\phi = \chi_{\{a_1=k\}}$ avec $k \in \{0,1\}$:

Corollaire 3.15. *Lebesgue presque tout point $x \in \mathbb{R}$ est normal.*

Le caractère normal de π n'est pas connu...

3.2. Fréquence dans le développement en fraction continue. La mesure de Gauss μ_G est ergodique. La densité de cette mesure étant strictement positive sur $[0,1]$ tout ensemble de $[0,1]$ est de mesure de Lebesgue totale si et seulement si il est de mesure totale pour la mesure de Gauss. En appliquant le théorème ergodique ponctuel au système ergodique $([0,1], \mathcal{B}, G, \mu_G)$ et à la fonction $\phi_k \in L^1(\mu_G)$ donnée par l'indicatrice $\chi_{\{a_1=k\}}$ pour $k \in \mathbb{N}^*$ on obtient :

Corollaire 3.16. *Pour Lebesgue presque tout $x \in \mathbb{R}$, on a pour tout entier $k \geq 1$*

$$\frac{1}{n}\sharp\{1 \leq l \leq n,\ a_l(x) = k\} \xrightarrow{n} \frac{\log\left(1 + \frac{1}{k(k+2)}\right)}{\log 2}.$$

En particulier Lebesgue presque tout point a un développement en fraction continue non borné.

4. Exercices

Exercice 15.
Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste. Soit θ un réel et p_θ la projection orthogonale de $L^2(\mu)$ sur le sous espace $\{\phi \in L^2(\mu) : \phi \circ f = e^{-i\theta}\phi\}$. Montrer que pour tout ϕ dans $L^2(\mu)$:

$$\frac{1}{n}\sum_{k=0}^{n-1} e^{ik\theta}\phi \circ f^k$$

converge dans $L^2(\mu)$ vers $p_\theta(\phi)$.

Exercice 16. *Loi forte des grands nombre.*
Déduire la loi forte des grand nombres du théorème ergodique ponctuel.

Exercice 17. *Théorème ergodique en moyenne L^p.*
1. Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste. Montrer que pour tout $\phi \in L^p(\mu)$ avec $p \in [1, +\infty[$,

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

2. On considère le système dynamique probabiliste (ergodique) donné par le doublement de l'angle sur le cercle muni de la mesure de Lebesgue. Soit ϕ une fonction réelle continue sur le cercle avec $\phi(0) \neq \frac{1}{2}(\phi(1/3) + \phi(2/3))$. Montrer que $\left(\frac{1}{n}\sum_{k=0}^{n-1} U_f^k \phi\right)_n$ ne converge pas dans $L^\infty(Leb)$.

Exercice 18. *Constante de Khintchine.*
Montrer qu'il existe une constante $K > 0$ tel que pour Lebesgue presque tout $x \in [0, 1)$ on a

$$\left(\prod_{0<i\leq n} a_i(x)\right)^{\frac{1}{n}} \xrightarrow{n} K$$

avec $(a_k(x))_k$ le développement en fraction continue de x. Cette constante, encore bien mystérieuse, est appelé la constante de Khintchine.

Exercice 19. *Récurrence de Khintchine.*
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 probabiliste, on se propose de montrer que pour 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

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

— En déduire que pour n assez grand on pour tout entier l

$$\mu(A)^2 < \epsilon + \frac{1}{n^2}\sum_{0\leq i,j\leq n-1} \mu(f^{-i}A \cap f^{-j-l}A),$$

— Conclure.

Exercice 20. *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 bien définie,
— Montrez que pour tout A et toute mesure de proba T-invariante ergodique 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 21. *Divergence des sommes de Birkhof.*
Soit $(X, \mathcal{A}, \mu, f)$ un système dynamique probabiliste ergodique et soit $\phi \in L^1$. On suppose que μ p.t. $x \in X$ on a $\sum_{n \in \mathbb{N}} \phi(f^n x) = +\infty$. Montrer qu'il existe un entier p tel que $A_p := \{x,\ \forall n \geq p\ \sum_{l=0}^n \phi(f^l x) \geq 1\}$ est de mesure positive. En déduire que
$$\int \phi \, d\mu > 0.$$
On pourra montrer que $\int \phi \, d\mu \geq \frac{\mu(A_p)}{p}$.

Exercice 22. *Exposants de Lyapunov en dimension* 1.
Soit $f : [0,1] \to [0,1]$ une fonction C^1 et μ une mesure f-invariante ergodique. Montrer qu'il existe $\chi(\mu) \in \mathbb{R} \cup -\infty$ tel que

$$\mu \text{ } p.p. \text{ } x,\ \frac{1}{n} \log |(f^n)'|(x) \text{ converge vers } \chi(\mu).$$

Chapitre 4

Minimalité et unique ergodicité

Sommaire

1. Mesures invariantes d'un système topologique

On considère dans tout ce chapitre un système dynamique topologique (X, f), i.e. X est un espace métrisable 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 topologie 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 $(\phi_n)_{n\in\mathbb{N}}$ est une partie dense dénombrable de $\mathcal{C}(X)$ alors la métrique d suivante induit la topologie faible-* :

$$\forall \mu, \nu \in \mathcal{M}(X), \; d(\mu, \nu) := \sum_{n\in\mathbb{N}} \frac{|\int \phi_n \, d\mu - \int \phi_n \, d\nu|}{2^{n+2}(1 + \|\phi_n\|_\infty)}.$$

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

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

1.1. Compacité. On s'intéresse aux propriétés topologiques de l'ensemble des mesures f-invariantes de $\mathcal{M}(X)$. Par définition de la topologie faible-*, l'application $f^* : \mathcal{M}(X) \to \mathcal{M}(X)$ est continue. L'ensemble

$\mathcal{M}(X,f) := \{\mu \in \mathcal{M}(X),\ \mu \text{ est } f\text{-invariante}\}$ est donc fermé comme ensemble des points fixes d'une application continue. De plus il n'est jamais vide :

LEMME 4.1. *L'ensemble $\mathcal{M}(X,f)$ est un compact non vide de $\mathcal{M}(X)$.*

DÉMONSTRATION. Pour $x \in X$ on note δ_x la mesure de Dirac en x. On fixe x et on note $\mu_n := \frac{1}{n}\sum_{k=0}^{n-1}\delta_{f^k x}$. On vérifie facilement,

$$\begin{aligned} f^*\mu_n &= \frac{1}{n}\sum_{k=1}^{n}\delta_{f^k x}, \\ &= \mu_n + \frac{1}{n}\left(\delta_{f^n x} - \delta_x\right), \end{aligned}$$

et donc

$$d(f^*\mu_n, \mu_n) \leq \frac{1}{n}.$$

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 $\mu \in \mathcal{M}(X,f)$.

$\square$

En fait le même argument permet de montrer l'énoncé suivant :

LEMME 4.2. *Avec les notations précédentes et $U_n^* := \frac{1}{n}\sum_{k=0}^{n-1}(U_f^*)^k$, la suite de compacts $(U_n^*(\mathcal{M}(X)))_n$ converge pour la topologie de Hausdorf vers $\mathcal{M}(X,f)$.*

DÉMONSTRATION. Les ensembles $O_n := \{\mu \in \mathcal{M}(X),\ d(f^*\mu,\mu) \leq \frac{1}{n}\}$, $n \in \mathbb{N}^*$, définissent une suite de compacts emboités satisfaisant $\bigcap_n O_n = \mathcal{M}(X,f)$. En particulier, cette suite $(O_n)_n$ converge pour la topologie de Hausdorf vers $\mathcal{M}(X,f)$.

Mais on a comme dans la preuve précédente $d(f^*\mu_n,\mu_n) \leq \frac{1}{n}$ pour toute mesure $\mu_n \in U_n^*(\mathcal{M}(X))$ et donc $\mathcal{M}(X,f) \subset U_n^*(\mathcal{M}(X)) \subset O_n$ pour tout n. La suite de compacts $(U_n^*(\mathcal{M}(X)))_n$ converge donc aussi vers $\mathcal{M}(X,f)$.

$\square$

On notera par la suite $\mathcal{M}_e(X,f)$ le sous ensemble de $\mathcal{M}(X,f)$ formé par les mesures ergodiques. Pour $\mu \in \mathcal{M}(X,f)$ et $x \in X$ on dit que x est générique pour μ lorsque

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

LEMME 4.3. *On suppose $\mu \in \mathcal{M}_e(X,f)$. Alors μ p.t. x est générique pour μ.*

DÉMONSTRATION. Soit $(\phi_p)_{p\in\mathbb{N}}$ une partie dense dénombrable de $\mathcal{C}(X)$. Il suffit de montrer qu'il existe un ensemble de mesure totale E tel que pour tout $x \in E$ et tout p, la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1}\phi_p(f^k x)\right)_n$ converge vers $\int \phi_p\, d\mu$ quand n tends vers l'infini. Mais on sait d'après le Théorème ergodique ponctuel que pour chaque p il existe un ensemble E_p de mesure totale tel que tout $x \in E_p$ satisfait cette propriété de convergence pour ϕ_p. Il suffit donc de prendre $x \in E := \bigcap_{p\in\mathbb{N}} E_p$. □

1.2. Structure convexe. L'ensemble $\mathcal{M}(X,f)$ est clairement un sous-ensemble convexe de $\mathcal{M}(X)$. On rappelle que 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[1]. Le théorème de Krein-Milman affirme que dans un espace localement convexe séparé (ici l'espace vectoriel des mesures Boréliennes signées) tout convexe compact est égale à l'adhérence de l'enveloppe convexe de ses points extrémaux :

$$C = \overline{conv(ex(C))}.$$

Les mesures ergodiques sont les points extrémaux du convexe $\mathcal{M}(X,f)$:

LEMME 4.4.

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

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 ponctuel appliqué à μ et μ_i pour $i = 1,2$, on a la convergence μ_i 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 $i = 1,2$ et tout $E \in \mathcal{B}$, mais $\mu_1 \neq \mu_2$.

1. Les segments ouverts sont les ensembles de la forme $\{t\mu + (1-t)\nu,\ 0 < t < 1\}$ pour $\mu \neq \nu \in C$

2. Puisque $\mu_i \ll \mu$, il existe d'après le théorème de Radon-Nikodyn $\phi_i \in L^1(\mu)$ tel que $\mu_i = \phi_i d\mu$. Nous affirmons que 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.

Montrons maintenant notre affirmation. Il suffit de vérifier que $E = E_t = \{\phi_i < t\}$ vérifie pour tout $t \in \mathbb{R}$ [2] $\mu(E \triangle f^{-1}E) = 0$ (pourquoi ?). Par invariance de μ on a $\mu(E \setminus f^{-1}E) = \mu(f^{-1}E \setminus E)$ puis

$$\begin{aligned} \mu_i(E) &= \int_{E\cap f^{-1}E} \phi_i \, d\mu + \int_{E\setminus f^{-1}E} \phi_i \, d\mu, \\ \mu_i(f^{-1}E) &= \int_{E\cap f^{-1}E} \phi_i \, d\mu + \int_{f^{-1}E\setminus E} \phi_i \, d\mu. \end{aligned}$$

Par invariance de μ_i on obtient donc

$$\int_{f^{-1}E\setminus E} \phi_i \, d\mu = \int_{E\setminus f^{-1}E} \phi_i \, d\mu.$$

Par définition de E, l'intégrant de droite est strictement plus petit l'intégrant de gauche. Puisque les ensembles sur lesquels on intègre sont de même mesure, cette mesure est nécessairement nulle et donc $\mu(E \triangle f^{-1}E) = 0$.

□

LEMME 4.5. *Soient μ et ν deux mesures ergodiques distinctes de $\mathcal{M}(X, T)$. 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 $(\mu(E), \nu(E)) \in \{0, 1\}^2 \setminus \{(0, 0), (1, 1)\}$.

□

Lorsque M est une mesure de probabilité borélienne sur $\mathcal{M}(X, f)$ 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 $\Phi : \mathcal{M}(X, T) \to \mathbb{R}$

$$\int \Phi(\nu) \, dM(\nu) = \Phi(bar(M)).$$

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

THÉORÈME 4.6. *Pour toute mesure $\mu \in \mathcal{M}(X, f)$ il existe une unique mesure de probabilité borélienne M_μ sur $\mathcal{M}(X, f)$ (appelée la décomposition ergodique de μ) supportée sur les mesures ergodiques telle que*

2. $\triangle$ désigne la différence symétrique : $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

$bar(M_\mu) = \mu$, en particulier pour tout $\phi \in \mathcal{C}(X)$:

$$\int \phi \, d\mu = \int \left(\int \phi \, d\nu \right) \, dM_\mu(\nu).$$

2. Récurrence, Transitivité, Minimalité

On se concentre ici sur des propriétés dynamiques des systèmes topologiques. Soit (X, f) un tel système.

Un point $x \in X$ est dit non errant lorsque pour tout voisinage U de x il existe un entier $n > 0$ tel que $f^n U \cap U \neq \emptyset$. L'ensemble $\Omega(f)$ est fermé et f-invariant.

Un point $x \in X$ est dit récurrent lorsqu'il existe une suite strictement croissante d'entiers $(n_k)_k$ telle que $f^{n_k} x \xrightarrow{k} x$. Les points périodiques sont récurrents. Les points récurrents sont non errants.

Proposition 4.7. *Soit $\mu \in \mathcal{M}(X, f)$. Alors μ-presque tout point est récurrent. En particulier tout système topologique (X, f) admet un point récurrent et donc $\Omega(f)$ est non vide. De plus toute mesure invariante est supportée par $\Omega(f)$.*

Démonstration. Soit $\mathcal{V} = (V_n)_{n \in \mathbb{N}}$ 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 Poincaré, $\mu(F_n) = 0$. Alors on vérifie facilement que tout $x \notin \bigcup_{n \in \mathbb{N}} F_n$ est récurrent. □

Un système (X, f) est dit *transitif* lorsqu'il existe $x \in X$ avec $X = \overline{\{f^n x, \ n \geq N\}}$ pour tout $N \in \mathbb{N}$. Dans le cas d'un système inversible, i.e. f est un homéomorphisme de X, cela équivaut à la densité de l'orbite de x, i.e. $X = \overline{\{f^n x, \ n \in \mathbb{N}\}}$.

Proposition 4.8. *Les conditions suivantes sont équivalentes :*

(1) (X, f) est transitif,

(2) pour tous U, V ouverts non vides de X il existe un entier $n > 0$ avec $U \cap f^{-n} V \neq \emptyset$.

Démonstration. Soit $x \in X$ avec $X = \overline{\{f^n x, \ n \geq N\}}$ pour tout $N \in \mathbb{N}$. Pour tout U, V ouverts de X il existe $q > p$ avec $f^p(x) \in U$ et $f^q(x) \in V$. Alors $f^p(x) \in U \cap f^{p-q} V$.

Montrons maintenant la réciproque. Soit $(U_l)_{l \in \mathbb{N}}$ une base dénombrable de voisinages ouverts de X. Par récurrence il existe une suite strictement croissante d'entiers $(k_n)_n$ tels que les ouverts $V_n = \bigcap_{1 \leq l \leq n} f^{-k_l} U_l$ sont non vides pour tout n. En particulier tout point x dans $\bigcap_n \overline{V_n}$ vérifie $X = \overline{\{f^n x, \ n \geq N\}}$ pour tout $N \in \mathbb{N}$. □

Un système (X, f) est dit *minimal* s'il n'admet pas de sous-système autre que lui-même.

PROPOSITION 4.9. *Les conditions suivantes sont équivalentes :*

(1) (X, f) *est minimal,*

(2) toute orbite est dense (en particulier tout système minimal est transitif, tout point X *est récurrent et* $\Omega(f) = X$*).*

(3) pour tout ouvert non vide U *de* X *il existe un entier* N *tel que*

$$\bigcup_{0 \leq n \leq N} f^{-n} U = X$$

DÉMONSTRATION. Pour tout $x \in X$ le compact

$$K_x := \overline{\{f^n x,\ n \in \mathbb{N}\}}$$

donné par l'adhérence de l'orbite de x est stable par f, i.e. il satisfait $f(K_x) \subset K_x$. Ceci entraine facilement l'équivalence des deux premiers points. Montrons que (1) entraîne (3). Le complémentaire de $\bigcup_{n\in\mathbb{N}} f^{-n}U$ est un compact stable par f. Par minimalité on a $\bigcup_{n\in\mathbb{N}} f^{-n}U = X$. On conclut (3) par compacité de X. Réciproquement si $K \neq X, \emptyset$ est un compact stable par f alors son complémentaire U satisfait $f^{-n}U \subset U$ pour tout entier positif n. En particulier $\bigcup_{n\in\mathbb{N}} f^{-n}U \neq X$, ce qui contredit (3). $\square$

PROPOSITION 4.10. *Tout système dynamique topologique* (X, f) *admet un sous-système minimal.*

Cela fournit une autre preuve de l'existence de points récurrents.

DÉMONSTRATION. Soit $(U_l)_{l\in\mathbb{N}}$ une base dénombrable d'ouverts de X. Pour tout k on note $V_k = \bigcup_{l\in\mathbb{N}} f^{-l}U_k$. D'après le critère (3) de la proposition précédente si (X, f) n'est pas minimal, l'ensemble $X_1 = X \setminus V_{k_0}$ est un compact non vide stable par f pour un certain k_0. On choisit alors k_0 minimal parmi les entiers vérifiant cette propriété. Puis par récurrence si X_p n'est pas minimal, alors on définit X_{p+1} comme $X_{p+1} = X_p \setminus V_{k_p} \neq \emptyset$ pour un k_p minimal. On vérifie facilement que $X_\infty = \bigcap_{p\geq 1} X_p$ est un sous système minimal à l'aide du critère (3). $\square$

3. Unique ergodicité

Le système topologique (X, f) est dit *uniquement ergodique* lorsqu'il existe une unique mesure invariante (ergodique). Un système uniquement ergodique n'est pas nécessairement minimal. Par exemple toute application contractante d'un compact dans lui-même est uniquement ergodique (l'unique mesure invariante étant la mesure de Dirac au point fixe) mais elle n'est pas minimale lorsque le compact n'est pas réduit à un singleton. De façon plus surprenante il existe des exemples (plus difficiles à construire) de dynamiques minimales mais pas uniquement ergodiques.

Pour les systèmes uniquement ergodiques on a le théorème ergodique *uniforme* suivant. Comme dans le chapitre précédent on notera pour tout entier $n > 0$:

$$U_n := \frac{1}{n}\sum_{k=0}^{n-1}(U_f)^k.$$

THÉORÈME 4.11. *Le système (X, f) est uniquement ergodique si et seulement si pour tout $\phi \in \mathcal{C}(X)$ la suite de fonctions $(U_n(\phi))_n$ converge (uniformément) vers une fonction constante égale à $\int \phi\, d\mu$ avec μ l'unique mesure invariante. Dans ce cas, tout point de X est en particulier générique pour μ.*

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

$$\int \phi\, d(U_n^*\delta_x) = \frac{1}{n}\sum_{k=0}^{n-1}\phi(f^k x) \xrightarrow{n} \int \phi\, d\mu \text{ uniformément en } x.$$

□

PROPOSITION 4.12. *Pour α irrationnel, $(\mathbb{R}/\mathbb{Z}, f_\alpha)$ est uniquement ergodique.*

Une preuve classique consiste à montrer la convergence uniforme des sommes de Birkhof associées aux monômes trigonométriques puis à conclure en utilisant la densité des polynômes trigonométriques dans $\mathcal{C}(\mathbb{S}^1)$. 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 $\phi \in \mathcal{C}(\mathbb{S}^1)$, l'application $\gamma \mapsto \int \phi\, d(f_\gamma^*\mu) = \int \phi(x+\gamma)\, d\mu(x)$ est continue par le théorème de continuité des intégrales à paramètre. □

On dit qu'une suite $(x_k)_k \in \mathbb{R}^{\mathbb{N}}$ est équidistribuée modulo 1 si et seulement si pour tout $\phi \in C(\mathbb{R}/\mathbb{Z})$ on a avec $\tilde{\phi} : \mathbb{R} \to \mathbb{C}$ la fonction continue $\mathbb{Z}$-périodique associée à ϕ :

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

PROPOSITION 4.13. *Si P est un polynôme de degré supérieur ou égal à 1 avec un coefficient dominant irrationnel, alors $(P(n))_n$ est équidistribuée modulo 1.*

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, on montre que $Leb \times Leb$ est une mesure ergodique pour le produit semi-direct, $S(x, y) = (x+\alpha, x+y)$ sur le tore $\mathbb{R}/\mathbb{Z}\times\mathbb{R}/\mathbb{Z}$ (voir Exercice 5). On utilise ensuite le lemme suivant :

LEMME 4.14. *Soit (X, f) un système uniquement ergodique et soit $\psi : X \to \mathbb{R}/\mathbb{Z}$ une fonction continue. On suppose que $\mu \times Leb$ est ergodique pour le produit semi-direct $f_\psi(x, y) = (f(x), y + \psi(x))$ avec μ l'unique mesure de probabilité f-invariante. Alors $(X \times \mathbb{R}/\mathbb{Z}, f_\psi)$ est uniquement ergodique.*

DÉMONSTRATION. Soit ν_0 une mesure de probabilité ergodique f_ψ-invariante. On note $(g_t)_{t\in\mathbb{R}}$ le flot de $X\times\mathbb{R}/\mathbb{Z}$ défini par $g_t(x, y) = (x, y+t)$. Il commute avec f_ψ si bien que $\nu_t = g_t^*\nu_0$ est f_ψ-invariante et ergodique. La mesure $\nu := \int \nu_t \; dLeb(t)$ est donc elle aussi invariante. On vérifie maintenant que ν coïncide avec la mesure produit $\mu \times Leb$. Pour toute fonction continue $\phi \in C(X \times \mathbb{R}/\mathbb{Z})$ on a

$$\begin{aligned}
\int \phi(x,y)\, d\nu(x,y) &= \int \left(\int \phi(x,y)\, d\nu_t(x,y)\right) dLeb(t),\\
&= \int \left(\int \phi\circ g_t(x,y)\, d\nu_0(x,y)\right)\, dLeb(t),\\
&= \int\int \phi(x,y+t)\, d\nu_0(x,y) dLeb(t),\\
&\overset{\text{Fubini}}{=} \int \left(\int \phi(x,y+t)\, dLeb(t)\right) d\nu_0(x,y).
\end{aligned}$$

Puis par invariance de Leb par translation, on obtient

$$\int \phi(x,y)\, d\nu(x,y) = \int \left(\int \phi(x,z)\, dLeb(z)\right)\, d\nu_0(x,y).$$

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

$$\begin{aligned}\int \phi(x,y)\, d\nu(x,y) &= \int \left(\int \phi(x,z)\, dLeb(z)\right) d\pi^*\nu_0(x,y),\\ &= \int \left(\int \phi(x,z)\, dLeb(z)\right) d\mu(x),\\ &= \int \phi(x,z)\, d\mu(x)\, dLeb(z).\end{aligned}$$

On en conclut que $\nu = \mu \times Leb$. Cette mesure étant supposée ergodique, a donc $\nu_t = \mu \times Leb$ pour Lebesgue presque tout t par unicité de la décomposition ergodique de ν[3]. En effet celle-ci est donnée d'une part par $\Theta^* Leb_{[0,1]}$ avec $\Theta : [0,1] \to \mathcal{M}_e(X, f)$ défini par $\Theta(t) = \nu_t$ pour $t \in [0,1]$ et d'autre part par la mesure de Dirac en $\mu \times Leb$. Pour Lebesgue presque tout t on a $g_{-t}^*\nu_t = \nu_0 = g_{-t}^*(\mu \times Leb) = \mu \times Leb$. □

On calcule facilement par récurrence : $S^n(0,0) = (n\alpha, n(n-1)\alpha/2)$. D'après le lemme précédent S est uniquement ergodique. En particulier le point $(0,0)$ est générique pour $\mu \times Leb$ et on a pour toute fonction réelle ϕ continue sur le cercle :

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}(1\times\phi)(S^n(0,0)) &= \frac{1}{n}\sum_{k=0}^{n-1}\phi\left(n(n-1)\alpha/2\right),\\ &\overset{n}{\to} \int \phi\; dLeb.\end{aligned}$$

4. Exercices

Exercice 23.
On considère deux systèmes dynamiques topologiques (X, T) et (X, S) définis sur le même espace métrique compact et satisfaisant $S \circ T = T \circ S$. Montrer qu'il existe une mesure T- et S-invariante à la fois. A-t-on $\mathcal{M}(X, T) = \mathcal{M}(Y, S)$?

Exercice 24. *Unicité de la mesure G-invariante absolument continue.* Montrer que la mesure de Gauss μ_G est l'unique mesure invariante ergodique par l'application de Gauss, qui est absolument continue par rapport à la mesure de Lebesgue.

3. On peut aussi remarquer que $\frac{1}{Leb(I)}\int_I \nu_t\; dLeb(t)$ est égale à $\mu \times Leb$ pour tout intervalle non trivial I car $\mu \times Leb$ est un point extrémal de $\mathcal{M}(X \times \mathbb{R}/\mathbb{Z}, f_\psi)$. Puis $\frac{1}{Leb(I)}\int_I \nu_t\; dLeb(t) \xrightarrow{I\to s} \nu_s$.

Exercice 25. (1) Soit (X, f) un système topologique transitif. Soit $\phi \in \mathcal{C}(X)$ avec $\phi \circ f = \phi$. Montrer que ϕ est constante.

(2) Obtenir la même conclusion sous l'hypothèse (X, f) uniquement ergodique.

(3) Montrer que si (X, f) est minimal alors les fonctions semi-continues supérieurement $\phi : X \to \mathbb{R}$ vérifiant $\phi \circ f = \phi$ sont les fonctions constantes. Montrer que la réciproque est vraie lorsque (X, f) est inversible, mais fausse en général pour des systèmes non inversibles.

(4) Donner des contre-exemples à la minimalité et à l'unique ergodicité si l'on se restreint aux fonctions continues, i.e. si l'on suppose seulement que toute fonction $\phi \in \mathcal{C}(X)$ vérifiant $\phi \circ f = \phi$ est constante.

Exercice 26.
Soit (X, T) une dynamique topologique. Montrer que $\{\phi \in \mathcal{C}(X),\ \forall \mu \in M(X,T)\ \int \phi \, d\mu = 0\}$ coïncide avec la fermeture des cobords continus, i.e.

$$\left\{\phi \in \mathcal{C}(X),\ \forall \mu \in M(X,T)\ \int \phi \, d\mu = 0\right\} = \overline{\{\psi \circ T - \psi,\ \psi \in \mathcal{C}(X)\}}.$$

En déduire que si $\phi \in \mathcal{C}(X)$ a ses sommes de Birkhof bornées en tout point alors ϕ est une limite de cobords.

Indications : On pourra raisonner par l'absurde et considérer une forme linéaire continue de $\mathcal{C}(X)$ non identiquement nulle sur $\{\phi \in \mathcal{C}(X),\ \forall \mu \in M(X,T)\ \int \phi \, d\mu = 0\}$, mais qui s'annule sur les cobords continus.

Exercice 27.
Soit $\alpha \notin \mathbb{Q}$. On considère la rotation R_α d'angle α sur le cercle $\mathbb{R}/2\pi\mathbb{Z}$. Montrer qu'il existe une fonction continue tel que $\int \phi \ dLeb = 0$ mais qui n'est pas un cobord continu, i.e. il n'existe pas de fonction continue $\psi : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C}$ telle que $\phi = \psi \circ R_\alpha - \psi$.

Indications : On rappelle que les meilleures approximations rationnelles de α (données par les fractions continues $\frac{p_n}{q_n}$ de α) vérifient $\min_{r \in \mathbb{Z}} |q_n \alpha - r| \leq \frac{1}{q_n}$. On pourra alors considérer une fonction ϕ donnée par un développement de Fourier de la forme $\sum_{n \in I \subset \mathbb{N}} a_n e^{i q_n x}$.

Exercice 28. *Gottschalk-Hedlund.*
Soit (X, T) une dynamique minimale et soit $\phi : X \to \mathbb{R}$ une fonction continue tels que les sommes de Birkhof $(\sum_{k=0}^{n-1} \phi(T^k x_0))_n$ soient bornées pour un certain $x_0 \in X$.

(1) On s'intéresse tout d'abord à la dynamique F de $X \times \mathbb{R}$ envoyant (x,t) sur $(Tx, t+\phi(x))$. En considérant l'ensemble des valeurs d'adhérence de $\{F^n(x_0,0),\ n \in \mathbb{N}\}$ montrer qu'il existe une partie minimale compacte N de F.

(2) Notons $\pi : X \times \mathbb{R} \to X$ la projection sur X. Montrez que $\pi(N) = X$.

(3) Soit T_λ la translation verticale de λ sur $X \times \mathbb{R}$, i.e. $T_\lambda(x,t) = (x, t+\lambda)$. Montrer que $T_\lambda(N) \cap N \neq \emptyset$ ssi $\lambda = 0$.

(4) Conclure que ϕ est un cobord continu.

Exercice 29. *Récurrence multiple topologique.*
On se propose de montrer le théorème de récurrence multiple topologique :

Théorème 4.15. *Soit (X,f) un système topologique. Pour tout entier $l \geq 1$ il existe $x \in X$ et une suite strictement croissante d'entiers $(n_k)_k$ tel que $f^{in_k}(x) \xrightarrow{k} x$ pour tout $i \leq l$.*

(1) Montrer que l'on peut supposer (X,f) minimal. Etablir le cas $l=1$.

(2) Soit un entier $l > 1$. On considère sur X^l les dynamiques $h(x_1, ..., x_l) = (f(x_1), .., f(x_l))$ et $g(x_1, ..., x_l) = (f(x_1), .., f^l(x_l))$. Vérifiez que h et g commutent et que h restreint à Δ est minimale.

(3) On suppose le théorème de récurrence multiple topologique vrai pour $l-1$. Montrer qu'il existe $z_0 \in \Delta$ et $z_0^k \in \Delta$ et une suite n_k tel que $g^{n_k}(z_0^k) \xrightarrow{k} z_0$.

(4) En utilisant la minimalité de h, montrer que pour tout $\epsilon > 0$ et tout $y \in \Delta$ il existe $y' \in \Delta$ et un entier strictement positif n tels que
$$d(g^n y', y) < \epsilon.$$

(5) Montrer que l'infimum de G défini sur la diagonale $\Delta := \{(x, ..., x) \in X^l,\ x \in X\}$ par $G(y) = \inf_n d(g^n y, y)$ est nul.

(6) On suppose que G ne s'annule pas. Montrer qu'il existe $a > 0$ tel que $\{G \geq a\}$ soit un fermé d'intérieur non vide. En déduire qu'il existe un entier positif m tel que $X = \bigcup_{l=0}^{m} h^{-l}\{G \geq a\}$. Obtenez une contradiction à l'aide de la question précédente.

(7) Conclure.

(8) Donner une autre preuve en utilisant le principe de récurrence multiple de Fürstenberg.

Exercice 30. *Szemeredi.*
On se propose de montrer le théorème suivant en utilisant la récurrence multiple de Fürstenberg.

THÉORÈME 4.16. *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.*

(1) 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 $\mathbb{A}$. Soit $(n_j)_j$ une suite strictement croissante telle que $\lim_j \frac{\sharp \mathbb{A}\cap[-n_j,n_j]}{2n_j+1} = \overline{d}(\mathbb{A})$. Montrer que toute limite faible μ de $(\mu_j)_j := \left(\frac{1}{2n_j+1}\sum_{|l|\le n_j} \delta_{\sigma^l(x^{\mathbb{A}})}\right)_j$ vérifie $\mu(E) = \overline{d}(\mathbb{A}) > 0$ avec $E = [1] := \{(x_n)_n,\ x_0 = 1\}$.
(2) En appliquant le principe de récurrence multiple de Fürstenberg à $(\{0,1\}^{\mathbb{Z}}, \sigma, \mu)$ donner une preuve du théorème de Szemeredi.

Exercice 31.
Montrer qu'un facteur topologique $\pi : (Y,S) \to (X,T)$ induit une surjection continue $\pi^* : \mathcal{M}(Y,S) \to \mathcal{M}(X,T)$ entre les ensembles de mesures invariantes.

Exercice 32. *Supremum des exposants de Lyapunov.*
Soit $f : [0,1] \to [0,1]$ une fonction C^1 montrer que

$$\lim_n \frac{\log^+ \|(f^n)'\|_\infty}{n} = \max\left(\sup_{\mu\in\mathcal{M}([0,1],f)} \chi(\mu), 0\right).$$

Exercice 33. *Minimalité et unique ergodicité.* (1) Montrer que si un système dynamique topologique possède une mesure de probabilité invariante et ergodique de support total (i.e., tout ouvert non-vide est de mesure non-nulle), alors le système est topologiquement transitif.
(2) Montrer que si un système dynamique est uniquement ergodique avec une mesure de probabilité invariante de support total alors le système est minimal.

Exercice 34. *Isométrie.*
Montrer qu'un système dynamique topologique défini par une isométrie bijective est uniquement ergodique ssi il est topologiquement minimal, puis ssi il est topologiquement transitif.

Exercice 35.
Soient α et β deux réels rationnellement indépendants, i.e. il n'existe pas de rationnels u, v tels que $u\alpha + v\beta = 0$. Montrer que $R_\alpha \times R_\beta$ est uniquement ergodique et minimal sur $\mathbb{S}^1 \times \mathbb{S}^1$.

Exercice 36. *Unique ergodicité de l'odomètre.*
Soit un entier $p > 1$. L'odomètre de base p est le sous-ensemble de $\Sigma_p := \prod_{k=1}^{\infty} \frac{\mathbb{Z}}{p^k\mathbb{Z}}$ formé des suites $(x_k)_k$ telles que $x_k = x_{k+1}$ $[mod\ p^k]$ pour tout k. On considère l'addition par 1, noté τ, sur Σ_p : on ajoute 1 à toutes les coordonnées (autrement dit $(\tau(x))_k = x_k + 1$ pour tout k). Montrer que (Σ_p, τ) est minimal et uniquement ergodique.

Chapitre 5

Homéomorphismes du cercle

Sommaire

1. Relèvement, degré, orientation,...

L'application $\pi : x \mapsto x\ [mod\ 1]$ de $\mathbb{R}$ dans $\mathbb{R}/\mathbb{Z}$ étant un revêtement et $\mathbb{R}$ étant simplement connexe, il existe pour toute application continue $f : \mathbb{R}/\mathbb{Z} \circlearrowleft$ et pour tout $a \in \mathbb{R}$ avec $[a] = f(\overline{0})$, une unique application continue $F : \mathbb{R} \circlearrowleft$ avec $F(0) = a$ relevant f, i.e. telle que $\pi \circ F = f \circ \pi$ (cf Rappels Théorème 8.19).

Définition 5.1. *Le degré de f est l'entier $F(x+1) - F(x)$ (qui ne dépend ni de x, ni du choix du relèvement F).*

Lemme 5.2. *L'application $\phi_F := F - Id_{\mathbb{R}}$ est $\mathbb{Z}$-périodique pour tout relèvement F de f.*

Démonstration. Cela vient directement de la définition du degré et de $deg(F) = 1$:

$$\begin{aligned} \phi_F(x+1) - \phi_F(x) &= F(x+1) - F(x) - (x+1) + x, \\ &= deg(F) - 1 = 0. \end{aligned}$$

□

On vérifie facilement la loi de composition $deg(f \circ g) = deg(f) \times deg(g)$. En particulier si f est un homéomorphisme $deg(f) = 1$ ou -1. Pour tout $a, b \in \mathbb{R}/\mathbb{Z}$ on note $[a, b]$ l'arc reliant a et b en parcourant le cercle positivement.

DÉFINITION 5.3. *Soit f un homéomorphisme local du cercle. Les conditions suivantes sont équivalentes :*

(1) l'homéomorphisme f préserve l'ordre d'un triplet de trois points distincts, i.e.

$$\exists a, b, c \text{ avec } b \in]a, c[\text{ et } f(b) \in]f(a), f(c)[;$$

(2) l'homéomorphisme f préserve l'ordre de tous les triplets, i.e.

$$\forall a, b, c \text{ avec } b \in [a, c], \ f(b) \in [f(a), f(c)],$$

autrement dit $f([a, c]) = [f(a), f(c)]$;

(3) $deg(f) > 0$;

(4) les relèvements F de f sont strictement croissants.

Lorsqu'une de ces conditions est satisfaite, on dit que f préserve l'orientation.

2. Nombre de rotation

Dans la suite de ce chapitre on supposera toujours que f est un homéomorphisme du cercle préservant l'orientation. Un relevé de f sera noté F.

LEMME 5.4.

$$\max_{\mathbb{R}} \phi_F - \min_{\mathbb{R}} \phi_F < 1$$

DÉMONSTRATION. Par $\mathbb{Z}$-périodicité et continuité de $\phi_F = F - Id$, il existe x_- et x_+ tels que

$$x_- \leq x_+ < x_- + 1$$

et

$$\begin{aligned} \phi_F(x_-) &= \min_{\mathbb{R}} \phi_F, \\ \phi_F(x_+) &= \max_{\mathbb{R}} \phi_F \end{aligned}$$

Or on a $F(x_+) \in F([x_-, x_- + 1[) = [F(x_-), F(x_- + 1)[= [F(x_-), F(x_-) + 1[$ et donc

$$F(x_+) - F(x_-) < 1.$$

Puisque $x_- - x_+ \leq 0$, on obtient en sommant les deux inégalités précédentes

$$\phi_F(x_+) - \phi_F(x_-) = \max_{\mathbb{R}} \phi_F - \min_{\mathbb{R}} \phi_F < 1.$$

□

LEMME 5.5. *La suite de fonctions $(\frac{F^k - Id}{k})_k$ converge uniformément sur $\mathbb{R}$ vers une constante $\rho(F)$. La classe modulo 1 de $\rho(F)$ est indépendante du choix du relèvement. Elle est appelée le nombre de rotation de f et est notée $\rho(f)$.*

DÉMONSTRATION. On a

$$\phi_{F^k} = F^k - Id = \sum_{l=0}^{k-1} \phi_F \circ F^l.$$

La fonction ϕ_F étant $\mathbb{Z}$-périodique, elle induit une (unique) application continue $\overline{\phi_F} : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ satisfaisant $\overline{\phi_F} \circ \pi = \phi_F$. Puisque $\pi \circ F^l = f^l \circ \pi$, on a donc $\phi_F \circ F^l = \overline{\phi_F} \circ \pi \circ F^l = \overline{\phi_F} \circ f^l \circ \pi$. On peut donc écrire

$$\overline{\phi_{F^k}} = \sum_{l=0}^{k-1} \overline{\phi_F} \circ f^l.$$

Pour toute mesure $\mu \in \mathcal{M}(\mathbb{R}/\mathbb{Z}, f)$, on obtient

$$\int \overline{\phi_{F^k}}\ d\mu = k \int \overline{\phi_F}\ d\mu.$$

En appliquant le lemme précédent à F^k (qui est un relèvement de l'homéomorphisme du cercle f^k qui préserve aussi l'orientation), on conclut que pour tout $x \in \mathbb{R}$

$$\begin{aligned} \left|\phi_{F^k}(x) - \int \overline{\phi_{F^k}}\ d\mu\right| &\leq \int |\overline{\phi_{F^k}}(\pi(x)) - \overline{\phi_{F^k}}(y)|\ d\mu(y), \\ \left|\phi_{F^k}(x) - k\int \overline{\phi_F}\ d\mu\right| &\leq \max \phi_{F^k} - \min \phi_{F^k} < 1. \end{aligned}$$

Ceci entraîne le résultat souhaité avec $\rho(F) = \int \overline{\phi_F}\ d\mu$. Si G est un autre relèvement de f on a $G = F + l$ pour un certain entier l et $G^k = F^k + lk$ pour tout k. Il s'en suit que $\rho(G) = \rho(F) + l$. $\square$

Pour la rotation R_α d'angle $\alpha \in \mathbb{R}/\mathbb{Z}$ la fonction réelle $x \mapsto x + \alpha$ fournit un relèvement F pour lequel on a $F^k(x) = x + k\alpha$ pour tout $x \in \mathbb{R}/\mathbb{Z}$ et tout $k \in \mathbb{Z}$. On a donc $\rho(R_\alpha) = \alpha$.

Remarquez aussi que pour tout entier k, on a $\rho(f^k) = k\rho(f)$.

PROPOSITION 5.6. *Soient f, g deux homéomorphismes du cercle préservant l'orientation et soit h une application du cercle (pas nécessairement un homéomorphisme) de degré 1 tels que $h \circ f = g \circ h$. Alors $\rho(f) = \rho(g)$.*

DÉMONSTRATION. Soient F et H des relèvements de f et h respectivement. Alors $H \circ F$ est un relèvement de $h \circ f = g \circ h$. On choisit un relèvement G de g tel que $G(H(0)) = H \circ F(0)$. On a alors $H \circ F = G \circ H$

car ces deux fonctions relèvent $h \circ f = g \circ h$ et coïncident en 0. Puis on obtient par récurrence $H \circ F^k = G^k \circ H$ pour tout k. De plus on a

$$\begin{aligned} H \circ F^k &= F^k + \phi_H \circ F^k, \\ &= G^k \circ H. \end{aligned}$$

Puis

$$\begin{aligned} \frac{F^k - Id}{k} + \frac{\phi_H \circ F^k}{k} &= \frac{G^k \circ H - Id}{k}, \\ &= \frac{G^k \circ H - H}{k} + \frac{\phi_H}{k}. \end{aligned}$$

La fonction ϕ_H étant bornée, on obtient $\rho(f) = \rho(g)$ en passant à la limite quand k tends vers l'infini.

□

Par conséquent le nombre de rotation est invariant par les conjugaisons topologiques qui préservent l'orientation (lorsque la conjugaison ne préserve pas l'orientation les nombres de rotation sont opposés).

3. Cas $\rho(f) \in \mathbb{Q}$

Proposition 5.7. *Soit f un homéomorphisme du cercle préservant l'orientation. Alors $\rho(f) \in \mathbb{Q}$ si et seulement si f a un point périodique.*

Démonstration. Soit x un point du cercle avec $f^q(x) = x$. Alors si $X \in \pi^{-1}(\{x\})$ on a $F^q(X) = X + p$ pour un entier p, puis $F^{kq}(X) = X + kp$ pour tout entier $k > 0$. En particulier $\rho(F) = \frac{p}{q}$. Réciproquement, supposons $\rho(f) \in \mathbb{Q}$. Il existe donc un relèvement F de f avec $\rho(F) = \frac{p}{q}$. On a vu que pour tout $\mu \in \mathcal{M}(\mathbb{R}/\mathbb{Z}, f)$

$$\int \overline{\phi_F} \, d\mu = \rho(F) = p/q.$$

Or on a $\min \phi_{F^q} \leq \int \overline{\phi_{F^q}} \, d\mu = p \leq \max \phi_{F^q}$ et donc d'après le théorème des valeurs intermédiaires il existe $X_q \in \mathbb{R}$ tel que $x_q = \pi(X_q)$ satisfait $\overline{\phi_{F^q}}(x_q) = \phi_{F^q}(X_q) = p$ et donc $f^q(x_q) = x_q$. □

En fait on a montré plus précisément que $\rho(f) = p/q$ avec $p \wedge q = 1$ si et seulement si f a un point périodique de période minimale q. Dans ce cas tous les points périodiques ont q pour période minimale.

Remarque 5.8. *En général, lorsque le nombre de rotation est rationnel, l'homéomorphisme f (même si f est de classe C^2) n'est pas topologiquement conjuguée à la rotation associée. En fait tout homéomorphisme f préservant l'orientation est topologiquement conjugué à $R_{p/q}$ si et seulement si $f^q = Id$.*

4. Semi-conjugaison pour $\rho(f) \notin \mathbb{Q}$

PROPOSITION 5.9. *Soit f un homéomorphisme du cercle préservant l'orientation avec $\rho(f) \notin \mathbb{Q}$. Alors il existe une application g continue (surjective) du cercle de degré 1 admettant un relevé $G : \mathbb{R} \to \mathbb{R}$ croissant et satisfaisant*

$$g \circ f = R_{\rho(f)} \circ g.$$

L'application g est surjective mais pas nécessairement injective, autrement dit le relèvement G est croissant mais pas forcément strictement croissant.

DÉMONSTRATION. Soit $\mu \in \mathcal{M}(\mathbb{R}/\mathbb{Z}, f)$ (sans atomes car f n'admet pas de points périodiques). On pose $G(x) = \mu([0, \pi(x)])$ pour $x \in [0, 1]$ et on prolonge G sur $\mathbb{R}$ de sorte que $G(\cdot + 1) = G(\cdot) + 1$. Ainsi définie la fonction G est croissante continue et c'est le relèvement d'une application continue g du cercle de degré 1. De plus elle vérifie $G \circ F(x) - G \circ F(0) = G(x)$ pour tout x. Il suffit de le montrer pour $x \in [0, 1[$ car g et f sont de degré 1. Mais dans ce cas $G \circ F(x) - G \circ F(0) = \mu([f(0), f(\pi(x))]) = \mu\left(f([0, \pi(x)])\right)$ qui est égal à $\mu([0, \pi(x)]) = G(x)$ par f^{-1}-invariance de la mesure μ. Puisque $G + G \circ F(0)$ est un relèvement de $R_{\alpha'} \circ g$ avec $\alpha' = G \circ F(0)$, on conclut que $g \circ f = R_{\alpha'} \circ g$. D'après la Proposition 5.6 on a nécessairement $\alpha' = \rho(R'_\alpha) = \rho(f)$. $\square$

REMARQUE 5.10. *L'application g est une conjugaison topologique si et seulement si la mesure μ est sans atome de support total (que $\rho(f)$ soit rationnel ou pas). En particulier si $f^q = Id$ c'est le cas de la mesure $\frac{1}{q} \sum_{k=0}^{q-1} f^{*k} Leb$. L'homéomorphisme f est alors conjugué à une rotation $R_{\rho(f)} = R_{p/q}$ pour un certain entier p.*

5. Théorème de Denjoy

Pour un nombre réel α, on note

$$\mathcal{N}(\alpha) := \{q \in \mathbb{N}^*, \ \exists p \in \mathbb{Z} \text{ with } p \wedge q = 1 \text{ and } |q\alpha - p| < \frac{1}{q}\}.$$

Par le principe des tiroirs de Dirichlet, l'ensemble $\mathcal{N}(\alpha)$ est infini pour tout irrationnel α. On peut aussi vérifier que les *réduites* $\frac{p_n}{q_n} = [a_0(\alpha), ..., a_n(\alpha)]$ de α satisfont $q_n \in \mathcal{N}(\alpha)$ pour tout n.

LEMME 5.11. *Soient $\alpha \notin \mathbb{Q}$ et $q \in \mathcal{N}(\alpha)$, alors il existe une permutation σ de $\{0, ..., q-1\}$ tel que $i\alpha \in [\frac{\sigma(i)}{q}, \frac{\sigma(i)+1}{q}[\subset \mathbb{R}/\mathbb{Z}$ pour tout $i = 0, ..., q-1$.*

DÉMONSTRATION. On a $|\alpha - p/q| < 1/q^2$. On peut aussi supposer $\alpha \geq p/q$, l'autre cas se traitant de la même façon. Pour tout $i = 0, ..., q-1$ on a $|i\alpha - ip/q| < 1/q$ et donc $i\alpha \in [ip/q, (ip+1)/q[$. Mais ces arcs de cercle sont deux à deux disjoints pour $i = 0, 1, ...q-1$: en effet si $ip/q = i'p/q + l$ avec l un entier non nul, alors $p \wedge q = 1$ entraine que p divise l et donc $i - i'$ est un multiple non nul de q, ce qui contredit $i, i' \in \{0, 1, ..., q-1\}$. □

LEMME 5.12. *Soit μ une mesure de probabilité borélienne sur $\mathbb{R}/\mathbb{Z}$ et soit $\phi : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ une fonction de classe C^1. Alors pour toute famille finie d'arcs disjoints $J_0, ..., J_{n-1}$ et pour tout $(x_k)_{0\leq k<n} \in \prod_{0\leq k<n} J_k$ on a avec $\mu_{J_k} = \frac{\mu(\cdot \cap J_k)}{\mu(J_k)}$*

$$\sum_{0\leq k<n} \int |\phi(y) - \phi(x_k)|\, d\mu_{J_k}(y) \leq \|\phi'\|_\infty.$$

DÉMONSTRATION. On note ν la mesure de probabilité donnée par le produit $\nu := \prod_{0\leq k<n} \mu_{J_k}$. Par l'inégalité des accroissements finis on a pour tout $y = (y_0, ..., y_{n-1}) \in \prod_{0\leq k<n} J_k$:

$$\begin{aligned}\sum_k |\phi(y_k) - \phi(x_k)| &\leq \sum_k |y_k - x_k| \|\phi'\|_\infty, \\ &\leq \sum_k |J_k| \|\phi'\|_\infty \leq \|\phi'\|_\infty.\end{aligned}$$

On conclut en intégrant la fonction $y \mapsto \sum_k |\phi(y_k) - \phi(x_k)|$ relativement à ν. □

PROPOSITION 5.13 (Inégalité de Denjoy-Koksma). *Soit f un homéomorphisme du cercle préservant l'orientation avec $\rho(f) \notin \mathbb{Q}$. On considère une fonction $\phi : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ de classe C^1 et $q \in \mathcal{N}(\rho(f))$. Alors on a pour tout $x \in \mathbb{R}/\mathbb{Z}$ et tout $\mu \in \mathcal{M}(\mathbb{R}/\mathbb{Z}, f)$:*

$$\left| q \int \phi \; d\mu - \sum_{i=0}^{q-1} \phi \circ f^i(x) \right| \leq \|\phi'\|_\infty.$$

DÉMONSTRATION. On considère la semi-conjugaison g donnée par la Proposition 5.9. On fixe $x \in \mathbb{R}/\mathbb{Z}$ et on note y_i un antécédent par g de $g(x) + i/q$. Soit σ la permutation donnée par le Lemme 6.23 associée à $\alpha = \rho(f)$ et $q \in \mathcal{N}(\rho(f))$. En écrivant la semi-conjugaison on a $g \circ f^i(x) = R_{i\alpha} \circ g(x) = i\alpha + g(x) \in g(x) + [\frac{\sigma(i)}{q}, \frac{\sigma(i)+1}{q}[$. Puisque g admet un relèvement croissant, on a donc $f^i(x) \in [y_\sigma(i), y_{\sigma(i)+1}[$. On écrit alors

$$\left| q \int \phi \; d\mu - \sum_{i=0}^{q-1} \phi \circ f^i(x) \right| \leq \sum_i \left| q \int_{[y_{\sigma(i)}, y_{\sigma(i)+1}[} \phi \; d\mu - \phi \circ f^i(x) \right|,$$

Or $\mu([y_\sigma(i), y_{\sigma(i)+1}[) = g(y_{\sigma(i)+1}) - g(y_{\sigma(i)}) = 1/q$. Avec $J_i = [y_\sigma(i), y_{\sigma(i)+1}[$ on a donc en passant la constante $\phi \circ f^i(x)$ sous l'intégrale :

$$\begin{aligned}\left| q\int \phi \, d\mu - \sum_{i=0}^{q-1} \phi\circ f^i(x)\right| &\leq \sum_i q \int_{[y_{\sigma(i)}, y_{\sigma(i)+1}[} \left|\phi(y) - \phi\circ f^i(x)\right| d\mu(y), \\ &\leq \sum_{i=0}^{q-1} \int \left|\phi(y) - \phi\circ f^i(x)\right| d\mu_{J_i}(y).\end{aligned}$$

Enfin on a d'après le Lemme 5.12 avec $x_i = f^i(x)$, $0 \leq i < q$:

$$\left| q\int \phi \, d\mu - \sum_{i=0}^{q-1} \phi\circ f^i(x)\right| \leq \|\phi'\|_\infty.$$

□

Il suit de l'inégalité de Denjoy-Koksma que $\int \phi \, d\mu = \int \phi \, d\nu$ pour toute fonction de classe C^1 et tout $\mu, \nu \in \mathcal{M}(\mathbb{R}/\mathbb{Z}, f)$. Par un argument de densité on en déduit que les deux mesures μ et ν coïncident.

Corollaire 5.14. *Tout homéomorphisme du cercle préservant l'orientation avec $\rho(f) \notin \mathbb{Q}$ est uniquement ergodique.*

L'inégalité de Denjoy-Koksma permet d'établir l'inégalité de distorsion suivante pour les difféomorphismes du cercle de classe C^2.

Proposition 5.15 (Inégalité de Denjoy). *Soit f un difféomorphisme du cercle de classe C^2 préservant l'orientation avec $\rho(f) \notin \mathbb{Q}$ et $q \in \mathcal{N}(\rho(f))$. Alors pour tout $x \in \mathbb{R}/\mathbb{Z}$, on a*

$$|(f^q)'(x)| \in [e^{-V}, e^V] \text{ avec } V = \|\left(\log(f')\right)'\|_\infty.$$

Démonstration. Soit μ l'unique mesure f-invariante. On applique l'inégalité de Denjoy-Koksma à $\phi = \log|f'|$:

$$(5.1)\ \forall q \in \mathcal{N}(\rho(f))\ \forall x \in \mathbb{R}/\mathbb{Z}, \quad \left|\int \log|f'| \, d\mu - \frac{\log|(f^q)'(x)|}{q}\right| < V/q.$$

De plus on a $\int_{\mathbb{R}/\mathbb{Z}} |(f^q)'(y)| dy = 1$ (théorème de changement de variable). En particulier pour tout q il existe x_q avec $|(f^q)'(x_q)| = 1$. L'inégalité (5.1) en $x = x_q$ pour $q \in \mathcal{N}(\rho(f))$ permet d'établir en faisant tendre $q \in \mathcal{N}(\rho(f))$ vers l'infini

$$\int \log|f'| \, d\mu = 0$$

et finalement

$$\forall x \in \mathbb{R}/\mathbb{Z}, \ |\log(|(f^q)'(x)|)| < V.$$

□

THÉORÈME 5.16. *Tout difféomorphisme f du cercle préservant l'orientation de classe C^2 avec $\rho(f) \notin \mathbb{Q}$ est conjugué à $R_{\rho(f)}$.*

DÉMONSTRATION. Il suffit de montrer que la semi-conjugaison g est injective ce qui équivaut à dire que l'unique mesure invariante est de support total. Par l'absurde on suppose que l'ouvert $(\mathbb{R}/\mathbb{Z}) \setminus supp(\mu)$ est non vide. Cet ouvert étant invariant, f permute ses composantes connexes, qui sont des arcs de cercles. Pour une telle composante connexe I, les ensembles $(f^n(I))_{n\in\mathbb{N}}$ sont deux à deux disjoints. Sinon on aurait $f^k(I) = I$ pour un certain I et un certain entier k. L'application f aurait alors un point k-périodique car toute application continue d'un intervalle dans lui-même a un point fixe. En particulier $\sum_n Leb(f^n I) < +\infty$. Mais ceci est impossible car d'après l'inégalité de Denjoy, on a $e^{-V} Leb(I) \leq Leb(f^q I) \leq e^V Leb(I)$ pour tout $q \in \mathcal{N}(\rho(f))$ et $\mathcal{N}(\rho(f))$ est infini. $\square$

6. Contre-exemple

THÉORÈME 5.17. *Pour tout $\alpha \notin \mathbb{Q}$ il existe $f \in Diff^1(\mathbb{R}/\mathbb{Z})$ tel que $\rho(f) = \alpha$ et f non minimal. En particulier f n'est pas conjugué à $R_{\rho(f)}$.*

DÉMONSTRATION. On considère une suite $(l_q)_{q\in\mathbb{Z}}$ à terme > 0 telle que $\frac{l_{q+1}}{l_q} \xrightarrow{|q|\to+\infty} 1$ et $\sum_q l_q = 1$ (on peut prendre $l_q = \frac{C}{(|q|+1)\log^2(|q|+1)}$ pour C bien choisi). On explose l'orbite $\mathcal{O} = \{q\alpha,\ q \in \mathbb{Z}\}$ de 0 par la rotation R_α de sorte qu'au dessus de $q\alpha$ on ait un intervalle ouvert $J(q)$ de longueur l_q. En recollant ces intervalles dans le même ordre que $\mathcal{O}$ on obtient un cercle (l'intervalle $J(q)$ correspond à l'arc de longueur l_q dont l'extrémité gauche est [1] $\sum_{p,\ \|p\alpha\|<\|q\alpha\|} l_p$). Sur ce cercle on définit un homéomorphisme f préservant l'orientation de sorte que f envoie $J(q)$ sur $J(q+1)$ en une application croissante. La condition $\frac{l_{q+1}}{l_q} \xrightarrow{|q|\to+\infty} 1$ permet en fait de construire un tel exemple de classe C^1 car on peut choisir f comme étant une translation au voisinage des extrémités de $J(q)$. Ainsi définie f se prolonge en un difféomorphisme de classe C^1 sur tout le cercle. De plus la fonction h envoyant $J(q)$ sur $q\alpha$ $[mod\ 1]$ se prolonge continûment à tout le cercle car h est d'image dense et préserve l'orientation. De plus on a $h \circ f = R_\alpha \circ h$. En effet ceci est vrai sur $\bigcup_q J(q)$ puis partout par densité de $\bigcup_q J(q)$ dans le cercle. D'après le Lemme 5.6 on a donc $\rho(f) = \alpha$. Enfin tout $x \in \bigcup_q J(q)$ est errant car $f^n J(q) \cap J(q) = \emptyset$ pour tout $n \neq 0$. En particulier, le support de μ n'est pas total et f n'est pas minimal. $\square$

1. On rappelle que $\|q\alpha\| = \min_{r\in\mathbb{Z}} |q\alpha - r|$.

7. Exercices

Exercice 37.
Soient f et g des homéomorphismes du cercle préservant l'orientation.

(1) Donner un exemple où $\rho(f\circ g) \neq \rho(f)+\rho(g)$.

(2) Montrer que si $f\circ g = g\circ f$, alors $\rho(f\circ g) = \rho(f)+\rho(g)$. on pourra considérer une mesure f et g invariante.

(3) Montrer que $\rho(F\circ G) \leq \rho(F)+\rho(G)+1$.

Exercice 38.
On munit $Homeo(\mathbb{R}/\mathbb{Z})$ de la topologie de la convergence uniforme induite par la distance $D(f,g) = \sup_{x\in\mathbb{R}/\mathbb{Z}} d(f(x),g(x))$ avec d la distance [2] sur $\mathbb{R}/\mathbb{Z}$. Montrer que pour deux homéomorphismes f,g du cercle préservant l'orientation il existe des relevés F,G tels que $\|F-G\|_\infty = D(f,g)$. En déduire que le nombre de rotation est une fonction continue de $(Homeo^+(\mathbb{R}/\mathbb{Z}),D)$ dans $\mathbb{R}/\mathbb{Z}$.

Exercice 39.
Montrer que un homéomorphisme du cercle préservant l'orientation est topologiquement conjugué à une rotation irrationnelle ssi il est transitif.

Exercice 40.
Soit f un homéomorphisme du cercle préservant l'orientation tel que $f^q = Id$. Montrer que f est topologiquement conjuguée à une rotation d'angle p/q pour un entier p.

Exercice 41.
Soit f un difféomorphisme C^2 apériodique du cercle préservant l'orientation. Montrer que f est conjugué à une rotation par un difféomorphisme C^1 du cercle ssi il existe $x\in\mathbb{R}/\mathbb{Z}$ avec $\sup_{n\in\mathbb{N}} \log|(f^n)'(x)| < +\infty$.

Exercice 42.
Pour deux réels positifs $\lambda_1 > 1$ et $\lambda_2 < 1$, définissons $a = a(\lambda_1,\lambda_2)$ comme l'unique nombre réel tel que $a\lambda_1 + (1-a)\lambda_2 = 1$. On considère l'unique homéomorphisme affine par morceaux $f = f_{\lambda_1,\lambda_2} : \mathbb{S}^1 \to \mathbb{S}^1$ tel que $f(a) = 0$, et sa dérivée est égal à λ_1 sur $(0,a)$ et λ_2 sur $(a,1)$. Montrer que :

2. $d(x,y) = \min_{\pi(X)=x,\ \pi(Y)=y} |X-Y|$.

(1) Pour $\sigma = \lambda_1/\lambda_2$, soit h_σ l'homéomorphisme de $[0,1]$ défini par $h(x) = (\sigma^x - 1)/(\sigma - 1)$. Montrer que $h^{-1} \circ f_{\lambda_1,\lambda_2} \circ h$ coïncide avec la rotation d'angle ρ avec $\sigma^\rho = \lambda_2$.

(2) Conclure que le nombre de rotation f_{λ_1,λ_2} est égal à $\frac{log\lambda_1}{\log\lambda_1 - \log\lambda_2}$.

Exercice 43.
Soit f un homéomorphisme du cercle préservant l'orientation tel que $\rho(f) \notin \mathbb{Q}$.

(1) Montrer qu'il existe des points récurrents par la droite et la gauche à la fois, i.e. des points x du cercle tels que $f^{n_k}x < x$ et $f^{n_k}x \xrightarrow{k} x$ pour une sous suite $(n_k)_k$ et $f^{m_k}x > x$ et $f^{m_k}x \xrightarrow{k} x$ pour une autre sous suite $(m_k)_k$.

(2) Vérifier que $\rho(R_\alpha \circ f) \geq \rho(f)$.

(3) Montrer que $\rho(R_\alpha \circ f) > \rho(f) \in \mathbb{R}/\mathbb{Z}$ pour tout $\alpha > 0$ assez petit.

Exercice 44.
Soit f un homéomorphisme du cercle préservant l'orientation et soit $(f_\alpha)_\alpha$ la famille à paramètre donnée par $f_\alpha = R_\alpha \circ f$ pour tout $\alpha \in \mathbb{R}$. On suppose que $f_\alpha^q \neq Id$ pour tout α et tout q. Montrer que pour tout rationnel p/q, l'ensemble $\{\alpha,\ \rho(f_\alpha) = p/q\}$ est d'intérieur non vide.

Exercice 45. *Langues d'Arnold.*
Soit $\mathcal{I}$ la région du plan définie par $\mathcal{I} := \mathbb{R}\times]0, \frac{1}{2\pi}[$. On considère la famille à paramètre $(\alpha,\epsilon) \in \mathcal{I}$ d'homéos du cercle définie par $f_{\alpha,\epsilon}(x) = x + \alpha + \epsilon\sin(2\pi x)$ pour tout $x \in \mathbb{R}/\mathbb{Z}$. Lorsque $\epsilon = 0$, on peut aussi définir l'homéomorphisme $f_{\alpha,0}$ comme la rotation d'angle α.

(1) Montrer à l'aide des deux exercices précédent que pour tout $\frac{1}{2\pi} > \epsilon > 0$ la fonction $\alpha \mapsto \rho(f_{\alpha,\epsilon})$ est un escalier du diable, i.e. c'est une fonction croissante non constante dérivable de dérivée nulle sur une partie dense.

(2) Pour tout rationnel $\beta = p/q$ (sous forme irréductible), montrer qu'il existe deux fonctions $\psi_\beta < \phi_\beta$:$]0, \frac{1}{2\pi}[$ continue telles que

$$\{(\alpha,\epsilon) \in \mathcal{I},\ \rho(f_{\alpha,\epsilon}) = \beta\} := \{(\alpha,\epsilon) \in \mathcal{I},\ \psi_\beta(\epsilon) \leq \alpha \leq \phi_\beta(\epsilon)\}.$$

Chapitre 6

Entropie

Sommaire

1. Information et entropie statique

Soit $(X, \mathcal{B}, \mu)$ un espace probabiliste. Pour α, β deux partitions finies (mesurables) de X on dit que α est plus fine que β, ce que l'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 comme

$$\alpha \vee \beta := \{A \cap B,\ A \in \alpha \text{ et } B \in \beta\}.$$

DÉFINITION 6.1. *L'information de α relativement à β est la fonction*[1]

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

L'information de α relativement à la partition triviale sera notée

$$I(\alpha) = -\sum_{A\in\alpha} 1_A \log \mu(A).$$

On notera aussi $\mu(A|\beta)$ pour $E[1_A|\beta]$. Remarquez que α et β sont indépendantes, i.e. $\mu(A\cap B) = \mu(A)\mu(B)$ pour tout $(A, B) \in \alpha \times \beta$ si et seulement si $E[1_A|\beta] = \mu(A)$ pour tout $A \in \alpha$, ce qui s'écrit aussi $I(\alpha|\beta) = I(\alpha)$.

1. On utilise la convention $0 \log 0 = 0$.

LEMME 6.2.

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

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

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

□

DÉFINITION 6.3. *L'entropie statique conditionnelle de α sachant β est*

$$\begin{aligned} H_\mu(\alpha|\beta) &= \int I(\alpha|\beta)\, d\mu, \\ &= -\sum_{A\in\alpha, B\in\beta} \mu\,(A \cap B) \log\left(\frac{\mu(A \cap B)}{\mu(B)}\right). \end{aligned}$$

Lorsque β est la partition triviale on parle de l'entropie statique de α :

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

On a alors

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

LEMME 6.4. *Soient α, β, γ trois partitions finies,*

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

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

(iii) $H_\mu(\alpha) \leq \log \sharp\alpha$ avec égalité si et seulement si $\mu(A) = 1/\sharp\alpha$ 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 concavité de $\phi : t \mapsto -t \log t$ sur $[0, 1]$

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

Enfin pour la dernière égalité il suffit d'appliquer l'inégalité de Jensen (et son cas d'égalité) à la variable aléatoire équidistribuée X sur l'ensemble

$\{\mu(A),\ A\in\alpha\}$ pour la fonction concave ϕ :

$$\begin{aligned} -\sum_{A\in\alpha}\mu(A)\log\mu(A) &= \sharp\alpha\times E[\phi(X)], \\ &\leq \sharp\alpha\phi(E[X]) = \sharp\alpha\times\phi\left(\frac{1}{\sharp\alpha}\right) = \log\sharp\alpha. \end{aligned}$$

□

REMARQUE 6.5. *Lorsque les partitions α et β sont indépendantes la première inégalité devient pour γ triviale*

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

REMARQUE 6.6. *On travaillera uniquement avec des partitions finies, mais tout les résultats présentés s'appliquent également à des partitions dénombrables α, pourvu que $H_\mu(\alpha)$ soit finie.*

2. KS-Entropie et théorème des générateurs

On considère maintenant un système dynamique probabiliste $(X, f, \mathcal{B}, \mu)$. Pour un entier $n > 0$ et une partition α finie de X on note α^n la partition n-itérée :

$$\alpha^n := \bigvee_{k=0}^{n-1} f^{-k}\alpha.$$

LEMME 6.7. *Soient α et β deux partitions finies de X. La suite $(H_\mu(\alpha^n|\beta^n))_n$ est sous-additive.*

DÉMONSTRATION. Pour simplifier on supposera β triviale. La preuve du cas général suit les mêmes lignes. Remarquez tout d'abord que $\alpha^{n+m} = \alpha^n\vee f^{-n}\alpha^m$ pour tout $m, n\in\mathbb{N}^*$. En appliquant le Lemme 6.4 on obtient alors :

$$\begin{aligned} H_\mu(\alpha^{n+m}) &= H_\mu(\alpha^n\vee f^{-n}\alpha^m), \\ &\overset{\textit{Lemme } 6.4\ (i)}{=} H_\mu(\alpha^n|f^{-n}\alpha^m) + H_\mu(f^{-n}\alpha^m), \\ &\overset{\textit{Lemme } 6.4\ (ii)}{\leq} H_\mu(\alpha^n) + H_\mu(f^{-n}\alpha^m). \end{aligned}$$

Enfin par f-invariance de la mesure μ, on a

$$\begin{aligned} H_\mu(f^{-n}\alpha^m) &= -\sum_{A\in\alpha^m}\mu(f^{-n}A)\log\mu(f^{-n}A), \\ &= -\sum_{A\in\alpha^m}\mu(A)\log\mu(A), \\ &= H_\mu(\alpha^m). \end{aligned}$$

□

La propriété de sous-additivité permet de définir (cf Rappels Théorème 8.20) $h(\mu,\alpha) := \lim_n \frac{H_\mu(\alpha^n)}{n} = \inf_n \frac{H_\mu(\alpha^n)}{n}$. L'entropie de Kolmogorov-Sinai $h(\mu)$ (également notée $h_f(\mu)$) de μ est alors obtenue en prenant le supremum sur toutes les partitions finies

$$h(\mu) = \sup_\alpha h(\mu,\alpha).$$

Lemme 6.8. *(i)* $h(\mu,\beta) - h(\mu,\alpha) \leq H_\mu(\beta|\alpha)$,

(ii) pour $(X, f, \mathcal{B}, \mu)$ *inversible on a*

$$h(\mu, \bigvee_{k=-N}^{N} f^{-k}\alpha) = h(\mu,\alpha).$$

Démonstration. D'après le Lemme 6.4 on a pour tout entier $n > 0$

$$H_\mu(\alpha^n) - H_\mu(\beta^n) \leq H_\mu(\alpha^n|\beta^n).$$

La suite $(H_\mu(\alpha^n|\beta^n))_n$ étant sous-additive d'après le Lemme 6.7, on a $H_\mu(\alpha^n|\beta^n) \leq nH_\mu(\alpha|\beta)$. On en conclut facilement le point (i).

Prouvons maintenant (ii). En posant $\alpha_N = \bigvee_{k=-N}^{N} f^{-k}\alpha$ on a $\alpha_N^n = f^N\alpha^{n+2N}$. Il suit alors du Lemme 6.4 et de l'invariance de la mesure μ pour f^{-1} que

$$H_\mu(\alpha^n) \leq H_\mu(\alpha_N^n) \leq H_\mu(f^{-N}\alpha^{n+2N}) = H_\mu(\alpha^{n+2N}).$$

En divisant par n puis en passant à la limite quand n tends vers $+\infty$ on obtient finalement $h(\mu,\alpha) = h(\mu,\alpha_N)$. □

On dit qu'une suite croissante $(\alpha_n)_n$ de partitions finies est génératrice lorsque $\bigvee_{n\in\mathbb{N}} \alpha_n = \mathcal{B}$, i.e. la plus petite tribu [2] contenant les éléments de α_n pour tout n est la tribu $\mathcal{B}$. Lorsque X est un espace métrique complet séparable et $\mathcal{B}$ sa tribu des Boréliens, on peut montrer qu'il existe toujours une suite génératrice de partitions finies. On se placera toujours dans ce cadre dans la suite de ce chapitre. Pour une famille finie $\mathcal{F}$ de sous-ensembles de X, on appelle diamètre de $\mathcal{F}$ le maximum des diamètres des éléments de $\mathcal{F}$. Si X est un compact métrisable et $\mathcal{B}$ sa tribu des boréliens alors toute suite croissante $(\alpha_n)_n$ de partitions finies boréliennes, telle que le diamètre de α_n tend vers 0, définit une suite génératrice. En effet toute fonction réelle continue sur X est limite presque sûre de fonctions (étagée) α_n-mesurable et est donc mesurable relativement à la tribu $\bigvee_{n\in\mathbb{N}} \alpha_n$.

2. à des ensembles de mesure nulle près.

LEMME 6.9. *Soit* $(\alpha_n)_n$ *une suite génératrice de partitions. Pour tout* $\epsilon > 0$ *et tout* $B \in \mathcal{B}$ *il existe un entier* m *et un élément*[3] $A_m \in \alpha_m$ *tel que* $\mu(A_m \Delta B) < \epsilon$.

DÉMONSTRATION. On vérifie que

$$\mathfrak{B} := \{B \in \mathcal{B}, \forall \epsilon > 0 \ \exists \ m \text{ et } A_m \in \alpha_m \ \text{ tel que } \mu(A_m \Delta B) < \epsilon\}$$

définit une tribu. Clairement $X \in \mathfrak{B}$. Vérifions que $\mathfrak{B}$ est stable par passage au complémentaire. Si B satisfait $\mu(A_m \Delta B) < \epsilon$ pour $A_m \in \alpha_m$ alors $(X \setminus A_m)\Delta(X \setminus B) = A_m \Delta B$ et $X \setminus A_m \in \alpha_m$. Enfin montrons que $\mathfrak{B}$ est stable par union dénombrable. Soient $(B_n)_n$ une union dénombrable d'éléments de $\mathfrak{B}$. Pour $\epsilon > 0$ fixé on considère un entier N tel que $\mu(\bigcup_n B_n) < \mu(\bigcup_{0 \le k < N} B_k) + \epsilon/2$. Pour tout $0 \le k < N$ il existe $m_k \in \mathbb{N}$ et $A_{m_k} \in \alpha_{m_k}$ avec $\mu(A_{m_k} \Delta B_k) < \epsilon/2N$. Alors $A := \bigcup_{0 \le k < N} A_{m_k}$ appartient à α_m avec $m = \max_k m_k$ et vérifie $A\Delta(\bigcup_{0 \le k < N} B_k) \subset \bigcup_{0 \le k < N} A_{m_k} \Delta B_k$. Donc

$$\mu\left(A\Delta\left(\bigcup_n B_n\right)\right) \ \le \ \sum_{0 \le k < N} \mu\left(A_{m_k} \Delta B_k\right) + \epsilon/2 < \epsilon.$$

Donc $\mathfrak{B}$ est une tribu et on a alors $\mathfrak{B} = \mathcal{B}$ car $(\alpha_m)_m$ est génératrice et $\alpha_m \subset \mathfrak{B}$ pour tout m. □

THÉORÈME 6.10 (Générateurs de Sinaï). *Soit* $(\alpha_n)_n$ *une suite génératrice de partitions, alors*

$$h(\mu) = \lim_n h(\mu, \alpha_n).$$

DÉMONSTRATION. D'après le Lemme 6.8 (i) il suffit de montrer que pour toute partition β la suite décroissante $(H_\mu(\beta|\alpha_N))_N$ tends vers 0 quand N tend vers l'infini.

$$\begin{aligned} H_\mu(\beta|\alpha_N) &= -\sum_{B \in \beta} \int 1_B \log \mu(B|\alpha_N) \ d\mu, \\ &= -\sum_{B \in \beta} \int \mu(B|\alpha_N) \log \mu(B|\alpha_N) \ d\mu, \\ &= \sum_{B \in \beta} \int \phi\left(\mu(B|\alpha_N)\right) d\mu, \end{aligned}$$

avec $\phi(t) = -t \log t$ pour $t \in [0,1]$. D'après le Lemme 6.9, pour tout $\epsilon > 0$ et pour tout $B \in \beta$ il existe un entier n et un élément $A_n \in \alpha_n$ avec $\mu(A_n \Delta B) < \epsilon^2$. L'espérance conditionnelle $\mu(B|\alpha_n)$ étant la projection orthogonale de 1_B sur $L^2(\alpha_n)$, on en déduit que $\|\mu(B|\alpha_n) - 1_B\|_2 \le$

3. On note encore α_m la tribu engendrée par la partition. Celle-ci est donnée par les unions finies d'éléments de la partition.

$\|1_{A_n} - 1_B\|_2 = \sqrt{\mu(A_n \Delta B)} < \epsilon$ et $\|\mu(B|\alpha_{m+1}) - 1_B\|_2 \leq \|\mu(B|\alpha_m) - 1_B\|_2$ pour tout entier m. Il s'en suit que pour tout $B \in \beta$, la suite $(\mu(B|\alpha_n))_n$ converge dans L^2 vers 1_B. Quitte à extraire une sous-suite on peut supposer qu'on a aussi une convergence presque sûre (pour tout B dans la partition finie β). Par convergence dominée on a alors $\int \phi\left(\mu(B|\alpha_N)\right) d\mu \to_N 0$. $\square$

Une partition finie α est dite génératrice pour un système inversible lorsque $(\bigvee_{k=-n}^{n} f^{-k}\alpha)_{n\in\mathbb{N}}$ est génératrice. Le corollaire suivant se déduit directement du Théorème 6.10 et du Lemme 6.8 (ii) :

Corollaire 6.11. *Pour toute partition génératrice α d'un système inversible $(X, f, \mathcal{B}, \mu)$ on a*

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

Remarque 6.12. *Pour un système probabiliste non nécessairement inversible, on peut définir une notion de générateur positif comme étant une partition α telle que la suite $\left(\bigvee_{0\leq k\leq n} f^{-k}\alpha\right)_n$ est génératrice. Dans ce cadre on a encore $h(\mu) = h(\mu, \alpha)$.*

On présente dans le lemme suivant les propriétés de l'entropie d'un produit, d'un facteur et d'une puissance.

Lemme 6.13. *Soient $(X, \mathcal{B}, f, \mu)$ et $(Y, \mathcal{C}, g, \nu)$ deux systèmes probabilistes, alors*

- *$h(\mu \times \nu) = h(\mu) + h(\nu)$, où $\nu \times \mu$ désigne la mesure produit sur $X \times Y$ invariante par $f \times g$,*
- *si $\pi : (Y, \nu) \to (X, \mu)$ est un facteur alors $h(\mu) \leq h(\nu)$,*
- *$h(f^k, \mu) = kh(f, \mu)$ pour tout $k \in \mathbb{N}$,*
- *si f est inversible, $h(f, \mu) = h(f^{-1}, \mu)$.*

Démonstration. Par indépendance, l'entropie pour la mesure produit d'un produit de deux partitions est donnée par la somme des entropies relatives à ces partitions. La formule sur l'entropie du produit suit alors du théorème des générateurs de Sinaï puisque les partitions produits engendrent la tribu produit (si $(\alpha_n)_n$ et $(\beta_n)_n$ sont respectivement des suites génératrices de $(X, \mathcal{B}, \mu)$ et $(Y, \mathcal{C}, \nu)$, alors $(\alpha_n \times \beta_n)_n$ est une suite génératrice du produit). Pour le deuxième point on vérifie facilement que pour toute partition α on a

$$\begin{aligned} h(\mu, \alpha) &= h(\pi^*\nu, \alpha), \\ &= h(\nu, \pi^{-1}\alpha), \end{aligned}$$

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

$\square$

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

3. Formule de Newhouse et Shanon-Mcmillan-Breiman

On considère dans ce paragraphe un système probabiliste ergodique $(X, f, \mathcal{B}, \mu)$. Pour un tel système on va donner deux définitions alternatives pour l'entropie de μ. Soit α une partition de X. Pour tout $x \in X$ on note $\alpha^n(x)$ l'élément de α^n contenant x. On s'intéresse au taux de décroissance exponentielle de la μ-mesure de $\alpha^n(x)$ pour $x \in X$:

$$s^{+/-}(x) := \overline{\underline{\lim}}_n - \frac{\log \mu\left(\alpha^n(x)\right)}{n}.$$

LEMME 6.14. *$s^{+/-}$ sont des constantes (presque partout).*

DÉMONSTRATION. Puisque $\int \phi \; d\mu = \int \phi \circ f \; d\mu$ pour tout $\phi \in L^1$ par invariance de la mesure, toute fonction mesurable intégrable vérifiant $\phi \circ f \leq \phi$ satisfait $\phi \circ f = \phi$ p.p. et donc ϕ est constante p.p. par ergodicité. Mais $\log \mu\left(\alpha^n(x)\right) \leq \log \mu\left(\alpha^{n-1}(fx)\right)$ implique $s^{+/-}(fx) \leq s^{+/-}(x)$. $\square$

On considère maintenant le taux de croissance exponentielle du nombre d'éléments de α^n recouvrant un ensemble de μ-mesure minorée. Pour tout $1 > \lambda > 0$ on pose

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

Ces deux quantités sont égales à l'entropie de μ relativement à la partition α.

THÉORÈME 6.15. *Pour tout $1 > \lambda > 0$, on a*

$$s^+ = s^- = n_\lambda = h(\mu, \alpha).$$

L'égalité $s^+ = s^- = h(\mu, \alpha)$ est attribuée à Shanon, Mcmillan et Breimman tandis que l'égalité $n_\lambda = h(\mu, \alpha)$ est due à Newhouse.

DÉMONSTRATION.

1. $\boxed{n_\lambda \leq s^-.}$ Soit $\epsilon > 0$. Pour μ p.p. x on note $n(x)$ le premier entier $n > 1/\epsilon$ tel que $\mu\left(\alpha^n(x)\right) \geq e^{-n(s^- + \epsilon)}$ (ainsi définie la fonction n est mesurable). 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 $\frac{\sharp\{0 \leq k < n,\ T^k x \in G\}}{n} > 1 - \epsilon$. Pour x fixé dans E et $n > N_1$ on découpe 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 $f^{k_i}x$ dans G et les longueurs $n(f^{k_i}x)$ associées. Puisque $n > 1/\epsilon$ le nombre de configurations possibles pour les symboles [et] est plus petit que $\binom{n}{\epsilon n}^2$. Maintenant dans les intervalles délimités par les crochets, le nombre de α-noms 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 G ou dans le dernier intervalle dépassant n (qui est de longueur au plus N_0), donc $n' > (1-\epsilon)n - N_0$. Par conséquent

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

puis

$$n_\lambda \leq s^- + \epsilon(1 + \log \sharp\alpha) + \overline{\lim}_n \frac{1}{n} \log \binom{n}{\epsilon n}.$$

A l'aide de le formule de Stirling on vérifie facilement que

$$\lim_{\epsilon \to 0} \overline{\lim}_n \frac{1}{n} \log \binom{n}{\epsilon n} = 0,$$

de sorte que nous obtenons $n_\lambda \leq s^-$, en prenant la limite quand ϵ tend vers 0 dans l'inégalité précédente.

2. $\boxed{n_\lambda \geq s^+.}$ Supposons par l'absurde que $n_\lambda < s^+$. Soient des réels a et b avec $n_\lambda < a < b < s^+$ 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 $\sharp C_n \leq Ce^{an}$ (pour un $C > 0$). Alors on a

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

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 vérifie $\mu(\alpha^n(x)) < e^{-bn}$ pour une infinité de n. C'est en particulier le cas pour μ p.t. $x \in E \subset \bigcap_{n \in \mathbb{N}} C_n$. Contradiction.

3. $\boxed{s^- \leq h(\mu, \alpha).}$ Pour tout n on a

$$\frac{H_\mu(\alpha^n)}{n} = \int \frac{I(\alpha^n)}{n} \, d\mu,$$

où $I(\alpha^n)$ désigne l'information de la partition α^n donnée par $I(\alpha^n) = -\sum_{A \in \alpha^n} 1_A \log \mu(A)$. Ainsi $\underline{\lim}_n \frac{I(\alpha^n)}{n} = s^-$ et le lemme de Fatou donne donc

$$\begin{aligned} h(\mu, \alpha) &= \lim_n \frac{H_\mu(\alpha^n)}{n}, \\ &\geq \int \underline{\lim}_n \frac{I(\alpha^n)}{n} \, d\mu = s^-. \end{aligned}$$

4. $\boxed{h(\mu, \alpha) \leq \lim_{\lambda \to 1} n_\lambda.}$ On va en fait montrer cette identité pour toute mesure invariante (pas nécessairement ergodique). Pour une infinité de n on dispose d'une collection $C_n \subset \alpha^n$ de mesure totale plus grande que λ telle que $\frac{\log \sharp C_n}{n} \simeq n_\lambda$. On a

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

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

□

4. Harmonicité de l'entropie

L'entropie de Kolmogorov-Sinaï, vue comme une fonction réelle définie sur l'ensemble $\mathcal{M}(X, f)$ des mesures de probabilité invariantes d'un système dynamique topologique (X, f), est une fonction affine :

THÉORÈME 6.16. *Soit (X, f) un système dynamique topologique et soient μ, ν deux mesures de probabilité f-invariantes, alors pour tout $\lambda \in [0, 1]$*

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

DÉMONSTRATION. Montrons que la fonction $\xi \mapsto h(\xi, \alpha)$ est affine pour toute partition (finie) α. Tout d'abord par concavité de $\phi(t) = -t \log t$ sur $[0, 1]$ on a pour tout entier $n > 0$

$$H_{\lambda \mu + (1 - \lambda) \nu}(\alpha^n) \geq \lambda H_\mu(\alpha^n) + (1 - \lambda) H_\nu(\alpha^n)$$

si bien que $h(\lambda \mu + (1 - \lambda) \nu, \alpha) \geq \lambda h(\mu) + (1 - \lambda) h(\nu)$.

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.

On en déduit le caractère affine de $\xi \mapsto h(\xi)$ en considérant une partition α satisfaisant $h(\mu) \simeq h(\mu, \alpha)$, $h(\nu) \simeq h(\nu, \alpha)$ et $h(\lambda\mu + (1-\lambda)\nu) \simeq h(\lambda\mu + (1-\lambda)\nu, \alpha)$.

□

En fait on peut montrer que l'entropie mesurée $h : \mathcal{M}(X, f) \to \mathbb{R}$ d'une système topologique (X, f) est harmonique, i.e. si $\mu = \int \nu \, dM_\mu(\nu)$ est la décomposition ergodique de μ alors

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

5. Entropie topologique

On considère dans cette partie un système dynamique topologique (X, f). Pour deux recouvrements finis d'ouverts $\mathcal{U}$ et $\mathcal{V}$ on dit que $\mathcal{V}$ est plus fin que $\mathcal{U}$, ce que l'on écrit $\mathcal{V} > \mathcal{U}$, lorsque tout élément de $\mathcal{V}$ est inclus dans un élément de $\mathcal{U}$. Aussi le recouvrement joint $\mathcal{U} \vee \mathcal{V}$ est défini comme

$$\mathcal{U} \vee \mathcal{V} := \{U \cap V, \ U \in \mathcal{U} \text{ et } V \in \mathcal{V}\}.$$

On note enfin $\mathcal{U}_V$ le recouvrement induit par $\mathcal{U}$ sur un sous-ensemble V de X. On définit $H(\mathcal{U})$ 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 $\mathcal{U}$. On établit facilement les propriétés suivantes :

Lemme 6.17. — *$H(\mathcal{U} \vee \mathcal{V}|\mathcal{W}) \leq H(\mathcal{U}|\mathcal{V} \vee \mathcal{W}) + H(\mathcal{V}|\mathcal{W})$,*
— *$H(f^{-1}\mathcal{U}|f^{-1}\mathcal{V}) \leq H(\mathcal{U}|\mathcal{V})$,*
— *si $\mathcal{V} > \mathcal{W}$ alors $H(\mathcal{U}|\mathcal{V}) \leq H(\mathcal{U}|\mathcal{W})$,*
— *$H(\mathcal{U}) \leq \log \sharp\mathcal{U}$.*

Pour un recouvrement fini d'ouverts $\mathcal{U}$ de X et pour tout entier $n > 0$ on note $\mathcal{U}^n$ le recouvrement n-itéré $\mathcal{U}^n := \bigvee_{j=0}^{n-1} f^{-j}\mathcal{U}$. Comme pour l'entropie mesurée, on montre facilement avec le lemme précédent que la suite $(H(\mathcal{U}^n|\mathcal{V}^n))_n$ est sous-additive et il s'en suit que $h(f,\mathcal{U}) = \lim_n \frac{H(\mathcal{U}^n)}{n} = \inf_n \frac{H(\mathcal{U}^n)}{n}$ est bien définie. Cette quantité sera aussi notée $h(\mathcal{U})$ par la suite. L'entropie topologique de (X, f), notée $h_{top}(f)$ est alors définie comme

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

De façon similaire au Lemme 6.8 on montre :

Lemme 6.18. — *$h(\mathcal{U}) - h(\mathcal{V}) \leq H(\mathcal{U}|\mathcal{V})$,*

— *pour un homéomorphisme f, on a*

$$h\left(\bigvee_{k=-N}^{N} f^{-k}\mathcal{U}\right) = h(\mathcal{U}).$$

Une suite de recouvrements d'ouverts $(\mathcal{U}_n)_n$ est dite génératrice lorsque le diamètre de $\mathcal{U}_n$ tend vers 0. De façon similaire au théorème des générateurs de Sinaï on a

Théorème 6.19. *Soit $(\mathcal{U}_n)_n$ une suite génératrice de recouvrements d'ouverts, alors*

$$h_{top}(f) = \lim_n h(\mathcal{U}_n).$$

Par compacité de X pour tout recouvrement ouvert $\mathcal{U}$ de X il existe $\epsilon_{\mathcal{U}} > 0$ tel que toute boule de diamètre inférieur à $\epsilon_{\mathcal{U}}$ est inclus dans un élément de $\mathcal{U}$. Le maximum des tels réels $\epsilon_{\mathcal{U}}$ est appelé *la constante de Lebesgue* de $\mathcal{U}$.

Démonstration. Comme pour le théorème des générateurs de Sinaï on a pour tout recouvrement d'ouverts $\mathcal{V}$

$$h(\mathcal{V}) - h(\mathcal{U}_N) \leq H(\mathcal{V}|\mathcal{U}_N).$$

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

$$H(\mathcal{V}|\mathcal{U}_N) = 0.$$

□

On dit qu'un homéomorphisme f de X est expansif s'il existe un recouvrement ouvert $\mathcal{U}$ tel que le diamètre de $\bigvee_{k=-n}^{n} f^k\mathcal{U}$ tendent vers 0 quand n tends vers l'infini. On dira alors que $\mathcal{U}$ est un générateur topologique. Si α est une partition (mesurable finie) dont le diamètre est strictement inférieur à la constante de Lebesgue d'un générateur topologique $\mathcal{U}$ alors α est un générateur (au sens mesuré) pour toutes les mesures f-invariantes. Remarquez enfin que la propriété d'expansivité est préservée par conjugaison topologique.

COROLLAIRE 6.20. *Soit $\mathcal{V}$ un générateur topologique pour un système topologique inversible (X, f), alors*

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

REMARQUE 6.21. *Un système topologique (X, f) non nécessairement inversible est dit positivement expansif s'il admet un générateur topologique positif, i.e. un recouvrement d'ouverts $\mathcal{U}$ de X tel que la suite $\left(\bigvee_{0\leq k\leq n} f^{-k}\mathcal{U}\right)_n$ est génératrice. De la même façon on a alors $h_{top}(f) = h(\mathcal{U})$.*

On présente dans le lemme suivant les propriétés de l'entropie topologique d'un produit, d'un facteur et d'une puissance. Les preuves faciles sont laissées au lecteur.

LEMME 6.22.
— *$h_{top}(f \times g) = h_{top}(f) + h_{top}(g)$,*
— *si $\pi : (Y, g) \to (X, f)$ est un facteur topologique alors $h_{top}(f) \leq h_{top}(g)$,*
— *$h_{top}(f^k) = kh_{top}(f)$ pour tout $k \in \mathbb{N}$,*
— *si f est inversible, $h_{top}(f^{-1}) = h_{top}(f)$.*

6. Entropie à la Bowen

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

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

Clairement si $\mathcal{U}$ est un recouvrement d'ouverts 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 (resp. pour un) $U_x^n \in \mathcal{U}^n$ contenant x. En particulier on a

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

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

Pour tout entier $n > 0$ on peut définir la distance d_n sur X par

$$d_n(x, y) := \max_{k=0,\ldots,n-1} d(f^k x, f^k y).$$

Une (n,ϵ)-boule dynamique est une boule de rayon ϵ pour la métrique d_n. Un sous-ensemble E_n de X est dit (n,ϵ)-séparé s'il est ϵ-séparé pour la métrique d_n, i.e.

$$\forall x \neq y \in E_n \quad \exists 0 \leq k \leq n-1 \text{ tel que } d(f^k x, f^k y) > \epsilon.$$

On montre alors facilement que

$$h_{top}(f) = \lim_{\epsilon \to 0} h'_{top}(f,\epsilon)$$

$$\text{avec } h'_{top}(f,\epsilon) := \limsup_n \frac{1}{n} \log \max\{\sharp E_n,\ E_n \text{ est } (n,\epsilon)\text{-séparé}\}.$$

En effet on a

(i) $\min\{\sharp C_n,\ \bigcup_{x\in C_n} B(x,n,\epsilon) = X\} \leq \max\{\sharp E_n,\ E_n \text{ est } (n,\epsilon)\text{-séparé}\}$,

(ii) $\max\{\sharp E_n,\ E_n \text{ est } (n,\epsilon)\text{-séparé}\} \leq \min\{\sharp C_n,\ \bigcup_{x\in C_n} B(x,n,\epsilon/2) = X\}$.

Le lemme combinatoire suivant permet de faire le lien entre les partitions itérées et les boules dynamiques.

Lemme 6.23. *Soit μ une mesure ergodique d'un système topologique (X,f). Soit α une partition finie avec $\mu(\partial\alpha) = 0$. Alors pour tout $\delta > 0$ il existe ϵ tel que pour μ p.t. 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 $\beta > 0$ tel que le β-voisinage $\partial^\beta\alpha$ de $\partial\alpha$ soit de mesure plus petite que δ. D'après le théorème ergodique ponctuel, pour μ p.t. x on a $\frac{1}{n}\sharp\{0 \leq k < n,\ f^k x \in \partial^\beta\alpha\} \to_n \mu(\partial^\beta\alpha) < \delta$. Soit $\epsilon = \min_{A\in\alpha} d\left(A \setminus \partial^\beta\alpha, X \setminus A\right)$. Alors $B(x,n,\epsilon)$ rencontre au plus $(\sharp\alpha)^{a_n}$ éléments de α^n avec $a_n = \sharp\{0 \leq k < n,\ f^k x \in \partial^\beta\alpha\}$. □

On peut toujours construire des partitions arbitrairement fines satisfaisant l'hypothèse du lemme précédent :

Lemme 6.24. *Soit X un espace métrisable compact et soit $\mu \in \mathcal{M}(X)$, alors il existe des partitions α de diamètre arbitrairement petit avec $\mu(\partial\alpha) = 0$.*

Démonstration. Il suffit de construire un recouvrement fini de X de diamètre arbitrairement petit avec la même propriété et de considérer la partition engendrée par ce recouvrement. Par compacité de X l'existence d'un tel recouvrement est assuré s'il existe une base de voisinages ouverts dont les bords sont de μ-mesure nulle. Le lemme suit alors de l'affirmation suivante : *pour tout $x \in X$ on a $\mu(\partial B(x,r)) = 0$ pour des* *réels $r > 0$ arbitrairement petits*. Pour $x \in X$, les bords des boules $B(x, r)$ pour $r > 0$ sont les sphères disjointes $\{y \in X,\ d(x, y) = r\}$ et donc

$$\{r \in \mathbb{R}^+,\ \mu\left(\partial B(x, r)\right) > 0\}$$

est un ensemble au plus dénombrable, la mesure μ étant finie. Cela prouve l'affirmation et conclut la démonstration du lemme.

□

7. Principe variationnel

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

THÉORÈME 6.25. *Soit (X, f) un système topologique, on a*

$$h_{top}(f) = \sup_{\mu \in \mathcal{M}(X,f)} h(\mu) = \sup_{\mu \in \mathcal{M}_e(X,f)} h(\mu).$$

DÉMONSTRATION. Montrons tout d'abord que $h(\mu) \leq h_{top}(f)$ pour toute mesure f-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 $h(\mu, \alpha) \simeq h(\mu)$ (voir Lemme 6.24). Pour δ fixé, d'après le Lemme 6.23 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. A l'aide de la formule de Newhouse on obtient donc

$$h(\mu, \alpha) \leq h_{top}(f, \epsilon/2) + \delta \leq h_{top}(f) + \delta.$$

Soit $\epsilon > 0$ tel que $h'_{top}(f, \epsilon) \simeq h_{top}(f)$ et soit α une partition de diamètre $< \epsilon$. Nous allons montrer qu'il existe une mesure f-invariante μ avec $h(\mu, \alpha) \geq h'_{top}(f, \epsilon)$. On considère pour tout n un ensemble (n, ϵ)-séparé E_n de cardinal maximal parmi les ensembles (n, ϵ)-séparés. On pose $\mu_n := \frac{1}{\sharp E_n} \sum_{x \in E_n} \delta_x$. 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 6.4 du chapitre précédent pour tout $0 \le j < p$ avec $n - j = ([n/p]-1)p + i_j$

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

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

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

où l'on a noté $\nu_n := \frac{1}{p([n/p]-1)} \sum_{k=0}^{p([n/p]-1)-1} f_*^k \mu_n$. On montre comme dans le Lemme 4.1 que toute limite faible ν de $(\nu_n)_n$ est f-invariante (ν_n ne l'était pas a priori). Soit $(n_k)_k$ une sous suite telle que $\lim_k \frac{\log E_{n_k}}{n_k} = h'_{top}(f, \epsilon)$. Soit ν une limite faible de $(\nu_{n_k})_k$. On peut choisir la partition α avec $\nu(\partial\alpha) = 0$ et donc $\nu(\partial\alpha^p) = 0$ de sorte que $\mu \mapsto H_\mu(\alpha^p)$ est une fonction réelle de $\mathcal{M}(X)$ continue en ν. On a alors en passant à la limite en n (à p fixé) :

$$\frac{H_\nu(\alpha^p)}{p} \ge h'_{top}(f, \epsilon),$$

et en passant à la limite en p on conclut que

$$h(\nu, \alpha) \ge h'_{top}(f, \epsilon).$$

□

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 variationnel.

Théorème 6.26. *On suppose que $f : X \to X$ soit un homéomorphisme expansif. Alors il existe une mesure d'entropie maximale.*

Démonstration. Soit $\mu_p \in \mathcal{M}(X, f)$ tel que $h(\mu_p) > h_{top}(f) - \frac{1}{p}$. Soit μ une limite faible de $(\mu_p)_p$. Vérifions que μ est une mesure d'entropie maximale. 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 déjà remarqué on a $h(\nu, \alpha) = h(\nu)$ pour tout $\nu \in \mathcal{M}(X, f)$. Pour conclure il suffit de vérifier que $\overline{\lim}_p h(\mu_p, \alpha) \le h(\mu, \alpha)$. Soit n un entier strictement positif tel que $h(\mu, \alpha) \simeq \frac{1}{n} H_\mu(\alpha^n)$. Puisque

$\mu(\partial\alpha) = 0 = \mu(\partial\alpha^n)$ la fonction $\nu \mapsto H_\nu(\alpha^n)$ est continue en μ. Pour p assez grand on a donc $h(\mu_p) \le \frac{H_{\mu_p}(\alpha^n)}{n} \simeq \frac{H_\mu(\alpha^n)}{n} \simeq h(\mu,\alpha) = h(\mu)$. $\square$

Remarque 6.27. *Dans la preuve précédente, nous avons en fait montré que la fonction $\mu \mapsto h(\mu)$ de $\mathcal{M}(X,f)$ dans $\mathbb{R}^+$ est semi-continue supérieurement.*

8. Exercices

Exercice 46. *Entropie nulle.*
1. Montrer que l'entropie de Kolmogorov-Sinai des mesures périodiques est nulle.
2. Montrer que toute isométrie est d'entropie topologique nulle.
3. Montrer que tout homéomorphisme du cercle est d'entropie topologique nulle.

Exercice 47.
Soit $\pi : (Y,S) \to (X,T)$ une extension topologique telle que

$$\sup_{x\in X} \sharp \pi^{-1} x < +\infty.$$

Montrer que $h_{top}(T) = h_{top}(S)$.

Exercice 48. *Entropie des mesures de Bernoulli.*
Montrer que l'entropie de la mesure de Bernoulli de paramètre $(p_1, ..., p_s)$ sur $\{1, ..., s\}^{\mathbb{Z}}$ est donnée par $-\sum_{i=1,...,s} p_i \log p_i$.

Exercice 49.
Montrer que l'entropie de la mesure de Gauss μ_G est donnée par $\int_{[0,1]} \log |G'| d\mu_G$ (on pourra utiliser la formule de Shanon-McMilman-Breiman).

Exercice 50.
Montrer que le doublement de l'angle f_2 sur le cercle est d'entropie topologique $\log 2$ et que la mesure de Lebesgue est l'unique mesure d'entropie maximale.

Exercice 51. *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^{τ_A} l'application de premier retour induite sur A et $\mu_A = \mu(. \cap A)/\mu(A)$. Montrer que

$$h_{f^{\tau_A}}(\mu_A) = \frac{h_f(\mu)}{\mu(A)}.$$

Exercice 52. *Générateur unilatère pour un système inversible.*
Soit $(X, \mathcal{B}, f, \mu)$ un système dynamique probabiliste inversible. On suppose qu'il existe une partition mesurable finie α telle que $(\bigvee_{0\leq k\leq n} f^{-k}\alpha)_n$ soit une suite génératrice. Montrer que $h(\mu) = 0$.

Exercice 53. $\times 2, \times 3$.
On note f_2, f_3 les applications $x \mapsto 2x$ et $x \mapsto 3x$ sur $\mathbb{R}/\mathbb{Z}$. Enfin note T l'application $x \mapsto x + 1/2$. Soit μ une mesure de probabilité f_2 et f_3-invariante. On suppose que $h_{f_2}(\mu) > 0$ et que μ est f_3-ergodique.

(1) Montrer que $T^*\mu$ est f_3-invariante et ergodique.
(2) Montrer en utilisant l'Exercice 52 que $T^*\mu = \mu$.
(3) Calculer $\int e^{2ik\pi x} d\mu(x)$ pour tout $k \in \mathbb{Z}$. Conclure que μ est la mesure de Lebesgue.

Exercice 54.
Soit $f : M \to M$ une application C^1 sur une variété Riemannienne compacte $(M, \|\|)$. Montrer que[4]

$$h_{top}(f) \leq \lim_n \frac{1}{n} \log \max_k \|\Lambda^k df^n\|.$$

Exercice 55. *Entropie des automorphisme linéaire du tore.*
Montrer que l'entropie de la mesure de Lebesgue pour un automorphisme linéaire f_A du tore est $\sum_i \log|\lambda_i|$ où les λ_i sont les valeurs propres de A de module > 1. En déduire que la mesure de Lebesgue est une mesure d'entropie maximale de f_A. On pourra utiliser l'Exercice 54 pour montrer que $h_{fop}(f_A) \leq \sum_i \log|\lambda_i|$, puis la formule de Shanon-McMilman-Breiman pour établir la majoration $h(Leb) \geq \sum_i \log|\lambda_i|$. Donner un exemple sur le tore $\mathbb{T}^3$, où la mesure d'entropie maximale n'est pas unique.

Exercice 56. *Inégalité de Ruelle sur l'intervalle.*
On considère une application f continue C^1 par morceaux de l'intervalle.

(1) Soit α une partition finie de $[0, 1]$ en intervalles de longueur $1/n$. On note $|f'(x)|_\alpha := \sup_{y\in\alpha(x)} |f'(y)|$, où $\alpha(x)$ désigne l'élément de α contenant x. Montrez que pour tout $\mu \in \mathcal{M}([0, 1], f)$ et pour tout entier $k > 0$, on a

$$H_\mu(\alpha^{k+1}|\alpha^k) \leq \log 3 + \int \log^+ |f'(x)|_\alpha d\mu(x).$$

4. Remarquer que le terme de droite de dépend pas de $\|\|$.

(2) En déduire que

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

Exercice 57. *Entropie des applications monotones.*
Soit f une application continue monotone par morceaux de l'intervalle et soit α la partition en branches monotones de f. Montrez que

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

Exercice 58. *Entropie des applications tentes.*
Pour tout $0 \leq a \leq 2$ on pose T_a l'application de l'intervalle définie par

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

Montrer que

$$h_{top}(T_a) = \log^+ a.$$

Exercice 59. *Exemple sans mesure d'entropie maximale.*
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 topologique X ainsi obtenue est donnée par $V(x)$ l'ensemble des ouverts de X_n contenant x lorsque $x \in X_n$ et l'ensemble $V(*)$ des complémentaires des compacts de $\bigcup_n X_n$. On considère la dynamique T sur X coïncidant 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.

Chapitre 7

Dynamique Symbolique

Sommaire

1. Minimalité et unique ergodicité

Nous avons défini dans l'introduction le système dynamique topologique donné par le décalage complet sur $\{0,1\}^{\mathbb{Z}}$. De même pour tout ensemble S fini on considère le décalage complet σ sur l'ensemble des suites à valeurs dans S. Cet ensemble s'identifie au produit $S^{\mathbb{Z}}$. Lorsque S est muni de la topologie discrète, la topologie produit fait de $S^{\mathbb{Z}}$ un espace compact (en fait un ensemble de Cantor). De plus la métrique d suivante est compatible avec cette topologie :

$$d\left((u_n)_n, (v_n)_n\right) = \sum_{n\in\mathbb{Z}} \frac{\delta_{u_n,v_n}}{2^{|n|}},$$

où $\delta_{x,y}$ désigne le symbole de Kronecker, $\delta_{x,y} = 0$ si $x = y$, sinon 1. Le décalage σ sur $S^{\mathbb{Z}}$ qui à $(u_n)_n$ associe la suite $(u_{n+1})_n$ est continue.

La partition de $S^{\mathbb{Z}}$ donnée par les cylindres d'ordre 1, c'est à dire les ensembles de la forme $[i] = \{(u_n)_n,\ u_0 = i\}$ pour $i \in S$, est appelée la partition de coordonnée zéro et sera notée α_0. Cette partition est aussi un recouvrement en ouvert-fermé de $S^{\mathbb{Z}}$.

Un sous-décalage est une dynamique induite sur un sous-ensemble Y fermé σ-invariant d'un décalage complet $S^{\mathbb{Z}}$. Pour tout entier n, un élément $i_0...i_{n-1}$ de S^n est appelé un n-mot de Y si $[i_0...i_{n-1}] \cap Y \neq \emptyset$. On note $\mathcal{L}_n(Y)$ l'ensemble des n-mots de Y et $\mathcal{L}(Y) = \bigcup_n \mathcal{L}_n(Y)$ l'ensemble des mots de Y. De même pour $u \in S^{\mathbb{Z}}$ on note $\mathcal{L}(u)$ l'ensemble des mots de u, i.e. les éléments de $\bigcup_k S^k$ de la forme $u_p u_{p+1} ... u_q$ pour $p \leq q \in \mathbb{Z}$. On peut aussi définir des sous-décalages de $S^{\mathbb{N}}$ mais on se restreindra ici à des dynamiques symboliques inversibles.

Pour les sous-décalages transitifs on a le critère suivant de minimalité :

PROPOSITION 7.1. *Soit (Y, σ) un sous-décalage transitif et soit $u \in Y$ d'orbite dense, i.e. $\overline{\{\sigma^l(u),\ l \in \mathbb{Z}\}} = Y$. Alors (Y, σ) est minimal si et seulement si tout mot de u apparait dans u avec des sauts bornés :*

$$\forall w \in \mathcal{L}(u)\ \exists k > 0 \text{ tel que}$$

$$\forall n \in \mathbb{Z}\ \exists m \text{ avec } |m - n| \leq k, \text{ on ait}$$

$$w = u_m u_{m+1} ... u_{|w|+m-1}$$

DÉMONSTRATION. Tout d'abord l'orbite de u étant dense, les cylindres $[w]_p := \{(x_n)_n \in Y,\ x_p ... x_{|w|+p-1} = w\}$ pour $p \in \mathbb{Z}$ associés aux mots $w \in \mathcal{L}(u)$ forme une base d'ouverts. Le fait que tout mot de u apparait dans u avec des sauts bornés peut se réécrire de la façon suivante : il existe un entier k tel que pour tout cylindre $[w]_p$

$$\bigcup_{0 \leq l \leq k} \sigma^{-l}[w]_p \supset \{\sigma^q(u),\ q \in \mathbb{Z}\}.$$

Les cylindres étant fermés et l'orbite de u étant dense, ceci est est équivalent à

$$\bigcup_{0 \leq l \leq k} \sigma^{-k}[w]_p = Y.$$

Il s'agit du dernier critère de minimalité de la Proposition 4.9. □

L'unique ergodicité se traduit de la façon suivante :

PROPOSITION 7.2. *Soit (Y, σ) un sous-décalage transitif et soit $u \in Y$ d'orbite dense. Le sous-décalage (Y, σ) est uniquement ergodique si et seulement si tout $w \in \mathcal{L}(u)$ apparait dans u avec une certaine fréquence uniforme, i.e. la suite*

$$\frac{1}{N} \sharp \{0 \leq n \leq N,\ u_{n+j} ... u_{n+j+|w|-1} = w\}$$

converge uniformément en $j \in \mathbb{N}$ quand $N \to +\infty$.

DÉMONSTRATION. Puisque les fonctions continues données par les indicatrices $\chi_{\sigma^{-l}[w]_p}$ pour $w \in \mathcal{L}(u)$ et $p, l \in \mathbb{N}$ engendrent un sous-espace vectoriel dense de $C(Y)$, l'unique ergodicité s'écrit dans ce cas

$$\forall w \in \mathcal{L}(u)\ \forall p, l \in \mathbb{N}\ \exists a_w \in \mathbb{R},\ \text{tels que}$$
$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{[w]_p}(\sigma^{k+l} y) \xrightarrow{n} a_w$$

uniformément en $y \in Y$. L'orbite de u étant dense, la convergence uniforme sur l'orbite de u est équivalent à la convergence uniforme sur Y. □

A l'aide de ces critères de minimalité et d'unique ergodicité, nous construisons une dynamique symbolique minimale mais pas uniquement ergodique :

i) *Construction du mot u :* On définit par récurrence des suites $(u_k)_k$ et $(v_k)_k$ de mot finis dans l'alphabet $\{0,1\}$. On pose $u_0 = 01$ et $v_0 = 001$ puis pour tout entier $k > 0$ on définit par concaténation [1]

$$\begin{aligned} u_{k+1} &= u_k^{m_k} v_k u_k^{m_k}, \\ v_{k+1} &= v_k^{n_k} u_{k+1} v_k^{n_k}, \end{aligned}$$

où $(m_k)_k$ et $(n_k)_k$ sont des suites d'entiers choisis de sorte que pour $k > 0$ la proportion de 0 dans u_k (resp. v_k) soit strictement inférieure à 5/9 (resp. strictement supérieure à 11/18). Pour $k > 1$ on voit les mots u_k et v_k commes des suites indexées sur des intervalle d'entiers centrés en 0. Comme ces suites finies coïncident sur l'intersection de leurs intervalles de définitions, on peut définir une suite $U \in \{0,1\}^{\mathbb{Z}}$ comme la limite terme à terme des suites $(u_k)_k$ et $(v_k)_k$.

ii) *Minimalité :* Si w est un mot de U, alors il apparait dans u_k et v_k pour un certain k. Mais les mots concaténés à u_k pour obtenir U après l'étape k de la récurrence sont tous des concaténations de u_k et v_k. Par conséquent w apparait avec des sauts bornés dans U.

iii) *Non unique ergodicité :* Si (Y, σ) avec $Y = \left\{\overline{\sigma^l(U),\ l \in \mathbb{Z}}\right\}$ était uniquement ergodique, alors d'après le théorème ergodique uniforme (Théorème 4.11) la suite $(S_n)_n$ définie pour tout $n \in \mathbb{N}$ par $S_n := \frac{1}{n}\sum_{k=0}^{n-1} \chi_{[0]}(\sigma^l U)$ pour tout $n \in \mathbb{N}^*$ devrait converger vers $\mu([0])$ avec μ l'unique mesure σ-invariante, quand n tend vers l'infini. Mais par construction on a $\underline{\lim}_n S_n \leq 5/9$ et $\overline{\lim}_n S_n \geq 11/18$.

2. Entropie

L'entropie d'une dynamique symbolique s'exprime facilement en fonction de la croissance du nombre de n-mots.

1. Pour un mot fini u et un entier $k > 0$, on note u^k la concaténation $\underbrace{u \cdots u}_{k\text{fois}}$.

PROPOSITION 7.3. *Soit (Y, σ) un sous-décalage de $S^{\mathbb{Z}}$. La partition de coordonnée zéro induite sur Y est un générateur topologique. En particulier tout sous-décalage est expansif.*

DÉMONSTRATION. La partition itérée $\bigvee_{k=-N}^{N} \sigma^{-k}\alpha_0$ est un recouvrement en ouvert-fermés formé par les *cylindres*

$$[i_{-N}...i_N] := \{(x_n)_n,\ x_k = i_k \text{ pour } k = -n, ..., n\}$$

Or le diamètre de ces cylindres pour la distance d ci-dessus est inférieur à $1/2^{N-1}$. $\square$

Puisque $\sharp\alpha_0^n = \mathcal{L}_n(Y)$ pour tout entier $n > 0$, il suit donc du Corollaire 6.20 :

COROLLAIRE 7.4. *Soit (Y, σ) un sous-décalage de $S^{\mathbb{Z}}$. Alors*

$$h_{top}(\sigma) = \lim_n \frac{\log \sharp\mathcal{L}_n(Y)}{n}.$$

Le décalage complet vérifie la propriété universelle suivante :

THÉORÈME 7.5 (Théorème des générateurs topologiques de Krieger). *Soit (X, f) un système topologique inversible, expansif, apériodique et zéro-dimensionnel,*[2] *alors pour tout S avec $h_{top}(f) < \log \sharp S$ il existe un sous-décalage Y de $S^{\mathbb{Z}}$ topologiquement conjugué à X.*

La preuve s'appuie sur le lemme suivant, appelé *lemme des marqueurs*.

LEMME 7.6. *Soit (X, f) un système topologique inversible, apériodique et zéro-dimensionnel. Alors pour tout entier $N > 0$ il existe un ouvert-fermé U tel que les ensembles f^kU pour $k = 0, ..., N$ sont deux à deux disjoints et $\bigcup_{k=-N}^{N} f^kU = X$.*

DÉMONSTRATION. Par apériodicité tout $x \in X$ admet un voisinage fermé U_x tel que f^kU_x pour $k = 0, ..., N$ sont deux à deux disjoints. Par compacité il existe un sous-recouvrement fini $(U_i)_{i=1,...,p}$ de $(U_x)_{x\in X}$. Puis on définit par récurrence $U_1' = U_1$ et $U_{i+1}' = U_i' \bigcup \left(U_{i+1} \setminus \left(\bigcup_{k=-N}^{N} f^kU_i'\right)\right)$ pour $i < p$. On vérifie facilement que f^kU_i' pour $k = 0, ..., N$ sont deux à deux disjoints et que $\bigcup_{k=-N}^{N} f^kU_i' \supset \bigcup_{1\le j\le i} U_i$. On pose alors $U = U_p'$. $\square$

PREUVE DE THÉORÈME 7.5 : Pour simplifier nous montrons seulement le théorème pour $\log(\sharp S - 1) > h_{top}(f)$. Soit $\mathcal{V}$ un générateur topologique de (X, f) formé d'ouvert-fermés. On choisit N tel que pour $n \geq N$

$$\sharp\mathcal{V}^{n+1} \leq (\sharp S - 1)^n.$$

2. Le compact X admet une base de voisinages constituée d'ouvert-fermé. Autrement dit X est un ensemble de Cantor, i.e. un compact totalement discontinu.

On fixe un point $s \in S$ et on note $S_* = S \setminus \{s\}$. Pour tout $n = N + 1, ..., 2N+1$ on considère une injection ϕ_n de $\mathcal{V}^n$ dans S_*^{n-1}. Pour $x \in X$, on définit une suite $y = \psi(x)$ dans $S^{\mathbb{Z}}$ comme suit. Tout d'abord $y_k = s$ si et seulement si $f^k x$ appartient à U. Puis si $k < l$ sont deux entiers consécutifs avec $y_k = y_l = s$ alors $l > k + N$. On pose ensuite $y_{k+1}...y_{l-1} = \phi_{l-k}(\mathcal{V}^{l-k}(f^k x))$. Les ensembles U et V étant ouvert-fermés l'application ψ est continue. De plus on a clairement $\psi \circ f = \sigma \circ \psi$. Enfin si $\psi(x) = \psi(y)$ alors les applications $(\phi_n)_n$ étant injectives, les points x et y appartiennent au même élément de $\bigvee_{k=-n}^{n} f^k \mathcal{V}$ pour tout n. Ainsi ψ définit une conjugaison topologique de (X, f) vers le sous-décalage $(\psi(X), \sigma)$.

□

En utilisant une construction similaire, on peut aussi montrer :

THÉORÈME 7.7 (Théorème des générateurs ergodiques de Krieger). *Soit S un alphabet fini. Tout système ergodique $(X, \mathcal{A}, f, \mu)$ est isomorphe à $(S^{\mathbb{Z}}, \mathcal{B}, \sigma, \nu)$ pour une mesure ν invariante par le décalage σ pourvu que $h(\mu) < \log \sharp S$.*

3. Sous-décalage de type fini

Un sous-décalage Y de $S^{\mathbb{Z}}$ est dit de type fini lorsqu'il existe un sous-ensemble fini F de $\bigcup_k S^k$ tel que Y coïncide avec l'ensemble des suites $u \in S^{\mathbb{Z}}$ avec $w \notin \mathcal{L}(u)$ pour tout $w \in F$.

Un tel décalage peut être représenté à l'aide d'un graphe fini. Soit $\mathcal{G}$ un graphe orienté avec un ensemble fini T de sommets. On considère l'ensemble $\Sigma(\mathcal{G})$ des suites $(u_n)_n \in T^{\mathbb{Z}}$ telles que $u_n \to u_{n+1}$ pour tout n. Cet ensemble de $T^{\mathbb{Z}}$ est fermé et invariant par le décalage σ et définit donc bien un sous-décalage. Il est de type fini, l'ensemble F des mots interdits correspondant aux 2-mots uv avec $u \nrightarrow v$. En fait on décrit ainsi tout les sous-décalages de type fini.

PROPOSITION 7.8. *Pour tout décalage de type fini (Y, σ), il existe un graphe orienté fini $\mathcal{G}$ tel que (Y, σ) et $(\Sigma(\mathcal{G}), \sigma)$ soient topologiquement conjugués.*

DÉMONSTRATION. Quitte à changer F on peut supposer que tous les mots interdits sont de même longueur $p \in \mathbb{N}$. On définit un graphe $\mathcal{G}$ avec $\mathcal{L}_p(Y) = S^p \setminus F$ pour ensemble de sommets. Les flèches sont définies de la façon suivante :

$$x_1...x_p \to x'_1...x'_p \Leftrightarrow x_1 x'_1...x'_p = x_1...x_p x'_p.$$

On pose $\phi(x) = x_0...x_{p-1} \in S^p \setminus F$ pour tout $x = (x_n)_n \in Y$. L'application $\psi : Y \to \Sigma(\mathcal{G})$ définie par $\psi((x_n)_n) = \big(\phi(\sigma^k x)\big)_k$ est alors une conjugaison topologique. Vérifions la surjectivité de ψ, les autres propriétés étant immédiates. Soit $(u_n)_n \in \Sigma(\mathcal{G})$ montrons que la suite $y = (y_n)_n$ avec y_n la première lettre de u_n est dans Y. Pour cela il suffit de voir que tous les p-mots de y n'appartiennent pas à F, ce qui découle de la définition de $\Sigma(\mathcal{G})$. $\square$

Un graphe fini orienté $\mathcal{G}$ avec un ensemble T de sommets est donné par une matrice $A = (a_{i,j})_{1 \leq i,j \leq s}$ avec $T = \{t_1, ..., t_s\}$ à coefficients dans $\{0, 1\}$ en posant $a_{i,j} = 1$ si et seulement $t_i \to t_j$. Cette matrice est appelée la *matrice d'adjacence* associée à $\mathcal{G}$ et $T = \{t_1, ..., t_s\}$. Pour une telle matrice $A = (a_{i,j})_{1 \leq i,j \leq s}$ on notera $\Sigma_A \subset \{1, ..., s\}^{\mathbb{Z}}$ le sous-décalage de type fini associé au graphe sur $\{1, ..., s\}$ avec $i \to j$ pour $a_{i,j} = 1$. Dans la suite on identifiera ces différents points de vue.

3.1. Transitivité, points périodiques et entropie topologique. Un graphe orienté $\mathcal{G}$ est *connexe* si toute paire (i, j) de sommets peut être relié par un chemin orienté de i vers j. Une matrice A carrée d'ordre s à coefficients positifs ou nuls est dite irréductible lorsque pour tout $1 \leq i, j \leq s$ il existe un entier $n \in \mathbb{N}^*$ avec $(A^n)_{ij} > 0$.

Proposition 7.9. *Les propriétés suivantes sont équivalentes :*

(1) $\mathcal{G}$ est connexe,

(2) A est irréductible,

(3) le sous-décalage de type fini (Y, σ) associé à $\mathcal{G}$ est topologiquement transitif.

Démonstration. Les deux premières assertions sont clairement équivalentes. Supposons (Y, σ) topologiquement transitif. Alors pour tout $i, j \in S$ il existe un entier $n > 0$ tel que $[i] \cap \sigma^{-n}[j] \neq \emptyset$. Il existe donc un chemin orienté de i vers j. Réciproquement supposons $\mathcal{G}$ connexe. Il suffit d'après la Proposition 4.8 de montrer que pour tous ouverts U, V de Y on a $U \cap \sigma^{-n}V$ pour un entier $n > 0$. On peut supposer que U et V sont des cylindres $U = [k_i, ..., k_j]$ et $V = [l_p, ..., l_q]$. Le graphe étant connexe, il existe un chemin orienté de k_j vers l_p. Si m désigne la longueur de ce chemin alors $U \cap \sigma^{i-j-m}V$ est non vide. $\square$

Pour tout entier $n > 0$ on notera $Per_n(\Sigma_A)$ l'ensemble des points n-périodiques du système (Σ_A, σ) :

$$Per_n(\Sigma_A) := \{x \in \Sigma_A, \ \sigma^n(x) = x\}.$$

Proposition 7.10.

$$\sharp Per_n(\Sigma_A) = tr(A^n).$$

Démonstration. Les orbites n-périodiques sont en bijection avec les chemins fermés dans le graphe de longueur n. Le nombre de tels chemins partant de i se lit sur le i^{eme} terme de la diagonale de A^n. En particulier $\sharp Per_n(\Sigma_A) = tr(A^n)$. $\square$

L'entropie topologique s'exprime aussi simplement en fonction de la matrice d'adjacence. On note $\rho(A)$ le rayon spectral de la matrice A, qui rappelons-le est donné par le maximum des modules des valeurs propres de A.

PROPOSITION 7.11.

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

DÉMONSTRATION. On considère la norme suivante sur $M_s(\mathbb{R})$

$$\|(b_{i,j})_{i,j}\| := \sum_{i,j} |b_{i,j}|.$$

On voit alors facilement que

$$\sharp \mathcal{L}_n(Y) = \|A^n\|.$$

Or d'après le théorème du rayon spectral, on a pour toute norme $\lim_n \frac{\log \|A^n\|}{n} = \log \rho(A)$. □

Lorsque A est primitive on a $\lim_n \frac{1}{n} \log tr(A^n) = \log \rho(A)$ et par conséquent l'entropie topologique de Σ_A coïncide avec le taux de croissance exponentielle du nombre de points périodiques.

COROLLAIRE 7.12.

$$h_{top}(\Sigma_A) = \lim_n \frac{1}{n} \log \sharp Per_n(\Sigma_A).$$

3.2. Mesure markovienne. Une matrice $P = (p_{ij})_{i,j} \in M_s([0,1])$ est dite stochastique lorsque

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

Une probabilité π sur $\{1, ..., s\}$, donnée par un vecteur ligne $\pi = (\pi_1, ..., \pi_s)$, est dite stationnaire lorsque $\pi P = \pi$. On appelle mesure markovienne, ou mesure de Markov, la mesure invariante μ_π associée à une probabilité stationnaire π définie pour tout cylindre par

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

On peut associer à toute matrice stochastique P le sous-décalage de type fini dont la matrice d'adjacence $A = (a_{ij})_{ij}$ est définie par $a_{ij} = 1$ si et seulement si $p_{ij} > 0$. La mesure de probabilité μ_π est une mesure invariante de (Σ_A, σ). Une matrice carrée à termes réels positifs est dite *primitive*, lorsque les termes d'une de ses puissances sont tous strictement positifs. En particulier une matrice primitive est irréductible.

THÉORÈME 7.13. *(Perron-Froebenius) Il existe toujours des probabilités stationnaires π relativement à P. De plus celle-ci est unique lorsque A_P est irréductible et $\frac{1}{n}\sum_{k=0}^{n-1}\psi P^k \xrightarrow{n} \pi$ pour tout vecteur de probabilité ψ. Si A_P est primitive, on a alors $\psi P^n \xrightarrow{n} \pi$ pour tout vecteur de probabilité ψ.*

DÉMONSTRATION. Toute limite d'une suite de la forme

$$\left(\frac{1}{n}\sum_{k=0}^{n-1}\psi P^k\right)_n,$$

avec ψ un vecteur de probabilité, est stationnaire. Lorsque A_P est irréductible on voit facilement que tout vecteur de probabilité stationnaire a ses coordonnées strictement positives. L'unicité de la probabilité stationnaire dans le cas irréductible (resp. la convergence forte dans le cas primitif) résulte alors du théorème de Perron-Froebenius [3](cf Rappels Théorème 8.21), qui garantit pour une matrice positive irréductible (resp. primitive) 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 7.14. *Soit P une matrice stochastique et π une mesure de probabilité stationnaire. Alors la mesure de Markov associée à (π, P) est :*
- *ergodique si et seulement si A_P est irréductible,*
- *mélangeante (même Bernoulli, en fait) si et seulement si A_P est primitive.*

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 7.13, dans le cas primitif, $(P^{n-j})_{k_jl_0}$ 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 si μ_π est mélangeante il existe un entier n tel que $\mu_\pi([m] \cap \sigma^{-n}[l]) = \pi_m(P^n)_{ml} > 0$ pour tous m, l. La matrice A_P est donc primitive.

3. On rappelle que $\rho(P) = \rho(P^t) = 1$ pour une matrice stochastique P.

Considérons enfin le cas ergodique. On suppose A_P irréductible. Pour établir l'ergodicité il suffit de vérifier d'après le Corollaire 3.5 que pour deux cylindres A et B on a

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu_\pi(A\cap\sigma^{-n}B)\xrightarrow{n}\mu_\pi(A)\mu_\pi(B)$$

Cela découle du calcul précédent et du fait que $\frac{1}{n}\sum_{k=0}^{n-1}\psi P^k\xrightarrow{n}\pi$ pour tout ψ. Enfin si μ_π est ergodique, en prenant $A=[m]$ et $B=[l]$, on a $\mu_\pi(A\cap\sigma^{-k}B)=\pi_m(P^k)_{ml}>0$ pour un certain k et donc A_P est irréductible. □

L'entropie d'une mesure de Markov est donnée par la formule suivante :

Proposition 7.15. *L'entropie d'une mesure de Markov avec probabilité stationnaire $(\pi_i)_i$ et matrice de transition $(p_{ij})_{i,j}$ est égale à $-\sum_{i,j}\pi_i p_{ij}\log p_{ij}$,*

Démonstration. On suppose la mesure de Markov ergodique (laissant le cas général en exercice). La partition de coordonnée zéro α_0 de (Σ_A,σ) étant génératrice, on a $h(\mu_\pi)=h(\mu_\pi,\alpha_0)$. D'après la formule de Shanon-McMillman-Breiman, on a pour μ_π presque tout $x\in\Sigma_A$

$$h(\mu_\pi,\alpha_0)=\lim_n -\frac{1}{n}\log\mu_\pi(\alpha_0^n(x)).$$

Puis par définition de μ_π on calcule avec $\alpha_0^n(x)=[x_0...x_{n-1}]$:

$$\begin{aligned}-\frac{1}{n}\log\mu_\pi(\alpha_0^n(x)) &= -\frac{1}{n}\log\left(\pi_{x_0}\prod_{k=0}^{n-2}p_{x_kx_{k+1}}\right),\\ &= -\frac{\log\pi_{x_0}}{n}+\frac{1}{n}\sum_{k=0}^{n-2}\phi(\sigma^k(x)),\end{aligned}$$

où l'on a noté $\phi:\Sigma_A\to\mathbb{R}$ la fonction qui à $x=(x_n)_n$ associe $-\log p_{x_0x_1}$. La fonction ϕ est bornée car elle prend un nombre fini de valeurs. Elle est donc intégrable relativement à μ_π. En appliquant le théorème ergodique ponctuel à ϕ, on obtient pour μ presque tout x

$$\begin{aligned}\lim_n\frac{1}{n}\sum_{k=0}^{n-2}\phi(\sigma^k(x)) &= \int\phi\ d\mu_\pi,\\ &= -\sum_{i,j}\pi_i p_{i,j}\log p_{i,j}.\end{aligned}$$

□

3.3. Mesure d'entropie maximale. Soit $A \in M_s(\{0,1\})$ une matrice irréductible. D'après le théorème de Perron-Froebenius, il existe $u = (u_1, ..., u_s)$ et $v = (v_1, ..., v_s)$ des vecteurs propres strictement positifs à gauche et à droite associés à la valeur propre maximale $\lambda = \rho(A) > 0$ (notez que u et v sont définis à une constante positive multiplicative près).

Définition 7.16. *La mesure de Parry du sous-décalage de type fini Σ_A est la mesure Markovienne μ_π associé à la matrice stochastique $P = (p_{ij})$ et à la probabilité stationnaire $\pi = (\pi_i)$ définie pour $1 \leq i, j \leq s$ par*

$$p_{ij} = \frac{A_{ij} v_j}{\lambda v_i},$$
$$\pi_i = \frac{u_i v_i}{\sum_j u_j v_j}.$$

On vérifie par un calcul élémentaire que cette définition fait bien sens, i.e. π définit bien une probabilité et P est une matrice stochastique. De plus P et π ne dépendent pas du choix de u et v. On détaille maintenant le calcul de son entropie. On rappelle que l'entropie topologique de Σ_A est donnée par $\log \lambda$.

Proposition 7.17. *La mesure de Parry μ_A associée à A est d'entropie maximale, i.e. $h(\mu_A) = \log \lambda$.*

Démonstration. D'après la Proposition 7.15, on a

$$\begin{aligned}
h(\mu_A) &= -\sum_{i,j} \pi_i p_{i,j} \log p_{i,j}, \\
&= -\frac{1}{\sum_j u_j v_j} \sum_{i,j} u_i v_i \frac{A_{ij} v_j}{\lambda v_i} \log\left(\frac{A_{ij} v_j}{\lambda v_i}\right), \\
&= -\frac{1}{\sum_j u_j v_j} \left(\sum_{i,j} u_i \frac{A_{ij} v_j}{\lambda} \log A_{i,j} + u_i \frac{A_{ij} v_j}{\lambda} \log\left(\frac{v_j}{\lambda v_i}\right)\right).
\end{aligned}$$

La première double somme est nulle car $A_{i,j}$ prend ses valeurs dans $\{0,1\}$.

$$\begin{aligned}
h(\mu_A) &= -\frac{1}{\sum_j u_j v_j} \sum_{i,j} u_i \frac{A_{ij} v_j}{\lambda} \log\left(\frac{v_j}{\lambda v_i}\right), \\
&= -\frac{1}{\sum_j u_j v_j} \left(\sum_{i,j} u_i \frac{A_{ij} v_j}{\lambda} (\log(v_j) - \log(v_i)) - \sum_{i,j} u_i \frac{A_{ij} v_j}{\lambda} \log \lambda \right).
\end{aligned}$$

Le vecteur v étant vecteur propre à droite, on a $\sum_j A_{ij} v_j = \lambda v_i$ pour tout i et de même $\sum_i u_i A_{ij} = \lambda u_j$. On en déduit que la première double somme est nulle, puis on obtient alors $h(\mu_A) = \log \lambda$. $\square$

La mesure de Parry μ_A satisfait aussi la propriété suivante, dite *de Gibbs.*

PROPOSITION 7.18. *Il existe* $0 < a < b$ *tels que pour tout cylindre* C_n *d'ordre* n *on a*

$$\frac{a}{\lambda^n} \leq \mu_A(C_n) \leq \frac{b}{\lambda^n}. \tag{3.1}$$

DÉMONSTRATION. On calcule facilement pour $C_n = [a_1...a_n]$:

$$\begin{aligned}\mu_A([a_1...a_n]) &= \pi_{a_1}\prod_{l=1}^{n-1} p_{a_l a_{l+1}},\\ &= \pi_{a_1}\frac{v_{a_n}}{\lambda^n v_{a_1}}.\end{aligned}$$

□

THÉORÈME 7.19. *La mesure de Parry* μ_A *est l'unique mesure d'entropie maximale de* (Σ_A, σ).

DÉMONSTRATION. On raisonne par l'absurde. Soit m autre mesure d'entropie maximale. On peut supposer m ergodique par harmonicité de l'entropie. D'après le Lemme 4.5 les mesures ergodiques m et μ_A sont orthogonales, i.e. il existe un ensemble borélien E avec $\mu_A(E) = 0$ et $m(E) = 1$. Par régularité de la mesure de probabilité borélienne m il existe un compact $K \subset E$ avec $m(K) \simeq 1$. Pour tout n on note E_n l'ensemble des cylindres de $\bigvee_{k=-n}^{n} \sigma^{-k}\alpha_0$ rencontrant K. En notant encore E_n l'union de ces cylindres on a alors $K = \bigcap_n^{\searrow} E_n$ et donc pour n grand on a $\mu_A(E_n) \simeq 0$ et $m(E_n) \simeq 1$. On a alors avec α_0 la partition génératrice de coordonnée zéro

$$\begin{aligned}h(m) &= \log\lambda,\\ &= h(m,\alpha_0),\\ &= \inf_n \frac{1}{2n+1} H_m\left(\bigvee_{k=-n}^{n} \sigma^k\alpha_0\right).\end{aligned}$$

Par conséquent on a pour tout n avec $Y = \Sigma_A$

$$\begin{aligned}
(2n+1)\log\lambda &\leq H_m(\{E_n, Y\setminus E_n\}) + H_m(\bigvee_{k=-n}^{n}\sigma^k\alpha_0|\{E_n, Y\setminus E_n\}),\\
&\leq H_m(\{E_n, Y\setminus E_n\}) + m(E_n)H_{m_{E_n}}\left(\bigvee_{k=-n}^{n}\sigma^k\alpha_0\right)\\
&\quad + m(Y\setminus E_n)H_{m_{Y\setminus E_n}}\left(\bigvee_{k=-n}^{n}\sigma^k\alpha_0\right).
\end{aligned}$$

Puis d'après la Proposition 7.18 on a

$$\frac{a}{\lambda^{2n+1}}\sharp E_n \leq \mu_A(E_n) \leq \frac{b}{\lambda^{2n+1}}\sharp E_n$$

et donc avec un $O(1)$ dépendant uniquement de a et b

$$\log\sharp E_n = (2n+1)\log\lambda + \log\mu_A(E_n) + O(1).$$

De même on obtient

$$\log\sharp\left((\bigvee_{k=-n}^{n}\sigma^k\alpha_0)\setminus E_n\right) = (2n+1)\log\lambda + \log\left(1-\mu_A(E_n)\right) + O(1).$$

Enfin en utilisant l'inégalité $H_{m_Z}(\beta) \leq \log\sharp\left(\beta\cap Z\right)$ pour tout $Z\subset X$ et pour toute partition finie β, nous obtenons

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

Ceci constitue une contradiction car en choisissant K et n convenablement ce dernier terme est strictement négatif. □

Les points périodiques de Σ_A s'équidistribuent le long de la mesure d'entropie maximale μ_A dans le cas primitif :

THÉORÈME 7.20. *On suppose A primitive. On a alors la convergence faible-$*$ suivante :*

$$\frac{1}{\sharp Per_n(\Sigma_A)}\sum_{x\in Per_n(\Sigma_A)}\delta_x \overset{n}{\to} \mu_A.$$

DÉMONSTRATION. Pour tout entier $n>0$ on note μ_n la mesure atomique σ-invariante équidistribuée sur les points n-périodiques. Alors pour tout cylindre $[i_1, ..., i_k]$ on a

$$\begin{aligned} \mu_n([i_1, ..., i_k]) &= \frac{\sharp\left(Per_n(\Sigma_A) \cap [i_1, ..., i_k]\right)}{\sharp Per_n(\Sigma_A)}, \\ &\overset{n \geq k}{=} \frac{\left(A^{n-k}\right)_{i_k, i_1}}{tr(A^n)} \end{aligned}$$

La matrice A étant primitive, l'espace propre associé à $\rho(A)$ est de dimension 1 et les autres valeurs propres sont de modules strictement inférieurs à 1. En particulier on a $tr(A^n) \sim^n \lambda^n$ et $\left(\frac{A}{\lambda}\right)^n$ converge vers une matrice $\tilde{A}$ de rang 1, quand n tend vers l'infini. Les colonnes de $\tilde{A}$ sont vecteurs propres de A donc toutes colinéaires à v. De même les vecteurs ligne de $\tilde{A}$ sont colinéaires à u. Autrement dit, on peut trouver deux vecteurs α et β tels que

$$\tilde{A} = v\alpha^t \text{ et } \tilde{A} = \beta u^t.$$

Mais on a clairement $\tilde{A}v = v$ puis $\beta u^t v = (\sum_l u_l v_l)\beta = v$ et donc pour tout i, j on a $(\tilde{A})_{i,j} = \frac{u_j v_i}{\sum_l u_l v_l}$. On en conclut que

$$\mu_n([i_1, ..., i_k]) \xrightarrow{n} \frac{u_{i_1} v_{i_m}}{\lambda^k \sum_l u_l v_l} = \mu_A([i_1, ..., i_k]).$$

$\square$

On aurait aussi pu montrer que toute limite faible de $\left(\frac{1}{\sharp Per_n(\Sigma_A)} \sum_{x \in Per_n(\Sigma_A)} \delta_x\right)_n$ est d'entropie $\log \lambda$.

4. Exercices

Exercice 60.
Montrer que l'ensemble des points périodiques d'un sous-décalage de type fini transitif est dense puis que l'ensemble des mesures périodiques est dense dans l'ensemble des mesures ergodiques.

Exercice 61.
Montrer directement avec la définition utilisant l'entropie statique que l'entropie d'une mesure markovienne quelconque associée à $\pi = (\pi_i)_i$ et $P = (p_{ij})_{i,j}$ est donnée par $-\sum_{i,j} \pi_i p_{ij} \log p_{ij}$.

Exercice 62.
Considérons la matrice

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

et le sous-décalage de type fini Σ_A associé.

(1) Calculer le nombre de points n-périodiques.
(2) Déterminer l'entropie topologique de Σ_A.
(3) Montrer qu'il existe une unique mesure markovienne μ associée à la matrice de transition

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Calculer $\mu([1])$, $\mu([2])$,$\mu([3])$ et $h(\mu)$.
(4) Soit μ_A la mesure de Parry de Σ_A. Calculer $\mu_A([1])$, $\mu_A([2])$,$\mu_A([3])$ et $h(\mu_A)$.

Exercice 63. *Code bloc.*
Soient X un sous-décalage et $\mathcal{A}$ un alphabet fini. Une application $c : X \to \mathcal{A}^{\mathbb{Z}}$ est appelé un code bloc lorsqu'il existe $k, l \in \mathbb{N}$ et une application $\alpha : \mathcal{L}_{l+k+1}(X) \to \mathcal{A}$ avec $c(x)_i = \alpha(x_{-k}...x_l)$ pour tout $i \in \mathbb{Z}$ et tout $x = (x_n)_{n\in\mathbb{Z}} \in X$.

(1) Montrer que tout code bloc c définit une semi-conjugaison de X dans $\mathcal{A}^{\mathbb{Z}}$.
(2) Montrer que toute semi-conjugaison d'un sous-décalage X dans un autre Y est donné par un code bloc.

Exercice 64. *Représentation par les arrêtes.*
On considère un graphe orienté fini $\mathcal{G}$ avec possiblement plusieurs arrêtes reliant deux sommets donnés. On note $\mathcal{E}$ l'ensemble des arrêtes et $\Sigma_e(\mathcal{G})$ l'ensemble des suites d'arêtes qui décrivent un chemin infini de $\mathcal{G}$. Le décalage sur $\Sigma_e(\mathcal{G})$ est appelé un décalage arrête.

(1) Montrer que le décalage arrête $\Sigma_e(\mathcal{G})$ est un sous-décalage de type fini.
(2) Réciproquement montrer que tout sous-décalage de type fini est topologiquement conjugué à un tel décalage arrête.

Exercice 65. *Equivalence de décalage forte.*
Pour un graphe orienté fini $\mathcal{G}$ sur $\{1, ..., n\}$ (avec possiblement plusieurs arrêtes reliant deux sommets donnés) on considère la matrice carré $B_{\mathcal{G}} = (a_{ij})_{i,j}$ d'ordre n avec $a_{i,j}$ le nombre d'arêtes reliant (directement) i à j. Soient deux graphes $\mathcal{G}$ et $\mathcal{G}'$ on suppose qu'il existe des matrices (pas nécessairement carrées) U et V satisfaisant $B_{\mathcal{G}} = UV$ et $B_{\mathcal{G}'} = VU$. Montrer que les décalages arrêtes associés à $\mathcal{G}$ et $\mathcal{G}'$ sont topologiquement conjugués.

Exercice 66. *Système sturmien.*
Soit $\mathcal{A}$ un ensemble fini. Pour une suite $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ et $n \in \mathbb{N}^*$ on note $\mathcal{L}_n(u)$ l'ensemble des mots de longueurs n dans u, i.e.

$$\mathcal{L}_n(u) := \{u_k u_{k+1} ... u_{k+n-1} \in \mathcal{A}^n,\ k \in \mathbb{N}\}.$$

On appelle complexité de u la suite $(p_u(n))_{n\in\mathbb{N}}$ définie pour tout n par le cardinal de $\mathcal{L}_n(u)$:

$$p_u(n) = \sharp\mathcal{L}_n(u).$$

On dit que la suite u est prépériodique lorsque il existe des entiers $l \geq 0$ et $p \geq 1$ tels que

$$\sigma^{l+p}(u) = \sigma^l(u).$$

(1) Montrer que pour tout $n, l \in \mathbb{N}$ on a

$$p_{\sigma^l(u)}(n) \leq p_u(n) \leq l + p_{\sigma^l(u)}(n).$$

En déduire que la complexité d'une suite prépériodique est une suite bornée.

(2) Dans cette question et la suivante on suppose que p_u n'est pas strictement croissante. Montrer que p_u est bornée puis qu'il existe des entiers l et n tel que $u_k u_{k+1} ... u_{k+n-1} \mapsto u_{k+1} ... u_{k+n}$ est une bijection de $\mathcal{L}_n(\sigma^l(u))$.

(3) En déduire que u est prépériodique. Conclure que pour toute suite non prépériodique u on a $p_u(n) \geq n + 1$ pour tout n.

Soit $\alpha \in [0, 1]$ irrationnel. On considère la suite $u = (u_n)_n \in \{0, 1\}^{\mathbb{N}}$ de terme général

$$u_n = [(n + 1)\alpha] - [n\alpha].$$

On note R la rotation d'angle α sur le cercle unité $\mathbb{R}/\mathbb{Z}$ et P la partition formée par les deux arcs de cercle $I_0 = [0, 1 - \alpha[$ et $I_1 = [1 - \alpha, 1[$.

(4) Vérifier que $u_n = 1$ si et seulement si $R^n(0) \in I_1$.

(5) Montrer que la partition $R^{-1}P \vee ... \vee R^{-n}P$ a $n + 1$ éléments.

(6) Conclure que $p_u(n) = n + 1$ pour tout n.

Chapitre 8

Partition de Markov pour les automorphismes linéaires hyperboliques

Sommaire

On considère dans ce chapitre un automorphisme *hyperbolique* linéaire f_A du tore $\mathbb{R}^n/\mathbb{Z}^n$, i.e. la matrice A n'a pas de valeur propre de module 1. Or le déterminant de A, qui est égal au produit des valeurs propres avec multiplicité, vaut 1. Il existe donc au moins une valeur propre de A de module strictement inférieur à 1 et une autre de module supérieur à 1.

1. Généralités

1.1. Normes adaptées. On notera E_s (resp. E_u) l'espace vectoriel réel donné par la somme des sous-espaces caractéristiques associés aux valeurs propres de A de module strictement inférieur à 1 (resp. strictement supérieur à 1). Rappelons que si Π_A désigne le polynôme minimal de A, alors $E_u = Ker(P(A))$ et $E_s = Ker(Q(A))$ avec $\Pi_A = QR$, où P et Q sont des polynômes réels non constants ayant des racines de module respectivement >1 et <1 (en particulier P et Q sont premiers entre eux et donc $\mathbb{R}^n = E_u \oplus E_s$). Ces deux sous-espaces vectoriels réels sont invariants par l'application linéaire A. On note $\lambda_s = \max\{|\lambda|,\ \lambda \text{ valeur propre de } A \text{ avec } |\lambda| < 1\} < 1$ et de même $\lambda_u = \min\{|\lambda|,\ \lambda \text{ valeur propre de } A \text{ avec } |\lambda| > 1\} > 1$. On a $\rho(A|_{E_s}) = \lambda_s$ et $\rho(A^{-1}|_{E_u}) = 1/\lambda_u$. En dimension deux, on a $\|Av_u\|(=) \geq \lambda_u \|v_u\|$ et $\|Av_s\|(=) \leq \lambda_s \|v_s\|$ pour tout $(v_u, v_s) \in E_u \times E_s$. En dimension supérieure de telles inégalités ne sont plus vraies pour la norme euclidienne $\|\cdot\|$ et il est alors commode de travailler avec une autre norme, qui dilate dans la direction de E_u et contracte dans celle de E_s.

Lemme 8.1. *Pour tout* $(1<)\lambda < \min(\lambda_u, 1/\lambda_s)$ *il existe une norme* $\|\cdot\|_\lambda \geq \|\cdot\|$ *sur* $\mathbb{R}^n$ *telle que*

$$\forall v_u \in E_u, \quad \|Av_u\|_\lambda \geq \lambda \|v_u\|_\lambda$$
$$\forall v_s \in E_s, \quad \lambda \|Av_s\|_\lambda \leq \|v_s\|_\lambda.$$

Démonstration. Il suffit de considérer (les séries suivantes convergent d'après le théorème du rayon spectral)

$$\forall v_u \in E_u, \quad \|v_u\|_u = \sum_{n\in\mathbb{N}} \lambda^n \|A^{-n} v_u\|$$
$$\forall v_s \in E_s, \quad \|v_s\|_s = \sum_{n\in\mathbb{N}} \lambda^n \|A^{n} v_s\|$$

puis finalement

$$\|v_u + v_s\|_\lambda = \|v_u\|_u + \|v_s\|_s.$$

□

On notera d_λ la distance induite sur le tore :

$$\forall x, y \in \mathbb{T}^n,\ d_\lambda(x,y) = \inf_{X,Y} \|X - Y\|_\lambda,$$

où l'infimum porte sur les points X et Y de $\mathbb{R}^n$ se projetant sur x et y respectivement dans $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

1.2. Rectangles de $\mathbb{R}^n$. Un rectangle de $\mathbb{R}^n$ est un produit de la forme $R = R_u \times R_s \subset E_u \times E_s$ avec $R_u \subset E_u$ et $R_s \subset E_s$, que l'on verra comme un sous-ensemble de $\mathbb{R}^n$ via l'identification $E_u \times E_s \simeq E_u \oplus E_s = \mathbb{R}^n$. Lorsque R est un rectangle, l'ensemble $A(R) = A(R_u) \times A(R_s)$ est aussi un rectangle. Les rectangles sont aussi stables par intersection finie.

Lemme 8.2. *On considère des rectangles* $R^{-N}, ..., R^N$ *de* $\mathbb{R}^n$. *Alors l'ensemble* $\bigcap_{i=-N}^{N} A^{-i}R^i$ *est un rectangle (éventuellement vide) de diamètre pour* $\|\cdot\|_\lambda$ *inférieur à*

$$\big(diam(R^{-N}) + diam(R^N)\big)\, \lambda^{-N}.$$

DÉMONSTRATION. L'ensemble $\bigcap_{i=-N}^{N} A^{-i}R^i = \left(\bigcap_{i=-N}^{N} A^{-i}R_u^i\right) \times \left(\bigcap_{i=-N}^{N} A^{-i}R_s^i\right)$ définit un rectangle contenu dans le produit $A^{-N}R_u^N \times A^N R_s^{-N}$ qui est de diamètre pour $\|\cdot\|_\lambda$ inférieur à $diam(A^{-N}R_u^N) + diam(A^N R_s^{-N}) \leq \left(diam(R^N) + diam(R^{-N})\right)\lambda^{-N}$. □

Pour deux rectangles $R = R_u \times R_s$ et $T = T_u \times T_s$, on dira que R traverse T lorsque $R_u \supset T_u$ et $R_s \subset T_s$.

LEMME 8.3. *On considère des rectangles $R^{-N}, ..., R^N$, tels que $A(R^i)$ traverse R^{i+1} pour tout $i = -N, ..., N-1$. Alors l'ensemble $\bigcap_{i=-N}^{N} A^{-i}R^i$ est un rectangle non vide.*

DÉMONSTRATION. En notant $R^i = R_u^i \times R_s^i$, pour tout $i = -N, ..., N-1$ on a :

$$A(R_u^i) \supset R_u^{i+1} \text{ et } A(R_s^i) \subset R_s^{i+1}$$

L'ensemble $\bigcap_{i=-N}^{N} A^{-i}R^i$ coïncide alors avec le produit $A^{-N}R_u^N \times A^N R_s^{-N}$. En particulier il est non vide. □

Dans la suite on appellera moyen, (resp. petit, minuscule) rectangle tout rectangle R avec un diamètre (pour la norme euclidienne) inférieur à $\frac{1}{2}$ (resp. $\min\left(\frac{1}{2\|A\|}, \frac{1}{2\|A^{-1}\|}\right)$, $\min\left(\frac{1}{8\|A\|}, \frac{1}{8\|A^{-1}\|}\right)$). Ainsi un moyen rectangle se projette bijectivement sur le tore $\mathbb{R}^n/\mathbb{Z}^n$. Pour un petit rectangle R, les ensembles $A(R)$ et $A^{-1}(R)$ sont des moyens rectangles. De plus deux minuscules rectangles d'intersection non vide sont inclus dans un même petit rectangle.

1.3. Rectangles de $\mathbb{T}^n$ et expansivité de f_A. On identifie tout moyen (resp. petit, minuscule) rectangle $R \subset \mathbb{R}^n$ avec son image R dans le tore $\mathbb{T}^n$, qu'on appelle un moyen (resp. petit, minuscule) rectangle du tore.

LEMME 8.4. *Pour tout $\bigcap_{k=-N}^{N} f_A^{-k}\mathsf{R}^k$ avec $\mathsf{R}^{-N}, ..., \mathsf{R}^N$ des petits rectangles du tore, il existe un unique élément $(l_k)_{k=-N,...,N} \in (\mathbb{Z}^n)^{2N+1}$ avec $l_0 = 0$ tel que le rectangle $\bigcap_{k=-N}^{N} A^{-k}(R^k + l_k)$ de $\mathbb{R}^n$ se projette bijectivement sur $\bigcap_{k=-N}^{N} f_A^{-k}\mathsf{R}^k$.*

DÉMONSTRATION. En effet pour deux petits rectangles $R, T \subset \mathbb{R}^n$ il existe au plus un élément $l \in \mathbb{Z}^n$ tel que $A(R) \cap (T + l) \neq \emptyset$ et de même pour A^{-1}. □

Pour deux moyens rectangles du tore, R et T, on dira que R traverse T s'il existe $l \in \mathbb{Z}^n$ tel que R traverse $T + l$ vus comme des rectangles dans $\mathbb{R}^n$. Dans le Lemme 8.2 et le Lemme 8.3, on peut alors remplacer A par f_A et $\|\cdot\|_\lambda$ par d_λ si l'on considère des petits rectangles du tore.

Corollaire 8.5. *On considère des petits rectangles* $\mathsf{R}^{-N}, ..., \mathsf{R}^{N}$ *du tore. Alors l'ensemble* $\bigcap_{i=-N}^{N} f_A^{-i}\mathsf{R}^i$ *est un petit rectangle du tore de diamètre pour* d_λ *inférieur à* λ^{-N}*. Si de plus,* $f_A(\mathsf{R}^i)$ *traverse* R^{i+1} *pour tout* $i = -N, ..., N-1$*, alors ce rectangle est non vide.*

En particulier, pour tout recouvrement fini $\mathcal{R}$ en petits rectangles ouverts, le diamètre du recouvrement itéré $\bigvee_{k=-N}^{N} f_A^{-k}\mathcal{R}$ est inférieur à λ^{-N}. Un tel recouvrement ouvert est donc un générateur topologique :

Proposition 8.6. *Les automorphismes linéaires hyperboliques du tore sont expansifs.*

2. Partition de Markov et sous-décalage de type fini associé

Une *partition de Markov* de $(\mathbb{T}^n, f_A)$ est une partition $\mathcal{R}$ en minuscules rectangles du tore $\mathbb{T}^n$, telle que pour tout $\mathsf{R} \in \mathcal{R}$ le moyen rectangle $f_A(\mathsf{R})$ traverse les éléments de $\mathcal{R}$ qu'il rencontre.

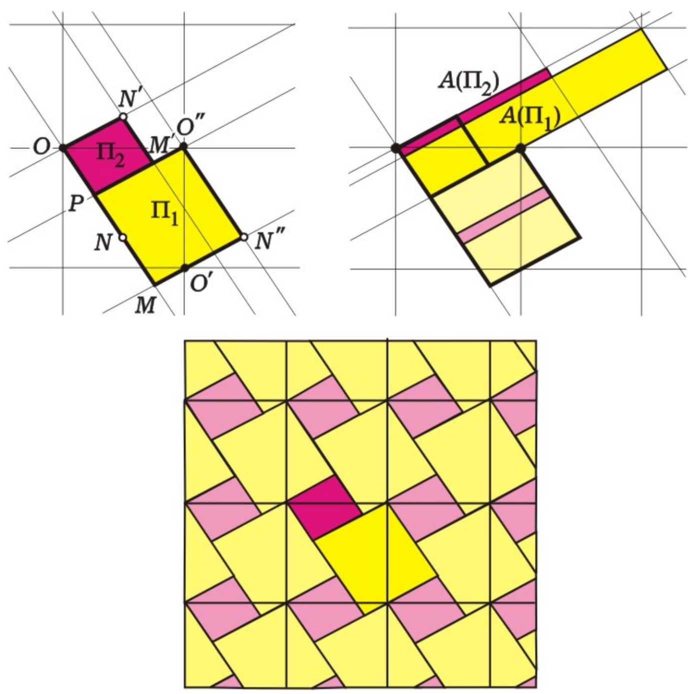


Figure 1: *Un exemple de partition de Markov de* $\mathbb{T}^2$.

A toute partition de Markov $\mathcal{R}$ de $(\mathbb{T}^n, f_A)$ on peut associer un sous-décalage de type fini $\Sigma(\mathcal{G}_\mathcal{R})$ de la façon suivante :
— Les sommets de $\mathcal{G}_\mathcal{R}$ sont les éléments de $\mathcal{R}$,
— $\mathsf{R} \to \mathsf{T}$ pour $\mathsf{R}, \mathsf{T} \in \mathcal{R} \quad \Leftrightarrow \quad f_A(\mathsf{R}) \cap \mathsf{T} \neq \emptyset$.

On considère le sous-graphe $\mathcal{G}'_{\mathcal{R}}$ de $\mathcal{G}_{\mathcal{R}}$, où l'on a enlevé tous les sommets $\mathsf{R} \in \mathcal{R}$ de mesure de Lebesgue nulle. On notera $\mathcal{R}' = \{\mathsf{R} \in \mathcal{R},\ Leb(\mathsf{R}) \neq 0\}$.

PROPOSITION 8.7. *La matrice d'adjacence associée à $\mathcal{G}'_{\mathcal{R}}$ est primitive.*

DÉMONSTRATION. Au Chapitre 2, on a vu que la mesure de Lebesgue était mélangeante pour les automorphismes linéaires du tore f_A dont la matrice A n'a pas de valeur propre racine de l'unité. Donc on a pour k assez grand et pour toute paire de rectangles $\mathsf{R}, \mathsf{T} \in \mathcal{R}$ de mesure de Lebesgue non nulle,

$$Leb(\mathsf{T} \cap f_A^{-k}\mathsf{R}) \sim Leb(\mathsf{T})Leb(\mathsf{R}) > 0.$$

Ceci implique que la k^{eme} puissance de la matrice d'adjacence associée à $\mathcal{G}'_{\mathcal{R}}$ est strictement positive. $\square$

3. Semi-conjugaison

THÉORÈME 8.8. *Soit $\mathcal{R}$ une partition de Markov d'un automorphisme linéaire hyperbolique $(f_A, \mathbb{T}^n)$. Alors l'application $\pi : \Sigma(\mathcal{G}_{\mathcal{R}}) \to \mathbb{T}^n$ qui à $(\mathsf{R}^n)_{n\in\mathbb{Z}}$ associe $\bigcap_{n\in\mathbb{Z}} \overline{f_A^n \mathsf{R}^n}$ vérifie :*
— $\pi \circ \sigma = f_A \circ \pi$,
— $\pi : (\Sigma(\mathcal{G}_{\mathcal{R}}), d) \to (\mathbb{T}^n, d_\lambda)$ *est Hölder-continue,*
— $\pi\left(\Sigma(\mathcal{G}'_{\mathcal{R}})\right) = \mathbb{T}^n$,
— $\sup_{x\in\mathbb{T}^n} \sharp\pi^{-1}(x) \leq \sharp\mathcal{R}^2$,

DÉMONSTRATION. D'après le Corollaire 8.5, pour tout $(\mathsf{R}^n)_n \in \Sigma(\mathcal{G}_{\mathcal{R}})$, le compact $\bigcap_{|n|\leq N} \overline{f_A^n \mathsf{R}^n}$ est non vide de diamètre inférieur à $1/\lambda^N$ pour la distance d_λ. L'intersection $\bigcap_{n\in\mathbb{Z}} \overline{f_A^n \mathsf{R}^n}$ est donc un singleton et l'application π vérifie pour tout $\mathsf{R} = (\mathsf{R}^n)_n$ et $\mathsf{T} = (\mathsf{T}^n)_n$

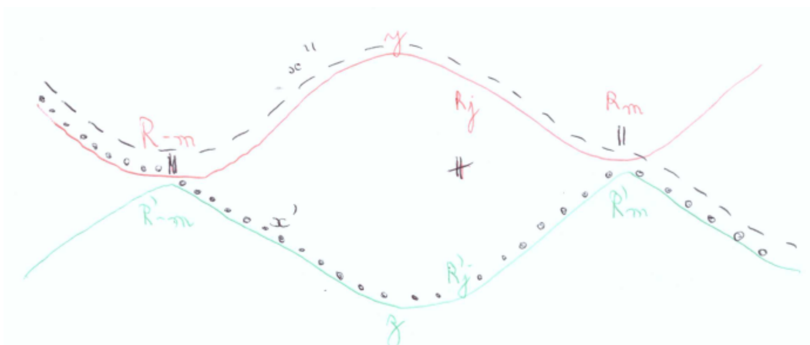
$$d_\lambda(\pi(\mathsf{R}), \pi(\mathsf{T})) \leq \frac{1}{\lambda^{n(\mathsf{R},\mathsf{T})}}$$

avec $n(\mathsf{R}, \mathsf{T}) := \max\{n \in \mathbb{N},\ \mathsf{R}^k = \mathsf{T}^k \text{ pour } k = -n, ..., n\}$. Or pour la distance d sur $\Sigma(\mathcal{G})$

$$d(\mathsf{R}, \mathsf{T}) = \sum_n \frac{\delta_{\mathsf{R}^n, \mathsf{T}^n}}{2^n} \geq 2^{-n(\mathsf{R},\mathsf{T})}$$

et donc $d_\lambda(\pi(\mathsf{R}), \pi(\mathsf{T})) \leq d(\mathsf{R}, \mathsf{T})^\theta$ avec $\theta = \frac{\log\lambda}{\log 2}$. L'application π est clairement une semi-conjugaison. Elle est surjective car elle admet un inverse (non continue) à droite donné par $\psi : x \mapsto \left(\mathcal{R}(f_A^n x)\right)_n$, où $\mathcal{R}(x)$ désigne l'élément de $\mathcal{R}$ contenant x. Notons $\mathcal{V}$ l'union des rectangles de $\mathcal{R}$ de mesure de Lebesgue nulle. Alors pour $x \notin \bigcup_n f_A^n \mathcal{V}$ on a $\psi(x) \in \Sigma(\mathcal{G}'_{\mathcal{R}})$. Or $\bigcup_n f_A^n \mathcal{V}$ est de mesure de Lebesgue nulle, en particulier d'intérieur vide. L'image par l'application continue π du compact $\Sigma(\mathcal{G}'_{\mathcal{R}})$ est donc un compact contenant $\bigcup_n f_A^n \mathcal{V}$. Par conséquent on a $\pi\left(\Sigma(\mathcal{G}'_{\mathcal{R}})\right) = \mathbb{T}^n$.

Supposons par l'absurde qu'il existe $x \in X$ avec $\sharp\pi^{-1}(x) > \sharp\mathcal{R}^2$. On considère un entier m positif et un sous-ensemble fini F_x de $\pi^{-1}(x)$ de cardinal strictement supérieur à $\sharp\mathcal{R}^2$ tel que deux points distincts de F_x diffèrent par leur k^{eme} coordonnée pour un entier k avec $-m < k < m$. L'application $\Phi : F_x \to \mathcal{R}^2$ qui à $(\mathsf{R}^n)_{n\in\mathbb{Z}}$ associe $(\mathsf{R}^{-m}, \mathsf{R}^m)$ n'est pas injective (par argument de cardinalité). Il existe donc $(\mathsf{R}^n)_n$ et $(\mathsf{R}'^n)_n$ dans $\Sigma(\mathcal{G}_{\mathcal{R}})$ avec $\pi((\mathsf{R}^n)_n) = \pi((\mathsf{R}'^n)_n)(= x)$, mais aussi $\mathsf{R}^{-m} = \mathsf{R}'^{-m}$, $\mathsf{R}^m = \mathsf{R}'^m$ et $\mathsf{R}^j \neq \mathsf{R}'^j$ pour un certain $-m < j < m$. Observez que $f_A^i(x) \in \overline{\mathsf{R}^i} \cap \overline{\mathsf{R}'^i} \neq \emptyset$ pour tout $-m < i < m$. On considère $y \in \bigcap_{i=-m}^m f_A^{-i}\mathsf{R}^i$ et $z \in \bigcap_{i=-m}^m f_A^{-i}\mathsf{R}'^i$.



Pour un moyen rectangle $R = R_u \times R_s$ de $\mathbb{R}^n$ et $t = (t_u, t_s) \in R$ on notera $W^u(t,R) = R_u \times \{t_s\}$ et $W^s(t,R) = \{t_u\} \times R_s$. Comme pour les moyens rectangles on identifie $W^u(t,R)$ et $W^s(t,R)$ avec leurs images $W^u(t,\mathsf{R})$ et $W^s(t,\mathsf{R})$ dans le tore. Alors on pose $f_A^{-m}(x') = W^u(f_A^{-m}y, \mathsf{R}^{-m}) \cap W^s(f_A^{-m}z, \mathsf{R}^{-m})$ et $f_A^m(x'') = W^u(f_A^m y, \mathsf{R}^m) \cap W^s(f_A^m z, \mathsf{R}^m)$. On a $x'' \neq x'$, car en utilisant la propriété de Markov de $\mathcal{R}$ on a $f_A^j(x') \in \mathsf{R}'^j$ et $f_A^j(x'') \in \mathsf{R}^j$ pour $-m < j < m$. En effet on a

$$f_A^j(x') \in f_A^{j+m} W^s(f_A^{-m}z, \mathsf{R}'^{-m} = \mathsf{R}^m) \subset W^s(f_A^j z, \mathsf{R}'^j) \subset \mathsf{R}'^j$$

et on raisonne de même pour x''. Deux minuscules rectangles, dont les clôtures s'intersectent, sont inclus dans un même petit rectangle. C'est donc le cas de R^i et R'^i, pour $i = -m, ..., m$. En particulier il existe une suite $(\mathsf{S}^n)_{n\in\mathbb{Z}}$ de petits rectangles du tore avec $f_A^n(x)$ et $f_A^n(x')$ dans S^n pour tout n. Le Corollaire 8.5 entraine alors la contradiction $x' = x''$. $\square$

Nous avons vu que l'entropie d'un facteur est toujours inférieur ou égal à l'entropie de l'extension. Lorsque les fibres sont finies, on a en fait égalité des entropies. Montrons dans le cas où l'extension est un sous-décalage.

Lemme 8.9. *Soient* (Y, σ) *un sous-décalage et* (X, f) *un système topologique. Si* $\pi : (Y, \sigma) \to (X, f)$ *est une extension topologique avec* $\sup_{x \in X} \sharp \pi^{-1}(x) < +\infty$. *Alors*

$$\forall \nu \in \mathcal{M}(Y, \sigma), \ h(\pi^* \nu) = h(\nu).$$

Démonstration. On note Q la partition en coordonnée zéro de Y. Pour tout $x \in X$ il existe $\alpha_x > 0$ tel que

$$\pi^{-1} B(x, \alpha_x) \subset \bigcup_{\pi(z)=x} Q(z).$$

En effet sinon on aurait une suite $(y_q)_q \in Y^{\mathbb{N}}$ avec $\pi(y_q) \xrightarrow{q} x$ et $y_q \notin \bigcup_{\pi(z)=x} Q(z)$ pour tout q. Mais toute limite y de y_q vérifie $\pi(y) = x$ et donc y_q appartient donc au voisinage ouvert $Q(y)$ de y pour q assez grand, ce qui contredit notre hypothèse.

Soit $\mathcal{U} = (B(x_i, \alpha_{x_i}))_{i \in I}$ un recouvrement fini de X. On considère une partition P de X de diamètre inférieur à la constante de Lebesgue de $\mathcal{U}$. Enfin notons R la partition $\pi^{-1}(P)$. Par construction, pour tout $A \in P$, l'ensemble $\pi^{-1}(A)$ rencontre au plus $C_\pi := \sup_{x \in X} \sharp \pi^{-1}(x)$ éléments de Q. Par conséquent pour tout $\nu \in \mathcal{M}(Y, \sigma)$ on a

$$\begin{aligned} h(\nu, R) &\leq h(\nu, \pi^{-1} P) + h(\nu, R | \pi^{-1} P), \\ &\leq h(\pi^* \nu, P) + \log C_\pi., \\ h(\nu, Q) &\leq h(\pi^* \nu) + \log C_\pi. \end{aligned}$$

Or $h(\nu, Q) = h(\nu) \leq h(\pi^* \nu) + \log C_\pi$. Pour tout $l \in \mathbb{N}^*$ l'application π définit encore une semi-conjugaison entre (Y, σ^l) et (X, f^l) on a donc

$$l h_\sigma(\nu) = h_{\sigma^l}(\nu) \leq h_{f^l}(\pi^* \nu) + \log C_\pi = l h_f(\pi^* \nu) + \log C_\pi.$$

Ceci étant vrai pour tout l, on obtient $h_f(\nu) \leq h_f(\pi^* \nu)$. L'inégalité inverse suit du Lemme 6.13.

□

Dans la prochaine section nous construisons une partition de Markov pour f_A.

Théorème 8.10. *Tout automorphisme linéaire hyperbolique du tore admet une partition de Markov, dont les éléments ont une frontière de mesure de Lebesgue nulle.*

Corollaire 8.11. *L'entropie topologique d'un automorphisme linéaire hyperbolique du* n*-tore* f_A *est donnée par*

$$h_{top}(f_A) = \log Jac(A|_{E_u}).$$

De plus la mesure de Lebesgue est l'unique mesure d'entropie maximale.

Démonstration. D'après la Proposition 8.8 et le Lemme 8.9, $(f_A, \mathbb{T}^n)$ a une unique mesure d'entropie maximale si c'est le cas de $(\sigma, \Sigma(\mathcal{G}'_{\mathcal{R}}))$. D'après le Lemme 8.7 la matrice d'adjacence est primitive et il résulte donc du Théorème 7.19 que la mesure de Parry est l'unique mesure d'entropie maximale de $(\sigma, \Sigma(\mathcal{G}'_{\mathcal{R}}))$. Donc $(f_A, \mathbb{T}^n)$ a une unique mesure d'entropie maximale ν. Notons $t = \sharp\mathcal{R}'$ et $\mathcal{R}' = \{\mathsf{R}^1, ..., \mathsf{R}^t\}$. Le vecteur $u = (Leb(R^i_u))_i$ est un vecteur propre à droite de la matrice d'adjacence $\mathcal{A}$ de $\mathcal{G}'_{\mathcal{R}}$ pour la valeur propre $Jac(A|_{E_u}) = \frac{1}{Jac(A|_{E_s})}$. En effet pour $R = R^i$ on a :

$$\begin{aligned} Leb(A(R_u)) &= Jac(A|_{E_u}) Leb(R_u), \\ &= Jac(A|_{E_u}) u_i, \\ &= \sum_{R'_u,\ R' \cap A(R) \neq \emptyset} Leb(R'_u), \end{aligned}$$

où la somme porte sur les rectangles $R' \in \{S + l,\ \mathsf{S} \in \mathcal{R}' \text{ et } \ l \in \mathbb{Z}^n\}$ dans $\mathbb{R}^n$. Comme nous l'avons déjà remarqué précédemment, les rectangles de $\mathcal{R}$ étant petits, il existe, pour $\mathsf{S} \in \mathcal{R}$ donné, au plus un élément $l \in \mathbb{Z}^n$ tel que $A(R) \cap (S + l) \neq \emptyset$. Par conséquent cette dernière somme est égale à

$$\sum_{R'_u,\ \mathsf{R}' \in \mathcal{R}' \text{ et } \mathsf{R} \to \mathsf{R}'} Leb(R'_u) = (\mathcal{A}u)_i \,.$$

En raisonnant de même pour A^{-1} on montre que v est un vecteur propre à gauche de la matrice d'adjacence $\mathcal{A}$ de $\mathcal{G}'_{\mathcal{R}}$ pour la valeur propre $\frac{1}{Jac(A|_{E_s})} = Jac(A|_{E_u})$. D'après le théorème de Perron-Froebenius, cette valeur propre coïncide avec le rayon spectral (c'est en effet la seule valeur propre admettant des vecteurs propres strictement positifs). En particulier $h_{top}(f_A) = h_{top}(\Sigma(\mathcal{G}'_{\mathcal{R}})) = \log Jac(A|_{E_u})$. De plus on a avec $\phi : (w_u, w_s) \in E_u \times E_s \mapsto w_u + w_s \in E_u + E_s$,

$$\begin{aligned} \sum_j u_j v_j &= \sum_j Leb(R^j_u) Leb(R^j_s), \\ &= \frac{1}{Jac(\phi)} \sum_j Leb(R^j_u \times R^j_s), \\ &= \frac{1}{Jac(\phi)} \sum_j Leb(R^j), \\ &= \frac{1}{Jac(\phi)}. \end{aligned}$$

Soit $\tilde{Leb}$ un relevé de Leb par $\pi : \Sigma(\mathcal{G}'_{\mathcal{R}}) \to \mathbb{T}^n$. Pour tout $i = 1, ..., n$, on a $\pi^{-1}(\overset{o}{\mathsf{R}^i}) \subset [i] \subset \pi^{-1}(\overline{\mathsf{R}^i})$. De $\pi^* \tilde{Leb} = Leb$ il suit $Leb(\overset{o}{\mathsf{R}^i}) =$

$\tilde{Leb}(\pi^{-1}(\overset{o}{\mathsf{R}^i})) \leq \tilde{Leb}[i] = \tilde{Leb}(\pi^{-1}\mathsf{R}^i) \leq \tilde{Leb}(\pi^{-1}(\overline{\mathsf{R}^i})) = Leb(\overline{\mathsf{R}^i})$ pour tout i. Le bord de R_i étant de mesure de Lebesgue nulle, on a alors $\tilde{Leb}[i] = Leb(\mathsf{R}^i) = Leb(R^i) = Jac(\phi)Leb(R^i_u)Leb(R^i_s) = \frac{u_i v_i}{\sum_j u_j v_j}$.

Enfin pour un cylindre $[i_0..., i_k]$ avec $k \geq 1$ de $\Sigma(\mathcal{G}_\mathcal{R})$ on note $(l_j)_{j=0}^k$ l'unique élément de $(\mathbb{Z}^n)^{k+1}$ avec $l_0 = 0$ tel que le rectangle $\bigcap_{j=0}^k A^{-j}(R^{i_j} + l_j)$ de $\mathbb{R}^n$ se projette bijectivement sur $\bigcap_{j=0}^k f_A^{-j}\mathsf{R}^{i_j}$. On a alors

$$\begin{aligned}
\tilde{Leb}[i_0..., i_k] &= Leb(\bigcap_{j=0}^{k} f_A^{-j}\mathsf{R}^{i_j}), \\
&= Leb(\bigcap_{j=0}^{k} A^{-j}(R^{i_j} + l_j)), \\
&= Leb(\bigcap_{j=0}^{k-1} A^{-j}(R^{i_j} + l_j)) \times Jac(A|_{E_s}) \times \frac{Leb(R_u^{i_k})}{Leb(R_u^{i_{k-1}})}, \\
&= \tilde{Leb}[i_0..., i_{k-1}] \times \frac{v_{i_k}}{v_{i_{k-1}} Jac(A|_{E_u})}.
\end{aligned}$$

La mesure $\tilde{Leb}$ est donc la mesure de Parry de $\Sigma(\mathcal{G}_\mathcal{R})$ associé à $\lambda = Jac(A|_{E_u})$. D'après le Lemme 8.9 la semi-conjugaison π préserve l'entropie des mesures invariantes il s'en suit que :

— f_A et de $\Sigma(\mathcal{G}_\mathcal{R})$ ont même entropie topologique, à savoir $\log Jac(A|_{E_u})$,

— Leb est l'unique mesure d'entropie maximale de f_A.

□

4. Construction de partitions de Markov

4.1. Première étape : Construction d'une extension Markovienne. Fixons $p \in \mathbb{N}^*$. On note $Lip(f_A)$ la constante de Lipschitz de l'application f_A sur le n-tore muni de la métrique d_λ. Soit q un entier positif $1/q < \frac{\lambda-1}{p\lambda(\max(Lip(f_A), Lip(f_{A^{-1}}))+1)}$. On choisit un autre entier positif q' de sorte que tout point de $\mathbb{R}^n$ est à distance d'au plus $1/q$ de l'ensemble $D_{q'} = \frac{1}{q'}\mathbb{Z}^n$ pour $\|\cdot\|_\lambda$. Pour tout $a \in D_{q'}$ on considère le rectangle $R^a = R^a_u \times R^a_s$ avec $R^a_{u/s}$ la boule de centre a et de rayon $1/p$ dans $(E_{u/s}, \|\cdot\|_{u/s})$ et on pose $\mathfrak{S} := \{R^a = R^a_u \times R^a_s,\ a \in D_{q'}\}$. Remarquez que tout élément de $\mathfrak{S}$ est de diamètre inférieur à $4/p$. En particulier pour p assez grand, les éléments de $\mathfrak{S}$ sont des rectangles minuscules. Pour tout $x \in E_{q'} := D_{q'}/\mathbb{Z}^n$ on note R^x le rectangle du tore associé à R^a pour $a \in D_{q'}$ se projetant dans le tore sur x (ce rectangle ne dépend pas d'un tel a). Soit $\mathcal{S}$ la collection finie de rectangles du tore associé à $\mathfrak{S}$, c'est à dire $\mathcal{S} := \{\mathsf{R}^x,\ x \in E_{q'}\}$. On définit alors un graphe $\mathcal{G}_\mathcal{S}$ comme suit :

— les sommets de $\mathcal{G}_\mathcal{S}$ sont les éléments de $\mathcal{S}$,

— $\mathsf{R}^x \to \mathsf{R}^y$ lorsque $f_A(\mathsf{R}^x)$ traverse R^y.

LEMME 8.12. *L'application* $\pi : \Sigma(\mathcal{G}_{\mathcal{S}}) \to \mathbb{T}^n$ *qui à* $(\mathsf{R}^n)_{n\in\mathbb{Z}}$ *associe* $\bigcap_n \overline{f_A^{-n}\mathsf{R}^n}$ *vérifie :*
— $\pi \circ \sigma = f_A \circ \pi$,
— π *est Hölder-continue,*
— π *est surjective.*

DÉMONSTRATION. Les deux premiers points se montrant facilement (comme pour le Théorème 8.8) on se concentre sur la surjectivité de π. Il suffit de voir que pour tout point y du n-tore et tout $\mathsf{R}^{x^0} \in \mathcal{S}$ contenant y avec $d_\lambda(y, x^0) < 1/q$ il existe $\mathsf{R}^{x^{-1}}$ et $\mathsf{R}^{x^1} \in \mathcal{S}$ contenant respectivement $f_A^{-1}(y)$ et $f_A(y)$ tels que $\mathsf{R}^{x^{-1}} \to \mathsf{R}^{x^0} \to \mathsf{R}^{x^1}$, $d_\lambda(x^1, f_A(y)) < 1/q$ et $d_\lambda(x^{-1}, f_A^{-1}(y)) < 1/q$. Montrons l'existence d'un tel x^1 le cas de x^{-1} étant similaire. Pour $a \in D_{q'}$ se projetant sur x^0, l'image de R^{a^0} est le rectangle $A(R_u^{a^0}) \times A(R_s^{a^0})$. Or $A(R_u^{a^0})$ contient une boule de rayon $\frac{\lambda}{p}$ pour $\|\cdot\|_u$ de centre $A(a_u^0)$ et de même $A(R_s^{a^0})$ est inclus dans une boule de rayon $\frac{1}{p\lambda}$ pour $\|\cdot\|_s$ de centre $A(a_s^0)$. Soit $x^1 \in E_{q'}$ avec $d_\lambda(x^1, f_A(y)) \leq 1/q$, on a $d_\lambda(x^1, f_A(x^0)) \leq d_\lambda(x^1, f_A(y)) + d_\lambda(f_A(x^0), f_A(y)) \leq 1/q + Lip(f_A)/q < \frac{\lambda-1}{p}$. Il existe donc $a^1 \in D_{q'}$ se projetant sur x^1 avec $R_u^{a^1} \subset A(R_u^{a^0})$ et $A(R_s^{a^0}) \subset R_s^{a^1}$, i.e. $A(R^{x^0})$ traverse R^{x^1} ou encore $f_A(\mathsf{R}^{x^0})$ traverse R^{x^1}. $\square$

4.2. Seconde étape : Recouvrement markovien en rectangles du tore. On considère maintenant le recouvrement de $\mathbb{T}^n$ donnée par $\mathcal{T} := \{\mathsf{T}^{\mathsf{S}},\ \mathsf{S} \in \mathcal{S}\}$ avec

$$\mathsf{T}^{\mathsf{S}} := \pi([\mathsf{S}]) = \{\pi\left((u_n)_n\right),\ u_0 = \mathsf{S}\} \subset \mathsf{S}.$$

LEMME 8.13. *Tout* $\mathsf{T} \in \mathcal{T}$ *est un rectangle. De plus si* $\mathsf{S} \to \mathsf{U}$ *dans* $\mathcal{G}_{\mathcal{S}}$ *alors* $f_A(\mathsf{T}^{\mathsf{S}})$ *traverse* T^{U}.

DÉMONSTRATION. Il suffit de voir que pour tout $(x, y) \in \mathsf{T}^{\mathsf{S}}$ l'intersection de $W^u(x, \mathsf{S})$ et $W^s(y, \mathsf{S})$ appartient à T^{S}. Si $x = \pi\left((u_n)_n\right)$ et $y = \pi\left((v_n)_n\right)$ avec $u_0 = v_0 = S$ alors la suite $(w_n)_n$ définie par $w_n = u_n$ pour $n \leq 0$ et $w_n = v_n$ pour $n \geq 0$ appartient à $\Sigma(\mathcal{G}_{\mathcal{S}})$ et $\pi((w_n)_n) \in W^u(x, \mathsf{S}) \cap W^s(y, \mathsf{S})$. $\square$

L'application A étant linéaire on peut supposer que T_u et T_s sont convexes. En effet quitte à les remplacer par leurs enveloppe convexes, la propriété markovienne $A(T_s^{\mathsf{S}}) \subset T_s^{\mathsf{U}}$ et $A(T_u^{\mathsf{S}}) \supset T_u^{\mathsf{U}}$ est préservée. En particulier les bords de T_u et T_s sont de mesure de Lebesgue nulle et il en est donc de même de $T = T_u \times T_s$ et T.

REMARQUE 8.14. *Si pour* $x \in \mathbb{T}^n$ *on a* $f_A(x) \in \mathsf{T}^{\mathsf{U}}$ *avec* $\mathsf{U} \in \mathcal{S}$ *alors il existe* $\mathsf{S} \in \mathcal{S}$ *avec* $x \in \mathsf{T}^{\mathsf{S}}$ *et* $\mathsf{S} \to \mathsf{U}$ *(donc* $f_A(\mathsf{T}^{\mathsf{S}})$ *traverse* T^{U}*d'après* *le Lemme ci-dessus). Cette propriété n'était a priori pas vérifiée par le recouvrement en rectangles $\mathcal{S}$.*

4.3. Dernière étape : Raffinement du recouvrement en une partition. On raffine maintenant le recouvrement $\mathcal{T}$ en une partition. Pour $\mathsf{T}, \mathsf{T}' \in \mathcal{T}$ avec $\mathsf{T} \cap \mathsf{T}' \neq \emptyset$, on note $\alpha(\mathsf{T}, \mathsf{T}')$ la partition en rectangle induite sur $\mathsf{T} \cup \mathsf{T}'$ par $\{T_u \setminus T'_u, T'_u \setminus T_u, T_u \cap T'_u\} \times \{T_s \setminus T'_s, T'_s \setminus T_s, T_s \cap T'_s\}$ puis

$$\mathcal{R} := \bigvee_{\mathsf{T} \cap \mathsf{T}' \neq \emptyset} \alpha(\mathsf{T}, \mathsf{T}').$$

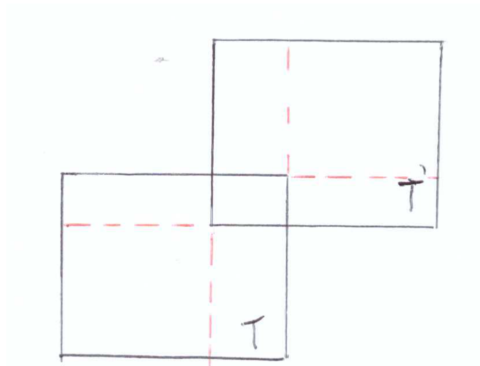


Figure 2: *La partition $\alpha(\mathsf{T}, \mathsf{T}')$ de $\mathsf{T} \cup \mathsf{T}'$.*

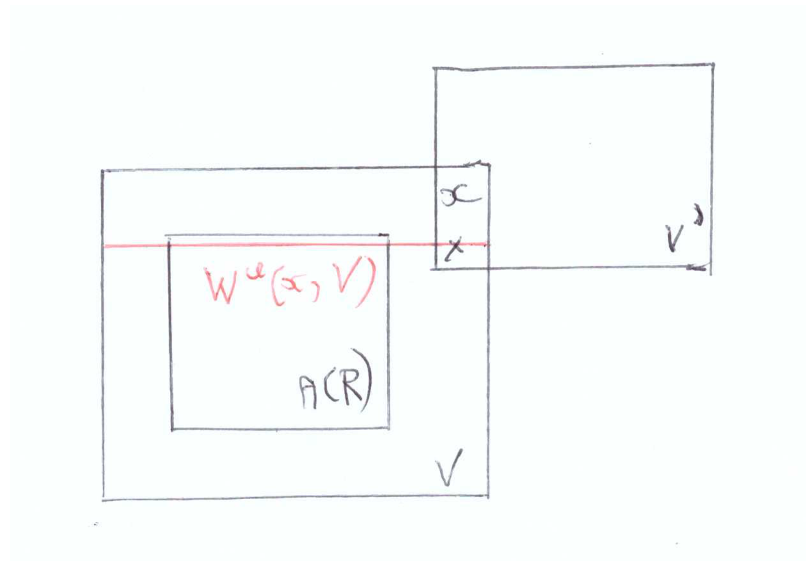
En général la partition engendrée par un recouvrement en rectangle n'est pas une partition en rectangle, toutefois c'est ici le cas de $\mathcal{R}$.

Lemme 8.15. *$\mathcal{R}$ est une partition de Markov, dont tous les ensembles ont un bord de mesure de Lebesgue nulle.*

Démonstration. Tout élément de $\mathcal{R}$ est un rectangle. Pour cela on doit vérifier que si un rectangle R de $\alpha(\mathsf{T}, \mathsf{T}')$ pour $\mathsf{T} \cap \mathsf{T}' \neq \emptyset$ rencontre $\mathcal{R}(x)$ alors $\mathcal{R}(x) \subset \mathsf{R}$ ou bien $\mathcal{R}(x) \cap \mathsf{R} = \emptyset$. On peut supposer qu'il existe T'' tel que $\mathsf{T} \cap \mathsf{T}'' \neq \emptyset$ ou $\mathsf{T}' \cap \mathsf{T}'' \neq \emptyset$ et au moins un des trois rectangles $\mathsf{T}, \mathsf{T}', \mathsf{T}''$ contient x, sinon $\mathcal{R}(x) \cap \mathsf{R} = \emptyset$. Si $x \in \mathsf{T} \cup \mathsf{T}'$ alors $\mathcal{R}(x) \cap \mathsf{R} = \emptyset$ excepté si R est l'élément de la partition en rectangle $\alpha(\mathsf{T}, \mathsf{T}')$ de $\mathsf{T} \cup \mathsf{T}'$ contenant x et on a alors $\mathcal{R}(x) \subset \mathsf{R}$. Considérons donc le dernier cas : $x \in \mathsf{T}''$, en particulier $\mathcal{R}(x) \subset \mathsf{T}''$. Si $\mathsf{R} \subset \mathsf{T}$ et $\mathsf{T} \cap \mathsf{T}'' = \emptyset$ on a alors $\mathsf{R} \cap \mathcal{R}(x) = \emptyset$. De même si $\mathsf{R} \subset \mathsf{T}'$ et $\mathsf{T}' \cap \mathsf{T}'' = \emptyset$. On peut donc supposer que $x \notin \mathsf{R} \subset \mathsf{T}$ et $\mathsf{T} \cap \mathsf{T}'' \neq \emptyset$. Donc il existe un rectangle $\mathsf{R}' \in \alpha(\mathsf{T}'', \mathsf{T})$ contenant x et donc $\mathcal{R}(x)$ tel que $\mathsf{R}' \cap \mathsf{R} = \emptyset$.

Montrons maintenant la propriété de Markov. Les rectangles de $\mathcal{R}$ étant minuscules on peut travailler dans $\mathbb{R}^n$ avec A au lieu de f_A. Par symétrie (considérer A^{-1} au lieu de A) il suffit de voir que $A(R_s) \subset S_s$ pour tout $\mathsf{R}, \mathsf{S} \in \mathcal{R}$ avec $f_A(\mathsf{R}) \cap \mathsf{S} \neq \emptyset$. Sinon il existe $\mathsf{V}, \mathsf{V}' \in \mathcal{T}$ avec

$\mathsf{V} \cap \mathsf{V}' \neq \emptyset$ et $f_A(\mathsf{R}) \cap \mathsf{V} \neq \emptyset$ tel que $A(R_s)$ rencontre deux éléments de $\{V_s \setminus V'_s, V'_s \setminus V_s, V_s \cap V'_s\}$. D'après la Remarque 8.14 il existe $\mathsf{T}'' \in \mathcal{T}$ contenant R, tel que $f_A(\mathsf{T}'')$ traverse V et donc $A(T''_s) \subset V_s$. Donc $A(R_s) \subset A(T''_s) \subset V_s$. Il suffit donc de montrer que si l'ensemble $A(R_s)$ rencontre V'_s alors il est inclus dans $V'_s \cap V_s$.



Dans ce cas il existe $x \in \mathsf{V} \cap \mathsf{V}'$ tel que $W^u(x, \mathsf{V}) \cap A(R) \neq \emptyset$. Par conséquent $A^{-1}(W^u(x, \mathsf{V})) \subset W^u(A^{-1}x, \mathsf{T}) \subset \mathsf{T}$ et $A^{-1}(W^u(x, \mathsf{V}')) \subset W^u(A^{-1}x, \mathsf{T}') \subset \mathsf{T}'$ pour certains $\mathsf{T}, \mathsf{T}' \in \mathcal{T}$ d'après la Remarque 8.14. Or $\mathsf{R} \cap A^{-1}(W^u(x, \mathsf{V})) \neq \emptyset$, donc $\mathsf{R} \cap \mathsf{T} \neq \emptyset$. De plus $A^{-1}x \in \mathsf{T} \cap \mathsf{T}'$ et donc $(A^{-1}y)_s \in T_s \cap T'_s$ pour tout $y \in W^u(x, \mathsf{V})$, en particulier $R_s \cap T_s \cap T'_s \neq \emptyset$. Puisque $\mathsf{R} \in \mathcal{R}$ on en déduit que $R_s \subset T'_s \cap T_s$ puis $A(R_s) \subset A(T'_s) \cap A(T_s) \subset V'_s \cap V_s$.

Enfin la propriété d'avoir un bord de mesure de Lebesgue nulle est stable par intersection finie et passage au complémentaire. Tout élément de $\mathcal{R}$ vérifie donc cette propriété puisque c'est le cas des éléments (convexes) de $\mathcal{T}$. $\square$

Un point du tore possède plusieurs antécédents par la semi-conjugaison π lorsque son orbite visite le bord de la partition de Markov :

$$\{x, \ \sharp \pi^{-1} x > 1\} = \bigcup_n f_A^n(\partial \mathcal{R}).$$

Puisque les bords de la partition de Markov sont de mesure de Lebesgue nulle, π est un isomorphisme de $(\Sigma(\mathcal{G}_{\mathcal{R}}), \sigma, \tilde{Leb})$ vers $(\mathbb{T}^n, f_A, Leb)$.

En dimension 2, les éléments de $\mathcal{T}$ sont des rectangles usuels, i.e. T_u et T_s sont des intervalles. Pour $\mathsf{R} = R_u \times R_s \in \mathcal{R}$ les ensembles R_u et

R_s sont alors des unions finies d'intervalles. La propriété de Markov est encore vérifiée par la partition $\mathcal{R}'$ donnée par les composantes connexes de $R \in \mathcal{R}$, qui sont toujours des rectangles, de sorte que $\mathcal{R}'$ définit aussi une partition de Markov. Les ensembles de $\mathcal{R}'$ sont alors des rectangles au sens usuel. En dimension supérieure, il n'est en général pas possible de construire une partition de Markov en parallélépipèdes.

Rappels

π-systèmes. Soit $(X, \mathcal{B})$ un espace mesurable. Une collection $\mathcal{A}$ d'ensembles mesurables stable par intersection finie est appelée un π-système.

Théorème 8.16. *Soit μ et ν deux mesures sur $(X, \mathcal{B})$ telles que $\mu(A) = \nu(A)$ pour tout $A \in \mathcal{A}$. Alors μ et ν coïncident sur la tribu engendrée par $\mathcal{A}$.*

En particulier si deux mesures boréliennes de $\mathbb{R}^n$ coïncident sur les pavés, elles sont égales.

Théorème de Radon-Nikodyn. Soit $(X, \mathcal{B})$ un espace mesurable.

Théorème 8.17. *Soient ν et μ une mesure signée finie et une mesure (positive) finie respectivement. On suppose que ν est absolument continue par rapport à μ, i.e. $\nu(A) = 0$ pour tout $A \in \mathcal{B}$ avec $\mu(A) = 0$. Alors il existe $f \in L^1(\mu)$ tel que pour tout $A \in \mathcal{B}$*

$$\nu(A) = \int_A f \; d\mu$$

Lois de consistance de Kolmogorov.

Théorème 8.18. *Soit $(\mu_n)_n$ une suite de probabilités boréliennes sur $\mathbb{R}^n$ telle que pour tout $\pi_n^* \mu_{n+1} = \mu_n$ pour tout $n \in \mathbb{N}$, où $\pi_n : \mathbb{R}^{n+1} \to \mathbb{R}^n$ désigne la projection sur les n-premiers facteurs. Alors il existe une unique mesure borélienne μ sur $\mathbb{R}^{\mathbb{N}}$ tel que $\mu(A \times \mathbb{R}^{\mathbb{N}}) = \mu_n(A)$ pour tout borélien de $\mathbb{R}^n$ et pour tout n.*

Espérance conditionnelle. Soient $\mathcal{G} \subset \mathcal{F}$ deux tribus sur un ensemble X et soit μ une mesure de probabilité sur $(X, \mathcal{F})$. Alors pour tout $f \in L^1(\mathcal{F})$ il existe une unique fonction $E[f|\mathcal{G}]$ de $L^1(\mathcal{G})$ telle que pour tout $G \in \mathcal{G}$:

$$\int_G f d\mu = \int_G E[f|\mathcal{G}] d\mu.$$

Dans le cas où $f \in L^2(\mathcal{F})$, l'espérance conditionnelle de f relativement à $\mathcal{G}$ est donnée par la projection orthogonale de f sur $L^2(\mathcal{G})$.

Théorème de différentiation de Lebesgue. Soit $f \in L^1_{loc}(\mathbb{R})$. On considère une collection d'intervalles $\mathcal{C}$ tel que tout point $x \in \mathbb{R}$ appartient à des intervalles de $\mathcal{C}$ de longueur arbitrairement petite. Alors le théorème de différentiation de Lebesgue affirme que pour Lebesgue presque tout $x \in \mathbb{R}$

$$f(x) = \lim_{|I| \to 0,\ x \in I \in \mathcal{C}} \frac{\int_I f(t)dt}{|I|}.$$

En particulier si on applique ceci à f l'indicatrice d'un borélien A de $\mathbb{R}$, on obtient que Lebesgue presque tout point $x \in A$ satisfait

$$\frac{Leb(I \cap A)}{Leb(I)} \xrightarrow{|I| \to 0,\ x \in I \in \mathcal{C}} 1.$$

Un tel point x est appelé point de densité de Lebesgue de A (relativement à $\mathcal{C}$).

Revêtement et théorème de relèvement. Une application continue surjective entre deux espace topologiques, $\pi : Y \to X$, est un revêtement lorsque tout point $x \in X$ admet un voisinage ouvert V tel que $\pi^{-1}V$ est une réunion disjointe d'ouverts qui s'envoie homéomorphiquement sur V par π.

Une version faible du théorème de relèvement s'écrit alors comme suit :

THÉORÈME 8.19. *Soit $\pi : Y \to X$ un revêtement et soit $f : Z \to X$ une application continue avec Z un espace topologique simplement connexe. Alors pour tout $z \in Z$ et tout $y \in Y$ avec $\pi(y) = f(z)$ il existe une unique application $F : Z \to Y$ telle que $\pi \circ F = f$ et $F(z) = y$.*

Suites sous-additives. Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite sous-additive lorsque pour tout $p, q \in \mathbb{N}$

$$u_{p+q} \leq u_p + u_q.$$

THÉORÈME 8.20. *Soit $(u_n)_n$ une suite sous-additive. Alors la suite $\left(\frac{u_n}{n}\right)_n$ est convergente et*

$$\lim_n \frac{u_n}{n} = \inf_n \frac{u_n}{n} \in \mathbb{R} \cup \{-\infty\}.$$

Théorie spectrale pour les matrices positives. On considère des matrices carrées réelles à coefficients positifs. Une telle matrice est dite irréductible lorsque pour toute coordonnée i, j il existe un entier positif n avec $(A^n)_{ij} > 0$. La période d'une matrice irréductible est le pgcd de des entiers m avec $(A^m)_{ii} > 0$ (on peut vérifier que cette quantité ne dépend pas de i).

THÉORÈME 8.21. *Soit A une matrice irréductible. Le rayon spectral est valeur propre. De plus l'espace propre associé est de dimension un et contient un vecteur à coordonnées strictement positives. Enfin tout vecteur propre à coordonnées strictement positives est associé à la valeur propre donnée par le rayon spectral.*

Une matrice irréductible est dite primitive lorsque l'on peut choisir l'entier n in dépendant de i, j, i.e. il existe un entier n positif tel que les coefficients de A^n sont tous strictement positifs. Une matrice irréductible est primitive si et seulement si sa période est égale à un.

THÉORÈME 8.22. *Soit A une matrice primitive. Les valeurs propres différentes du rayon spectral ont un module strictement inférieur au rayon spectral.*

Bibliographie

Général

Michael Brin and Garrett Stuck, *Introduction to Dynamical Systems* (2nd ed.). Cambridge University Press, 2002

Yves Coudène, *Théorie ergodique et systèmes dynamiques*, EDPS-ciences, 2013

Anatole Katok and Boris Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems,*Volume 54 de Encyclopedia of Mathematics and its Applications, Éditeur Cambridge University Press, 1997

Ricardo Mané, *Ergodic Theory and Differentiable Dynamics* (Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics) Softcover reprint of the original 1st ed. 1987 Edition

Homéomorphismes du cercle

Wellington De Melo and Sebastian Van Strien, *One Dimensional Dynamics.* Springer-Verlag, 1993

Michel Herman, *Sur la conjugaison des difféomorphismes du cercle à des rotations*, Mémoires de la SMF tome 46, 1976

Entropie

Tomasz Downarowicz, *Entropy in dynamical systems*, New Mathematical Monographs, Vol. 18, Cambridge University Press, Cambridge, 2011

Peter Walters,*An introduction to ergodic theory*, Graduate texts in Mathematics 79, 1982

Dynamique symbolique

Bruce Kitchens, *Symbolic Dynamics : One-sided, Two-sided and Countable State Markov Shifts,*, Éditeur Springer Science and Business Media, 2012

Douglas Lind and Brian Marcus, *Symbolic dynamics and coding*, Cambridge University press, 1995

Partitions de Markov

Roy Adler, *Symbolic dynamics and Markov partitions,* https ://arxiv.org/abs/math/9607214, 1996

Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms,* Lecture Note in Mathematics 470, 1975