Introduction aux systèmes dynamiques

Classe Franco-Chinoise - printemps 2024

University of Science and Technology of China

Jérôme BUZZI

VERSION 0.5 EN COURS DE RELECTURE

Table des matières

Avant-propos

Ce cours a pour but d'introduire les idées fondamentales de la théorie des systèmes dynamiques, en montrant les interactions entre les points de vue mesurés, topologiques et différentiables et en essayant de donner quelques outils incontournables pour l'étudiant(e) qui s'intéresse à la théorie pour elle-même ou bien pour ses applications à d'autres domaines des mathématiques.

Il a donc fallu trouver un équilibre entre théorèmes fondamentaux, applications et le temps toujours limité ainsi qu'entre les divers types de systèmes dynamiques hyperboliques vs elliptiques.

La version présente est éminemment provisoire. Elle est même en cours de relecture et on y trouvera certainement un grand nombre de typos, sans doute quelques erreurs mathématiques, en petit nombre j'espère. Je remercie les lecteurs et lectrices qui voudront bien m'en faire part.

Chapitre 1

Qu'est-ce que la dynamique ?

Dans cette introduction, nous allons définir quelques types de systèmes dynamiques illustrés par quelques exemples-clés. Puis nous décrirons quelques problématiques centrales de la théorie.

1. Un peu d'histoire et quelques applications

La théorie des systèmes dynamiques prend racine dans les problèmes de la mécanique céleste. Le problème de la stabilité du système solaire est déjà posé par Newton et considéré par une lignée de mathématiciens, encore ininterrompue aujourd'hui.

Au-delà de l'étude des flots définis par les équations différentielles ainsi que de l'itération, la théorie considère aujourd'hui des actions de groupes plus généraux afin de traiter des questions extrêmement diverses. En voici quelques exemples :

(1) étant donné un réel x, son développement décimal contient-il toutes les suites finies de chiffres ? avec une fréquence définie ? laquelle ?

(2) répartition modulo 1 des valeurs d'un polynôme sur les entiers ?

(3) stabilité et robustesse de modélisations

(4) $\liminf_{n\to\infty} n \cdot d(n\alpha, \mathbb{Z}) \cdot d(n\beta, \mathbb{Z}) = 0$ (conjecture de Littlewood) ?

(5) statistique du développement en fraction continue $x = \frac{1}{a_1+\frac{1}{a_2+\frac{1}{a_3+\dots}}}$

(6) $\overline{\{Q(x) : x \in \mathbb{Z}^d\}}$ pour Q une forme quadratique

(7) répartition des géodésiques fermées sur une surface hyperbolique

(8) longueurs des composantes connexes de $[0,1] \setminus \{\sqrt{k} : k = 1, \dots, n\}$

(9) problèmes autour du codage d'une source aléatoire stationnaire

(10) thermodynamique sur réseau

1.1. Définitions de systèmes dynamiques. Informellement, un système dynamique est la donnée de (T, X, α) où

— T est un semigroupe, par exemple : $\mathbb{N}$ (ensemble des entiers naturels éventuellement nuls), $\mathbb{Z}$ (ensemble des entiers), $\mathbb{R}_+$ (réels positifs éventuellement nuls), $\mathbb{R}$ (réels), $GL_d(\mathbb{R})$ (matrices $d \times d$ inversibles),...

— X est un espace par exemple un espace topologique, un espace de probabilité,...

— $\alpha : T \times X \to X$ est une action de T, i.e., $\alpha^g \alpha^h = \alpha^{g \cdot h}$ avec $\alpha^g := \alpha(g, \cdot)$ un endomorphisme de X. On demande parfois une certaine régularité de α.

L'orbite de $x \in X$ est alors la famille $(\alpha^t(x))_{t \in T}$.

Dans la catégorie topologique, on s'intéresse particulièrement au cas suivant :

1.1. Définition. Un *système dynamique topologique compact* se définit par (T, X, α) où T est un semigroupe, X est un espace topologique métrisable et compact, et $\alpha : T \times X \to X$ est continue avec $\alpha^t : X \to X$ une application continue pour chaque $t \in T$.

On définit un *morphisme mesuré* $(\Omega, \mathcal{F}, \mu)$ vers un autre $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mu})$ comme $h : \tilde{\Omega}' \to \tilde{\Omega}'$ vérifiant :

— $\Omega' \in \mathcal{F}$, $\tilde{\Omega}' \in \tilde{\mathcal{F}}$ et $\mu(\Omega \setminus \Omega') = \tilde{\mu}(\tilde{\Omega} \setminus \tilde{\Omega}') = 0$

— h est mesurable : $h^{-1}(\tilde{\mathcal{F}}) \subset \mathcal{F}$

— h envoie μ sur $\tilde{\mu}$: $h_*(\mu) := \mu \circ h^{-1} = \tilde{\mu}$

Notons que h étant mesurable, $h_*(\mu) := \mu \circ h^{-1}$ envoie la mesure μ sur Ω sur une mesure sur $\tilde{\Omega}$.

1.2. Définition. Un *système dynamique probabiliste* se définit par (T, X, α) où T est un semigroupe, $X = (\Omega, \mathcal{F}, \mu)$ est un espace de probabilité, et $\alpha : T \times X \to X$ est mesurable avec $\alpha^t : X \to X$

1.3. Remarque. En temps discret ($T = \mathbb{N}$ ou $\mathbb{Z}$) l'action α est définie d'une façon évidente par son générateur α^1.

2. Exemples et constructions

2.1. Exemples algébriques.

2.1.1. *Un groupe métrique compact : le tore.* L'espace $\mathbb{R}^d$ est un groupe. Il agit donc sur lui-même par translation : $x \mapsto x + v$ pour tout $v \in \mathbb{R}^d$. Pour se ramener à un espace compact, on considère le quotient comme suite.

2.1. Définition. Le tore $\mathbb{T}^d$ de dimension $d \geq 1$ est défini comme le groupe topologique quotient $\mathbb{R}^d/\mathbb{Z}^d$ avec la projection canonique $\pi : \mathbb{R}^d \to \mathbb{T}^d$, $x \mapsto x + \mathbb{Z}^d$. On le munit de :

— la distance $d_{\mathbb{T}^d}(\pi(x), \pi(y)) := \inf\{\|x - y - n\| : n \in \mathbb{Z}^d\}$;

— la mesure borélienne $m_{\mathbb{T}^d}(B) := \mathrm{Leb}(\pi^{-1} \cap [0, 1[^d)$;

2.2. Proposition. *Pour chaque $d \geq 1$, on a :*

(1) la topologie quotient est compacte et compatible avec cette distance ;

(2) $\pi : B_{\mathbb{R}^d}(x, 1/2) \to B_{\mathbb{T}^d}(\pi(x), 1/2)$ est un homéomorphisme et une isométrie ;

(3) la distance est invariante par toutes les translations : $d(x + V, y + V) = d(x, y)$ pour tout $V \in \mathbb{R}^d$;

(4) $m_{\mathbb{T}^d}$ est l'unique probabilité borélienne invariante pour toutes les translations : $m_{\mathbb{T}^d}(B + x) = m_{\mathbb{T}^d}(B)$ pour tout $x \in \mathbb{R}^d$.

La propriété (2) permet de définir une structure de variété sur $\mathbb{T}^d$.

La propriété (4) se reformule en disant que $m_{\mathbb{T}^d}$ est la mesure de Haar normalisée du groupe $\mathbb{T}^d$.

2.3. Corollaire. *Pour tout $d \geq 1$ et $\alpha \in \mathbb{R}^d$:*
- *— $(\mathbb{T}^d, \mathbb{Z}, \tau_\alpha)$ définit un système dynamique topologique compact ;*
- *— $((\mathbb{T}^d, \mathcal{B}_{\mathbb{T}^d}, m_{\mathbb{T}^d}), \mathbb{Z}, \tau_\alpha)$ définit un système dynamique probabiliste.*

2.1.2. *Les translations du tore.*

2.4. Définition. Pour tout $\alpha \in \mathbb{R}^d$, la translation de vecteur α est : $\tau_\alpha : \mathbb{T}^d \to \mathbb{T}^d$, $x \mapsto x + \alpha$.

Notons que τ_α est continue (et même C^∞).

La translation τ_α définit :
- — une dynamique topologique compacte en temps $\mathbb{Z}$: $(\mathbb{T}^d, \mathbb{Z}, \tau_\alpha)$;
- — une dynamique probabiliste en temps $\mathbb{Z}$: $((\mathbb{T}^d, \mathcal{B}_{\mathbb{T}^d}, m_{\mathbb{T}^d}), \mathbb{Z}, \tau_\alpha)$.

Il est typique que le même générateur puisse être utilement considéré dans différentes catégories, notamment topologique et probabiliste.

2.5. Remarque. Une question fondamentale et classique est de comprendre comment l'orbite de 0 sous la translation τ_α ($\alpha \in \mathbb{R}^d$) :

$$(n \cdot \alpha + \mathbb{Z}^d)_{n \in \mathbb{Z}}$$

se répartit dans le tore : où s'accumule-t-elle ? y est-elle uniformément répartie ?

2.6. Exercice. Montrer que l'ensemble des points périodiques est vide ou égale à $\mathbb{T}^d$ selon la valeur de $\alpha \in \mathbb{R}^d$. Déterminez l'ensemble des $\alpha \in \mathbb{R}^d$ pour lequel tous les points du tore sont périodiques.

2.1.3. *Les endomorphismes du tore.*

2.7. Définition. Pour toute matrice $A \in M_d(\mathbb{Z})$ avec $\det(A) \neq 0$, on définit $T_A : \mathbb{T}^d \to \mathbb{T}^d$ par $T_A(\pi(x)) = \pi(Ax)$.

2.8. Lemme. *L'application T_A est bien définie et donc continue.*
C'est de plus un homéomorphisme si et seulement si $\det(A) = \pm 1$.

2.9. Lemme. *Pour tout $A \in M_d(\mathbb{Z})$, l'application $\pi(x) \mapsto \pi(Ax)$ préserve la mesure uniforme $m_{\mathbb{T}^d}$ si et seulement si $\det(A) \neq 0$.*

Démonstration. Si $\det A \neq 0$, $A : \mathbb{R}^d \to \mathbb{R}^d$ est une bijection, en particulier une surjection. On calcule $\tau_\alpha \circ T_A(\pi(x)) = \pi(Ax+\alpha)$ et $T_A \circ \tau_\beta(\pi(x)) = \pi(Ax+A\beta)$.

On abrège $m = m_{\mathbb{T}^d}$. En posant $\beta := A^{-1}\alpha \in \mathbb{R}^d$, on trouve $\tau_\alpha \circ T_A = T_A \circ \tau_\beta$ donc, en servant de cette identité et de l'invariance de m par tout translation :

$$(\tau_\alpha)_*((T_A)_*(m)) = (\tau_\alpha \circ T_A)_*(m) = (T_A \circ \tau_\beta)_*(m) = (T_A)_*((\tau_\beta)_*(m)) = (T_A)_*(m)$$

Donc $(T_A)_*(m)$ est invariante par toutes les translations donc doit coïncider avec la mesure uniforme m.

Si $\det(A) = 0$, il existe une hyperplan H de $\mathbb{R}^d$ tel que $A^{-1}(H) = \mathbb{R}^d$. On en conclut que :

$$(T_A)_*(m)(\pi(H)) = m(\pi(A^{-1}(H))) = 1$$

Or $\pi(H)$ est une union dénombrable de compact de dimension $d-1$. Donc $m(\pi(H)) = 0$. Le morphisme T_A ne préserve donc pas la mesure m. □

2.10. Exercice. Retrouver, en utilisant le théorème de changement de variables que $T : \mathbb{T} \to \mathbb{T}$ défini par $T(\pi(x)) = \pi(d \cdot x)$ pour tout $x \in \mathbb{R}$ préserve bien la mesure $m_{\mathbb{T}}$.

2.11. Corollaire. *Pour tout $d \geq 1$ et $A \in M_d(\mathbb{Z})$ avec* $\det A$*, le morphisme T_A du tore définit :*

- *une dynamique topologique compacte en temps $\mathbb{Z}$: $(\mathbb{T}^d, \mathbb{Z}, T_A)$;*
- *une dynamique probabiliste en temps $\mathbb{Z}$: $((\mathbb{T}^d, \mathcal{B}_{\mathbb{T}^d}, m_{\mathbb{T}^d}), \mathbb{Z}, T_A)$.*

2.12. Exercice. Supposons que $\det(A) = \pm 1$ et que A n'a aucune valeur propre racine de l'unité. Montrer que $\mathrm{Per}(T_A) = \mathbb{Q}^d/\mathbb{Z}^d$. En déduire que $\mathrm{Per}(T_A)$ est dense dans $\mathbb{T}^d$.

2.13. Exercice. Soient un entier $D \geq 2$ et $T_D : \mathbb{T} \to \mathbb{T}$, $\pi(x) \mapsto \pi(D \cdot x)$. Montrer que si $0.c_1c_2c_3\dots$ le développement propre[1] en base D de x, alors
$$\forall n \geq 1 \quad c_n = E(D \cdot T_D^{n-1} x)$$
où $E(t)$ est la partie entière ($E(t) \leq t < E(t) + 1$)

2.2. Exemples symboliques. La *dynamique symbolique* se situe entre dynamique et combinatoire.

2.2.1. *Le décalage.*

2.14. Définition. Pour $d \geq 1$, on pose $\mathcal{A}_d := \mathbb{Z}/d\mathbb{Z}$ le groupe cyclique à d éléments. Le *décalage inversible* sur d symboles est
$$\Sigma_d := \mathcal{A}_d^{\mathbb{Z}} \text{ et } \sigma_d : \Sigma_d \to \Sigma_d,\ (x_n)_{n\in\mathbb{Z}} \mapsto (x_{n+1})_{n\in\mathbb{Z}}.$$
On munit Σ_d de la topologie produit des topologies discrètes.

On écrit $[\![a, b]\!] := [a, b] \cap \mathbb{Z}$.

2.15. Définition. L'ensemble des *mots* sur l'alphabet $\mathcal{A}_d$ est :
$$\mathcal{A}_d^* := \bigcup_{-\infty < p \leq q < +\infty} \mathcal{A}_d^{[\![p,q]\!]}.$$
Le *cylindre* défini par $a \in \mathcal{A}_d^{[\![p,q]\!]}$ est :
$$[a] := \{x \in \Sigma_d : \forall k \in [\![p, q]\!]\ x_k = a_k\}$$
Pour tous $p < q$ finis, on a la partition évidente :
$$\Sigma_d = \bigsqcup_{a \in \mathcal{A}_d^{[\![p,q]\!]}} [a] \tag{1}$$

2.16. Lemme. *L'ensemble des cylindres est une base de fermés-ouverts pour la topologie de Σ_d.*

Démonstration. On note $\pi_k : \Sigma_d \to \mathcal{A}_d$, $x \mapsto x_k$. La topologie produit $\mathcal{T}$ est engendrée par la collection $\mathcal{C}$ des ensembles de la forme suivante :
$$\bigcap_{i \in I} \pi_i{-}1(U_i) \qquad (I \Subset \mathbb{Z},\ U_i \text{ ouvert de } \mathcal{A}_d).$$
La collection $\mathcal{C}$ contient le vide et Σ_d et est stable par intersection finie. On vérifie facilement que les ouverts de $\mathcal{T}$ sont exactement les unions quelconques d'éléments de $\mathcal{C}$.

Pour tout $a \in \mathcal{A}_d^*$, le cylindre : $[a] = \bigcap_{p \leq k \leq q} \pi_k^{-1}(a_k)$ On note que $[a] \in \mathcal{C}$. En fait les éléments de $\mathcal{C}$ sont des unions finies de cylindres. Les cylindres sont donc

1. C'est-à-dire que $c_n \in \{0, \dots, d-1\}$ pour tout $n \geq 1$, $c_n \neq d-1$ pour une infinité de n, et $x = \sum_{n\geq 1} d^{-n} c_n$.

des ouverts. Vu (2), le complémentaire d'un cylindre est une union finie de cylindre. Les cylindres sont donc également fermés. □

2.17. LEMME. *La formule* $d_\Sigma(x,y) := d^{-\inf\{|k| : x_k \neq y_k\}}$ *définit une distance compatible avec la topologie sur* Σ_d.

Une suite $x^k \in \Sigma_d$ *converge vers* $x \in \Sigma_d$ *si et seulement si elle converge terme à terme :*

$$\forall n \in \mathbb{Z}\ \exists K\ \forall k \geq K \qquad x_n^k = x_n.$$

DÉMONSTRATION. On voit facilement que d_Σ est une distance sur Σ_d.

On vérifie que chaque boule ouverte pour la distance $B(x,r) := \{y \in \Sigma_d : d_\Sigma(x,y) < r\}$ coïncide avec un cylindre $[x_{-n} \dots x_n]$ si $d^{-n-1} < r \leq d^{-n}$ donc est ouverte. On voit également que si x est un point d'un ouvert U, alors il existe un cylindre $x \in [a] \subset U$, donc $B(x,r) \subset U$ pour $r > 0$ assez petit. La distance est donc compatible avec la topologie produite.

En considérant les voisinages de x de la forme $[x_n]$, $n \in \mathbb{Z}$, on constate l'implication vers la gauche. Réciproquement, pour tous $p < q$ finis, la convergence terme à terme fournit $N = \max_{p \leq i \leq q} N(i)$ tel que $x^k \in [x_p \dots x_q]$. Comme les cylindres engendrent la topologie, $x^k \to x$. □

2.18. PROPOSITION. *On suppose* $d \geq 2$.

L'espace topologique Σ_d *est compact, sans points isolés et totalement discontinu. En particulier* Σ_d *est homéomorphe à l'espace de Cantor.*

L'application σ_d *est un homéomorphisme.*

2.19. REMARQUE. Tous les Σ_d, $d \geq 2$, sont homéomorphes.

DÉMONSTRATION. Tout produit d'espaces compacts est compact. On peut le démontrer directement par un argument diagonal de Cantor. On vérifie aisément qu'il est non-vide, sans point isolé, et totalement discontinu (il admet une base de fermés-ouverts). Ceci implique l'homéomorphisme avec l'espace de Cantor.

Il est clair que σ_d est continue, chaque $\pi_i(\sigma_d(x))$ ne dépendant que d'un nombre fini de coordonnées de x (en fait $\pi_{i+1}(x)$). Il a un inverse continu : $(x_n)_{n\in\mathbb{Z}} \mapsto (x_{n-1})_{n\in\mathbb{Z}}$. L'application σ_d est donc un homéomorphisme. □

2.20. REMARQUE. On peut démontrer que la dimension de Hausdorff de l'espace métrique (Σ_d, d_Σ) est égale à d. En particulier, si $d \neq d'$, alors aucun homéomorphisme entre Σ_d et $\Sigma_{d'}$ n'est bi-lipschitzien.

2.21. PROPOSITION. *Pour toute probabilité* P *sur* $\mathcal{A}_d$, *la mesure* m_P *définie comme le produit infini* $P^{\otimes\mathbb{Z}}$ *est une mesure e borélienne de probabilité sur* Σ_d *et invariante par* σ_d.

DÉMONSTRATION. On cherche μ_P avec :

$$\forall a \in \mathcal{A}_d^* \quad \mu_P([a]) = d^{-|a|}$$

L'existence de la mesure produit découle par exemple du théorème d'extension de Kolmogorov (voir Théorème 2.22). La condition de compatibilité est une conséquence immédiate de :

$$[b] = \bigsqcup_{a \in \mathcal{A}_d^{[\![p,q]\!]\cup[\![r,s]\!]}} [abc] \text{ avec } d^{-|b|} = \sum_{a \in \mathcal{A}_d^{[\![p,q]\!]\cup[\![r,s]\!]}} d^{-|abc|}.$$

L'invariance quant à elle découle de :

$$\sigma_d^{-1}[a] = [a'] \text{ où } a' = a'_{p-1} \dots a'_{q-1} \text{ avec } a'_k = a_{k+1}$$

avec la remarque $|a'| = |a|$. L'invariance se prolonge à la tribu engendrée par les cylindres, donc à tous les boréliens : $(\sigma_d)_*(m_P) = m_P$. □

Voici le théorème d'extension de Kolmogorov. En fait une version simplifiée qu'on peut trouver dans [**?**, p.XXX])

2.22. Théorème (Kolmogorov). *Soit $(X_i, \mathcal{B}_i)$, $i \in \mathbb{N}$, une suite d'espaces mesurables (ensembles munis d'une tribu).*

Supposons donnée pour toute partie finie $I \subset \mathbb{N}$ une mesure de probabilité μ_I sur $\prod_{i\in I} X_i$ muni de la tribu produit des $\mathcal{B}_i$. Supposons la condition de compatibilité $\mu_J = (\pi_{IJ})_*(\mu_I)$ pour toutes les parties finies $J \subset I \subset N$, $\pi_{IJ} : \prod_{i\in I} X_i \to \prod_{j\in J} X_j$ étant la projection canonique.

Alors il existe une mesure de probabilité μ sur $\prod_{i\in\mathbb{N}} X_i$ muni de la tribu produit vérifiant, pour toute partie finie $I \subset \mathbb{N}$,

$$\mu_I = (\pi_I)_*(\mu)$$

où $\pi_I : \prod_{i\in\mathbb{N}} X_i \to \prod_{i\in I} X_i$ est la projection canonique. De plus cette mesure est unique.

2.23. Corollaire. *Pour tout $d \geq 1$ et toute probabilité P sur $\mathcal{A}_d$:*

— *$(\Sigma_d, \mathbb{Z}, \sigma_d)$ définit une dynamique topologique compacte ;*

— *$((\Sigma_d, \mathcal{B}_{\Sigma_d}, \mu_P), \mathbb{Z}, \sigma_d)$ définit une dynamique probabiliste.*

2.2.2. *Le décalage unilatère.* Il s'agit d'une variante non-inversible du décalage bilatère. Listons ses propriétés (leurs preuves est une simple adaptation du cas bilatère).

2.24. Définition. Le *décalage non-inversible* sur d symboles est

$$\Sigma_d^+ := \mathcal{A}_d^{\mathbb{N}} \text{ et } \sigma_d^+ : \Sigma_d^+ \to \Sigma_d^+,\ (x_n)_{n\in\mathbb{N}} \mapsto (x_{n+1})_{n\in\mathbb{N}}.$$

On munit Σ_d^+ de la topologie produit des topologies discrètes.

2.25. Définition. L'ensemble des *mots positifs* sur l'alphabet $\mathcal{A}_d$ est :

$$\mathcal{A}_d^+ := \bigcup_{0\leq p\leq q<+\infty} \mathcal{A}_d^{[\![p,q]\!]}.$$

Le *cylindre positif* défini par $a \in \mathcal{A}_d^{[\![p,q]\!]}$ est :

$$[a]_+ := \{x \in \Sigma_d^+ : \forall k \in [\![p,q]\!]\ x_k = a_k\}$$

Pour tous $0 \leq p < q$ finis, on a la partition évidente :

$$\Sigma_d^+ = \bigsqcup_{a\in\mathcal{A}_d^{[\![p,q]\!]}} [a]_+ \tag{2}$$

2.26. Lemme. *L'ensemble des cylindres positifs est une base de fermés-ouverts pour la topologie de Σ_d^+.*

2.27. Lemme. *La formule* $d_\Sigma^+(x,y) := d^{-\inf\{k : x_k \neq y_k\}}$ *définit une distance compatible avec la topologie sur* Σ_d^+*.*

Une suite $x^k \in \Sigma_d^+$ *converge vers* $x \in \Sigma_d^+$ *si et seulement si elle converge terme à terme :*

$$\forall n \in \mathbb{N}\ \exists K\ \forall k \geq K \qquad x_n^k = x_n.$$

2.28. Proposition. *On suppose* $d \geq 2$.

L'espace topologique Σ_d^+ *est compact, sans points isolés et totalement discontinu. En particulier* Σ_d^+ *est homéomorphe à l'espace de Cantor.*

L'application σ_d^+ *est une application continue. Chaque point a exactement* d *antécédents :* $(\sigma_d^+)^{-1}(x) = \{ix_0x_1 \ldots : i \in \mathcal{A}_d\}$.

2.29. Remarque. Tous les Σ_d^+ et Σ_d, $d \geq 2$, sont homéomorphes.

On peut démontrer que la dimension de Hausdorff de l'espace métrique (Σ_d^+, d_Σ^+) est égale à d. En particulier, si $d \neq d'$, alors aucun homéomorphisme entre Σ_d^+ et $\Sigma_{d'}^+$ n'est bi-lipschitzien.

2.30. Proposition. *Pour toute probabilité* P *sur* $\mathcal{A}_d$*, la mesure* m_P^+ *définie comme le produit infini* $P^{\otimes\mathbb{N}}$ *est une mesure e borélienne de probabilité sur* Σ_d^+ *et invariante par* σ_d^+*.*

Démonstration. Il suffit d'adapter la preuve de l'invariance en remarquant :

$$(\sigma_d^+)^{-1}([a]_+) = \bigsqcup_{u \in \mathcal{A}_d} [ua]_+$$

Mais, à gauche, il y a d cylindres ayant une mesure réduite d'un facteur d. On donc bien l'invariance pour les cylindres. □

2.2.3. *L'odomètre.* Rappelons-nous que $\Sigma_d^+ = (\mathbb{Z}/d\mathbb{Z})^{\mathbb{N}}$.

2.31. Définition. L'*odomètre à d symboles* est l'application $S_d : \Sigma_d^+ \to \Sigma_d^+$ défini par :

$$S_d(x) = y \iff \forall n \in \mathbb{N} y_n = \begin{cases} x_n + 1 \mod d & \text{si } n < \inf\{k \geq 0 : x_k \neq d-1 \\ x_n & \text{sinon.} \end{cases}$$

On peut identifier Σ_d^+ à l'ensemble des entiers d-adiques et S_d à l'opération $x \mapsto x + 1$ sur ce groupe.

On rappelle que $\mu_d := \mu_P$ avec P la mesure $(1/d, \ldots, 1/d)$ sur $\mathcal{A}_d$.

2.32. Proposition. *L'application* $S_d : \Sigma_d^+ \to \Sigma_d^+$ *est un homéomorphisme.*

Elle préserve une unique mesure borélienne de probabilité : la mesure uniforme μ_d *sur* Σ_d^+*.*

Démonstration. Les n premières coordonnées de $S_d(x)$ ne dépendent que des n premières coordonnées de x. S_d est donc continu. Il admet un inverse du même type donc égalemenrt continu. L'odomètre est donc un homéomorphisme.

Le fait essentiel est le suivant. Pour $a \in \mathcal{A}_d^{[\![0,n]\!]}$, on définit $N(a)$ comme l'entier s'écrivant $a_n a_{n-1} \ldots a_1 a 0$ en base d.

$$\forall a \in \mathcal{A}_d^{[\![0,n]\!]}\ S_d([a]_+) = [b]_+ \text{ si } N(b) = N(a) + 1 \mod d^n \tag{3}$$

En effet, si $a \neq (d-1)^{[\![0,n]\!]}$, alors tout $x \in [a]$ s'écrit $x = ay$ avec y parcourant Σ_d^+ et

$$S_d(ay_0y_1 \ldots) = by_0y_1 \ldots$$

et (3) est vérifiée.

Sinon, $S_d((d-1)^n y_0 y_1 \dots) = 0^n z_0 z_1 \dots$ avec $z = S_d(y)$. On déduit alors de $S_d(\Sigma_d^+) = \Sigma_d^+$, l'équation (3).

On vérifie d'abord que la mesure uniforme m_d^+ est invariante pour S_d.

On montre maintenant que c'est la seule. Soit μ une mesure borélienne de probabilité sur Σ_d^+. Supposons-la Il suffit de montrer que tous les cylindres $[a]_+$, $a \in \mathcal{A}_d^{[\![0,n]\!]}$ ont la même mesure. $\square$

2.3. Flots d'une équation différentielle. Le théorème de Cauchy-Lipschitz à paramètres permet de définir (partiellement) un flot :

2.33. Théorème (Cauchy-Lipschitz). *Soit M une variété différentielle sans bord et $V : M \to M$ un champ de vecteur supposé localement Lipschitz et de classe C^p, $p \geq 0$.*

Il existe un unique $F : U \times M \to M$ avec U ouvert de $\mathbb{R} \times M$ contenant $\{0\} \times M$, différentiable par rapport à la première variable et vérifiant pour tout $x_0 \in M$, si on pose :

$$f_{x_0} : I_{x_0} \to M,\ t \mapsto F(t, x_0) \text{ avec } U \cap (\mathbb{R} \times M) = I_{x_0} \times M$$

alors f_{x_0} est l'unique solution maximale du problème de Cauchy :

$$x(0) = x_0,\ x'(t) = V(x(t)).$$

De plus F et $\partial_t F$ sont C^p.

On renvoie à ***** pour une preuve.

Un champ de vecteurs, même régulier, peut ne pas admettre de solution maximale définie sur $\mathbb{R}$ tout entier. Ceci peut amener à de délicats problèmes entre analyse de la théorie des équations différentielles et dynamique proprement dite. L'outil fondamental pour étudier ce genre de question est le résultat suivant.

2.34. Corollaire. *Si $(t, x), (t+s, x) \in U$, alors $F(t, F^s(x)) = F(t+s, x)$.*

2.35. Théorème (Explosion). *Soit M une variété différentielle sans bord et $V : M \to M$ un champ de vecteur supposé localement Lipschitz et de classe C^p, $p \geq 0$.*

Si $T^+ := \sup I_{x_0} < \infty$, alors, pour tout compact $K \subset M$, il existe $t_n \in I_{x_0}$ avec $t_n \to T^+$ et $F(t_n, \cdot) \notin K$.

De même si $\inf I_{x_0} > -\infty$.

2.36. Corollaire. *Soit M une variété différentielle sans bord et $V : M \to M$ un champ de vecteur supposé localement Lipschitz et de classe C^p, $p \geq 0$. On définit $F : U \to M$ comme ci-dessus.*

Si M est une variété compacte et sans bord, alors F est défini pour tous les temps, ie, sur $U = \mathbb{R} \times M$. C'est une action de $\mathbb{R}$ sur M. Elle définit ainsi une dynamique topologique compacte en temps $\mathbb{R}$.

Si, de plus, le champ de vecteurs V est à divergence nulle : $\sum_{i=1}^d \partial F_i \partial x_i = 0$, alors le flot préserve la mesure uniforme $m_{\mathbb{T}^d}$. On définit ainsi une dynamique probabliste en temps $\mathbb{R}$.

2.37. Exercice. Démontrer le corollaire ci-dessus.

3. Propriétés asymptotiques topologiques et mesurées

La dynamique s'attache aux propriétés asymptotiques des orbites.

Orbites périodiques.

3.1. DÉFINITION. Soit (X, T, F) une dynamique topologique en temps $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+$ ou $\mathbb{R}$.

Si $x = F^T(x)$ avec $T \neq 0$, on dit que x est périodique et que T est une période de x.

Si $\{t \in T : F^t(x) = x\} = T\mathbb{Z}$ on dit que x est exactement T-périodique.

3.2. EXERCICE. Soit (X, T, F) une dynamique topologiquer. Si x est périodique, vérifiez que $\mathcal{O}_F(x) = \mathcal{O}_F^+(x) = \omega_F(x)$. La réciproque est-elle vraie ?

3.3. EXERCICE. On considère la dynamique compacte $(\mathbb{T}^d, \mathbb{Z}, T_A)$ où $d \geq 2$, $A \in GL_d(\mathbb{Z})$ et la condition (*) : aucune valeur propre de A n'est racine de l'unité.

Sous la condition (*), montrer que $\mathrm{Per}(T_A) = \mathbb{Q}^d/\mathbb{Z}^d$. La condition (*) est-elle nécessaire ?

Que se passe-t-il si $A \in M_d(\mathbb{Z})$ avec $|\det A| \geq 2$?

Asymptotique topologique : ensembles omega-limites.

3.4. DÉFINITION. Soit (X, T, F) une dynamique topologique et $x \in X$. L'*ensemble omega-limite* de x sous F est :

$$\omega_F(x) := \{\lim_k F^{t_k}(x) : t_k \nearrow \infty, \text{ pourvu que la limite existe}\}$$

La terminologie vient de ce que ω est la dernière lettre de l'alphabet grec.

On cherchera à comprendre ces ensembles selon la condition initiale x, quitte à exclure des cas exceptionnels si nécessaire.

3.5. PROPOSITION. *Pour toute dynamique topologique, chaque ensemble ω-limite est un fermé et est invariant par F : $\forall t \in T\ F^t(\omega_F(x)) = \omega_F(x)$.*

Si X est compact, alors les ensembles omega-limites sont compacts et non-vides.

DÉMONSTRATION. L'ensemble limite est fermé par construction.

Si $y \in \omega_F(x)$, il existe $t_k \nearrow \infty$ avec $y = \lim_k F^{t_k}(x)$. F étant continue, $F^s(y) = lim_k F^{t_k+s}(x)$: $F^s(y) \in \omega_F(x)$ donc $F^s(\omega_F(x)) \subset \omega_F(x)$.

Si $y = \lim_k F^{t_k}(x) \in \omega_F(x)$ et $s \in T$, alors $z = \lim_k F^{t_k - s}(x) \in \omega_F(x)$ ($t_k - s$ est bien défini pour t_k grand et $t_k - s \to \infty$) et $F^s(z) = y$. Donc $\omega_F(x) \subset F^s(\omega_F(x))$. D'où l'égalité. $\square$

3.6. DÉFINITION. Soit (X, T, F) une dynamique topologique. On dit que $x \in X$ est récurrent pour F s'il existe $t_k \nearrow \infty$ vérifiant $x = \lim_k F^{t_k}(x)$.

Autrement dit, x est récurrent si et seulement si $x \in \omega_F(x)$.

3.7. EXERCICE. Soit $(\mathbb{T}^d, \tau_\alpha)$, $\alpha \in \mathbb{R}^d$. On omet l'indice α.

(1) pour tous $x, y \in \mathbb{T}^d$, montrer que $\omega_\tau(x) = \omega_\tau(y) + (x - y)$;
(2) montrer que soit aucun $x \in \mathbb{T}^d$ n'est récurrent pour τ, soit tout $x \in \mathbb{T}^d$ est récurrent ;
(3) montrer que $\omega_F(0)$ est un sous-groupe fermé de $\mathbb{T}^d$.

3.8. EXERCICE. On considère le décalage bilatère $(\Sigma_d, \mathbb{Z}, \sigma_d)$. On omet les indices d. On considère $x \in \Sigma$.

— Trouver x tel que $\omega_\sigma(x) \neq \overline{\mathcal{O}^+(x)}$;
— Montrer que $\omega_\sigma(x) = \Sigma$ si et seulement si $\overline{\mathcal{O}^+(x)} = \Sigma$.
— Montrer que $\omega_\sigma(x) = \Sigma$ si et seulement si tout mot fini sur $\mathcal{A}_d$ apparaît dans $(x_n)_{n \geq 0}$.

Asymptotique mesurée : mesures empiriques.

Un peu d'analyse fonctionnelle. On souhaite munir l'ensemble des mesures de probabilités d'une topologie compacte. On utilise pour cela l'identification suivante :

3.9. Théorème (Riesz). *Soit X un espace métrique localement compact et dont tout ouvert est une union dénombrable de compacts muni de sa tribu borélienne $\mathcal{B}$.*

Il y a une bijection entre :

- *les mesures boréliennes positives $\mu : \mathcal{B} \to [0, \infty]$ finies sur les compacts et ;*
- *les formes linéaires positives $\Lambda : C_c(X) \to \mathbb{R}$ avec $C_c(X)$ l'ensemble des fonctions continues à support compact ;*

via l'identification : $\forall \phi \in C_c(X)\ \Lambda(\phi) = \mu(\phi)$.

De plus, les mesures ci-dessus sont régulières : pour tout $B \in \mathcal{B}$,

$$\mu(B) = \sup\{\mu(K) : K \subset B \text{ et } K \text{ compact}\} = \inf\{\mu(U) : U \supset B \text{ et } U \text{ ouvert}\}.$$

3.10. Définition. On dira dans ce cours que X est un *espace du théorème de Riesz* si c'est un espace métrique localement compact et dont tout ouvert est une union dénombrable de compacts

3.11. Définition. Les mesures boréliennes positives, finies sur les compacts et régulières sont appelées *mesures de Radon*. On note $\mathrm{Rad}(X)$ l'ensemble de ces mesures et on le munit de la topologie vague, induite par la topologie * faible du dual $C_c(X)^*$ et de la norme : $\|\mu\| := \sup\{\mu(\phi) : \phi \in C_c(X) : \sup \|\phi\| \leq 1\}$.

A préciser : La norme ci-dessus est-elle la variation totale ?

La topologie vague de $\mathrm{Rad}(X)$ est engendrée par les ensembles :

$$U(\mu, \phi, \epsilon) := \{\nu \in \mathrm{Rad}(X) : |\nu(\phi) - \mu(\phi)| < \epsilon\} \quad (\mu \in \mathrm{Rad}(X), \phi \in C_c(X), \epsilon > 0)$$

En restriction à la boule unité $S\mathbb{P}(X) := \{\mu \in \mathrm{Rad}(X) : \|\mu\| \leq 1\}$, cette topologie est métrisable. On peut obtenir une distance compatible de la façon suivante.

3.12. Exercice. Soit X un espace du théorème de Riesz. Construire une suite de fonctions $\phi_1, \phi_2, \cdots \in C_c(X)$ avec $\sup_{i,x} |\phi_i(x)| \leq 1$ dont les combinaisons linéaires soient dense dans $C_c(X)$.

Indication : on pourra utiliser le lemme d'Uryshon, selon lequel dans tout espace métrique X (donc normal), pour tous fermés disjoints A, B, il existe une fonction continue $f : X \to [0, 1]$ avec $f|_A = 0$ et $f|_B = 1$.

En déduire que la formule suivante définit une distance sur $S\mathbb{P}(X)$ compatible avec la topologie * faible :

$$\forall \mu, \nu \in S\mathbb{P}(X) \quad d(\mu, \nu) := \sum_{i \geq 1} 2^{-i} |\mu(\phi_i) - \nu(\phi_i)|.$$

La formule précédente s'étend-elle à $\mathrm{Rad}(X)$ tout entier ?

3.13. Théorème. *Soit X un espace du théorème de Riesz. Alors $S\mathbb{P}(X)$ muni de la topologie vague est un compact métrisable.*

Démonstration. L'exercice précédent a construit une distance.

Le théorème de Banach-Acaoglu montre que, dans tout espace vectoriel normé E, la boule unité fermée $\{x \in E : \|x\| \leq 1\}$ est compact dans la topolgoie * faible. D'où l'affirmation. □

On s'intéresse tout particulièrement à l'ensemble des mesures de probabilité :

$$\mathbb{P}(X) := \{\mu \in \mathrm{Rad}(X) : \mu(X) = 1\}.$$

L'exemple des mesures de Dirac aux points $n \to \infty$, $\delta_n \in \mathbb{P}(\mathbb{R})$, montre que $\mathbb{P}(X)$ n'est pas compact sous les hypothèses précédentes. La remarque triviale selon laquelle $1 \in C_c(X)$ dans le cas où X est compact, montre que dans ce cas $\mathbb{P}(X)$ est fermé donc compact :

3.14. Corollaire. *Si X est un espace métrique compact, alors $\mathbb{P}(X)$ est compact pour la topologie vague.*

Mesures empiriques et quasi-empiriques. On essaie de définir des statistiques à partir d'une orbite donnée pour un système dynamique topologique $(X, \mathbb{N}, f)$ où X est un espace du théorème de Riesz. Les définitions ci-dessous se généralisent aisément aux temps $\mathbb{Z}$ (sans changement), $\mathbb{R}_+$ ou $\mathbb{R}$ (en utilisant des intégrales sur le temps à la place des sommes).

3.15. Définition. Un point $x \in X$ admet la *mesure empirique* $\mu_x \in \mathbb{P}(X)$ si

$$\mu_x := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k(x)}$$

où la limite est au sens de la topologie vague.

L'ensemble des *mesures quasi-empiriques* est

$$M_x := \left\{ \lim_i \frac{1}{n_i} \sum_{k=0}^{n_i-1} \delta_{f^k(x)} \in \mathbb{P}(X) : n_i \to \infty \right\}.$$

3.16. Lemme. *Soit (X, T, f) une dynamique probabiliste en temps $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+, \mathbb{R}$.*

- *Si X est un espace du théorème de Riesz, alors M_x est un fermé pour tout $x \in X$. De plus tout $\mu \in M_x$ est invariante : $f_*(\mu) = \mu$;*
- *Si X est un espace métrique compact, alors, pour tout $x \in X$, M_x est une partie compacte non-vide de $\mathbb{P}(f) := \{\mu \in \mathbb{P}(X) : f_*(\mu) = \mu\}$, l'ensemble des mesures boréliennes de probabilité invariantes.*

Démonstration. Un argument diagonal de Cantor permet de montrer la fermeture de M_x, ensemble de points d'accumulation. Remarquons que $\mu \mapsto f_*(\mu)$ est continue pour la topologie vague dans $S\mathbb{P}(X)$. Pour tout $\phi \in C_c(X)$, on a :

$$(f_*(\mu) - \mu)(\phi) = \lim_i \frac{1}{n_i} \sum_{k=0}^{n_i-1} (\phi(f^{k+1}x) - \phi(f^k x)) = \lim_i \frac{\phi(f^{n_i}x) - \phi(x)}{n_i} = 0$$

Mais $C_c(X)$ est dense dans $L^1(\mu + f_*(\mu))$ donc l'identité précédente s'étend à $\phi \in L^1(\mu + f_*(\mu))$ et donc aux fonctions indicatrices mesurables. D'où $f_*(\mu) = \mu$. □

3.17. Exercice. Montrer par un exemple que la convergence vague $\mu_k \to \mu$ dans un espace métrique compact n'implique pas que pour toute fonction ϕ bornée et mesurable, on ait $\mu_k(\phi) \to \mu(\phi)$.

3.18. Exercice. Montrer que le décalage bilatère $(\Sigma_d, \mathbb{Z}, \sigma_d)$ admet un point $x \in \Sigma_d$ tel que $\#M_x > 1$ et donc sans mesure empirique.

Trouver un G_δ dense [2] $X \subset \Sigma_d$ tel que, pour tout $x \in X$, $M_x = \mathbb{P}(f)$.

2. Un G_δ désigne une intersection dénombrable d'ouverts. Voir **********

4. Conjugaison, classification et stabilité

4.1. Généralités. Comme toute théorie mathématique, la dynamique cherche à classifier ses objets. La notion d'isomorphisme naturelle est la conjugaison.

Conjugaison.

4.1. Définition. Une *conjugaison d'un type donné* entre deux systèmes dynamiques (X, T, F) et (Y, T, G) en un même temps T est un isomorphisme du type donné entre les espaces $h : X \to Y$ vérifiant :

$$\forall t \in T \quad h \circ F^t = G^t \circ h.$$

La notion de conjugaison est à décliner en fonction des structures considérées. On sera souvent amener à étudier des isomorphismes préservant une structure plus faible que la structure initiale (par exemple la conjugaison de type topologique entre des dynamiques différentiables).

Ceci généralise la notion de changement de variable. En particulier, on ne peut conjuguer que des dynamiques de même temps.

Notons que si $T = \mathbb{N}$ ou $\mathbb{R}$, l'identité ci-dessus est équivalente à

$$h \circ f = g \circ h$$

en notant f, g les générateurs de F, G.

Invariant et 'invariant complet. La stratégie usuelle pour obtenir, ou approcher une classification, est d'introduire des *invariants de conjugaison* et, éventuellement de les enrichir dans l'espoir d'obtenir des *invariants complets*

4.2. Définition. Etant donné une classe $\mathcal{D}$ de dynamiques et une notion de conjugaison sur cette classe, un *invariant de conjugaison* est une application $I : \mathcal{D} \to \mathcal{O}$ vérifiant :

$$\forall F, G \in \mathcal{D} \quad F \text{ conjuguée à } G \implies I(F) = I(G)$$

L'invariant est *complet* si, de plus, la réciproque a lieu :

$$\forall F, G \in \mathcal{D} \quad F \text{ conjuguée à } G \iff I(F) = I(G).$$

L'invariant prend des valeurs *égales* sur des dynamiques *conjuguées*. Il n'a d'intérêt que s'il est "raisonnable", contrairement, e.g., à l'invariant complet suivant :

$$I : F \in \mathcal{D} \longmapsto \{G \in \mathcal{D} : F, G \text{ conjuguées}\}$$

Foreman, Rudolph et Weiss ont formalisé cette notion d'invariant raisonnable et montré qu'il n'existe pas d'invariant complet raisonnable dans bien des cas naturels. Nous verrons plusieurs situations dans lesquelles il existe bien des invariants (voire des invariants complets) raisonnables.

Stabilité. C'est une version locale du problème de classification.

4.3. Définition. Soit $\mathcal{D}$ une classe de dynamiques muni d'une certaine topologie. Fixons une notion de conjugaison d'un certain type entre les éléments de $\mathcal{D}$.

On dit que $f \in \mathcal{D}$ est *stable dans $\mathcal{D}$ pour le type* si tout $g \in \mathcal{D}$ assez proche de f admet une conjugaison du type donné avec f.

Semiconjugaison. Parfois on doit accepter d'oublier certains détails avant de pouvoir ramener une dynamique à une autre. Cela correspond à la notion suivante :

4.4. DÉFINITION. Une *semiconjugaison d'un type donné* entre deux systèmes dynamiques (X, T, F) et (Y, T, G) en un même temps T est un *endomorphisme* du type donné entre les espaces $h : X \to Y$ vérifiant :

$$\forall t \in T \quad h \circ F^t = G^t \circ h.$$

On dit alors que F est une *extension* et G est un *facteur*.

Comme la semiconjugaison h n'est plus un isomorphisme mais un endomorphisme, on s'attend à ce que la dynamique F soit plus riche que la dynamique G. Cela arrive souvent lorsqu'on cherche à ramener une dynamique lisse ou continue à un modèle algébrique.

4.2. Conjugaison topologique.

4.5. DÉFINITION. Soient (X, T, F) et (Y, T, G) deux dynamiques topologiques en un même temps T. Elles sont *topologiquement conjuguées* s'il existe un homéomorphisme $h : X \to Y$ vérifiant, pour tout $x \in X$,

$$\forall t \in T \quad h \circ F^t(x) = G^t \circ h(x)$$

Elles sont *topologiquement semiconjuguées* s'il existe une application continue et surjective $h : X \to Y$ satisfaisant l'équation précédente.

4.6. EXERCICE. Soient $(\Sigma_d^+, \mathbb{N}, \sigma_d^+)$ le décalage unilatère et $(\mathbb{T}, \mathbb{N}, T_d)$ la multiplication par d modulo 1.

Montrer que ces deux systèmes ne sont pas topologiquement conjugués.

Montrer qu'ils sont topologiquement semiconjugués. On pourra considérer l'application $\pi : \Sigma_d^+ \to \mathbb{T}$ définie par :

$$\pi : \alpha \longmapsto \sum_{n \geq 0} \frac{\alpha_n}{d^{n+1}} + \mathbb{Z}$$

4.3. Conjugaison en mesure.

4.7. DÉFINITION. Soient $(X = (\Omega, \mathcal{B}, \mu), T, F)$ et $(Y = (\Omega', \mathcal{B}', \mu'), T, G)$ deux dynamiques mesurées en un même temps T. Elles sont *conjuguées* s'il existe un isomorphisme $h : X \to Y$ vérifiant,

$$\mu - \forall x \in X \; \forall t \in T \quad h \circ F^t(x) = G^t \circ h(x).$$

4.8. REMARQUE. Si T est dénombrable, il est équivalent de demander :

$$\forall t \in T \; \mu - \forall x \in X \quad h \circ F^t(x) = G^t(x) \circ h(x).$$

A préciser : Si T ne l'est pas, les conditions sont distinctes.

4.9. EXERCICE. Soient $((\Sigma_d^+, \mathcal{B}_d, m_d), \mathbb{N}, \sigma_d^+)$ le décalage unilatère muni de la tribu borélienne et de la mesure uniforme et $((\mathbb{T}, \mathcal{B}_\mathbb{T}, m_\mathbb{T}), \mathbb{N}, T_d)$ la multiplication par d modulo 1 munie de la tribu borélienne et de la mesure de Lebesgue.

Montrer que ces deux systèmes sont conjugués en mesure.

Invariants spectraux. On considère une dynamique probabiliste $((X,\mathcal{B},\mu),\mathbb{Z},f)$.

4.10. Définition. L'opérateur de Koopman associé est :

$$U_f : L^2(\mu) \to L^2(\mu), \quad \phi \longmapsto \phi \circ f$$

Comme f préserve μ et est inversible, l'opérateur de Koopman est unitaire et inversible sur l'espace de Hilbert $L^2(\mu)$. On vérifie aisément que si h est une conjugaison mesurée de (f,μ) avec une autre dynamique probabiliste (g,ν), alors $U_h i : L^2(\mu) \to L^2(\nu)$, $\phi \mapsto \phi \circ h$ est un isomorphisme unitaire entre les deux opérateurs de Koopman :

$$U_h \circ U_f = U_g \circ U_h$$

En particulier, le spectre discret de f :

$$\sigma(f) := \{\lambda \in \mathbb{C} : \exists \phi \in L^2(\mu) \setminus 0 \; U_f(\phi) = \lambda\phi\}$$

doit alors coïncider avec celui de g :

4.11. Lemme. *Le spectre discret est un invariant de conjugaison mesurée entre dynamiques probabilistes en temps $\mathbb{Z}$.*

4.12. Exercice. Soit $d \geq 1$. Pour $\alpha \in \mathbb{R}^d$, τ_α désignera aussi bien l'application du tore que le système dynamique probabiliste engendré par τ_α et la mesure uniforme m. On considère les vecteurs "totalement irrationels" :

$$A_d := \{\alpha \in \mathbb{R}^d : (1,\alpha_1,\dots,\alpha_d) \text{ famille } \mathbb{Q}\text{-libre}\}.$$

(1) On suppose que τ_α, τ_β $(\alpha,\beta \in \mathbb{R}^d)$ sont conjugués en mesure. Calculer $\sigma(\tau_\alpha)$ en utilisant les séries de Fourier.

(2) En déduire qu'il existe $M \in M_d(\mathbb{Z})$ et $A \in \mathbb{Z}^d$ tels que : $\alpha = M\beta + A$. Puis que $\beta = N\alpha + B$ avec $V \in M_d(\mathbb{Z})$ et $B \in \mathbb{Z}^d$.

(3) On suppose que $\alpha \in A_d$. Montrer que $\mathrm{Id} = MN$. En déduire que $M, N \in GL_d(\mathbb{Z})$. *[Indication : on pourra utiliser l'isomorphisme $\mathbb{Q}$-linéaire $(u_0,\dots,u_d) \mapsto \sum_{i=0}^d \alpha_i u_i$ où $\alpha_0 := 1$.]*

(4) En déduire que la conjugaison en mesure de τ_α, τ_β avec $\alpha \in A_d$ implique que $\alpha = Q\beta$ avec $Q \in GL_d(\mathbb{Z})$ et la conjugaison topologique de τ_α, τ_β par l'automorphisme T_Q.

(5) Montrer que $\sigma(U_{\tau_\alpha})$ est un invariant complet pour la conjugaison mesurée des dynamiques probabilistes τ_α, $\alpha \in A_d$.

4.13. Exercice. Soit (X,T,F) une dynamique topologique compacte. On suppose qu'il existe un groupe agissant sur X par homéomorphimes notés $\alpha^g := (x \mapsto g.x)$ et commutant avec la dynamique : (i) $\forall x,y \in X \; \exists g \in G \; g.x = y$; (ii) $\forall g \in G$ $\forall t \in T \; F^t \circ \alpha^g = \alpha^g \circ F^t$.

Montrer que tous les points y de la G-orbite de x ont des mesures quasi-empiriques conjuguées : pour tous $g \in G$, les ensembles M_x et $M_{g.x}$ sont en bijection et les mesures correspondantes sont conjuguées en mesure.

Chapitre 2

Constructions

1. Produits et cocycles

Donnons l'idée générale :

1.1. Méta-Définition. [1] Le *produit (direct)* de deux dynamiques (X, T, F) et (Y, T, G) de la même catégorie (topologique, probabiliste, etc.) en un même temps T, est le système $(X \times Y, T, F \times G)$ avec

$$F \times G : \mathbb{R} \times (X \times Y) \to X \times Y, \quad (t, x) \longmapsto (F(t, x), G(t, x)).$$

1.2. Définition. Si X, Y sont des espaces topologiques, alors $X \times Y$ est le produit ensembliste muni de la topologie produit. Si X, Y sont des espaces métriques, alors on munit le produit de la distance

$$d((x, y), (x', y')) := \max(d(x, x'), d(y, y')).$$

1.3. Définition. Si X, Y sont des espaces probabilistes, alors $X \times Y$ est le produit ensembliste muni de la mesure produit sur la tribu produit.

A préciser : produit des tribus boréliennes et tribu borélienne du produit

1.4. Exemple. *Soient $\alpha \in \mathbb{R}^a$, $\beta \in \mathbb{R}^b$. On vérifie que le produit de $\tau_\alpha : \mathbb{T}^a \to \mathbb{T}^a$, et de $\tau_\beta : \mathbb{T}^b \to \mathbb{T}^b$ est $\tau_{\alpha\beta}$ où $\alpha\beta \in \mathbb{R}^{a+b}$. Ces trois translations peuvent être vues comme dynamiques topologiques ou comme dynamiques probabilistes.*

1.5. Méta-Définition. Un *cocycle* au-dessus d'une dynamique (X, T, F) est la donnée $(G_x^t)_{t \in T, x \in X}$ d'une famille “régulière” d'endomorphismes d'un espace Y vérifiant :

$$(i)\ G_x^0 = \mathrm{Id}_Y,\ (ii)\ \forall t, s \in T\ G_x^{t+t} = G_{F^s(x)}^t \circ G_x^s.$$

L'action induite sur $X \times Y$ est $H : T \times (X \times Y) \to X \times Y$ est appelée *produit tordu (ou gauche)* :

$$(x, y) \longmapsto (F^t(x), G_x^t(y)).$$

1.6. Remarques. On considère un cocycle au sens de la méta-définition précédente. On vérifie facilement :

(1) L'application $H : T \times (X \times Y) \to X \times Y$ est une action de T sur $X \times Y$;
(2) Si $G_x^t = G^t$ est indépendant de x, alors la dynamique induite par le cocycle est le produit direct $F \times G$. En particulier, toute dynamique s'identifie à un cocycle au-dessus de la dynamique triviale (point fixe).

1. Par méta-définition, je veux dire : un modèle de définition dans lequel on doit préciser les structures et régularités.

(3) Pour les temps classiques discrets $T = \mathbb{N}$ ou $\mathbb{Z}$, un cocycle se définit par une famille "régulière" d'endomorphismes ou d'isomorphismes g_x de l'espace Y.

Donnons maintenant les définitions dans les catégories pertinentes :

1.7. Définition. Un *cocycle topologique*[2] est un cocycle vérifiant :

(i) (X, T, F) est une dynamique topologique ;

(ii) chaque $G_x^t : Y \to Y$ est une application continue d'un espace topologique Y (voire compact) dans lui-même ;

(iii) $(t, x, y) \mapsto G_x^t(y)$ est continue[3].

1.8. Définition. Un *cocycle probabiliste* est un cocycle vérifiant :

(i) (X, T, F) est une dynamique probabiliste ;

(ii) pour μ-p.t. $x \in X$ et tout $t \in T$, $G_x^t : Y \to Y$ est un endomorphisme d'un espace de probabilité Y dans lui-même ;

(iii) $(t, x, y) \mapsto G_x^t(y)$ est mesurable pour la tribu produit[4].

1.9. Remarques. On vérifie facilement les propriétés suivantes :

(1) L'action induite par un cocycle topologique, compact ou probabiliste est une dynamique du même type ;

(2) Si $\alpha : X \to \mathbb{R}^d$ est une application mesurable, la famille $g_x = \tau_{\alpha(x)}$ d'isomorphismes de l'espace de probabilité $(\mathbb{T}^d, \mathcal{B}_{\mathbb{T}^d}, m_{\mathbb{T}^d})$ engendre un cocycle probabiliste au-dessus de la dynamique $(X, \mathbb{Z}, F)$.

1.10. Exercice. Pour tout $\alpha \in \mathbb{R}$, écrire $\alpha \cdot n^2 \mod 1$ sous la forme $\phi \circ T^n(x_0)$ où $\phi : \mathbb{T}^2 \to \mathbb{T}$ est une fonction continue, $x_0 \in \mathbb{T}^2$ et T est une extension par le groupe $\mathbb{T}$ d'une rotation de $\mathbb{T}$.

Généraliser en remplaçant $\alpha \cdot n^2$ par un polynôme réel arbitraire P.

2. Extensions naturelles : passage à une dynamique inversible

On suppose que T est un temps classique : $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+$ ou $\mathbb{R}$.

L'extension naturelle est en un sens que nous allons préciser "la plus petite des extensions en temps T d'un système donné en temps $T_+ := T \cap \mathbb{R}_+$". On la construit comme l'action naturelle sur l'ensemble des *orbites complètes* de F qu'on projette sur leur position à l'instant 0.

Notations. On désigne par $F|T_+$ la restriction $F|T_+ : T_+ \times X \to X$. Il est commode d'utiliser X, F ou (X, T, F) comme synonymes les uns des autres.

Dans le cadre topologique, on a la définition suivante :

2.1. Définition. Soit $T = \mathbb{Z}, \mathbb{R}$ et $T_+ := T \cap \mathbb{R}_+$. L'*extension naturelle* d'une dynamique topologique (X, T_+, F) est la dynamique topologique $(\widehat{X}, T, \widehat{F})$ donnée par :

— $\widehat{X} := \{(x_s)_{s \in \mathbb{Z}} \in X^{\mathbb{Z}} : \forall s \in \mathbb{Z} \forall t \in \mathbb{N}\ F^t(x_s) = x_{s+t}\}$ muni de la topologie produit ;

2. Si les espaces X et Y sont métriques ou compacts, c'est alors le cas de l'espace produit qui est défini comme dans le cas du produit direct.

3. C'est la régularité mentionnée dans la méta-définition.

4. C'est la régularité mentionnée dans la méta-définition.

— $\forall t \in T\ \widehat{F}^t((x_s)_{s\in\mathbb{Z}}) = (F^{t-E(t)}(x_{s+E(t)}))_{s\in\mathbb{Z}}$;

— $\widehat{\pi} : \widehat{X} \to X,\ (x_s)_{s\in\mathbb{Z}} \longmapsto x_0$.

Si X est un espace métrique, on pose

$$\forall \widehat{x}, \widehat{y} \in \widehat{X} \quad d(\widehat{x}, \widehat{y}) := \sum_{n\in\mathbb{N}} 2^{-n} \min(d(x_{-n}, y_{-n}), 1).$$

On vérifie que ceci définit bien une dynamique topologique avec les propriétés annoncées.

2.2. Proposition. *Soit F une dynamique topologique en temps T_+ supposée surjective : pour tout $t \in T_+$, $F^t(X) = X$.*

Son extension naturelle $\widehat{F}$ est une dynamique topologique et $\widehat{\pi}$ est une semiconjugaison topologique faisant de $\widehat{F}$ une extension de F. Si de plus X est compact, resp. métrique, alors $\widehat{X}$ est compact, resp. métrique.

Si $\widehat{\pi}$ est 'injective au niveau des orbites complètes" dans le sens suivant :

$$\widehat{x} = (\widehat{\pi} \circ \widehat{F}^n(\widehat{x}))_{n\in\mathbb{Z}}. \tag{4}$$

Démonstration. On se place dans le cas $T := \mathbb{Z}$. On reprend les notations de la définition : (X, T_+, F), $(\widehat{X}, T, \widehat{F})$ et $\widehat{\pi} : \widehat{X} \to X$.

Comme F^t est continu pour chaque $t \in T$, $\widehat{X}$ est un fermé de $X^{\mathbb{Z}}$. Si X est compact, alors $X^{\mathbb{Z}}$ et donc $\widehat{X}$ le sont.

L'application $\widehat{F}^t : \widehat{X} \to \widehat{X}$ est bien définie par la formule : $(F^{t-E(t)}(x_{s+E(t)}))_{s\in\mathbb{Z}}$ car $t - E(t) \geq 0$. On vérifie que c'est une action :

$$\begin{aligned}\widehat{F}^u \circ \widehat{F}^v(\widehat{x}) =& (F^{u-E(u)+v-E(v)}(x_{s+E(u)+E(v)}))_{s\in\mathbb{Z}} \\ =& (F^{u+v-E(u+v)+k}(x_{s+E(u+v)+k}))_{s\in\mathbb{Z}} = \widehat{F}^{u+v}(\widehat{x}),\end{aligned}$$

où $k = E(u) + E(v) - E(u+v) \in \mathbb{Z}$ et en utilisant l'identité : $F^k(x_{s-k}) = x_s$.

L'application $F^t : \widehat{X} \to \widehat{X}$ est continue d'après la définition de la topologie produit.

On a montré que $\widehat{F}$ est une dynamique topologie [métrique ou compacte si F l'est]. Montrons que c'est une extension topologique de F par l'application $\widehat{\pi}$.

L'application $\widehat{\pi} : \widehat{X} \to X$ est continue d'après la définition de la topologie produit.

Elle est surjective car la surjectivité de F permet de construire une préimage $\widehat{x} \in \widehat{\pi}^{-1}(x)$ pour tout $x \in X$ de la façon suivante. Pour $n \in \mathbb{N}$, on pose $x_n = F^n(x)$. On définit par récurrence le passé : pour $n \in \mathbb{N}$, $x_{-n-1} \in F^{-1}(x_{-n})$ (non-vide car $F^t : X \to X$ est surjective).

Enfin, pour tout $t \in T_+$,

$$\widehat{\pi} \circ \widehat{F}^t(\widehat{x}) = \widehat{\pi}((F^{t-E(t)}x_{s+E(t)})_{t\in T} = F^{t-E(t)}(x_{E(t)}) = F^t(x_0) = F^t \circ \widehat{\pi}(\widehat{x})$$

donc $\widehat{\pi}$ continue et surjective est une semiconjugaison topologique. □

On traite la compatibilité de la métrique à part.

2.3. Lemme. *Si X est métrique, alors $\widehat{X}$ l'est pour la distance*[5] *$\hat{d}$ dont on vérifie la compatibilité avec la topologie produit.*

5. On restreint la somme aux temps entiers négatifs, y compris dans le cas $T = \mathbb{R}$.

Démonstration. La topologie produit est engendrée par les cylindres definis par les familles finies $(t_i, U_i)_{i\in I}$ ($t_i \in T$, U_i ouvert de X) selon :

$$C((t_i, U_i)_{i\in I}) := \{x \in \widehat{X} : \forall i \in I \;\; x_{t_i} \in U_i\}$$

Montrons que pour tout x appartenant à ce cylindre, une boule de rayon suffisamment petit centrée en x est incluse dans ce cylindre. Pour chaque $i \in I$, soit $n_i := E(t_i)$, la partie entière. Par continuité de $F^{t_i - n_i}$, il existe $\delta_i > 0$ tel que $d(x_{n_i}, y_{n_i}) < \delta \implies y_{t_i} \in U_i$. On voit donc que $B(x, \min 2^{-n_i}\delta_i) \subset C((t_i, U_i)_{i\in I})$.

Réciproquement si $x \in \widehat{X}$ et $\epsilon > 0$, on a

$$C((-k, B(x_{-k}, \epsilon/2))_{0\leq k\leq N}) \subset B(x, \epsilon)$$

si $2^{-N+1} < \epsilon/2$. □

2.4. Proposition. *Soit F une dynamique topologique en temps T_+ supposée surjective : pour tout $t \in T_+$, $F^t(X) = X$.*

Si $p : \widehat{G}|\mathbb{N} \to F$ est une semiconjugaison topologique avec $\widehat{G}$ une dynamique topologique en temps $\mathbb{Z}$ alors $p = \widehat{\pi} \circ \widehat{p}$ pour une unique application continue $\widehat{p} : \widehat{G} \to \widehat{F}$ telle que $\widehat{p} \circ \widehat{G} = \widehat{F} \circ \widehat{p}$.

Si $\widehat{G}$ est compacte, alors $\widehat{p} : \widehat{G} \to \widehat{F}$ est surjective et donc une semiconjugaison topologique.

Démonstration. Soit $(\widehat{Y}, \mathbb{Z}, \widehat{G})$ comme dans l'énoncé. Supposons qu'il existe $\widehat{p} : \widehat{Y} \to \widehat{X}$ vérifiant $\widehat{p} \circ \widehat{G}^t = \widehat{F}^t \circ \widehat{p}$ et $p = \widehat{\pi} \circ \widehat{p}$. On a donc, pour tout $\widehat{y} \in \widehat{Y}$:

$$\forall t \in T \quad \widehat{\pi} \circ \widehat{F}^t(\widehat{p}(\widehat{y})) = \widehat{\pi} \circ \widehat{p} \circ \widehat{G}^t(\widehat{y}) = p(\widehat{G}^t(\widehat{y}))$$

Vu (4), ceci implique

$$\widehat{p}(\widehat{y}) = (p(\widehat{G}^n(\widehat{y})))_{n\in\mathbb{Z}}.$$

La factorisation $\widehat{p}$ est donc unique.

Vérifions réciproquement que l'application $\widehat{p}$ définie par cette formule a les propriétés demandées. D'abord, elle est bien définie de $\widehat{Y}$ dans $\widehat{X}$ car $F^k(p(\widehat{G}^n(\widehat{y}))) = p(\widehat{G}^{n+k}(\widehat{y}))$ pour tous $n \in \mathbb{Z}$, $k \in \mathbb{N}$. On vérifie la relation de commutation :

$$\forall t \in T \quad \widehat{p} \circ \widehat{G}^t(\widehat{y}) = (p(\widehat{G}^{n+t}(\widehat{y}))_{n\in\mathbb{Z}}) = (F^{t-E(t)}(p(\widehat{G}^{n+E(t)}(\widehat{y})))_{n\in\mathbb{Z}}) = \widehat{F}^t \circ \widehat{p}(\widehat{y})$$

C'est bien une semiconjugaison faisant de $\widehat{F}$ un facteur topologique de $\widehat{G}$. On a la factorisation :$p = \widehat{\pi} \circ \hat{p}$.

Pour tout $\widehat{x} \in X$ et $n \gg 1$, la surjectivité de p donne $\widehat{y}$ tel que $p(\widehat{y}) = x_{-n}$. On a donc $\hat{d}(\widehat{p}(\widehat{G}^n\widehat{y}), \widehat{x}) \leq 2^{-n}$. L'image de $\widehat{p}$ est dense dans $\widehat{X}$. Quand $\widehat{Y}$ est compact, son image est fermée et donc coïncide avec $\widehat{X}$. L'application $\widehat{p} : \widehat{Y} \to \widehat{X}$ est bien une semiconjugaison. □

2.5. Exemple. *On considère les restrictions des décalages (unilatères ou bilatères) à $X = \{0,1\}^{\mathbb{N}}$ et à $\widehat{Y} = \{\widehat{y} \in \mathbb{N}^{\mathbb{Z}} : \forall n \in \mathbb{Z} \; y_{n+1} = y_n + 1 \text{ ou } 0\}$. Soit $\widehat{X}$ l'extension naturelle de X identifiée à $\{0,1\}^{\mathbb{Z}}$. Montrer que $p : \widehat{y} \in \widehat{Y} \mapsto x \in X$ avec $x_n = 0$ ssi $y_n = 0$ est une semiconjugaison mais que l'application $\widehat{p} : \widehat{Y} \to \widehat{X}$ de la dans la preuve ci-dessus n'est pas surjective.*

2.6. Corollaire. *Si $\widehat{F}_1$ et $\widehat{F}_2$ sont des dynamiques topologiques vérifiant le point (2) du lemme précédent, alors $\widehat{F}_1$ et $\widehat{F}_2$ sont topologiquement conjugués.*

Démonstration. Supposons qu'il existe deux extensions vérifiant le point (2) : $\widehat{\pi}_i : \widehat{F}_i|\mathbb{N} \to F$. On applique le point (2) d'abord avec $\widehat{G} = \widehat{F}_1$ et $\widehat{\pi} = \widehat{\pi}_2$. On obtient $\hat{p}_1 : \widehat{F}_1 \to \widehat{F}_2$ tel que $\widehat{\pi}_1 = \widehat{\pi}_2 \circ \hat{p}_1$. On obtient de même $\hat{p}_2 : \widehat{F}_2 \to \widehat{F}_1$ tel que $\widehat{\pi}_2 = \widehat{\pi}_1 \circ \hat{p}_2$. En conséquence : $\widehat{\pi}_1 = \widehat{\pi}_1 \circ (\hat{p}_1 \circ \hat{p}_2)$. Mais on a bien sûr $\widehat{\pi}_1 = \widehat{\pi}_1 \circ \mathrm{Id}$. L'unicité du point (2) appliqué à $\widehat{G} = \widehat{F}_1$ et $\widehat{\pi} = \widehat{\pi}_1$, donne $\hat{p}_1 \circ \hat{p}_2 = \mathrm{Id}$. De même, $\hat{p}_2 \circ \hat{p}_2 = \mathrm{Id}$. □

Extension naturelle probabiliste. La théorie est entièrement parallèle. Le point délicat est celui de la construction de la mesure. On recourt au théorème d'extension de Kolmogorov qui requiert certaines hypothèses sur l'espace de probablité. Nous supposerons un espace de probabilité standard, c'est-à-dire isomorphe à un espace polonais (espace topologique séparable complet pour une distance compatible). Le reste est entièrement parallèle au cas topologique et les preuves en seront laissées au lecteur diligent.

A préciser : vérifier si cela fonctionne pour $T = \mathbb{R}$

2.7. Définition. Soit $((X, \mathcal{B}, \mu), T_+, F)$ une dynamique probabiliste défini sur un espace de probabilité standard. Son *extension naturelle* est $((\widehat{X}, \widehat{\mathcal{B}}, \widehat{\mu}), T, \widehat{F})$ donnée par :

- $\widehat{X} := \{\widehat{x} := (x_n)_{s\in\mathbb{Z}} \in X^{\mathbb{Z}} : \forall n \in \mathbb{Z} \forall k \in \mathbb{N}\ x_{n+k} = F^k(x_n)\}$;
- $\forall t \in T\ \widehat{F}^t(\widehat{x}) = (F^{t-E(t)} x_{n+E(t)})_{n\in\mathbb{Z}}$;
- $\widehat{\pi} : \widehat{X} \to X,\ \widehat{x} \longmapsto x_0$;
- $\widehat{\mathcal{B}}$ la tribu produit où $\widehat{\pi}_n := \widehat{\pi} \circ \widehat{F}^{-n}$;
- $\widehat{\mu}$ une mesure de probabilité sur $(\widehat{X}, \widehat{\mathcal{B}})$ invariante par $\widehat{F}$ et se projettant $\widehat{\pi}_*(\widehat{\mu}) = \mu$.

2.8. Lemme. *L'extension naturelle $((\widehat{X}, \widehat{\mathcal{B}}, \widehat{\mu}), T, \widehat{F})$ d'un système dynamique probabiliste $((X, \mathcal{B}, \mu), T_+, F)$ est bien définie comme système dynamique probabiliste en temps T. C'est une extension probabiliste de F.*

Démonstration. Par rapport au cas topologique, le point clé est la construction de la mesure $\widehat{\mu}$ sur $X^{\mathbb{Z}}$ avec marginales prescrites :

$$\forall n \in \mathbb{Z} \quad (\widehat{\pi}_n)_*(\widehat{\mu}) := \widehat{\mu} \circ \widehat{F}^n \circ \widehat{\pi}^{-1} = \mu \tag{5}$$

Le théorème d'extension de Kolmogorov requiert que X soit un espace polonais et que les conditions marginales soient compatibles. Vu que les tribus croissent : $\widehat{\pi}_n^{-1}(\mathcal{B}) \subset \widehat{\pi}_{n+1}^{-1}(\mathcal{B})$, il suffit de prendre $\widehat{\pi}_n^{-1}(B) = \widehat{\pi}_{n+1}^{-1}(F^{-1}(B))$. Par F-invariance de μ, on a : $\mu(B) = \mu(F^{-1}(B)$. D'où la compatibilité.

On vérifie maintenant l'invariance de $\widehat{\mu}$ pour $\widehat{F}$, i.e., $\widehat{\mu}(\widehat{F}^{-1}(\hat{B})) = \widehat{\mu}(\hat{B})$ pour tout $\hat{B} \in \widehat{\mathcal{B}}$. On peut se restreindre à $\hat{B} \in \widehat{\pi}_n^{-1}(\mathcal{B})$, $n \geq 1$, i.e., $\hat{B} = \widehat{F}^n \circ \widehat{\pi}^{-1}(B)$ avec $B \in \mathcal{B}$. On calcule :

$$\begin{aligned}
\widehat{\mu}(\widehat{F}^{-1}(\hat{B})) &= \widehat{\mu}(\widehat{F}^{-1} \circ \widehat{\pi}_n^{-1}(B) && \text{vu } \hat{B} = \widehat{\pi}_n^{-1}(B) \\
&= \widehat{\mu}(\widehat{\pi}_n^{-1} \circ F^{-1}(B)) && \text{vu } F \circ \widehat{\pi}_n = \widehat{\pi}_n \circ F \\
&= \mu(F^{-1}B) = \mu(B) = \widehat{\mu}(\hat{B}) && \text{par construction de } \widehat{\mu} \text{ et invariance de } \mu
\end{aligned}$$

□

Comme dans le cas topologique, on montre :

2.9. Proposition. *Soit F une dynamique probabiliste en temps $\mathbb{N}$.*

(1) Son extension naturelle $\widehat{F}$ est une dynamique probabiliste et $\widehat{\pi}$ est une semiconjugaison en mesure faisant de $\widehat{F}$ une extension de F.

(2) Si $p : \widehat{G}|\mathbb{N} \to F$ est une semiconjugaison en mesure avec $\widehat{G}$ une dynamique probabiliste en temps $\mathbb{Z}$ alors $p = \widehat{\pi} \circ \widehat{p}$ pour une unique semiconjugaison $\widehat{p} : \widehat{G} \to \widehat{F}$.

3. Entre temps discret et temps continu

On examine deux constructions qui permettent de réduire l'étude du temps continu à celle du temps discret :

— l'application de premier retour qui passe d'un flot à l'itération ;

— la suspension qui passe de l'itération à un flot.

On présente ces deux constructions dans le cadre topologique (bien qu'on puisse les définir également dans le cas probabiliste).

3.1. Application de premier retour. Cette construction permet de passer du temps continu au temps discret.

3.1. Définition. Soient $(X, \mathbb{R}, F)$ une dynamique topologique en temps $\mathbb{R}$ et $S \subset X$, appelée section. Le *temps de premier retour* dans S est

$$\tau_F^S : S \to (0, \infty], \quad x \mapsto \inf\{t > 0 : F^t(x) \in S\}.$$

L'*application de premier retour* de F sur S est $f^S : S' \to S'$, $x \mapsto F^{\tau_F^S(x)}(x)$.

3.2. Définition. La section S est dite :

• *topologiquement transverse* si τ_F^S et $\tau_{F^{-1}}^S$ sont strictement positives (peut-être infinies) sur S et semicontinues supérieurement et continues en les points où elles sont finies.

• *admissible* si elle est topologiquement transverse et si τ_f^S et $\tau_{F^{-1}}^S$ sont finies sur S, continues et strictement positives.

• *globale* si toute orbite de F l'intersecte.

3.3. Définition. La *partie récurrente* de S est :

$$S' := \{x \in S : \exists s_k, t_k \to \infty\ F^{-s_k}(x_0), F^{t_k}(x_0) \to x \text{ as } k \to \infty\}.$$

3.4. Exercice. Si S est admissible alors $S = S'$.

Cette construction est classique et se retrouve dans les travaux de Poincaré (on l'appelle aussi application de Poincaré). Dans ce cadre, on a le résultat local général suivant, mais l'existence d'une section admissible et plus encore compacte ou globale n'est pas acquise.

3.5. Exercice. Soit M une variété de classe C^r avec un champ de vecteur $V : M \to TM$ de classe C^k, $1 \leq k \leq r$. Supposons que V définisse un flot complet F et que V ne s'annule jamais. On fixe S une hypersurface de classe C^r et transverse à V [6]

Montrer que S est topologiquement transverse puis que tout point récurrent de S appartient à S'.

6. En tout $x \in S$, $V(x) \oplus T_xS = T_xM$.

3.6. Théorème. *Soit* $(X, \mathbb{R}, F)$ *une dynamique topologique et* $S \subset X$ *une section topologiquement transverse.*

L'application de premier retour $f^S : S' \to S'$ *définit une dynamique topologique en temps* $\mathbb{Z}$ *dont les orbites sont exactement les intersections des orbites de* F *avec* S' *(sauf lorsque cette intersection est vide).*

Démonstration. S étant topologiquement transverse, les temps de retour τ_F^S et $\tau_{F^{-1}}^S$ sont continus et strictement positifs sur S'. Il s'ensuit que F^S et $(F^{-1})^S$ sont des applications continues de S' dans S' et inverses l'une de l'autre. Ce sont donc des homéomorphismes.

Soit $x \in S'$. Notons d'abord que tout point de $\mathcal{O}_F(x) \cap S$ est dans S' vu la définition de S'. Le temps de retour dans S étant strictement positif, $\{t \in \mathbb{R} : F^t(x') \in S'\}$ est une partie de $\mathbb{R}$ sans point d'accumulation. On peut donc l'ordonner sous la forme d'une suite strictement croissante $(t_n)_{n \in \mathbb{Z}}$ en fixant $t_0 = 0$. Il est alors clair que $t_{n+1} = t_n + \tau_F^S(F^{t_n}(x))$ et donc $F^S(x_n) = x_{n+1}$ où on note : $x_n := F^{t_n}(x)$. Autrement dit les points x_n, $n \in \mathbb{Z}$, qui sont les points de $\mathcal{O}_F(x) \cap S'$ sont les points de l'orbite de x sous F^S. □

Il est commode d'introduire les sommes de Birkhoff d'une fonction ϕ pour une dynamique f en temps $T = \mathbb{Z}$. Elles sont définies, pour chaque $n \geq 0$, par :

$$S_n^F \phi(x) \ := \ \phi(x) + \phi(f(x)) + \cdots + \phi(f^{n-1}x) \tag{6}$$

$$S_{-n}^F \phi(x) := -\phi(f^{-1}x) - \phi(f^{-2}x) - \cdots - \phi(f^{-n}x). \tag{7}$$

3.7. Exercice. Montrer que si $S_n^F \phi$ est définie par (6) quand $n \geq 0$ et satisfait la condition : pour tous $m, n \in \mathbb{Z}$,

$$S_{m+n}^F \phi(x) = S_m^F \phi(x) + S_n^F \phi(f^m x),$$

alors l'équation (7) est satisfaite pour tout $n \geq 0$.

On s'intéresse aux sommes de Birkhoff de la fonction τ. On les note $\tau_n := S_n^F \tau$.

3.8. Exercice. Pour tout $n \in \mathbb{Z}$ et $x \in S$, si $\phi : X \to \mathbb{R}$ est une fonction mesurable positive, alors :

$$\int_0^{\tau_n(x)} \phi(F^s(x))\, ds = S_n^{F^S} \Phi(x)$$

où $S_n^{F^S} \Phi(x)$ est la somme de Birkhoff d'ordre n pour la dynamique induite F^S et la fonction

$$\Phi(x) = \int_0^{\tau(x)} \phi(F^s x)\, ds.$$

Si $\phi = 1$, alors on obtient $(F^S)^n(x) = F^{\tau_n(x)}(x)$ où le membre de gauche est bien défini si et seulement si le membre de droite est bien défini et fini.

3.2. Suspension d'un espace par un homéomorphisme. On considère maintenant le procédé de suspension qui permet de reconstituer le flot à partir de l'application de premier retour et du temps de retour correspondant lorsque la section est globale. Pour ce faire, il faut d'abord définir un espace adéquat.

3.9. Définition. La *suspension* d'un espace X par un homéomorphisme $f : X \to X$ sous une fonction continue $\tau : X \to (0, \infty)$ avec $\inf \tau > 0$ est obtenue à partir de $X_{f,\tau} := \{(x,t) \in X \times \mathbb{R} : 0 \leq t \leq \tau(x)\}$ par les identifications : $(x, \tau(x)) \sim (f(x), 0)$.

3.10. Théorème. *Si $f : X \to X$ est un homéomorphisme d'un espace métrique X et $\tau : X \to (0, \infty)$ est une fonction continue, alors la suspension $X_{f,\tau}$ de X par f est également un espace métrique.*

L'application $\pi_{f,\tau} : X \times \mathbb{T} \to X \times \mathbb{R}$ définie par $(x, s) \mapsto (f^k(x), s - \tau_k(x))$ où $k = k(x, s)$ est l'unique entier tel que

$$\tau_k(x) \leq s < \tau_{k+1}(x)$$

est une surjection continue et un homéomorphisme local.[7]

3.11. Remarque. On peut démontrer l'analogue dans le cadre différentiable :

Si $f : X \to X$ est un difféomorphisme d'une variété X et $\tau : X \to (0, \infty)$ est une fonction différentiable avec $\inf \tau > 0$, alors la suspension $X_{f,\tau}$ de X par f est également une variété. La variété X_f est aussi régulière que le difféomorphisme.

Démonstration. En section 5 on considère $X_{f,1} := X \times [0, 1]$ où on identifie $(x, 1) \sim (f(x), 0)$. On y définit la distance de Bowen-Walters et on vérifie que c'est bien une distance compatible avec la topologie. On montre en particulier que l'application :

$$\pi_f : X \times \mathbb{R} \longrightarrow X_{f,1}, \quad (x, s) \longmapsto (f^{E(s)}(x), s - E(s)).$$

est surjective et localement un homéomorphisme.

On identifie ensuite $X_{f,\tau}$ à X_f par la bijection suivante :

$$i_{f,\tau} : X_{f,1} \longrightarrow X_{f,\tau}, \quad (x, s) \longmapsto (x, s \cdot \tau(x)).$$

On définit alors

$$\pi_{f,\tau} : X \times \mathbb{R} \longrightarrow X_{f,\tau}, \quad (x, s) \longmapsto (f^k(x), s - \tau_k(x))$$

où l'entier $k := k(x, s) \in \mathbb{Z}$ est uniquement défini par $\tau_k(x) \leq s < \tau_{k+1}(x)$ (τ ne s'annulant pas). On vérifie que

$$\pi_{f,\tau}(x, s) = i_{f,\tau} \circ \pi_f \circ R \text{ avec } R(x, s) := \left(x, k + \frac{s - \tau_k(x)}{\tau(f^k x)}\right)$$

pour le même entier $k = k(x, s)$. On vérifie que $R : X \times \mathbb{R} \to X \times \mathbb{R}$ est bien définie, surjective et un homéomorphisme local, τ étant continue et strictement positive. □

3.3. Suspension d'une dynamique à temps discret. La suspension d'une dynamique à temps $\mathbb{Z}$ sous une fonction $\tau : X \to (0, \infty)$ continue avec $\inf \tau > 0$ est la dynamique à temps $\mathbb{R}$ définit comme suit sur l'espace de suspension $X_{f,\tau}$ défini précédemment.

3.12. Définition. La suspension en temps $\mathbb{R}$ d'une dynamique topologique $(X, \mathbb{Z}, f)$ sous une fonction $\tau : X \to (0, \infty)$ continue avec $\inf \tau > 0$ est l'application $U_{f,\tau} : \mathbb{R} \times X_{f,\tau} \to X_{f,\tau}$:

$$U^t_{f,\tau} : X_{f,\tau} \longrightarrow X_{f,\tau}, \quad (x, s) \longmapsto (f^k x, s + t - \tau_k(x))$$

où k est l'unique entier tel que $\tau_k(x) \leq s + t < \tau_{k+1}(x)$.

3.13. Proposition. *L'application $U : \mathbb{R} \times X_{f,\tau} \longrightarrow X_{f,\tau}$ définit une dynamique topologique en temps $\mathbb{R}$ sur $X_{f,\tau}$.*

7. Tout $p \in X \times \mathbb{R}$ admet un voisinage U tel que $\pi(U)$ est un ouvert et $\pi : U \to \pi(U)$ un homéomorphisme.

DÉMONSTRATION. Le flot vertical $V_X : \mathbb{R} \times (X \times \mathbb{R}) \mapsto X \times \mathbb{R}$ défini par $V_X^t(x,s) = (x, s+t)$ est bien une action de $\mathbb{R}$ continue et par homéomorphisme.

On vérifie facilement que $\pi_{f,\tau} \circ V_X^t = U_{f,\tau}^t \circ \pi_{f,\tau}$.

$\square$

3.14. EXERCICE. Soient $(X, \mathbb{Z}, f)$ une dynamique métrique compacte et $\tau : X \to (0,\infty)$ une fonction continue. Soit $(X_f, \mathbb{R}, V_{f,\tau})$ la suspension de l'homéomorphisme f sous la fonction τ.

Montrer que $Y := X \times \{0\}$ est une section globale et admissible au sens des temps de retour et que le temps de premier retour sur Y coïncide avec τ (plus précisément : $\tau_V^Y(y,0) = \tau(y)$).

Montrer que l'application de premier retour dans Y est topologiquement conjuguée à $(X, \mathbb{Z}, f)$.

4. Induction (en temps discret)

On présente cette construction dans le cadre probabiliste en temps discret. On introduit d'abord quelques notions auxiliaires.

4.1. Récurrence. Etant donnée $Y \subset X$, sa *partie récurrente* Y' est :

$$Y' := \{y \in Y : \exists t_k \to \infty \; f^{t_k}(y) \in Y\} \qquad \text{(si } T = \mathbb{N}, \mathbb{R}_+)$$
$$Y' := \{y \in Y : \exists s_k, t_k \to \infty \; f^{-s_k}(x), f^{t_k}(y) \in Y\} \qquad \text{(si } T = \mathbb{Z}, \mathbb{R}).$$

On verra que c'est la partie sur laquelle la dynamique induite est naturellement définie.

4.1. EXERCICE. Vérifier que si f est un homéomorphisme, alors Y' est un G_δ de Y (une intersection dénombrable d'ouverts de Y). Trouver deux exemples de dynamiques topologiques compactes où, respectivement, $Y' = \emptyset$ ou de complémentaire dense.

4.2. THÉORÈME (Récurrence de Poincaré). *Soit $((X, \mathcal{B}_X, \mu), T, f)$ une dynamique probabiliste en temps classique ($T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+, \mathbb{R}$).*

Alors, pour toute partie mesurable $Y \subset X$ de partie récurrente Y', on a : $\mu(Y \setminus Y') = 0$.

Bien sûr, l'énoncé précédent ne dit rien si $\mu(Y) = 0$.

DÉMONSTRATION. Supposons d'abord $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+$ ou $\mathbb{R}$. On considère la fonction :

$$M : X \to \mathbb{N} \cup \{\infty\}, \quad x \longmapsto \sup\{k \geq 0 : F^k(x) \in Y\}.$$

Remarquons que c'est une fonction mesurable : pour tout $n \in \mathbb{N}$,

$$M^{-1}(n) = F^{-n}(Y) \setminus \bigcup_{k>n} F^{-k}(Y) = F^{-n}(Y_1) \text{ où } Y_1 := Y \setminus \bigcup_{k \geq 1} F^{-k}(Y).$$

Les préimages $F^{-n}(Y_1) = M^{-1}(n)$ sont disjointes. Par invariance de la mesure μ, elles ont mesure égale. Comme la mesure de l'espace est finie, ces préimages et donc leur union dénombrable sont de mesure nulle.

On a montré que pour μ-p.t. $x \in X$, ou bien l'orbite future de x ne visite jamais l'ensemble Y ou bien elle le visite lors d'une infinité de temps entiers positifs. En particulier $\mu(Y \setminus Y') = 0$.

Si $T = \mathbb{Z}$ ou $\mathbb{R}$, on peut aussi appliquer le raisonnement précédent à la dynamique inverse et prendre l'intersection des parties récurrentes de Y pour F^1 et pour F^{-1}. Elle est encore de mesure totale dans Y. $\square$

4.2. Application de Poincaré. Le *temps de retour* à Y est

$$\tau_f^Y : Y \to \{1, 2, \ldots, \infty\}, \quad y \mapsto \inf\{t > 0 : t \in T,\ f^t(y) \in Y\}$$

4.3. Définition. Soit (X, T, f) une dynamique probabiliste en temps $T = \mathbb{Z}$ ou $\mathbb{N}$.

Pour tout mesurable $Y \subset X$ avec $\mu(Y) > 0$, la *dynamique induite* par f sur Y est $((Y, \mathcal{B}_Y, \mu|_Y), T, f^Y)$ où :

— L'espace Y est muni de la tribu induite : $\mathcal{B}_Y := \{B \cap Y : B \in \mathcal{B}_X\}$;

— La mesure induite est $\mu|_Y := \mu(\cdot \cap Y)/\mu(Y)$;

— L'application est $f^Y : Y' \to Y'$, $y \mapsto f^{\tau_f^Y(y)}(y)$.

4.4. Proposition. *Le système induit* $((Y, \mathcal{B}_Y, \mu|Y), T, f^Y)$ *est encore une dynamique probabiliste.*

4.5. Remarque. On peut facilement définir une suspension en temps discret (ou tour de Kakutani-Rokhlin). En particulier, à partir de la dynamique induite sur une partie Y de mesure non-nulle et du temps de retour à Y, on peut reconstruire à conjugaison mesurée près la restriction de la dynamique initiale sur l'orbite $\bigcup_{n \geq 0} f^{-n}(Y)$ de la base Y.

Démonstration. Le choix de Y' implique immédiatement que $f^Y : Y' \to Y'$ est bien définie comme application mesurable définie μ-p.p. D'après le théorème de récurrence de Poincaré, $\mu(Y \setminus Y') = 1$: f^Y est définie μ^Y-presque partout.

Si $T = \mathbb{Z}$, on vérifie que f^Y a un inverse mesurable en remarquant que : $(f^Y)^{-1} = (f^{-1})^Y : Y' \to Y'$. Il reste à montrer l'invariance par f^Y de la mesure de probabiité $m|_Y$.

Supposons d'abord que $T = \mathbb{Z}$. Ceci permet de définir, pour tout $n \geq 1$:

$$Y_n^+ := \{y \in Y' : \tau_f^Y(y) = n\} \text{ et } Y_n^- := \{y \in Y' : \tau_{f^{-1}}^Y(y) = n\} \quad (n \geq 1)$$

qui constituent deux partitions de Y' avec $(f^Y)^{-1} : Y_n^- \to Y_n^+$, $y \mapsto f^{-n}(y)$. Soit $B \subset Y'$ mesurable. On calcule :

$$\begin{aligned} \mu\left((f^Y)^{-1}(B)\right) &= \mu\left(\bigcup_{n \geq 1} f^{-n}(B \cap Y_n^-)\right) = \sum_{n \geq 1} \mu\left(f^{-n}(B \cap Y_n^-)\right) \\ &= \sum_{n \geq 1} \mu\left(B \cap Y_n^-\right) = \mu(B). \end{aligned}$$

On en déduit que $(f_Y)_*(\mu^Y) = \mu^Y$ dans le cas $T = \mathbb{Z}$.

Supposons maintenant $T = \mathbb{N}$. Soit $((\widehat{X}, \widehat{\mathcal{B}}, \widehat{\mu}), \mathbb{Z}, \widehat{f})$ l'extension naturelle du système probabiliste $((X, \mathcal{B}, \mu), \mathbb{N}, f)$ selon la semiconjugaison $\widehat{\pi} : \widehat{X} \to X$.

On considère le système induit $(\widehat{\mu}^Y, \widehat{f}^Y)$ par $(\widehat{\mu}, \widehat{f})$ sur $\widehat{Y} := \widehat{\pi}^{-1}(Y)$ et $\widehat{Y}'$ sa partie récurrente. On va montrer que $\widehat{\pi}$ envoie $\widehat{\mu}_{|\widehat{Y}}$ sur $\mu_{|Y}$ et en déduire l'invariance.

Remarquons d'abord que, pour tout $\widehat{x} \in \widehat{Y}$

$$\begin{aligned}\tau_{\widehat{f}}^{\widehat{Y}}(\widehat{x}) &= \inf\{n \geq 1 : \widehat{f}^n(\widehat{x}) \in \widehat{\pi}^{-1}(Y)\} = \inf\{n \geq 1 : \widehat{\pi} \circ \widehat{f}^n(\widehat{x}) \in Y\} \\ &= \tau_f^Y \circ \widehat{\pi}(\widehat{x})\end{aligned}$$

car $\widehat{\pi} \circ \widehat{f}^n(\widehat{x}) = f^n \circ \widehat{\pi}(\widehat{x})$. Ensuite, en posant $n := \tau^{\widehat{Y}}(\widehat{x}) = \tau^Y(\widehat{\pi}(\widehat{x}))$:

$$\widehat{\pi} \circ \widehat{f}^{\widehat{Y}}(\widehat{x}) = \widehat{\pi} \circ \widehat{f}^n(\widehat{x}) = f^n \circ \widehat{\pi}(\widehat{x}) = f^Y \circ \widehat{\pi}(\widehat{x}).$$

Enfin, l'application $\widehat{\pi}$ étant une semiconjugaison probabiliste, $\widehat{\pi}_*(\widehat{\mu}) = \mu$.

Montrons d'abord que $\mu_{|Y} = \widehat{\pi}_*(\widehat{\mu}_{|\widehat{Y}})$:

$$\widehat{\pi}_*(\widehat{\mu}_{|\widehat{Y}}) = \frac{\widehat{\mu}(\widehat{\pi}^{-1}(\cdot) \cap \widehat{\pi}^{-1}(Y))}{\widehat{\mu}(\widehat{Y})} = \frac{\widehat{\mu}(\widehat{\pi}^{-1}(\cdot \cap Y))}{\widehat{\mu}(\widehat{Y})} = \frac{\mu(\cdot \cap Y)}{\mu(Y)} = \mu_{|Y}.$$

Montrons ensuite son invariance par f^Y :

$$f_*^Y(\mu^Y) = f_*^Y(\mu_{|Y}) = f_*^Y \circ \widehat{\pi}_*(\widehat{\mu}_{|\widehat{Y}}) = \widehat{\pi}_* \circ \widehat{f}_*^{\widehat{Y}}(\widehat{\mu}_{|\widehat{Y}} = \widehat{\pi}_*(\widehat{\mu}_{|\widehat{Y}}) = \mu_{|Y}.$$

La dynamique induite $(\mu_Y, \mathbb{N}, f)$ est bien probabiliste. □

4.6. Exercice. Démontrer le théorème de Kac : pour toute dynamique probabiliste $((X, \mathcal{B}, \mu), T, f)$ et toute partie $Y \in \mathcal{B}$, on a :

$$\int_Y \tau^Y \, d\mu = \mu\left(\bigcup_{n \geq 0} f^{-n} Y\right).$$

On pourra s'inspirer de la preuve de l'invariance de $\mu_{|Y}$ sous f^Y.

5. Appendice : La distance de Bowen-Walters

Le but de cet appendice est de munir la suspension d'un espace métrique d'une métrique compatible avec sa topologie. On montrera :

5.1. Proposition. *Soient X un espace métrique et $f : X \to X$ un homéomorphisme.*

La suspension (constante) de l'espace X par l'homéomorphisme f, $X_f := X_{f,1}$, a une structure d'espace métrique. De plus, l'application $\pi_f : X \times \mathbb{R} \to X_f$ définie par

$$\pi_f : (x, t) \longmapsto (f^k(x), t - k) \text{ où } k \leq t < k + 1 \text{ est la partie entière.}$$

est une surjection telle que tout $(x, t) \in X \times \mathbb{R}$ admet un voisinage U tel que $\pi_f : U \to \pi_f(U)$ est un homéomorphisme avec $\pi_f(U)$ ouvert de X_f.

On démontre cette proposition en exhibant une métrique compatible due à Bowen et Walters.

On appelle *chemin de Manhattan* γ sur X_f une suite finie $(x_0, t_0), \ldots, (x_n, t_n)$ où pour chaque i, $t_i = t_{i+1}$ ou x_i, x_{i+1} appartiennent à la même orbite de f. Le *poids* $p(\gamma)$ d'un tel chemin est la somme du poids de ces arêtes $(y, t)(y', t')$ défini par :

$$|(y, s)(y', s')| := \begin{cases} s\, d(f(y), f(y')) & \\ \qquad +(1 - s) d(y, y') & \text{si } s = s' \\ |u| & \text{si } (y', s') = (f^k(y), s - k + u) \text{ avec } k \in \mathbb{Z} \\ \infty & \text{sinon.} \end{cases}$$

La *distance de Bowen-Walters* sur $X_f = X \times [0, 1[$ est définie comme la borne inférieur des poids des chemins Manhattan :

$$d_f((x,t),(x',t')) := \inf\Big\{p(\gamma) : \gamma \text{ chemin de Manhattan de } (x,t) \text{ à } (x',t')\Big\}.$$

5.2. Lemme. *La formule précédente définie une distance $d_f : X_f \times X_f \to [0, \infty)$ sur X_f.*

Démonstration. On vérifie les axiomes d'une distance pour d_f. D'abord, l'application est symétrique. En suite, il existe toujours un chemin de Manhattan entre deux points quelconques, donc l'application est à valeurs dans $[0, \infty[$.

Si $d_f((x,t),(x',t')) = 0$, le poids de chaque segment doit être nul : $(x,t) = (x',t')$.

L'inégalité triangulaire vient de l'inégalité suivante :

$$\begin{aligned} d_f(p,r) &= \inf\{p(\gamma) : \gamma \text{ de } p \text{ à } r\} \\ &\leq \inf\{p(\gamma) : \gamma \text{ de } p \text{ à } r \text{ passant par } q\} \\ &= d_f(p,q) + d_f(q,r). \end{aligned}$$

L'inégalité triangulaire sur $\mathbb{R}$ donne, pour tout chemin de Manhattan :

$$\sum_{i=1}^{n} |(x_i,t_i)(x_{i+1},t_{i+1})| \geq \sum_{i=1}^{n} |t_i - t_{i+1} - n_i| \geq \min_{n\in\mathbb{Z}} |t - t' - n|.$$

□

******** A REDIGER ********

Chapitre 3

Dynamique topologique

La dynamique topologique étudie les propriétés des systèmes dynamiques de la forme (X, T, F) définie par une action par homéomorphisme de T su un espace topologique X.

On dira que la dynamique est *compacte* si X est de plus métrisable et compact On dira que son temps est classique si le semigroupe T, est $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{R}_+$, ou $\mathbb{R}$.

1. Minimalité et récurrence

1.1. Ensembles omega-limites et points récurrents. On considère (X, T, F) une dynamique topologique métrisable en temps $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+$ ou $\mathbb{R}$.

1.1. Définition. L'*ensemble omega-limite* de $x \in X$ est l'ensemble

$$\omega_F(x) := \{\lim_k F^{t_k}(x) : t_k \in T \to \infty\}$$

Parfois on peut classifier les ensembles omega-limites :

1.2. Exercice. Soit $\tau_\alpha : \mathbb{T} \to \mathbb{T}$ une rotation du tore de dimension 1 d'angle $\alpha \in \mathbb{R}$. Montrer que :

— si α est rationnel, alors chaque $\omega_\tau(x)$ coïncide avec l'orbite de x qui est périodique ;

— si α est irrationnel, alors chaque $\omega_\tau(x)$ coïncide avec le tore $\mathbb{T}$.

En déduire que chaque $x \in \mathbb{T}$ est récurrent.

1.3. Proposition. *Pour tout $x \in X$, l'ensemble $\omega_F(x)$ est un fermé vérifiant : $F^t(\omega_F(x)) \subset \omega_F(x)$ pour tout $t \in T$.*

Si T est un groupe ou si X est compact, alors $\omega_F(x)$ est un compact non-vide et $F^t(\omega_F(x)) = \omega_F(x)$ pour tout $t \in T$.

Démonstration. L'ensemble $\omega_F(x)$ est fermé comme ensemble de points d'accumulation. Si X est compact, il existe toujours des points d'accumulation et donc $\omega_F(x) \neq \emptyset$. De plus un fermé dans un espace compact est compact.

Montrons l'inclusion. Soit $y \in F^t(\omega_F(x))$. Par définition $y = F^t(z)$ avec $z = \lim_k F^{t_k}(x)$ et $t_k \in T \to \infty$. Par continuity de F^t, il vient $y = \lim_k F^{t_k+t}(z)$ et $t_k + t \in T \to \infty$. Donc $y \in \omega_F(x)$.

Réciproquement, soit $z \in \omega_F(x)$, i.e., $z = \lim_k F^{s_k}(x)$ et $s_k \in T \to \infty$. Soit $y = \lim_k F^{s_k - t}(x)$: cette limite existe, quitte à passer à une sous-suite, grâce à l'hypothèse de compacité. On a bien $F^t(y) = z$: c'est l'inclusion réciproque et donc l'égalité. $\square$

1.4. Exercice. Trouver une application continue f d'un espace métrique X tel que $f(\omega_f(x)) \subsetneq \omega_f(x)$. En particulier, $f^{-1}(\omega_f(x)) \not\subset \omega_f(x)$.

La notion suivante est fondamentale :

1.5. Définition. Un point $x \in X$ est *récurrent* pour F si $x \in \omega_F(x)$. On note

$$\mathrm{Rec}(F) := \{x \in X : x \text{ récurrent pour } F\}.$$

Souvent, l'ensemble des points récurrents a une structure compliquée :

1.6. Exercice. Soit $(\Sigma_d, \mathbb{Z}, \sigma_d)$ le décalage sur d symboles.
Montrer que l'ensemble des points récurrents est dense mais non fermé.

1.2. Théorème de Birkhoff et minimalité.

1.7. Théorème (Récurrence de Birkhoff). *Pour toute dynamique compacte (X, T, F), il existe au moins un point récurrent.*

1.8. Exemple. *Analyser l'homéomorphisme $f : \mathbb{T} \to \mathbb{T}$ défini par $f(x + \mathbb{Z}) = x^2 + \mathbb{Z}$ pour $x \in [0, 1]$.*

Pour démontrer ce théorème on introduit une propriété topologique globale.

Rappelons qu'une partie $Y \subset X$ d'une dynamique (X, T, F) est *invariante* si $F^t(Y) \subset Y$ pour tout $t \in T$. La partie Y est *strictement invariante* si $F^t(Y) = Y$ pour tout $t \in T$.

A préciser : $(F^t)^{-1}(Y) = Y$? ? ?

1.9. Définition. Un système dynamique topologique (X, T, F) est *(topologiquement) minimal* s'il n'existe pas de fermé nontrivial $\emptyset \subsetneq Y \subsetneq X$ qui soit invariant : $\forall t \in T \ F^t(Y) \subset Y$.

On dit qu'un fermé invariant nonvide $Y \subset X$ est *(topologiquement) minimal* si $(Y, T, F|Y)$ est (topologiquement) minimal.

Notons qu'une partie minimale est strictement invariante : $F^t(Y) \subsetneq Y$ contredirait sinon la minimalité de Y.

On peut reformuler cette définition de la façon suivante : un système dynamique topologique compact est minimal si et seulement si elle n'admet pas de sous-sytème topologique compact non-trivial.

1.10. Exercice. Une orbite périodique, l'odomètre à $d \geq 2$ ou une translation irrationnelle sur $\mathbb{T}$ sont des dynamiques compactes minimales.

1.11. Proposition. *La minimalité de (X, T, F) dynamique compacte est équivalente à chacune des propriétés suivantes :*

(1) ses seules parties fermées invariantes ($\forall t \in T \ F^t(Y) \subset Y$) sont $\emptyset$ et X ;
(2) ses seules parties fermées strictement invariantes ($\forall t \in T \ F^t(Y) = Y$) sont $\emptyset$ et X ;
(3) pour tout $x \in X$, $\omega_F(x) = X$;
(4) pour tout $x \in X$, $\overline{\mathcal{O}_F^{T_+}(x)} = X$;
(5) pour tout $x \in X$, $\overline{\mathcal{O}_F^{T}(x)} = X$.

Démonstration. Remarquons que (1) est la définition de la minimalité.
(1) $\Longrightarrow$ (2) : trivial.
(2) $\Longrightarrow$ (3) : tout ensemble omega-limite est une partie fermée non-vide et strictement invariante d'après la Proposition 1.3.

(3) $\Longrightarrow$ (4) $\Longrightarrow$ (5) : trivial.

(5) $\Longrightarrow$ (1) : Soient $Y \subset X$ nonvide, fermée et invariante et $y \in Y$. On considère $z \in \omega_F(y)$. Comme $\omega_F(y)$ est strictement invariant et fermé, il contient $\overline{\mathcal{O}_F^T(z)}$. Donc (5) implique que $Y = X$. □

1.12. Corollaire. *Si (X, T, F) est une dynamique compacte minimale, alors tout point de X est récurrent.*

Le théorème de Birkhoff découle de ce corollaire grâce au lemme suivant.

1.13. Lemme. *Si (X, T, F) est une dynamique compacte, alors il existe un fermé strictement invariant $\emptyset \neq Y \subset X$ topologiquement minimal.*

Démonstration. Soit

$$\mathcal{K} := \{Y \subset X : Y \text{ fermé, non-vide et invariant}\}$$

muni de l'ordre partiel défini par l'inclusion. Une chaîne de $\mathcal{K}$ est une partie $\mathcal{C} \subset \mathcal{K}$ vérifiant :

— $\forall C, C' \in \mathcal{C}$ $C \subset C'$ ou $C' \subset C$ (ordre total) ;

— si $\mathcal{D} \subset \mathcal{K}$ est une partie totalement ordonnée, alors $\mathcal{C} \subset \mathcal{D}$ implique $\mathcal{C} = \mathcal{D}$ (maximalité).

Si $\mathcal{C} \subset \mathcal{K}$ est une chaîne, montrons que

$$K_0 := \bigcap_{C \in \mathcal{C}} C \text{ est le minimum de } \mathcal{C}.$$

Montrons que $K_0 \neq \emptyset$. Sinon, $X = \bigcup_{C \in \mathcal{C}} (X \setminus C)$. Par compacité : $X = \bigcap_{i=1}^n (X \setminus C_i)$ avec $C_1, \ldots, C_n \in \mathcal{C}$. On peut supposer $C_1 \subset \cdots \subset C_n$ vu l'ordre total sur $\mathcal{C}$. On a donc $X = X \setminus C_n$, ce qui est absurde car $C_n \neq \emptyset$.

K_0 est compact, respectivement invariant, comme intersection de compacts, respectivement d'invariants. On conclut : $K_0 \in \mathcal{K}$ avec $K_0 \subset C$ pour tout $C \in \mathcal{C}$. En particulier, $\mathcal{C} \cup \{K_0\}$ est encore totalement ordonné et donc $K_0 \in \mathcal{C}$ est le minimum de $\mathcal{C}$.

On a montré que toute chaîne dans $\mathcal{K}$ admet un minimum. Le lemme de Zorn implique alors que $\mathcal{K}$ contient (au moins) un élément minimal Y. Ceci veut exactement dire que Y est topologiquement minimal pour F. □

1.14. Exercice. Soit (X, T, F) une dynamique topologique sur un compact métrisable. On suppose que tout point est récurrent. La dynamique F est-elle nécessairement minimale ? L'union de dynamiques minimales ?

1.3. Récurrence uniforme.

1.15. Définition. Soit (X, T, F) une dynamique topologique en temps classique[1]. On dit que $x \in X$ est uniformément récurrent si, pour tout voisinage U de x, il existe $L \geq 1$ tel que

$$\forall t \in T \; \exists 0 \leq s < L \quad F^{t+s}(x) \in U.$$

1.16. Théorème. *Soit (X, T, F) une dynamique topologique compacte et métrisable en temps classique.*

1. $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+, \mathbb{R}$.

— *Si la dynamique F est topologiquement minimale, alors tout point est uniformément récurrent et le nombre $L = L(\epsilon)$ dans la définition ci-dessus peut être choisi indépendamment de x.*

— *Si $x \in X$ est uniformément récurrent, alors $\overline{\mathcal{O}_F^+(x)}$ est topologiquement minimal.*

1.17. Corollaire. *Soit (X, T, F) une dynamique topologique compacte et métrisable en temps classique.*

Un point récurrent $x \in X$ est uniformément récurrent si et seulement si $\omega_F(x)$ est topologiquement minimal.

Démonstration. Montrons d'abord (1) par l'absurde : on suppose qu'il existe $\epsilon_* > 0$ tel que, pour tout $n \geq 1$, il existe $x_n \in X$ et $t_n \in T$ vérifiant :

$$\forall s \in T \quad 0 \leq s \leq 2n \implies F^s(x_n) \notin B_X(x_n, 2\epsilon). \tag{8}$$

Quitte à extraire, on peut supposer que $\lim_n x_n = x_*$ et $x_n \in B(x_*, \epsilon)$. On obtient donc :

$$\forall s \in T \quad 0 \leq s \leq 2n \implies F^s(x_n) \notin B_X(x_*, \epsilon).$$

On considère le compact :

$$Y := \bigcap_{n \in T} F^{-t}(\overline{X \setminus B(x_*, \epsilon)}).$$

Il est invariant : $F^t(Y) \subset Y$ pour tout $t \in T$. Il n'est pas égal à X puisqu'il ne contient pas x_*. Pour obtenir la contradiction recherchée, montrons qu'il n'est pas vide.

Soit y un point d'accumulation de la forme :

$$y = \lim_k F^{t_{n_k} + n_k}(x_{n_k}) \text{ avec } n_k \to \infty.$$

Vu l'hypothèse (8), pour tout $t \in T$,

$$F^t(y) = \lim_k F^{t_{n_k} + n_k + t}(x_{n_k})$$

Pour k grand, $0 \leq n_k + t \leq 2n_k$ et donc $F^t(y) \notin B(x_*, \epsilon)$, i.e., $y \in Y$ qui n'est pas vide.

On démontre maintenant la réciproque. Soit $x \in X$ uniformément récurrent, il faut voir que $\omega_F(x)$ est minimal.

Soit $y \in \omega_F(x)$, i.e., $y = \lim_k F^{t_k}(x)$ avec $t_k \in T$, $t_k \to \infty$. Il suffit de montrer que $x \in \overline{\mathcal{O}^+(y)}$, on aura alors $\overline{\mathcal{O}^+(y)} = \omega_F(x)$ et donc la minimalité. Montrons donc que pour tout $\epsilon > 0$, il existe $t \in T$, $t > 0$ avec $F^t(y) \in B(x, \epsilon)$.

Par récurrence uniforme de x, il existe pour tout k, $s_k \in T$ avec $S := \sup_k |s_k - t_k| < \infty$. Par continuité uniforme de $F : [0, S] \times X \to X$. il suffit donc d'extraire une sous-suite telle que $F^{t_k}(x)$ converge pour que $F^{s_k}(x)$ converge aussi. □

1.18. Exercice. Exhiber un point $x \in \Sigma_d$ récurrent et $\epsilon > 0$ tel que

$$\exists t_k \in T \to \infty \quad \forall t \in [t_k, t_k + k] \; d(x, f^t(x)) \geq \epsilon.$$

En déduire qu'un tel point n'appartient pas à un fermé minimal.

1.19. Exercice. Montrer que le décalage unilatère Σ_d^+ $(d \geq 2)$ n'est pas une union de dynamiques compactes et minimales.

1.4. Facteurs et extensions.

1.20. Proposition. *Si (X, T, F) est une extension topologique de (Y, T, G) par une semiconjugaison topologique $\pi : X \to Y$, alors : $\forall x \in X$ $\pi(\omega_F(x)) \subset \omega_G(\pi(x))$.*

Si $\pi : X \to Y$ est une conjugaison topologique ou si X est compact, alors $\omega_G(\pi(x)) = \pi(\omega_G(x))$ pour tout $x \in X$.

Cet énoncé est généralise la proposition 1.3. En effet, F^t est toujours une semiconjugaison de (X, T, F) avec lui-même et même une conjugaison dans le cas inversible (si le temps T est un groupe).

En particulier :

1.21. Lemme. *Si (X, T, F) est une extension topologique de (Y, T, G) par une semiconjugaison topologique $\pi : X \to Y$, alors :*

(1) l'image d'une partie topologiquement minimale est topologiquement minimale.

(2) l'image d'un point récurrent est un point récurrent ;

(3) l'image d'un point uniformément récurrent est un point uniformément récurrent.

Les implications réciproques sont fausses en général. On va démontrer quelques réciproques partielles importantes.

1.22. Exercice. Ecrire la preuve de la proposition 1.20 en relisant la preuve de la proposition 1.3.

1.23. Exercice. Dans la situation de la proposition précédente avec X compact, peut-on avoir $x \in X$ non F-récurrent avec $\pi(x)$ G-récurrent ?

Si on a une extension par un groupe compact alors on peut "relever" la propriété de récurrence :

1.24. Proposition. *Soit (X, T, F) une dynamique topologique en temps $\mathbb{Z}$, K un groupe topologique compact, et $\alpha : X \to K$ une application continue définissant le produit gauche (Y, T, G) avec $Y = X \times K$ et $G(x, y) = (F(x), \alpha(x)y)$.*

Si $x \in X$ est F-récurrent, alors tout point de $\{x\} \times K$ est G-récurrent :

$$\mathrm{Rec}(G) = \pi^{-1}(\mathrm{Rec}(F)).$$

Démonstration. Soit $x_0 \in X$ un point F-récurrent : $x_0 = \lim_j F^{t_j}(x_0)$ avec $t_j \to \infty$. Soit e l'identité de K. K étant compact, quitte à extraire, on a $(x_0, k_*) = \lim_j G^{t_j}(x_0, e)$, i.e., $(x, k_*) \in \omega_G(x_0, e)$.

Remarquons que $h_k : X \times K \to X \times K$, $(x, y) \mapsto (x, yk)$ est un homéomorphisme de $X \times K$ avec $h_k \circ G^t = G^t \circ h_k$ et $h_k(\{x_0\} \times K) = \{x_0\} \times K$. On en déduit : $(x_0, k_*^{n+1}) \in \omega_G(x_0, k_*^n)$ pour tout $n \in \mathbb{N}$.

On vérifie aisément la propriété de transitivité suivante grâce à la continuité de G :

$$y \in \omega_G(x) \text{ et } z \in \omega_G(y) \implies z \in \omega_G(x).$$

On en déduit :

$$\forall n \geq 1 (x_0, k_*^n) \in \omega_G(x_0, e).$$

Par compacité de K, il existe $(k_1^{n_i})_{i \geq 1}$ convergente avec $n_i \nearrow \infty$. Donc

$$k_1^{n_{i+1} - n_i} \to e.$$

Comme $\omega_G(x_0, e)$ est fermé, $(x_0, e) \in \omega_G(x_0, e)$: on a trouvé un point récurrent dans $\pi^{-1}(x_0)$.

On conclut en remarquant que $(x_0, e) \in \omega_G(x_0, e)$ implique que, pour tout $k \in K$,

$$(x_0, k) = h_k(x_0, e) \in h_k(\omega_G(x_0, e)) = \omega_G(x_0, k).$$

Tous les points de $\pi^{-1}(x_0)$ sont récurrents dès que x_0 est récurrent. La proposition 1.24 donne l'inclusion réciproque, d'où l'égalité annoncée. $\square$

1.25. Exercice. Montrer que pour tout $\alpha \in \mathbb{R}$, tous les points de $\mathbb{T}^2$ sont récurrents sous la transformation $(x, y) \mapsto (x + \alpha, y + 2x)$.

1.26. Question. *Existe-t-il un produit gauche compact $F \ltimes_\phi G$ avec $x \in X$ récurrent sans qu'aucun des points de $\{x\} \times Y$ ne soit récurrent ? (Trouver un exemple ou montrer que c'est impossible.)*

1.27. Proposition. *Soit (X, T, F) une dynamique topologique compacte et métrisable en temps classique, K un groupe topologique compact, et $\alpha : X \to K$ une application continue définissant le produit gauche (Y, T, G) avec $Y = X \times K$ et $G(x, y) = (F(x), \alpha(x)y)$.*

Si $x_0 \in X$ est uniformement récurrent pour F, alors tout point de $\{x_0\} \times K$ est uniformément récurrent pour l'extension G.

Démonstration. Soit $x_0 \in X$ uniformément récurrent. Prenons $y_0 \in K$ arbitraire. Quitte à restreindre F au compact invariant $\overline{\mathcal{O}(x_0)}$, on peut supposer F minimal. D'après le lemme de Birkhoff, $X \ltimes K$ admet une partie Z topologiquement minimale. Comme $X \ltimes K$ est une extension topologique de X, $\pi(Z)$ est fermée, invariante et non-vide. Par minimalité de X, $\pi(Z) = X$. Il existe donc $(x_0, y_1) \in Z$. L'auto-conjugaison topologique $h : (x, y) \mapsto (x, yy_1^{-1}y_0)$ envoie $(x_0, y_1) \in Z$ sur $(x_0, y_0) \in h(Z)$. Mais l'image d'un compact minimal par une conjugaison topologique est un compact minimal. Donc $(x_0, y_0) \in h(Z)$ est uniformément récurrent. $\square$

2. Généricité topologique

Théorème de Baire. Le théorème suivant fournit un analogue topologique à la notion probabiliste d'ensemble de mesure totale :

2.1. Théorème (Baire). *Si X est un espace métrique complet, alors toute intersection dénombrable d'ouverts denses est dense.*

Pour mettre en évidence ce parllèle, disons qu'un ensemble est *gras (au sens de Baire)* s'il contient une intersection dénombrable d'ouverts denses. On parle parfois d'ensemble de deuxième catégorie, d'ensemble résiduel ou d'ensemble topologiquement générique.

2.2. Corollaire. *Une intersection dénombrable d'ensembles gras est encore un ensemble gras.*

Attention les notions de généricité topologique et probabiliste sont indépendantes :

2.3. Exercice. L'ensemble des nombres de Liouville :

$$\mathcal{L} := \{x \in \mathbb{R} : \forall n \geq 1 \exists (p, q) \in \mathbb{Z} \times \mathbb{N}, q \geq 2\ |x - (p/q)| < 1/q^n\}$$

est un ensemble gras mais de mesure de Lebesgue nulle.

2.4. Lemme. *Si X est un espace métrique complet et $\phi : X \to \mathbb{R}$ est semicontinue supérieurement, alors l'ensemble des points de continuité de ϕ contient un ensemble gras au sens de Baire.*

En appliquant le lemme à $-\phi$ on obtient le même énoncé pour les fonctions semicontinues inférieurement.

Démonstration. Remarquons qu'on peut supposer que la fonction ϕ est également positive. En effet, on peut, si nécessaire, remplacer ϕ par la fonction $e^{\phi} : X \to \mathbb{R}$ qui sera encore semicontinue supérieurement. Pour chaque entier $N \geq 1$, on pose :

$$A_N := \{x \in X : \exists x_n \to x,\ \exists x'_n \to x \text{ s.t. } \phi(x_n) < \phi(x'_n) - 1/N\}$$

Notons que tout point de discontinuité appartient à A_N avec $N \geq 1$ assez grand. Il suffit de montrer que chaque A_N est un fermé d'intérieur vide : cela impliquera que le complémentaire de leur union est bien un G_δ dense donc gras.

Chaque A_N est un fermé comme ensemble de limites d'un ensemble "stable" [2] de suites.

Montrons par l'absurde que chaque A_N est d'intérieur vide. Supposons donc qu'il existe $y_0 \in X$ et $\delta_0 > 0$ tel que $B_X(y_0, \delta_0) \subset A_N$.

En particulier, il existe $x_n, x'_n \to y_0$ avec $\phi(x_n) < \phi(x'_n) - 1/N$.

La semicontinuité supérieure implique que pour n assez grand $\phi(x'_n) < \phi(y_0) - 1/(2N)$. Pour ces mêmes n,

$$\phi(x_n) < \phi(x'_n) - 1/N < \phi(y_0) - 1/(2N)$$

Fixons un tel n assez grand pour avoir aussi $y_1 := x_n \in B(y_0, \delta_0/2)$ et donc $B_X(y_1, \delta_1) \subset A_N$ pour $\delta_1 := \delta_0/2$.

Par une récurrence dont la formalisation est laissée au lecteur diligent, on obtient une suite $(y_n)_{n\geq 0}$ vérifiant : $\phi(y_n) \leq \phi(y_0) - n/(2N) \to -\infty$.

Ceci contredit l'hypothèse selon laquelle ϕ est minorée. □

Un cadre naturel en dynamique, au-delà du cas compact, est le suivant. On renvoie à [**?**] pour une introduction à la théorie de ces espaces du point de vue de leurs transformations boréliennes (théorie descriptive des ensembles) :

2.5. Définition. Un *espace polonais* est un espace topologique séparable métrisable par une distance complète.

2.6. Exemples. *La plupart des espaces rencontrés dans ce cours sont des espaces polonais :*

- *les espaces métriques compacts ;*
- *$\mathbb{R}^d$ ou tout espace de Banach séparables ;*
- *$\mathbb{P}(F)$ l'ensemble des mesures de probabilité invariantes par une dynamique continue sur un espace du thèorème de Riesz 3.9 (métrique, localement compact dont tout ouvert est une union dénombrable de compacts) ;*
- *toute partie borélienne d'un espace polonais.*

2. Appliquer l'argument diagonal de Cantor fournit une nouvelle suite du même type. ******

Non-errance et récurrence.

2.7. Définition. Soit (X, T, F) une dynamique topologique. Un point est dit *non-errant* si pour tout voisinage $U \neq \emptyset$ de x il existe des temps $t_k \to \infty$ tels que $F^{-t}(U) \cap U \neq \emptyset$.

La dynamique est elle-même qualifiée de non-errante si tout point est non-errant.

Tout point récurrent est non-errant. Mais la réciproque est fausse en général :

2.8. Exercice. Montrer que σ_d^+ $(d \geq 2)$ est non-errant. En déduire qu'il existe des points non-errants qui ne sont pas récurrents.

2.9. Exercice. Soit (X, T, F) une dynamique topologique métrisable en temps classique. Montrer que l'ensemble des points récurrents est de la forme :

$$\mathrm{Rec}(F) := \bigcap_{m,n \geq 1} \bigcup_{t \geq n} \{x \in X : d(x, F^t(x)) < 1/m\}.$$

En déduire que $\mathrm{Rec}(F)$ est un G_δ (une intersection dénombrable d'ouverts), non nécessairement dense.

2.10. Exercice. Un point $x \in X$ est non-errant si pour tout $\epsilon > 0$, il existe $y \in B(x, \epsilon)$ et $t \geq 1$ tel que $F^t(y) \in B(x, \epsilon)$.

En particulier, l'ensemble des points non-errants de F est un fermé.

2.11. Proposition. *Supposons $T = \mathbb{Z}$ ou $\mathbb{N}$. Si F est-elle même est non-errante, alors l'ensemble des points récurrents est gras au sens de Baire.*

Démonstration. On définit $\phi : X \to [0, \infty[$ par $\phi(x) := \inf_{n \geq 1} d(x, f^n x)$. $F(x) = 0$ si et seulement si $x \in \mathrm{Rec}(F)$. De plus, chacune des fonctions $x \mapsto d(x, f^n x)$ étant continue, ϕ est semicontinue supérieurement. Elle admet donc un ensemble gras de points de continuité.

Supposons qu'un point de continuité x_0 ne soit pas récurrent. On en déduit $\delta, \rho > 0$ tels que

$$\phi(x) > \delta \text{ pour tout } x \in B_X(x_0, \rho).$$

Quitte à les remplacer par leur minimum, on suppose $\rho = \delta > 0$. Ceci implique que pour tout $x \in B(x_0, \rho/2)$, $d(f^n x, x_0) \geq d(f^n x, x) - d(x, x_0) > \rho/2$. En particulier,

$$F^{-t}(B(x_0, \rho/2)) \cap B(x_0, \rho/2) = \emptyset,$$

contredisant l'hypothèse selon laquelle tous les points sont non-errants. □

3. Transitivité et orbites denses

L'objectif de cette section est d'introduire une notion d'irréductibilité plus générale (plus faible) que la minimalité : l'existence d'orbites denses.

3.1. Définition et premiers exemples.

3.1. Définition. Une dynamique topologique (X, T, F) est dite *topologiquement transitive* s'il existe une orbite dense, i.e., il existe $x_0 \in X$ avec $\overline{\mathcal{O}_F(x)} = X$.

On se place dans le cadre suivant :

Condition (*) :

(1) X est polonais et

(2) tout segment d'orbite $\{F(t,x) : t \in [0,1] \cap T\}$ ($x \in X$) est d'intérieur vide.

Si la condition (1) n'a pas lieu, alors ou bien il n'y a pas d'orbite dense, ou bien la dynamique est restreinte à une unique orbite périodique. En temps discret ($T = \mathbb{N}$ ou $\mathbb{Z}$), la condition (2) signifie que l'espace X n'a pas de point isolé.

3.2. EXERCICE. Caractériser les nombres $\alpha \in \mathbb{R}$ tels que la translation $\tau_\alpha : \mathbb{T} \to \mathbb{T}$ soit topologiquement transitive.

3.3. EXERCICE. Montrer que les décalages σ_d et σ_d^+ sont topologiquement transitifs pour tout entier $d \geq 1$.

3.2. Caractérisations. Commençons par le fait suivant :

3.4. LEMME. *Soit (X,T,F) une dynamique topologique.*
Si $x \in X$ est un point récurrent, alors $\omega(x) = \overline{\mathcal{O}_F^+(x)}$.

DÉMONSTRATION. Clairement $\omega(x) \subset \overline{\mathcal{O}_F^+(x)}$. L'ensemble omega-limite étant fermé, il suffit de vérifier que $\mathcal{O}_F^+(x) \subset \omega(x)$. Prenons donc $y := F^t(x)$ avec $t \in T_+$. Par hypothèse, $x = \lim_n F^{t_n}(x)$ avec $t_n \to \infty$ donc $y = \lim_n F^{t+t_n}(x) \in \omega(x)$. □

3.5. PROPOSITION. *propTransit Supposons (X,T,F) une dynamique topologique en temps classique satisfaisant la condition (*).*
La transitivité est alors équivalente à chacune des propriétés suivantes :

(1) il existe $x \in X$ tel que $\overline{\mathcal{O}(x)} = X$;

(2) il existe un ensemble gras de points de X tels que $\overline{\mathcal{O}(x)} = X$;

(3) il existe $x \in X$ tel que $\overline{\mathcal{O}^+(x)} = X$;

(4) il existe un ensemble gras de points de X tels que $\omega(x) = \overline{\mathcal{O}^+(x)} = X$;

(5) pour toute paire d'ouverts non-vides U,V de X, il existe $t \in T^+$ arbitrairement grand tel que $U \cap F^{-t}(V) \neq \emptyset$.

DÉMONSTRATION. Notons que, clairement, (4) $\Longrightarrow$ (2) $\Longrightarrow$ (1). Il suffit donc de montrer l'équivalence entre (1), (3), (4) et (5).

(4) $\Longrightarrow$ (3) $\Longrightarrow$ (1) : clair.

(1) $\Longrightarrow$ (5) : c'est l'implication clé. Soit x comme dans (1). Soit U,V comme dans (5) et $0 < T_0 < \infty$ et cherchons $t \in T^+$, $t > T_0$, tel que $U \cap F^{-t}(V) \neq \emptyset$. D'après (1), il existe $t_1 \in T$ tel que $F^{t_1}(x) \in U$.

Supposons d'abord le système non-inversible. Il existe donc $t_2 \in T$ tel que $F^{t_2}(x) \in V \setminus F^{[0,t_1+T_0]}(x)$ (ce dernier ouvert est non-vide vu (*)). $t = t_2 - t_1 > T_0$ convient et établit (5) dans ce premier cas.

Supposons maintenant le système inversible. Il existe $t_2 \in T$ tel que $F^{t_2}(x) \in V \setminus F^{[t_1-T_0,t_1+T_0]}(x)$.

Cas $t_2 > t_1 + T_0$ **:** on conclut comme ci-dessus.

Cas $t_2 < t_1 - T_0$ **:** fixons $\epsilon > 0$ tel que $d(y, F^{t_2}(x)) < \epsilon \implies F^{t_1-t_2}(y) \in U$. Il existe $t_3 \in T$ tel que $F^{t_3}(x) \in B(F^{t_2}(x), \epsilon) \cap V \setminus F^{[2t_2-t_1-T_0,t_1+T_0]}(x)$.

Sous-cas $t_3 < 2t_2 - t_1 - T_0$: $F^{t_3+(t_1-t_2)}(x) \in U$ et $F^{t_2}(x) \in V$ donc $t = 2t_2 - t_3 - t_1 > T_0$ convient.

Sous-cas $t_3 > t_1 + T_0$: $F^{t_3}(x) \in V$ et $F^{t_1}(x) \in U$ et $t = t_3 - t_1 > T_0$ convient et établit (5) dans tous les cas.

(5) $\Longrightarrow$ (4) : X étant un espace polonais, sa topologie admet une base dénombrable $(U_n)_{n\geq 1}$ (il suffit de considérer les boules ouvertes centrées en un point d'une partie dénombrable et dense et de rayon rationnel). Pour chaque $n \in \mathbb{N}$, on considére l'ouvert

$$X_n := \{x \in X : \mathcal{O}^+(x) \cap U_n \neq \emptyset\}.$$

Fixons $n \in \mathbb{N}$. D'après (5), pour tout $m \in \mathbb{N}$, $U_m \cap F^{-t}(U_n) \neq \emptyset$ pour un certain $t \in T^+$, i.e., $U_m \cap X_n \neq \emptyset$: X_n est donc dense dans X. Le théorème de Baire dit alors que l'intersection $D := \bigcap_{n\in\mathbb{N}} X_n$ est un ensemble gras de X. Enfin, remarquons que, pour tout $x \in D$, $\mathcal{O}^+(x)$ intersecte chaque U_n donc : $\overline{\mathcal{O}^+(x)} = X$.

Il reste à montrer que tout $x \in D$ est un point récurrent et d'appliquer le lemme. On fixe $n \geq 1$ et on considère l'ouvert $W := B(x, 1/n) \setminus F^{[0,n]}(x)$. Il est non-vide grâce à (*). On a donc $W \cap \mathcal{O}_+(x) \neq \emptyset$: $F^{t_n}(x) \in B(x, 1/n)$ pour un certain $t_n \in T_+$. Clairement, $t_n \geq n$. On en déduit $x = \lim_n F^{t_n}(x)$ avec $t_n \to \infty$, $t_n \in T_+$: x est bien récurrent. (4) est bien démontré. $\square$

3.6. Exercice. Soit (X, T, F) une dynamique topologique en temps classique satisfaisant la condition (*). Montrer par un exemple qu'on peut avoir $\omega_F(x) \subsetneq \overline{\mathcal{O}_F^+(x)}$.

Démontrer qu'en général, pour tout $x \in X$, $\overline{\mathcal{O}_F^+(x)} = X$ si et seulement si $\omega_F(x) = X$.

3.3. Transitivité et minimalité. La transitivité topologique requiert un point d'orbite dense (en fait, un ensemble gras sous la condiiton (*)). La minimalité requiert que tous les points soient d'orbite dense. La minimalité implique donc la transitivité.

3.7. Exercice. Montrer par un exemple que la transitivité topologique n'implique pas la minimalité topologique.

On a toutefois une réciproque partielle importante :

3.8. Proposition. *Si (X, T, F) est une dynamique topologique agissant par isométries et vérifiant la condition (*), alors la transitivité est équivalente à la minimalité.*

3.9. Corollaire. *Si G est un groupe topologique compact métrisable et si $\tau : x \mapsto gh$ pour un certain $g \in G$, alors la transitivité est équivalente à la minimalité.*

Démonstration du corollaire. Soit d une distance sur l'espace topologique G. Soit m_G la mesure de Haar invariante par les les translations à droite. Comme G est compact, cette mesure est finite. On pose :

$$\delta(x, y) := \int_G d(hx, hy)\, dm_G(h).$$

C'est bien une distance : $\delta(x, y) = 0 \implies x = y$; $\delta(x, y) = \delta(y, x)$ et pour tous $x, y \in G$, il existe $h \in G$ avec

$$\delta(x, z) = \int_G d(hx, hy)\, dm_G(h) \leq \int_G d(hx, hy) + d(hy, hz)\, dm_G(h) = \delta(x, y) + \delta(y, z).$$

On vérifie :

$$\delta(gx, gy) = \int_G d(hgx, hgy)\, dm_G(h) = \int_G d(kx, ky)\, dm_G(k) = \delta(x, y).$$

En utilisant l'invariance de m_G par la translation $x \mapsto xg$.

Les translations à gauche sont donc des isométries pour δ. Il suffit d'appliquer la proposition pour conclure. C'est possible car la condition (*) est vérifiée. $\square$

Démonstration de la proposition. Supposons que F est topologiquement transitive. D'après la proposition **??**, il existe x_0 tel que $\omega_F(x_0) = X$. Il suffit de vérifier que pour tout $x, y \in X$ et $\epsilon > 0$, il existe $t \in T$ tel que $d(F^t(x), y) < \epsilon$.

Par hypothèse, il existe $u, v \in T$, $d(F^u(x_0), x) < \epsilon/2$ et $d(F^v(x_0), y) < \epsilon/2$ avec $v > u$. On en déduit que

$$\begin{aligned} d(F^{v-u}(x), y) &< d(F^{v-u}(x), F^{u-v}(F^u(x_0)) + d(F^{u-v}(F^u(x_0), y) \\ &= d(x, F^u(x_0)) + d(F^v(x_0), y) < \epsilon. \end{aligned}$$

$\square$

3.10. Remarque. Si la dynamique est inversible ($T = \mathbb{Z}$ ou $\mathbb{R}$), on n'a pas besoin de la condition (*), on peut en effet travailler avec une orbite dense directement.

3.4. Transitivité et semiconjugaison. Le fait suivant est facile à vérifier :

3.11. Proposition. *Si π est une semiconjugaison topologique projetant (X, T, F) sur (Y, T, G), alors x est d'orbite dense implique que $\pi(x)$ est d'orbite dense. En particulier, tout facteur d'une dynamique topologique transitive est lui-même transitif.*

3.12. Exercice. Montrer que pour tout entier $d \geq 2$, $\tau_d : \mathbb{T} \to \mathbb{T}$ définie par $\tau_d(x) = d \cdot x \mod 1$ est topologiquement transitive.

3.13. Proposition. *Si (X, T, F) est une extension par un groupe topologique compact K de (Y, T, G) dynamique topologiquement minimale, alors les propriétés suivantes sont équivalentes :*

(1) la dynamique (X, T, F) est topologiquement minimale ;

(2) la dynamique (X, T, F) est topologiquement transitive.

Démonstration. (1) $\Longrightarrow$ (2) : la minimalité implique que toute orbite est dense donc la transitivité topologique (les dynamiques sont supposées non-vides).

(2) $\Longrightarrow$ (1) : Supposons que F est topologiquement transitive. Soit $Z \subset X$ une partie compacte non-vide et invariante pour F. Son image $\pi(Z)$ est un compact non-vide. Si $z \in Z$, alors $g(\pi(z)) = \pi(f(z)) \in \pi(Z)$: l'image est invariante. Par minimalité de G, $\pi(Z) = Y$.

En faisant agir les auto-conjugaisons topologiques $H_g : (y, k) \mapsto (y, kg)$ pour chaque $g \in K$, on observe que $H_g(Z)$, $g \in K$, donne une partition de X en compacts topologiquement minimaux. Par hypothèse, il existe un point d'orbite dense pour F. Il appartient à un compact topologiquement minimal qui doit donc être tout l'espace. On a montré (1). $\square$

4. Théorème ergodique uniforme (1)

4.1. Définition et premiers exemples. Soit (X, T, F) une dynamique topologique sur un espace métrique compact. L'ensemble des mesures de probabilité boréliennes $\mathbb{P}(X)$ sur l'espace X est un espace compact métrisable pour la topologie *-faible. Le théorème de Krylov-Bogolioubov en déduit l'existence d'au moins une probabilité borélienne invariante : $\#\mathbb{P}(F) \geq 1$. On s'intéresse au cas où cette inégalité est une égalité :

4.1. Définition. Une dynamique topologique est dite *uniquement ergodique* si elle admet une unique mesure de probabilité invariante et si $\mathbb{P}(X)$ est compact.

4.2. Exercice. Montrer que les décalages σ_d et σ_d^+ sont topologiquement transitifs sans être topologiquement minimaux (sauf si $d = 1$).

4.3. Proposition. *Soit G un groupe topologique métrique et compact et $g \in G$. On dénote par $\tau : x \mapsto g.x$ la translation à gauche définie par g.*

Les propriétés suivantes sont équivalentes :

(1) τ topologiquement minimal ;

(2) τ est topologiquement transitive ;

(3) $\{g^n : n \in \mathbb{Z}\}$ est dense dans G ;

(4) τ est uniquement ergodique.

Démonstration. Montrons les implications successives :

(1) $\Longrightarrow$ (2) : clair.

(2) $\Longrightarrow$ (3) : il existe un point $g_0 \in G$ d'orbite dense : pour tout $h \in G$, il existe $n_i \in \mathbb{Z}$ vérifiant : $\tau^{n_i}(g_0) = g^{n_i} g_0 \to h g_0$, donc en multipliant par g_0^{-1} à droite : $g^{n_i} \to h$.

(3) $\Longrightarrow$ (4) : Soit $\mu \in \mathbb{P}(\tau)$. Pour tout $\phi \in C(G)$ et $h \in G$:

$$(\tau_h)^*(\mu)(\phi) = \mu(x \mapsto \phi(h^{-1}x)) = \lim_i \mu(x \mapsto \phi(g^{-n_i}x)) = \mu(\phi)$$

car $\phi(g^{n_i}x) \to \phi(hx)$ par continuité de ϕ et la fonction ϕ est bornée. On a montré que μ est invariante par toutes les translations à gauche. Par unicité de la mesure de Haar normalisée : $\mathbb{P}(\tau)$ est réduite à $\{m_G\}$: τ est uniquement ergodique.

(4) $\Longrightarrow$ (1) : Si τ n'est pas topologiquement minimale, il existe $\emptyset \neq H \subsetneq G$ une partie fermée invariante. On peut donc y restreindre la dynamique topologique. Le théorème de Krylov-Bogolioubov implique qu'il existe $\nu \in \mathbb{P}(F|_H)$. En particulier, $\nu(H) = 1$ alors que $m_{\mathbb{T}}(H) < 1$ car $\mathbb{T} \setminus H$ est ouvert non-vide.

On a montré l'équivalence des quatre propriétés. $\square$

Deux applications immédiates :

4.4. Exercice. Montrer qu'une rotation du cercle $\tau_\alpha : \mathbb{T} \to \mathbb{T}$, $\alpha \in \mathbb{R}$, est uniquement ergodique si et seulement si elle est d'angle irrationnel.

4.5. Exercice. Montrer qu'un odomètre à $d \geq 1$ symboles est uniquement ergodique.

On peut généraliser l'argument de la preuve de l'implication (4) $\Longrightarrow$ (1) ci-dessus en utilisant la notion de support :

4.6. Définition. Le *support* d'une mesure borélienne μ est l'union des ouverts de mesure nulle. On le note $\operatorname{supp}(\mu)$. On dit que μ est de support total si $\operatorname{supp}(\mu) = X$.

La topologie admettant une base dénombrable, $\mu(\operatorname{supp}(\mu)) = 1$.

4.7. Exercice. Soit (X, T, F) une dynamique topologique compacte. Supposons qu'elle admette une ns unique mesure invariante μ et que cette mesure est de support total. Alors F est topologiquement minimal.

4.8. Remarque. Il existe des dynamiques compactes topologiquement minimales qui ne sont pas uniquement ergodiques.

4.2. Théorème ergodique uniforme. Un théorème ergodique est un énoncé reliant moyennes temporelles et intégrales par rapport à une mesure. L'unique ergodicité donne un résultat particulièrement fort grâce à une démonstration particulièrement simple.

On rappelle la définition des sommes de Birkhoff si $T = \mathbb{N}$ ou $T = \mathbb{Z}$:

$$S_n^F \phi(x) := \sum_{k=0}^{n-1} \phi \circ F^k(x) \text{ et } S_{-n}^F \phi(x) := -\sum_{k=1}^{n} \phi \circ F^{-k}(x) \qquad (t \in T_+)$$

et, si $T = \mathbb{R}_+$ ou $T = \mathbb{R}$:

$$S_t^F \phi(x) := \int_{s=0}^{t} \phi \circ F^s(x)\, ds \text{ et } S_{-t}^F \phi(x) := -\int s = 0^t \phi \circ F^{-s}(x)\, ds \qquad (t \in T_+)$$

4.9. Théorème (théorème ergodique uniforme). *Soit (X, T, F) une dynamique topologique sur un espace métrique compact.*

Si F est uniquement ergodique, alors pour toute fonction $\phi \in C(X)$

$$\lim_{t\to\infty} \frac{1}{t} S_F^t \phi(x) \to \mu(\phi) := \int_X \phi \, d\mu.$$

De plus la convergence est uniforme par rapport à x au sens suivant :

$$\lim_{L\to\infty} \sup_{x\in X} \sup_{t\in T, t\geq L} \left| \frac{1}{t} S_F^t \phi(x) - \mu(\phi) \right| = 0 \tag{9}$$

La preuve est modelée sur celle du théorème de Krylov-Bogolioubov (théorème **??**).

Démonstration. On procède par l'absurde en niant l'équation (9), la conclusion la plus forte ci-dessus. On obtient ainsi l'existence de $\phi_0 \in C(X)$ et $\epsilon_0 > 0$ et pour chaque $n \geq 1$, $x_n \in X$, $t_n \in T_+$ avec $t_n \geq n$ vérifiant :

$$\frac{1}{t_n} S_{t_n}^F \phi(x_n) \notin B(\mu(\phi_0), \epsilon_0)$$

On considère les mesures

$$\mu_n := \frac{1}{t_n} S_{t_n}^F \delta(x_n) \in \mathbb{P}(X)$$

Par compacité de $\mathbb{P}(X)$, il existe une sous-suite convergente dans $\mathbb{P}(X)$. Vu $t_n \to \infty$, la limite est une mesure invariante. Mais la limite ne peut être égale à μ. $\square$

4.10. Lemme. *Soit (X, T, F) une dynamique métrique compacte.*

Elle est uniquement ergodique si et seulement si, pour toute fonction $\phi \in C(X)$, la limite suivant existe et indépendante de $x \in X$:

$$\lim_{n\to\infty} \frac{1}{n} S_n\phi(x)$$

Démonstration. Si (X, T, F) est uniquement ergodique, alors l'assertion ci-dessus est une conséquence du théorème ergodique uniforme. Montrons la réciproque. Soit ν, ϕ comme ci-dessus.

On suppose que $T = \mathbb{N}$ ou $\mathbb{Z}$. On laisse les cas $T = \mathbb{R}_+$ ou $\mathbb{R}$ au lecteur diligent. On fixe $x_0 \in X$. L'argument de Krylov-Bogolioubov donne $t_n \to \infty$ tel que les mesures de probabilité

$$\frac{1}{t_n} \sum_{k=0}^{t_n-1} \delta_{F^k x_0} \to \mu_0 \in \mathbb{P}(F).$$

L'hypothèse nous donne que la limite $\Lambda(\Phi) := \lim_n \frac{1}{n} S_n\phi(x)$ existe et est indépendante de $x \in X$. On calcule :

$$\nu(\phi) = \int \frac{1}{n} S_n\phi(x)\, d\nu = \int \lim_n \frac{1}{n} S_n\phi(x)\, d\nu = \Lambda(\phi)$$

grâce au théorème de convergence dominé qui s'applique car :

$$\forall n \geq 1 \qquad \left|\frac{1}{n} S_n\phi(x)\right| \leq \|\phi\|_{\sup}.$$

On voit que $\nu(\phi) = \Lambda(\phi)$ ne dépend pas de $\nu \in \mathbb{P}(F)$. Une mesure de probabilité sur les boréliens d'un espace métrique compact étant définie par ses valeurs sur les fonctions continues d'après le théorème de Riesz, on en déduit que $\mathbb{P}(F)$ contient un seul élément. □

4.11. Proposition. *Si (X, T, F) est une dynamique compacte métrique agissant par isométrie alors les propriétés suivantes sont équivalentes :*

— *F topologiquement transitive ;*

— *F uniquement ergodique et l'unique mesure associée est de support* $\operatorname{supp}(\mu) = X$.

Démonstration. Si F est topologiquement transitive alors, il existe $x_0 \in X$ d'orbite dense.

Soient $\phi \in C(X)$ et $\epsilon > 0$. Par continuité uniforme de ϕ sur le compact X, il existe $\delta > 0$ tel que $d(x, y) < \delta \implies |f(x) - f(y)| < \epsilon/2$. Par hypothèse, il existe $p \in \mathbb{Z}$, $d(f^p(x_0), x) < \delta$. Pour tout $n \geq 1$, on note que $d(f^{k+p}(x_0), f^k(x)) = d(f^p(x_0), x) < \delta$ donc

$$|S_n\phi(x) - S_n\phi(f^p(x_0))| < \epsilon n/2$$

Pour n assez grand, on a donc :

$$\left|\frac{1}{n} S_n\phi(x) - \frac{1}{n} S_n\phi(x_0)\right| < \epsilon.$$

□

Chapitre 4

Théorie ergodique

1. Cadre et notations

1.1. Dynamiques probabilistes standard. Sauf mention contraire,

$$(X, \mathcal{B}, \mu, T, F)$$

est, dans ce chapitre, une *dynamique probabiliste standard* en temps classique :

— $T = \mathbb{N}, \mathbb{Z}, \mathbb{R}_+, \mathbb{R}$;

— $(X, \mathcal{B}, \mu)$ a est un espace de probabilité standard : un espace polonais X muni de sa tribu borélienne $\mathcal{B}$ et d'une mesure de probabilité borélienne μ.

— $F : T \times X \to X$ est mesurable avec $F^t : X \to X$ un endomorphisme de l'espace de probabilité $(X, \mathcal{B}, \mu)$ pour chaque $t \in T$.

On désigne cette dynamique indifféremment par (X, T, F), (μ, F), ou F si les termes manquants sont clairs d'après le contexte.

1.1. Remarque. Si on remplace la probabilité μ par une mesure positive m on dit que la dynamique $(X, \mathcal{B}, m, T, F)$ est *mesurée*.

A préciser : espace du théorème de Riesz ? ? ?

On rappelle les notations :

— $\mathbb{P}(X)$ est l'ensemble des mesures de probabilité boréliennes sur X muni de la tribu des boréliens $\mathcal{B}$;

— $\mathbb{P}(F)$ est l'ensemble des mesures $\nu \in \mathbb{P}(X)$ vérifiant $(F^t)_*(\nu) = \nu$ pour tout $t \in T$.

1.2. Remarque. Chaque $\nu \in \mathbb{P}(F)$ définit une nouvelle dynamique probabiliste standard :

$$(X, \mathcal{B}, \nu, T, F)$$

souvent abrégée en (ν, F).

1.2. Fonctions mesurables, égalité presque partout et espaces de Lebesgue. On dénote par $\mathbb{K}$ le corps des réels ou des complexes.

Deux fonctions mesurables $\phi, \psi : X \to \mathbb{K}$ sont égales μ-presque partout si

$$\mu(\{x \in X : \phi(x) \neq \psi(x)\}) = 0$$

On le note $\phi \overset{\mu}{\equiv} \psi$.

On note l'ensemble des fonctions mesurables (indépendant de μ) :

$$\mathcal{L}^0(\mu, \mathbb{K}) := \{\phi : X \to \mathbb{K} : f^{-1}(\mathcal{B}_\mathbb{R}) \subset \mathcal{B}\}$$

et pour $0 < p < \infty$, on note les espaces de Lebesgue d'exposant p (dépendants de μ) :

$$L^p(\mu, \mathbb{K}) := \{\phi \in L^0(\mu, \mathbb{K}) : \int |\phi|^p \, d\mu < +\infty\}$$

et

$$L^\infty(\mu, \mathbb{K}) := \{\phi \in L^0(\mu, \mathbb{K}) : \exists t < \infty \; \mu(\phi^{-1}([-t, t]) = 1\}$$

On définit $L^p(\mu, \mathbb{K}) := \mathcal{L}^p(\mu, \mathbb{K}) / \overset{\mu}{\equiv}$ pour tout $0 \leq p \leq \infty$. Pour $1 \leq p \leq \infty$, $L^p(\mu, \mathbb{K})$ est l'espace de Lebesgue d'exposant p une fois muni de la norme :

$$\|\phi\|_p := \left(\int_X |\phi|^p \, d\mu \right)^{1/p} \quad \text{et } \|\phi\|_\infty := \inf\{t \in (-\infty, \infty] : \mu(\phi^{-1}([-t, t])) = 1\}.$$

Le quotient par l'égalité μ-presque partout assure l'implication : $\|\phi\|_p = 0 \implies \phi = 0$ et la complétude.

On rappelle que chaque $L^p(\mu)$, $1 \leq p \leq \infty$, est un espace de Banach. Ces espaces sont séparables si $1 \leq p < \infty$ tandis que $L^2(\mu)$ est un espace de Hilbert.

2. Récurrence et induction

**** récurrence
**** Théorème de Kac
(**** suites de Poincaré)

3. Ergodicité

On peut ramener l'étude de nombre de problèmes au cas systèmes indécomposables, en introduisant un analogue formel de la minimalité topologique.

3.1. Définition de l'ergodicité. On rappelle que deux parties mesurables $A, B \in \mathcal{B}$ sont égales modulo μ si et seulement si leur *différence symétrique* :

$$A \Delta B := (A \setminus B) \cup (B \setminus A)$$

est de mesure nulle. On écrit alors $A \overset{\mu}{\equiv} B$.

3.1. Définition. La dynamique probabiliste ou mesurée $(X, \mathcal{B}, \mu, T, F)$ est *ergodique* si toute partie mesurable invariante modulo μ est triviale, ie :

$$\forall Y \in \mathcal{B} \quad \left(\forall t \in T \; F^{-t}(Y) \overset{\mu}{\equiv} Y \right) \implies Y \overset{\mu}{\equiv} \emptyset \text{ ou } Y \overset{\mu}{\equiv} X.$$

On peut réécrire cette dernière condition en $\mu(Y) = 0$ ou $\mu(X \setminus Y) = 0$, ou encore, si $\mu(X) = 1$, $\mu(Y) = 0$ ou $\mu(Y) = 1$.

Si la dynamique n'est pas ergodique alors elle se décompose en deux sous-systèmes de façon évidente. Il s'agit d'un analogue formel de la notion de partie topologiquement minimale.

3.2. Caractérisations. On considère les opérateurs de composition pour chaque $0 \leq p \leq \infty$:

$$U_F^t : L^p(\mu, \mathbb{K}) \to L^p(\mu, \mathbb{K}), \quad \phi \mapsto \phi \circ F^t.$$

3.2. LEMME. *Pour tout $0 \leq p \leq \infty$ et $t \in T$, les opérateurs $U_F^t : L^p(\mu, \mathbb{K}) \to L^p(\mu, \mathbb{K})$ sont bien définis. Pour $1 \leq p \leq \infty$, $|U_F^t(\phi)\|_p = \|\phi\|_p$. Ces opérateurs linéaires sont donc isométriques (en particulier continus). Ils sont inversibles si $T = \mathbb{Z}$ ou $\mathbb{R}$.*

DÉMONSTRATION. $\phi \in L^p(\mu, \mathbb{K})$, alors $\phi \circ F^t$ est un élément bien défini de $L^p(\mu, \mathbb{K})$.

En effet si $\phi \overset{\mu}{\equiv} \psi$ alors, pour tout $t \in T$: $\phi \circ F^t$ est bien définie μ-presque partout car F^t est un endomorphisme de $(X, \mathcal{B}, \mu)$.

$$\mu(\{\phi \circ F^t \neq \psi F^t\}) = \mu(\{\phi \neq \psi\}) = 0$$

et, si $1 \leq p < \infty$:

$$\|\phi \circ F^t\|_p = \left(\int_X |\phi \circ F^t|^p \, d\mu\right)^{1/p} = \left(\int_X |\phi|^p \, d\mu\right)^{1/p} = \|\phi\|_p$$

par invariance de μ. On démontre sans plus de difficultés le cas $p = \infty$.

Si $T = \mathbb{Z}$ ou $\mathbb{R}$, alors $(U_F^t)^{-1} = U_F^{-t}$. $\square$

3.3. LEMME. *Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste et $0 \leq p \leq \infty$. L'ergodicité de F est équivalente à la propriété suivante :*

$$\forall \phi \in L^p(\mu, \mathbb{K}) \quad \big(\forall t \in T \ \phi = \phi \circ F^t\big) \implies \phi \text{ est constante.}$$

Soulignons que la condition ci-dessus est en terme d'égalités entre classes de fonctions (on peut la réécrire en terme de fonctions et les égalités deviennent alors μ-presque partout).

DÉMONSTRATION. Si F n'est pas ergodique, il existe une partie Y invariante modulo μ avec $0 < \mu(Y) < 1$. Il suffit de considérer la fonction indicatrice de Y. Elle est mesurable, non triviale modulo μ, bornée et donc dans tout $L^p(\mu, \mathbb{R}) \subset L^p(\mu, \mathbb{C})$, la mesure étant une probabilité.

Réciproquement soit $\phi \in L^0(\mu, \mathbb{C})$, non triviale modulo μ et invariante modulo μ. On va en déduire que F n'est pas ergodique en exhibant une partie mesurable invariable modulo μ et non triviale modulo μ.

Remarquons qu'on peut supposer ϕ à valeurs réelles quitte à la remplacer par sa partie réelle ou imaginaire. En effet, ces parties réelles et imaginaires sont des fonctions à valeurs réelles et sont invariantes modulo μ. Si elles étaient toutes les deux triviales modulo μ, ϕ le serait aussi.

On affirme que qu'il existe $c \in \mathbb{R}$ tel que $Y := \{\phi \leq c\}$ vérifie $0 < \mu(Y) < 1$. Sinon, $c \mapsto \mu(\{\phi \leq c\})$ est une fonction croissante qui ne prend que les valeurs 0 et 1. On identifie la valeur "critique" :

$$c := \sup\{t \in \mathbb{R} : \mu(\{\phi \leq t\}) = 0\}.$$

Si $c = -\infty$, $\mu(\{\phi \leq t\}) = 1$ pour tout $t \in \mathbb{R}$ et donc la σ-additivité de μ implique que : $\mu(\{\phi = -\infty\}) = 1$ contredisant que ϕ soit à valeurs réelles. De même $c = +\infty$ impliquerait que $\mu(\{\phi = +\infty\}) = 1$, une autre contradiction. On a donc bien $c \in \mathbb{R}$.

Si $t < c$, $\mu(\{\phi \leq t\}) = 0$ et pour $t > c$, $\mu(\{\phi \leq t\}) = 1$. La valeur en $t = c$ n'est pas immédiate. On calcule en utilisant encore la σ-additivité :

$$\mu(\{\phi = c\}) = \mu(\bigcap_{n\geq 1}\{c-1/n < \phi \leq c-1/n\}) = \lim_n \mu(\{\phi \leq c+1/n\}) - \mu(\{\phi \leq c-1/n\}) = 1.$$

Ceci contredit la non-trivialité modulo μ de ϕ. Il existe donc bien $c \in \mathbb{R}$ tel que : $\{phi \leq c\}$ est une partie mesurable non-triviale modulo μ. Elle est invariante car, pour tout $t \in T$:

$$F^{-t}(\{\phi \leq c\}) = \{\phi \circ F^t \leq c\} \overset{\mu}{\equiv} \{\phi \leq c\}$$

grâce à l'invariance modulo μ de ϕ. La dynamique n'est donc pas ergodique. □

3.4. Exercice. Etendre le lemme précédent aux fonctions à valeurs complexes.

3.5. Corollaire. *La dynamique (μ, F) est ergodique si et seulement s'il n'existe pas de $\phi \in L^2(\mu, \mathbb{C})$ qui soit invariante modulo μ.*

3.3. Exemples.

3.6. Proposition. *Soit $\tau : \mathbb{T} \to \mathbb{T}$ la rotation $x \mapsto x + \alpha$ avec $\alpha \in \mathbb{R}$ préservant la mesure uniforme $m_{\mathbb{T}}$. Cette dynamique probabiliste est ergodique si et seulement si l'angle $\alpha \notin \mathbb{Q}$.*

Démonstration. D'après le lemme précédent, l'ergodicité recherchée est équivalente à l'assertion :

$$\forall \phi \in L^2(\mu) \quad (\forall t \in T\ \phi \circ F^t = \phi) \implies \phi \text{ triviale.}$$

On rappelle que $L^2(\mu)$ est l'espace de Hilbert des classes de fonctions à valeurs complexes et dont le carré du modulo est intégrable.

On utilise la théorie des séries de Fourier : les fonctions $e_n : x \mapsto e^{2i\pi nx}$, $n \in \mathbb{Z}$, forment une base hilbertienne de $L^2(m_{\mathbb{T}})$.

On suppose $\phi \in L^2(m_{\mathbb{T}})$ invariante. □

3.7. Exercice. Soit $T : \mathbb{T} \to \mathbb{T}$, $x \mapsto d \cdot x$ la multiplication par un entier $d \geq 2$ préservant la mesure uniforme $m_{\mathbb{T}}$. En utilisant les séries de Fourier, montrer que cette dynamique probabiliste est ergodique.

3.8. Lemme. *Si $T = \mathbb{N}$ ou $\mathbb{Z}$ et $\phi : X \to \mathbb{K}$ une fonction mesurable et invariante modulo μ : $\forall t \in T \quad \phi \circ F^t \overset{\mu}{\equiv} \phi$ alors ϕ est égale modulo μ à une fonction (exactement) invariante.*

Démonstration. On définit une nouvelle fonction $\psi : X \to \mathbb{R}$ en posant :

$$\psi(x) := \limsup_{n\to\infty} \frac{1}{n} S_n^F \phi(x).$$

Notons les propriétés suivantes de ψ : (i) définie partout et mesurable comme ϕ ; (ii) ψ est (exactement) invariante : $\psi \circ F^t = \psi$. Finalement,

$$\psi(x) \neq \phi(x) \implies \exists t \in T_+ \phi(F^t(x)) \neq \phi(x).$$

Or ceci n'a lieu que pour un ensemble de points $x \in X$ de mesure nulle. □

4. Théorème ergodique ponctuel – cas particulier

4.1. Théorème. *Soit $(X, \mathcal{B}, \mu, \mathbb{N}, F)$ une dynamique probabiliste supposée ergodique. Soit $\phi \in L^1(\mu)$. On a la convergence suivante :*

$$\frac{1}{n} S_n \phi(x) \longrightarrow \mu(\phi) \text{ } \mu\text{-presque partout.}$$

Le point clé est l'existence de la limite μ-presque partout de

$$\tilde{\phi}(x) := \lim_n \frac{1}{n} S_n \phi(x).$$

Le reste s'en déduit relativement facilement :

4.2. Exercice. Déduire le théorème 4.1 de l'existence de $\tilde{\phi}$ μ-presque partout.

Le lemme suivant implique que $\limsup_{n\to\infty} \frac{1}{n} S_n^F \phi(x)$ est une fonction invariante.

4.3. Lemme. *Si $\phi \in L^1(\mu)$, alors $\lim_{n\to\infty} \frac{1}{n}\phi(F^n x) = 0$.*

4.4. Exercice. Montrer que cette propriété est nécessaire, ie, vérifier que c'est une conséquence (facile) du théorème 4.1.

On va déduire le lemme du résultat suivant, classique en théorie des probabilités :

4.5. Lemme (Borel-Cantelli). *Soit $(X, \mathcal{B}, \mu)$ un espace de probabilité quelconque et $X_1, X_2, \dots$ une suite de parties mesurables. On a l'implication :*

$$\sum_{n\geq 1} \mu(X_n) < \infty \implies \text{pour } \mu\text{-p.t. } x \in X \ \#\{n \geq 1 : x \in X_n\} < \infty.$$

Si les $X_1, X_2, \dots$ sont indépendants[1]*, alors*

$$\sum_{n\geq 1} \mu(X_n) < \infty \iff \text{pour } \mu\text{-p.t. } x \in X \ \#\{n \geq 1 : x \in X_n\} < \infty.$$

4.6. Exercice. Démontrer la première implication en utilisant le théorème de Tonelli (Fubini pour les fonctions positives) : Si $(X, \mu), (Y, \nu)$ sont deux espaces de probabilités et si ϕ est une fonction mesurable et positive sur leur produit, alors

$$\int \phi \, d\mu \otimes \nu = \int \left(\int \phi(x, y) \, d\mu(x) \right) \, d\nu(y) = \int \left(\int \phi(x, y) \, d\nu(y) \right) \, d\nu(x)$$

et ces trois expressions sont ou bien toutes finies ou bien toutes infinies.

Démontrer la réciproque en estimant $\prod_{k\geq n}(1 - \mu(X_k))$ pour $n \to \infty$.

Démonstration du lemme 4.3. Il suffit de montrer que

$$\limsup_{n\to\infty} \frac{1}{n} \phi(F^n x) \leq 0 \text{ pour } \mu\text{-p.t. } x \in X \tag{10}$$

puis d'appliquer ce premier résultat à $-\phi$ qui vérifie les mêmes hypothèses. Pour ce faire, on montre, pour tout $\epsilon > 0$, que :

$$\text{pour } \mu\text{-p.t. } x \in X, \ \#\{n \geq 1 : \phi(F^n(x)) > \epsilon \cdot n\} < \infty. \tag{11}$$

1. C'est-à-dire : $\mu(X_{n_1} \cap X_{n_2} \cap \dots X_{n_r}) = \mu(X_{n_1})\mu(x_{n_2})\dots\mu(x_{n_r})$ pour toute suite finie $1 \leq n_1 < n_2 < \dots < n_r$.

Remarquons, qu'en notant $[\cdot]$ la partie entière et $x^+ := \max(x, 0)$ la partie positive, on a :

$$\sum_{n\geq 1} \mu(\{\phi(F^n x) > \epsilon \cdot n\}) = \sum_{n\geq 1} \mu(\{\phi(x) > \epsilon \cdot n\}) = \sum_{n\geq 1} \int_X 1_{\phi>\epsilon n}\, d\mu$$
$$\leq \int_X \left[\frac{\phi(x)^+}{\epsilon}\right] d\mu \leq \int_X \frac{1}{\epsilon}|\phi|\, d\mu = \frac{1}{\epsilon}\|\phi\|_{L^1(\mu)} < \infty.$$

où on s'est servi de l'invariance de μ par F et du théorème de Tonelli sur l'échange des intégrations lorsque la fonction est positive. Le lemme de Borel-Cantelli dit que les événements $\{\phi(F^n x) < \epsilon \cdot n\}$ étant de mesures sommables, ils ne se produisent, μ-p.p., qu'un nombre fini de fois. L'équation (11) est donc démontrée. $\square$

Démonstration du théorème 4.1. Il suffit de montrer la convergence de $\phi - \mu(\phi)$ vers 0, ie, on peut supposer $\mu(\phi) = 0$. On démontre d'abord :

$$\bar{\phi}(x) := \limsup_n \frac{1}{n} S_n \phi(x) \leq 0 \ \mu\text{-pp} \tag{12}$$

La limite sup ci-dessus définit une fonction mesurable. Le lemme 4.3 implique que :

$$\lim_{n\to\infty} \left| \frac{1}{n} S_n^F \phi(Fx) - \frac{1}{n} S_n^F \phi(x) \right| = \lim_{n\to\infty} \frac{1}{n}|\phi(F^n x)| + \frac{1}{n}|\phi(x)| = 0$$

C'est à dire que la fonction $\bar{\phi}$ est également invariante modulo μ. Par ergodicité, elle est égale, μ-p.p. à une certaine constante. Supposons cette constante strictement positive et recherchons une contradiction.

On définit une fonction mesurable et finie μ-p.p. en posant :

$$N(x) := \min\{n \geq 1 : \frac{1}{n} S_n \phi(x) > \alpha\}$$

Par σ-additivité, il existe un entier M tel que

$$C := \{N > M\} \text{ vérifie } \int_C |\phi| + \alpha\, d\mu < \frac{\alpha}{3}.$$

On va en déduire que les sommes de Birkhoff sont excessives par un découpage de l'intervalle entier $[\![0, n[\![:= \{0, 1, 2, \ldots, n-1\}$. Pour μ-p.t. $x \in X$, on définit par récurrence b_j, $j \geq 0$, et a_j, $j \geq 1$ en posant $b_0 := 0$ et :

$$a_{j+1} := \inf\{k \geq b_j : F^k(x) \notin C\} \quad b_{j+1} := a_{j+1} + N(F^{a_{j+1}} x)$$

On pose $r := \min\{n \geq 0 : b_{n+1} \geq n\}$. Notons qu'on peut avoir $a_{r+1} \geq n$ ou non. Dans ce dernier cas, $[\![a_{r+1}, b_{r+1}[\![$ intersecte mais n'est pas contenu dans $[\![0, n[\![$ et on a alors : $[\![a_{r+1}, n[\![\subset [\![n - M, n[\![$.

On a la partition :

$$[\![0, n[\![= [\![0, a_1[\![\cup [\![a_1, b_1[\![\cup [\![b_1, a_2[\![\cup \cdots \cup [\![b_r, \min(a_{r+1}, n)[\![\cup [\![\min(a_{r+1}, n), n[\![$$

On a :

$$[\![0, n[\![\setminus \bigcup_{i=1}^r [\![a_i, b_i[\![\subset \{k \in [\![0, n[\![: F^k(x) \in C\} \cup [\![n - M, n[\![.$$

On en déduit :

$$\sum_{k\in\bigcup_{i=1}^r [\![a_i, b_i[\![} \phi(F^k x) \geq \alpha \sum_{i=1}^r |b_i - a_i| = \alpha \cdot n - S_n(\alpha \cdot 1_C)(x) - M\alpha.$$

Donc :
$$S_n\phi(x) \geq \alpha n - S_n((|\phi| + \alpha)1_C)(x) - S_M(|\phi| + \alpha)(F^{n-M}x).$$
En intégrant par rapport à μ :
$$0 = n\int_X \phi\, d\mu = \int_X S_n\phi\, d\mu \geq n\cdot\alpha - n\int 1_C\cdot(|\phi|+\alpha)\, d\mu - M\int_X (|\phi| + \alpha)$$
En divisant par $n\cdot\mu(B)$:
$$0 \geq \alpha - \|1_C\cdot(|\phi|+\alpha)\|_{L^1} - \frac{M}{n}\left(\|\phi\|_{L^1} + \alpha\right)$$
Le choix de M implique que le second terme est majoré par $\alpha/3$. Pour n assez grand, le troisième et dernier terme l'est aussi. On obtient $\alpha \leq 0$, une contradiction.

L'inégalité (17) est établie. En l'appliquant à l'opposé de ϕ, on obtient
$$\liminf_n \frac{1}{n}S_n\phi \geq 0 \ \mu\text{-pp}$$
d'où la convergence μ-p.p. Le théorème est démontré. □

4.7. Corollaire. *Si $(X, \mathcal{B}, \mu, \mathbb{N}, F)$ est une dynamique probabiliste ergodique et $\phi \in L^p(\mu)$ avec $1 \leq p < \infty$.*
$$\frac{1}{n}S_n\phi(x) \longrightarrow \mu(\phi) \text{ au sens de la topologie } L^p(\mu).$$

Les deux cas essentiels sont $p = 1$ et $p = 2$.

Démonstration. Fixons $1 \leq p < \infty$ et $\phi \in L^p(\mu)$. Remarquons que μ étant une probabilité, $L^p(\mu) \subset L^1(\mu)$. On peut donc appliquer le théorème 4.1 et obtenir la convergence μ-presque partout.

On se place d'abord dans le cas particulier où ϕ est borné : $\phi \in L^\infty(\mu)$. Pour tout $n \geq 1$,
$$\left|\frac{1}{n}S_n\phi - \mu(\phi)\right|^p \leq 2^p\|\phi\|^p_{L^\infty}\cdot\mathbf{1} \in L^1(\mu).$$
Le théorème de convergence dominée s'applique :
$$\frac{1}{n}S_n\phi \overset{n\to\infty}{\longrightarrow} \mu(\phi) \text{ dans } L^p(\mu).$$
Le résultat est démontré dans le cas borné.

On considère maintenant $\phi \in L^p(\mu)$. On suppose ϕ à valeurs dans $\mathbb{R}_+$, le cas général s'y réduisant sans difficulté. On se donne $\epsilon > 0$.

On pose :
$$\phi_M(x) := \min(\phi(x), M).$$
Vu $0 \leq (\phi - \phi_M)^p \leq \phi^p \in L^1(\mu)$, le théorème de convergence dominée donne M_0 tel que :
$$\|\phi_{M_0} - \phi\|_{L^p} < \epsilon/3 \text{ et } |\mu(\phi) - \mu(\phi_{M_0})| < \epsilon/3.$$

Le cas particulier des fonctions bornées traité ci-dessus s'applique à ϕ_{M_0} et donne n_0 tel que :
$$\forall n \geq n_0 \quad \left\|\frac{1}{n}S_n\phi_{M_0} - \mu(\phi_{M_0})\mathbf{1}\right\|_{L^p} < \epsilon/3.$$

L'inégalité triangulaire donne

$$\forall n \geq 1 \quad \left\| \frac{1}{n} S_n \phi - \frac{1}{n} S_n \phi_{M_0} \right\|_{L^p} \leq \| \phi - \phi_{M_0} \|_{L^p} .$$

On en déduit que pour tout $n \geq n_0$:

$$\begin{aligned} \left\| \frac{1}{n} S_n \phi - \mu(\phi) \mathbf{1} \right\|_{L^p} \leq & \left\| \frac{1}{n} S_n \phi - \frac{1}{n} S_n \phi_{M_0} \right\|_{L^p} \\ & + \left\| \frac{1}{n} S_n \phi_{M_0} - \mu(\phi_{M_0}) \mathbf{1} \right\|_{L^p} + |\mu(\phi) - \mu(\phi_{M_0})| < \epsilon \end{aligned}$$

□

Chapitre 5

Applications du théorème ergodique

1. Loi forte des grands nombres

1.1. THÉORÈME. *Supposons que $X_1, X_2, \dots$ sont des variables aléatoires à valeurs réelles indépendantes et identiquement distribuées vérifiant $\mathbb{E}(|X_1|) < \infty$. Alors, presque sûrement,*

$$\lim_{n\to\infty} \frac{1}{n}(X_1 + \cdots + X_n) = \mathbb{E}(X_1).$$

DÉMONSTRATION. Soit $P \in \mathbb{P}(\mathbb{R})$ la loi de n'importe lequel des X_n.

Soit $X := \mathbb{R}^{\mathbb{N}^*}$ muni de la loi produit $\mu := P^{\otimes \mathbb{N}^*}$ et $S : X \to X$, $S((x_n)_{n\geq 1}) = (x_{n+1})_{n\geq 1}$. Notons que $\mu \in \mathbb{P}(S)$. Soit $\pi : \mathbb{R}^{\mathbb{N}^*} \to \mathbb{R}$, $(x_n)_{n\geq 1} \mapsto x_0$. On pose :

$$\forall n \geq 1 \ \tilde{X}_n := \pi \circ S^n$$

On constate que pour tout $N \geq 1$, la loi de $(\tilde{X}_1, \dots, \tilde{X}_N)$ coïncide avec la loi de $(X_1, \dots, X_n)$. L'assertion du théorème exprime dont dans le langage de la théorie des probabilités la convergence :

$$\lim_n \frac{1}{n}(\tilde{X}_1 + \cdots + \tilde{X}_n) = \mu(X_1) \ \mu\text{-p.p.}$$

Cela découle du théorème de Birkhoff pourvu qu'on montre l'ergodicité de (S, μ).On va montrer un peu plus en appliquant le théorème ******* : (S, μ) est fortement mélangeante .

Soit

$$\mathcal{B}^* := \{A_p, \dots, A_{p+\ell-1}) \in \mathcal{B}_{\mathbb{R}}^{\ell} : p \in \mathbb{Z},\ \ell \in \mathbb{N}\}.$$

Pour $A = (A_p, \dots, A_{p+\ell-1}) \in \mathcal{B}^*$, on pose $1_A(x) = \prod_{i=p}^{p+\ell-1} S^{-i}(A_i)$ et $\mathcal{D} := \{1_A : A \in \mathcal{B}^*, \}$, l'ensemble des fonctions indicatrices de cylindres. Comme les cylindres forment une base de la topologie, $\mathcal{D}$ engendre un espace vectoriel dense dans $L^2(\mu)$.

Soient $(A_p, \dots, A_{p+\ell-1}), (B_q, \dots, B_{q+k-1}) \in \mathcal{B}^*$. Notons que

$$\forall n \in \mathbb{Z} \quad 1_B \circ \sigma^n = 1_{B_{q+n}, \dots, B_{q+n+k-1}}$$

De sorte que, dès que $n > p - q + \ell$

$$\mu(1_A.1_B \circ S^n) = \mu(1_A)\mu(1_B)$$

Le mélange est démontré. L'ergodicité s'ensuit. □

2. Nombres normaux au sens de Borel

On choisit un entier $d \geq 2$. On sait que $T : \mathbb{T} \to \mathbb{T}$, $x \mapsto bx \mod 1$ préserve la mesure uniforme m sur $\mathbb{T}$ et (T, m) est ergodique. On en déduit :

2.1. Théorème. *Lebesgue-presque tout réel $x \in \mathbb{R}$, a un développement normal en toute base $b \geq 2$: si $x = n + \sum_{i \geq 1} c_i^b \cdot b^{-i}$ avec $c_i^b \neq b-1$ pour une infinité de $i \geq 1$, alors, pour toute suite finie $a \in \{0,1,\dots,b-1\}^{\{1,2,\dots,\ell\}}$, $\ell \geq 1$, on a :*

$$\lim_{n\to\infty} \frac{1}{n} \#\{0 \leq i < n : c_{i+1}c_{i+2}\dots c_{i+\ell} = a_1a_2\dots a_\ell\} = b^{-\ell}.$$

Démonstration. On a vu précédemment que :

$$c_i^b = C(T_b^{i-1}(\pi(x))) \text{ avec } C^b(\pi(t)) = E(bt) \text{ si } 0 \leq t < 1.$$

On en déduit que $c_i^b + 1 \dots c_{i+n}^b = a$ si et seulement si

$$T_b^i(\pi(x)) \in \left[\sum_{i=1}^{\ell} a_i b^{-i}, \sum_{i=1}^{\ell} a_i b^{-i} + b^{-\ell}\right[+ \mathbb{Z}.$$

Le théorème de Birkhoff appliqué au système ergodique $(T_b, m_{\mathbb{T}})$ donne alors que la limite recherchée existe et vaut

$$m_{\mathbb{T}}\left(\left[\sum_{i=1}^{\ell} a_i b^{-i}, \sum_{i=1}^{\ell} a_i b^{-i} + b^{-\ell}\right[+ \mathbb{Z}\right) = b^{-\ell}$$

□

2.2. Remarque. Le résultat précédent ne dit *rien* sur les réels spécifiques, par exemple $\sqrt{2}$ ou π.

3. Transitivité topologique

3.1. Proposition. *Si (X,T,F) est une dynamique topologique standard, et $\mu \in \mathbb{P}_{\mathrm{erg}}(F)$, alors pour μ-p.t. $x \in X$, on a :*

$$x \in \operatorname{supp}(\mu) \text{ et } \omega_F(x) = \operatorname{supp}(\mu).$$

Démonstration. On traite le cas du temps discret. Le temps discret est laissé au lecteur.

Si $y \in \operatorname{supp}(\mu)$ et $r > 0$, alors $\mu(B(y,r)) > 0$. On applique le théorème ergodique à (μ, F) et $1_{B(y,r)}$. On obtient une partie mesurable $X' \subset X$ tel que pour tout $x \in X'$,

$$\lim_n \frac{1}{n}\{0 \leq k < n : F^k(x) \in B(y,r)\} = \mu(B(y,r)) > 0.$$

En particulier il existe $k \to \infty$ $F^k(x) \in B(y,r)$ pour tout $x' \in X$.

Par définition (X étant séparable et métrique), $\mu(\operatorname{supp}(\mu)) = 1$, on peut donc remplacer X' par $X' \cap \operatorname{supp}(\mu)$. □

4. Mesures empiriques

On va définir la mesure empirique au temps $t \in T$, $t > 0$, $\mu_{x,t}^F$ de sorte que $\mu_{x,t}^F(\phi) = \frac{1}{t} S_t^F \phi(x)$ pour toute fonction $\phi \in C(X)$. Ceci définit bien une mesure de probabilité (non nécessairement invariante) en appliquant le théorème de Riesz.

Plus explicitement, en temps discret, on rappelle que δ_x désigne la mesure de probabilité sur $\mathcal{B}$ définie par

$$\forall B \in \mathcal{B} \quad \delta_x(B) = \begin{cases} 1 & \text{si } x \in B \\ 0 & \text{sinon.} \end{cases}$$

On appelle δ_x la *masse de Dirac* au point x. On a alors :

$$\mu^F_{x,t} := \sum_{s=0}^{t-1} \delta_{F^s x}$$

4.1. Définition. Soit (X, T, F) un système dynamique topologique standard et $\mu \in \mathbb{P}(F)$. Un point $x \in X$ est dit *générique* pour la mesure $\mu \in \mathbb{P}(F)$ si

$$\lim_{t\to\infty} \mu^F_{x,t} = \mu \text{ au sens de la topologie faible.}$$

L'ensemble de ces points est appelé *bassin ergodique* de μ et noté $\mathfrak{B}(\mu)$.

4.2. Remarque. La notion de généricité suppose le choix d'une topologie sur X. Elle n'est donc pas définie *a priori* sur un espace de probabilité.

4.3. Corollaire. *Soit (X, T, F) une dynamique topologique compacte métrique. Soit $\mu \in \mathbb{P}(F)$. Si μ est ergodique, alors μ-p.t. $x \in X$ est générique pour μ :*

$$\forall \mu \in \mathbb{P}_{\mathrm{erg}}(F) \quad \mu(\mathcal{B}(\mu)) = 1.$$

Démonstration. D'après la définition de la topologie faible, il suffit de trouver une partie mesurable $X' \subset X$ vérifiant :

$$\forall x \in X' \; \forall \phi \in C(X) \quad \lim_{n\to\infty} \frac{1}{n} S^F_n \phi(x) \to \mu(\phi). \tag{13}$$

Remarquons que l'ensemble de mesure totale X' doit convenir à toutes les fonctions continues. Cette "inversion des quantificateurs" est rendue possible par la séparabilité de $C(X)$:

Fixons $\phi_1, \phi_2, \ldots$ une suite dense dans $C(X)$ muni de la topologie uniforme.

Le théorème de Birkhoff donne, pour chaque entier $N \geq 1$, un mesurable X_N avec $\mu(X_N) = 1$ et pour tout $x \in X_N$,

$$\lim_{n\to\infty} \frac{1}{n} S^F_n \phi_N(x) \to \mu(\phi_N).$$

On en déduit que, pour tout $x \in X' := \bigcap_{N\geq 1} X_N$ et tout $\phi \in C(X)$ on a bien (13) : pour tout $\epsilon > 0$, il existe $N \geq 1$ avec $\|\phi - \phi_N\|_{\sup} < \epsilon$ et donc

$$\forall x \in X' \quad \limsup_{n\to\infty} \frac{1}{n} S^F_n \phi(x) \leq \limsup_{n\to\infty} \frac{1}{n} S^F_n \phi_N(x) + \epsilon \leq \mu(\phi_N) + \epsilon \leq \mu(\phi) + 2\epsilon.$$

Or $\epsilon > 0$ est arbitraire donc : $\limsup_{n\to\infty} \frac{1}{n} S^F_n \phi(x) \leq \mu(\phi)$. De même

$$\liminf_{n\to\infty} \frac{1}{n} S^F_n \phi(x) \geq \mu(\phi).$$

D'où (13). □

5. Mesures ergodiques comme points extrémaux

La limite μ-presque partout des moyennes de Birkhoff suffit à identifier la mesure μ :

5.1. Lemme. *Soit (X, T, F) une dynamique mesurable standard. Soient $\mu \in \mathbb{P}(F)$ et $m \in \mathbb{P}(X)$. Soit $\mathcal{D} \subset \mathcal{L}^\infty(\mu)$ dense dans $L^1(\mu)$:*

$$\left[\forall \phi \in \mathcal{D} \; \mu\text{-p.p.} \; \lim_{n\to\infty} \frac{1}{n} S^F_n \phi \text{ existe et vaut } m(\phi)\right] \implies \mu = m.$$

DÉMONSTRATION. Soit $\phi \in \mathcal{D}$, une (vraie) fonction mesurable et bornée. On note $\|\frac{1}{n}S_n^F\phi\|_\infty \leq \|\phi\|_\infty$ vu l'invariance de μ. Le théorème de convergence dominé permet d'affirmer :

$$m(\phi) = \int_X \lim_n \frac{1}{n} S_n^F \phi \, d\mu = \lim_n \frac{1}{n} \int_X S_n^F \phi \, d\mu = \mu(\phi)$$

En particulier, pour toute partie mesurable $A \subset X$, on a $\mu(A) = m(A)$.[1] $\square$

5.2. PROPOSITION. *Soient $\mu \in \mathbb{P}_{\mathrm{erg}}(F)$ et $\nu \in \mathbb{P}(F)$. Si $\nu \ll \mu$ alors $\nu = \mu$.*

DÉMONSTRATION. Le théorème ergodique appliqué à (μ, F) et $\phi \in L^1(\mu)$ donne :

$$\mu\text{-p.p.} \lim_{n\to\infty} \frac{1}{n} S_n^F \phi = \mu(\phi).$$

Vu $\nu \ll \mu$, ceci est également vrai ν-p.p. Le lemme 5.1 permet d'en déduire $\nu = \mu$. $\square$

5.3. EXERCICE. Montrer par un exemple qu'on peut avoir, avec les notations de la proposition, $\mu \ll \nu$ et $\nu \neq \mu$.

Le fait suivant est clair :

5.4. LEMME. *Le compact $\mathbb{P}(F)$ est un convexe :*

$$\mu, \nu \in \mathbb{P}(F) \text{ et } 0 \leq t \leq 1 \implies t \cdot \mu + (1-t) \cdot \nu \in \mathbb{P}(F)$$

5.5. DÉFINITION. Soit un convexe K dans un espace vectoriel quelconque.
$p \in K$ est un *point extrémal* de K si pour tous $0 < t < 1$ et $m, n \in K$:

$$p = t \cdot m + (1-t) \cdot n \implies p = m = n.$$

5.6. PROPOSITION. *Une mesure $\mu \in \mathbb{P}(f)$ est ergodique si et seulement si c'est un point extrémal du convexe compact $\mathbb{P}(F)$.*

DÉMONSTRATION. Si $\mu \in \mathbb{P}(f)$ n'est pas ergodique, il existe une partie mesurable $A \subset X$, non-triviale $(0 < \mu(A) < 1)$ et invariante modulo μ, on peut donc poser :

$$\mu = \mu(A) \cdot \nu|_A + \mu(X \setminus A) \cdot \nu|_{X\setminus A}$$

et $\mu|_A, \mu|_{X\setminus A} \in \mathbb{P}(F)$ et $0 < \mu(A) < 1$. On constate que μ n'est pas extrémale.

Reciproquement, si $\mu \in \mathbb{P}(F)$ n'est pas extrémale, il existe $0 < t < 1$, $\nu_1 \neq \nu_2$ deux éléments de $\mathbb{P}(F)$ avec :

$$\mu = t \cdot \nu_1 + (1-t) \cdot \nu_2$$

On voit que $\nu_1, \nu_2 \ll \mu$. Si μ était ergodique, alors $\nu_1 = \nu_2 = \mu$ d'après le lemme 5.2, une contradiction. Donc μ n'est pas ergodique. $\square$

5.7. PROPOSITION. *Soit (X, T, F) une dynamique topologique, métrique et compact. Alors $\mathbb{P}_{\mathrm{erg}}(F) \neq \emptyset$.*

---

1. Rétrospectivement on voit que $m(\phi) = m(\psi)$ si $\phi \overset{\mu}{\equiv} \psi$ et donc $m(\phi)$ est bien définie pour $\phi \in L^1(\mu)$.

DÉMONSTRATION. On fixe $\phi_1, \phi_2, \dots$ une suite dense dans $C(X)$.

On définit par récurrence $K_0 := \mathbb{P}(F)$ et

$$K_{n+1} := \{\mu \in K_n : \mu(\phi_n) = \sup\{\nu(\phi_n) : \phi \in K\}$$

On note que K_{n+1} est une partie compacte et non vide K_n. Donc

$$\mathcal{K} := \bigcap_{n \geq 1} K_n \neq \emptyset.$$

Notons que $\mathcal{K}$ contient exactement une mesure qu'on dénote par μ. En effet, si $\mu' \in \mathcal{K}$, on obtient $\mu' \in K_n$ et :

$$\mu'(\phi_{n+1}) = \sup\{\nu(\phi_{n+1}) : \nu \in K_n\} = \mu(\phi_n) \text{ et } \mu' \in K_{n+1}.$$

Comme la suite $\phi_1, \phi_2, \dots$ est dense dans $C(X)$, on obtient : $\mu' = \mu$. Le compact non-vide $\mathcal{K}$ est donc réduit à un élément unique μ.

Montrons par contradiction que μ est un point extrémal de $\mathbb{P}(F)$. Sinon

$$\mu = t \cdot \nu_1 + (1-t) \cdot \nu_2 \text{ avec } 0 < t < 1 \text{ et } \nu_1, \nu_2 \in \mathbb{P}(F).$$

Par une récurrence facile, on en déduit que, pour tout $n \geq 0$, $\nu_1, \nu_2 \in K_n$ et donc :

$$\nu_1(\phi_{n+1}) = \nu(\phi_{n+1}) = \mu(\phi_{n+1}) \text{ donc } \nu_1, \nu_2 \in K_{n+1}.$$

Donc $\nu_1, \nu_2 \in K_{n+1}$. Mais ceci implique $\nu_1 = \nu_2$. □

5.8. EXERCICE. Montrer que si F n'est pas uniquement ergodique, alors il existe au moins deux mesures de probabilité invariantes et ergodiques. *Indication : on pourra appliquer la construction précédente en choisissant bien la fonction ϕ_1.*

6. Propriétés de mélange

6.1. DÉFINITION. Une dynamique probabiliste $(X, \mathcal{B}, \mu, T, F)$ est *fortement mélangeante* si pour tous $A, B \in \mathcal{B}$,

$$\lim_{t\to\infty} \mu(A \cap F^{-t}(B)) = \mu(A)\mu(B). \tag{14}$$

6.2. REMARQUE. Si $\mu \in \mathbb{P}(F)$ est fortement mélangeante et de support total alors F est topologiquement mélangeante.

6.3. EXERCICE. Montrer qu'une rotation sur le cercle $\mathbb{T}$ bien qu'ergodique n'est jamais fortement mélangeante.

Ce mélange correspond à une propriété d'indépendance asymptotique :

6.4. PROPOSITION. *Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste. Elle est fortement mélangeante si et seulement si :*

$$\forall \phi, \psi \in L^2(\mu) \quad \lim_{t\to\infty} \int \phi \cdot \psi \circ F^t \, d\mu = \left(\int \phi \, d\mu\right)\left(\int \psi \, d\mu\right). \tag{15}$$

L'équivalence subsiste si on remplace $\phi, \psi \in L^2(\mu)$ par $\phi, \psi \in \mathcal{D}$ où $\mathcal{D}$ est une partie engendrant un sous-espace linéaire dense de $L^2(\mu)$.

DÉMONSTRATION. La condition (15) implique la définition (14) en prenant $\phi = 1_A$, $\psi = 1_B$. Etablissons la réciproque. On suppose le mélange (14). Si ϕ, ψ sont des fonctions simples, ie : $f = \sum_i c_i \cdot 1_{A_i}$ et $g = \sum_j d_j \cdot 1_{B_j}$, (15) est claire.

Fixons $\phi, \psi \in L^2(\mu)$. Il existe des nombres c_i, d_j et des parties mesurables tels que, pour $f = \sum_i c_i \cdot 1_{A_i}$ et $g = \sum_j d_j \cdot 1_{B_j}$, on a :

$$\|\phi - f\|_{L^2} < \epsilon,\ \mu(\phi) = \mu(f),\ \|\psi - g\|_{L^2} < \epsilon \text{ et } \mu(\psi) = \mu(g).$$

On calcule :

$$\begin{aligned}\int \phi \cdot \psi \circ F^t \, d\mu &= \int (f + (\phi - f)) \cdot (g + (\psi - g)) \, d\mu \\ &= \sum_{i,j} c_i d_j \mu(A_i \cap F^{-t} B_j) \\ &\quad + \int (\phi - f) \cdot \psi \circ F^t \, d\mu + \int \phi \cdot (\psi - g) \circ F^t \, d\mu \\ &\quad + \int (\phi - f) \cdot (\psi - g) \, d\mu\end{aligned} \tag{16}$$

Par définition du mélange, le premier terme converge vers :

$$\begin{aligned}\sum_{i,j} c_i d_j \mu(A_i)\mu(B_j) &= \left(\sum_i c_i \mu(A_i)\right)\left(\sum_j d_j \mu(B_j)\right) \\ &= \left(\int_X f \, d\mu\right)\left(\int_X g \, d\mu\right) = \left(\int_X \phi \, d\mu\right)\left(\int_X \psi \, d\mu\right)\end{aligned}$$

Les trois derniers termes de (16) peuvent être lus comme des produits scalaires (ou hermitien) donc bornés par le produit des normes :

$$\begin{aligned}&\|\phi - f\|_{L^2}\|\psi\|_{L^2} + \|\psi - g\|_{L^2}\|\phi\|_{L^2} + \|\phi - f\|_{L^2} \cdot \|\psi - g\|_{L^2} \\ &\qquad \leq \epsilon(\|\phi\|_{L^2} + \|\psi\|_{L^2} + 2\epsilon) + \epsilon^2.\end{aligned}$$

Donc, pour tout $\epsilon > 0$,

$$\limsup_{t \to \infty} \int \phi \cdot \psi \circ F^t \, d\mu \leq \mu(\phi)\mu(\psi) + \epsilon(\|\phi\|_{L^2} + \|\psi\|_{L^2} + 2\epsilon) + \epsilon^2,$$

soit

$$\limsup_{t \to \infty} \int \phi \cdot \psi \circ F^t \, d\mu \leq \mu(\phi)\mu(\psi)$$

De même pour la limite inférieure et donc la limite.

Si l'équation (15) est vérifiée pour $\phi, \psi \in \mathcal{D}$ qui engendre un sous-espace vectoriel dense de $L^2(\mu)$, alors le même calcul que ci-dessus permet d'étendre (15) à $L^2(\mu)$. □

6.5. Proposition. *Le mélange fort implique l'ergodicité.*

Démonstration. Soit $(X, \mathcal{B}, \mu, T, F)$ est fortement mélangeante. Soit $A \in \mathcal{B}$ invariant modulo μ. On a :

$$\mu(A)^2 = \lim_{t \to \infty} \mu(A \cap F^{-t}(A)) = \mu(A).$$

On a donc : $\mu(A) = 0$ ou 1 : A est trivial modulo μ, F est ergodique. □

6.6. Exercice. Soit $(\Sigma_d, \mathbb{Z}, \sigma_d)$ le décalage bilatère sur $d \geq 2$ symboles. Soit $\Pi_d := \{p \in \mathbb{R}_+ : p_1 + \cdots + p_d = 1\}$ l'ensemble des vecteurs de probabilités de dimension d. Soit $m_P := P^{\mathbb{Z}}$ la mesure de Bernoulli correspondante.

Montrer que (σ_d, m_P) est mélangeante donc ergodique.

En déduire la loi forte des grands nombres pour les variables aléatoires prenant un nombre fini de valeurs.

Chapitre 6

Tribu des invariants, énoncé général du thorème de Birkhoff

Dans le chapitre précédent nous avons étudié le cas particulier mais fondamental des dynamiques probabilistes ergodiques, c'est-à-dire indécomposables. En particulier, la limite des moyennes de Birkhoff ne pouvait être que constante presque partout, ce qui a facilité son analyse.

Une approche "naïve" serait de décomposer une dynamique arbitraire en dynamiques indécomposables et appliquer le cas ergodique du théorème à chacune de celles-ci. La difficulté ici est que la décomposition n'est pas nécessairement finie ou même dénombrable. Pour contourner cette difficulté, nous devrons introduire :

- la tribu des invariant qui formalise d'une certaine façon le "découpage" d'un système quelconque en "sous-systèmes indécomposables" ;
- l'espérance conditionnelle qui remplace une fonction intégrable par sa moyenne sur chaque "atome de la tribu".

1. Deux exemples simples

1.1. Exercice. Soient $p_1, p_2, \dots$ des réels strictement positifs et de somme 1. Soit $X := \bigsqcup_{n\geq 1}(n, \mathbb{T}^2)$ muni de la mesure

$$\mu(A) = \sum_{n\geq 1} p_n m_{\mathbb{T}^2}(\pi(A \cap X_n)) \text{ avec } \pi(n,x) = x$$

On définit $T : X \to X$ par $T(n,x) = (n, A^n x)$ pour tout $n \geq 1$ avec $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Montrer que pour μ-p.t. $(n,x) \in X$,

$$\lim_{t\to\infty} \frac{1}{t} S_t^T \phi(n,x) = \int_{\mathbb{T}^2} \phi(n,\cdot)\, dm_{\mathbb{T}}.$$

1.2. Exercice. Soient $X = [0,1] \times \mathbb{T}$ muni de la mesure borélienne produit des mesures uniformes sur $[0,1]$ et sur $\mathbb{T}$ et $F(x,y) = (x, y+\alpha)$ avec $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ fixé.

Montrer que pour μ-p.t. $x \in X$,

$$\lim_{t\to\infty} \frac{1}{t} S_t^F \phi(x,y) = \int_{\mathbb{T}} \phi(x,\cdot)\, dm_{\mathbb{T}}.$$

2. Tribu des invariants

Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste en temps discret $T = \mathbb{Z}$ ou $\mathbb{N}$.

2.1. Définition. On définit la *tribu des invariants* de F comme l'ensemble des boréliens invariants modulo μ :

$$\mathcal{I}(F,\mu) := \{Y \in \mathcal{B} : \forall t \in T\ \mu(F^{-t}(Y)\Delta Y) = 0\}$$

(On omettre μ ou même F et μ lorsque ça ne crée pas de risque de confusion.)

2.2. EXERCICE. Vérifier que $\mathcal{I}(F)$ est bien une tribu.

2.3. EXERCICE. Montrer que dans l'exercice 1.1, $\mathcal{I}(S)$ est la tribu engendrée modulo par les parties $\{n\} \times \mathbb{T}^2$, i.e.,

$$\mathcal{I}(S) = \{B \in \mathcal{B} : \exists I \subset \mathbb{N}^* \; B \overset{\mu}{\equiv} I \times \mathbb{T}^2\}.$$

2.4. EXERCICE. Montrer que dans l'exercice 1.2, $\mathcal{I}(T)$ est la tribu

$$\mathcal{I}(T) = \{B \in \mathcal{B} : \exists C \in \mathcal{B}_{[0,1]}; B \overset{\mu}{\equiv} C \times \mathbb{T}^2\}.$$

2.5. LEMME. *Une fonction $\phi \in L^0(\mu, \mathbb{K})$ est invariante modulo μ si et seulement si elle $\mathcal{I}(F)$-mesurable.*

DÉMONSTRATION. Il suffit de démontrer le lemme dans le cas $\mathbb{K} = \mathbb{R}$. Soit $\phi \in L^0(\mu, \mathbb{R})$. On a vu que ϕ était invariante modulo μ si et seulement si les parties $\phi^{-1}((-\infty, c])$ sont invariantes modulo μ. Mais cette dernière assertion est équivalente à la mesurabilité de ϕ par rapport à $\mathcal{I}(F)$ $\square$

Le fait suivant est maintenant clair :

2.6. LEMME. *(X, μ, F) est ergodique si et seulement si $\mathcal{I}(F)$ est la tribu triviale modulo μ, i.e., $A \in \mathcal{I}(F) \implies \mu(A) = 0$ ou 1.*

3. Espérance conditionnelle

3.1. DÉFINITION. Etant donnée une sous-tribu $\mathcal{C} \subset \mathcal{B}$, l'*espérance conditionnelle* de $\phi \in L^1(\mu)$ est une fonction $\bar{\phi} : X \to \mathbb{R}$ vérifiant : (i) $\bar{\phi}$ est $\mathcal{C}$-mesurable ; (ii) pour tout $C \in \mathcal{B}$,

$$\int_C \bar{\phi} \, d\mu = \int_C \phi \, d\mu$$

3.2. THÉORÈME. *Si $\phi \in L^1(\mu, \mathbb{K})$ et $\mathcal{B}$ est une sous-tribu de $\mathcal{B}$, alors l'espérance conditionnelle $\mathbb{E}_\mu(\phi|\mathcal{C}) \in L^1(\mu, \mathbb{K})$ existe et est unique (en tant que classe de fonction définie à égalité modulo μ).*

On peut déduire le théorème précédent de l'énoncé suivant. Rappelons que deux mesures μ, ν sont mutuellement singulières s'il existe un mesurable A tel que $\mu(A) = 0$ et $\nu(X \setminus A) = 0$.

3.3. THÉORÈME (Radon-Nikodym). *Soit μ, ν deux mesures boréliennes positives et σ-finies. Il existe ρ une mesure de sous-probabilité borélienne singulière par rapport à μ et $h \in L^1(\mu, \mathbb{R}_+)$ telle que*

$$\nu = \rho + h \cdot \mu \quad \textit{avec } (h \cdot \mu)(A) = \int_A h \, d\mu.$$

De plus, ρ et h sont uniques.

4. Théorème ergodique ponctuel - Cas général en temps discret

4.1. THÉORÈME (convergence presque partout). *Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste en temps discret ($T = \mathbb{N}, \mathbb{Z}$). Soit $\phi \in L^1(\mu, \mathbb{K})$. Alors*

$$\lim_{t \to \infty} \frac{1}{t} S_t^F \phi = \mathbb{E}_\mu(\phi|\mathcal{I}(F, \mu))$$

μ-presque partout.

Démonstration. Comme dans le cas ergodique, le point-clé est la convergence presque partout.

Soit $\phi \in L^1(\mu)$. On a $\mathbb{E}_\mu(\phi|\mathcal{I}) \in L^1(\mu)$. On peut donc poser :

$$\psi := \phi - \mathbb{E}_\mu(\phi|\mathcal{I}) \in L^1(\mu).$$

On a alors :

$$\mathbb{E}_\mu(\psi|\mathcal{I}) = \mathbb{E}_\mu(\phi - \mathbb{E}_\mu(\phi|\mathcal{I})|\mathcal{I}) = \mathbb{E}_\mu(\phi|\mathcal{I}) - \mathbb{E}_\mu(\mathbb{E}_\mu(\phi|\mathcal{I})|\mathcal{I}) = \mathbb{E}_\mu(\phi|\mathcal{I}) - \mathbb{E}_\mu(\phi|\mathcal{I}) = 0.$$

Il suffit de montrer la convergence presque partout de $\frac{1}{n} S_n^F \psi$ vers 0. On démontre d'abord :

$$\bar\phi(x) := \limsup_n \frac{1}{n} S_n \phi(x) \leq 0 \ \mu\text{-pp.} \tag{17}$$

La limite sup ci-dessus définit une fonction $\bar\phi$ mesurable. Le lemme 4.3 implique que :

$$\lim_{n\to\infty} \left| \frac{1}{n} S_n^F \phi(Fx) - \frac{1}{n} S_n^F \phi(x) \right| = \lim_{n\to\infty} \frac{1}{n} |\phi(F^n x)| + \frac{1}{n} |\phi(x)| = 0.$$

C'est à dire que la fonction $\bar\phi$ est également invariante modulo μ. Supposons, par l'absurde, qu'il existe $\alpha > 0$ tel que $B := \{x \in X : \bar\phi(x) > \alpha\}$ soit de mesure non-nulle. Notons que B est invariant modulo μ.

On définit une fonction mesurable et finie pour μ-p.t. $x \in B$ en posant :

$$N : B \to \mathbb{N},\ x \longmapsto \min \left\{ n \geq 1 : \frac{1}{n} S_n \phi(x) > \alpha \right\}.$$

Par σ-additivité, il existe un entier M tel que

$$C := \{x \in B : N(x) > M\} \text{ vérifie } \int_C |\phi| + \alpha \, d\mu < \frac{\alpha}{3} \mu(B).$$

On va en déduire que les sommes de Birkhoff sont excessives par un découpage de l'intervalle entier $[\![0, n[\![:= \{0, 1, 2, \dots, n-1\}$. Pour μ-p.t. $x \in B$, on définit par récurrence b_j, $j \geq 0$, et a_j, $j \geq 1$ en posant $b_0 := 0$ et :

$$a_{j+1} := \inf\{k \geq b_j : F^k(x) \notin C\} \quad b_{j+1} := a_{j+1} + N(F^{a_{j+1}} x)$$

On pose $r := \min\{n \geq 0 : b_{n+1} \geq n\}$. Notons qu'on peut avoir $a_{r+1} \geq n$ ou non. Dans ce dernier cas, $[\![a_{r+1}, b_{r+1}[\![$ intersecte mais n'est pas contenu dans $[\![0, n[\![$ et on a alors : $[\![a_{r+1}, n[\![\subset [\![n - M, n[\![$.

On a la partition :

$$[\![0, n[\![= [\![0, a_1[\![\cup [\![a_1, b_1[\![\cup [\![b_1, a_2[\![\cup \cdots \cup [\![b_r, \min(a_{r+1}, n)[\![\cup [\![\min(a_{r+1}, n), n[\![$$

On a :

$$[\![0, n[\![\setminus \bigcup_{i=1}^r [\![a_i, b_i[\![\subset \{k \in [\![0, n[\![: F^k(x) \in C\} \cup [\![n - M, n[\![.$$

On en déduit :

$$\sum_{k \in \bigcup_{i=1}^r [\![a_i, b_i[\![} \phi(F^k x) \geq \alpha \sum_{i=1}^r |b_i - a_i| = \alpha \cdot n - S_n(\alpha \cdot 1_C)(x) - M\alpha.$$

Donc :

$$S_n \phi(x) \geq \alpha n - S_n((|\phi| + \alpha) 1_C)(x) - S_M(|\phi| + \alpha)(F^{n-M} x).$$

En intégrant par rapport à μ sur $B \in \mathcal{I}$ et en utilisant $\mathbb{E}_\mu(\phi|\mathcal{I}) = 0$,

$$\begin{aligned} 0 = n \int_B \mathbb{E}_\mu(\phi|\mathcal{I})\, d\mu = n \int_B \phi\, d\mu = \int_B S_n\phi\, d\mu \\ \geq \alpha n \cdot \mu(B) - n \int_B 1_C \cdot (|\phi| + \alpha)\, d\mu - M \int_B (|\phi| + \alpha) \end{aligned}$$

En divisant par $n \cdot \mu(B)$ et en utilisant le choix de M :

$$0 \geq \alpha - \frac{1}{3}\alpha - \frac{M}{n} \frac{\|\phi\|_{L^1} + \alpha}{\mu(B)}.$$

Pour n assez grand, le troisième et dernier terme est aussi majoré par $\alpha/3$. On obtient $\alpha \leq 0$, une contradiction.

L'inégalité (17) est établie. En l'appliquant à l'opposé de ϕ, on obtient

$$\liminf_n \frac{1}{n} S_n\phi \geq 0 \ \mu\text{-pp}$$

d'où la convergence μ-p.p. Le théorème est démontré. □

4.2. Théorème (convergence en moyenne). *Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste en temps discret ($T = \mathbb{N}, \mathbb{Z}$). Soit $\phi \in L^p(\mu, \mathbb{K})$ avec $1 \leq p < \infty$. Alors*

$$\lim_{t\to\infty} \frac{1}{t} S_t^F \phi = \mathbb{E}_\mu(\phi|\mathcal{I}(F, \mu)) \ \textit{dans}\ L^p(\mu, \mathbb{K}).$$

Démonstration. Fixons $1 \leq p < \infty$ et $\phi \in L^p(\mu)$. On note que

$$\|\mathbb{E}_\mu(\phi|\mathcal{I})\|_{L^p(\mu)} \leq \|\phi\|_{L^p(\mu)}.$$

Remarquons que μ étant une probabilité, $L^p(\mu) \subset L^1(\mu)$. On peut donc appliquer le théorème précédent et obtenir la convergence μ-presque partout.

On se place d'abord dans le cas particulier où ϕ est borné : $\phi \in L^\infty(\mu)$ de sorte que $\mathbb{E}_\mu(\phi|\mathcal{I}) \in L^\infty(\mu)$ aussi. Pour tout $n \geq 1$,

$$\left| \frac{1}{n} S_n\phi - \mathbb{E}_\mu(\phi|\mathcal{I}) \right|^p \leq 2^p \|\phi\|_{L^\infty}^p \cdot \mathbf{1} \in L^1(\mu).$$

Le théorème de convergence dominée s'applique :

$$\frac{1}{n} S_n\phi \xrightarrow{n\to\infty} \mathbb{E}_\mu(\phi|\mathcal{I}) \text{ dans } L^p(\mu).$$

Le résultat est démontré dans le cas borné.

On considère maintenant $\phi \in L^p(\mu)$. On suppose ϕ à valeurs dans $\mathbb{R}_+$, le cas général s'y réduisant sans difficulté. On se donne $\epsilon > 0$.

On pose :

$$\phi_M(x) := \min(\phi(x), M).$$

Vu $0 \leq (\phi - \phi_M)^p \leq \phi^p \in L^1(\mu)$, le théorème de convergence dominée donne M_0 tel que :

$$\|\mathbb{E}_\mu(\phi|\mathcal{I}) - \mathbb{E}_\mu(\phi_{M_0}|\mathcal{I})\|_{L^p} \leq \|\phi_{M_0} - \phi\|_{L^p} < \epsilon/3.$$

Le cas particulier des fonctions bornées traité ci-dessus s'applique à ϕ_{M_0} et donne n_0 tel que :

$$\forall n \geq n_0 \quad \left\| \frac{1}{n} S_n\phi_{M_0} - \mathbb{E}_\mu(\phi_{M_0}|\mathcal{I}) \right\|_{L^p} < \epsilon/3.$$

L'inégalité triangulaire donne

$$\forall n \geq 1 \quad \left\| \frac{1}{n} S_n \phi - \frac{1}{n} S_n \phi_{M_0} \right\|_{L^p} \leq \|\phi - \phi_{M_0}\|_{L^p} .$$

On en déduit que pour tout $n \geq n_0$:

$$\begin{aligned} \left\| \frac{1}{n} S_n \phi - \mathbb{E}_\mu(\phi|\mathcal{I}) \right\|_{L^p} \leq & \left\| \frac{1}{n} S_n \phi - \frac{1}{n} S_n \phi_{M_0} \right\|_{L^p} \\ & + \left\| \frac{1}{n} S_n \phi_{M_0} - \mathbb{E}_\mu(\phi_{M_0}|\mathcal{I}) \right\|_{L^p} \\ & + \|\mathbb{E}_\mu(\phi|\mathcal{I}) - \mathbb{E}_\mu(\phi_{M_0}|\mathcal{I})\|_{L^p} < \epsilon \end{aligned}$$

□

5. Temps continu

Dans cette section, $(X, \mathcal{B}, \mu, T, F)$ est une dynamique probabiliste en temps continu $T = \mathbb{R}$ ou $\mathbb{R}_+$. Pour essayer de se ramener au temps discret, on considère pour chaque $\tau > 0$, le temps τ du flot, c'est-à-dire l'itération de $F^\tau : X \to X$.[1] Il découle immédiatement des définitions que :

— F^τ préserve la mesure μ ;

— un ensemble est F-invariant modulo μ si et seulement s'il est F^τ-invariant pour tout $\tau > 0$.

5.1. Ergodicité en temps continu. On voit immédiatement que l'ergodicité d'*un seul* temps τ d'un flot implique l'ergodicité du flot tout entier. Mais l'exemple suivant montre qu'il n'est pas vrai que l'ergodicité du flot implique celle de tous ses temps.

5.1. Exemple. *Le flot linéaire $\tau : \mathbb{R} \times \mathbb{T} \to \mathbb{T}$, $(t, x) \mapsto x + t$:*

— *préserve la mesure uniforme $m_{\mathbb{T}}$;*

— *est ergodique par rapport à $m_{\mathbb{T}}$;*

— *$\{t \in \mathbb{R} : \tau^t$ n'est pas ergodique$\} = \mathbb{Q}$.*

5.2. Remarque. On peut montrer que cependant, comme dans l'exemple précédent, toute dynamique probabiliste standard en temps $\mathbb{R}$, si elle est ergodique, admet des temps ergodiques.

5.3. Exercice. Soient $\alpha \in \mathbb{R}^d$, $d \geq 1$, et $(\mathbb{T}^d, \mathcal{B}, m_{\mathbb{T}^d}, \mathbb{R}, \Phi)$ la dynamique probabiliste en temps $T = \mathbb{R}$ définie par l'action $\Phi^t = \tau_{t\alpha}$ où $\tau_\beta : \mathbb{T}^d \to \mathbb{T}^d$ est la translation de vecteur $\beta \in \mathbb{R}^d$.

Montrer que $(\Phi, m_{\mathbb{T}^d})$ est ergodique si et seulement si $(\alpha_1, \ldots, \alpha_d)$ est une famille $\mathbb{Q}$-libre.

Vérifier que l'ergodicité de $(\tau_\alpha, m_{\mathbb{T}^d})$ implique celle de $(\Phi, m_{\mathbb{T}^d})$. Montrer par un exemple que la réciproque n'a pas lieu.

1. Le temps τ du flot n'est pas une section du flot.

5.2. Théorème de Birkhoff en temps continu.

5.4. Théorème. *Soit $(X, \mathcal{B}, \mu, T, F)$ une dynamique probabiliste en temps continu $T = \mathbb{R}$ ou $\mathbb{R}_+$. Soit $\phi \in L^1(\mu)$. Alors*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \phi(F^s(x))\,ds \to \mathbb{E}_\mu(\phi|\mathcal{I}) \ \mu\text{-p.p. et dans } L^1(\mu).$$

De plus, si $\phi \in L^p(\mu)$ avec $1 \le p < \infty$ alors la convergence a lieu également dans $L^p(\mu)$.

Démonstration. On va appliquer le théorème de Birkhoff en temps discret. Plus précisément on considère la dynamique $f := F^1$ qui préserve évidemment la même mesure μ et la fonction $\psi : X \to \mathbb{R}$, $\psi(x) = \int_0^1 \phi(F^s x)\,ds$, mesurable vu le théorème sur les intégrales à paramètres.

Montrons que $\psi \in L^p(\mu)$ avec $1 \le p < \infty$. En utilisant l'inégalité de Jensen, puis le théorème de Fubini-Tonelli et enfin l'invariance de μ par F on obtient :

$$\begin{aligned}\int_X |\psi(x)|^p\,d\mu &= \int_X \left|\int_0^1 \phi(F^t x)\,dt\right|^p d\mu(x) \le \int_X\int_0^1 |\phi(F^t x)|^p\,dt\,d\mu(x)\\ &\le \int_0^1 \left(\int_X |\phi(F^t x)|^p\,d\mu(x)\right) dt = \int_X |\phi|^p\,d\mu < \infty.\end{aligned}$$

Le même calcul justifie que $\hat\psi(x) := \int_0^1 |\phi(F^s x)|\,ds$ est intégrable. En particulier, on a :

$$\lim_{t\to\infty}\frac{1}{t}\int_{E(t)}^t |\phi(F^s x)|\,ds = 0 \ \mu\text{-p.p. et dans } L^p(\mu), \tag{18}$$

(la convergence dans $L^p(\mu)$ est évidente car $\| \int_{E(t)}^t |\phi(F^s x)|\,ds\|_{L^p} \le \|\hat\psi\|_{L^p}$ et la convergence p.p. découle du lemme 4.3).

Le théorème de Birkhoff appliqué à (f, μ) et $\psi \in L^p(\mu)$, donne la convergence μ-p.p. et dans $L^p(\mu)$:

$$\lim_{n\to\infty}\frac{1}{n}\int_0^n \phi(F^s x)\,ds = \mathbb{E}_\mu(\psi|\mathcal{I}(f)).$$

L'eq. (18) montre que ceci implique les mêmes convergences pour :

$$\bar\phi(x) := \lim_{t\to\infty}\frac{1}{t}\int_0^t \phi(F^s x)\,ds = \mathbb{E}_\mu(\psi|\mathcal{I}(f)).$$

Il reste à vérifier que cette limite coïncide avec $\mathbb{E}_\mu(\phi|\mathcal{I}(F))$. Tout $B \in \mathcal{I}(F)$ appartient à $\mathcal{I}(f)$ et B est invariant par le flot donc :

$$\int_B \bar\phi\,d\mu = \int_B\int_0^1 \phi(F^s x)\,ds\,d\mu = \int_B \phi\,d\mu.$$

Pour vérifier que $\bar\phi$ est $\mathcal{I}(F)$-mesurable il suffit de vérifier son invariance par le flot. Pour tout $u \in \mathbb{R}$,

$$\bar\phi(F^u x) = \lim_{t\to\infty}\frac{1}{t}\int_0^t \phi(F^{s+u}x)\,ds = \lim_{t\to\infty}\frac{t+u}{t}\frac{1}{t+u}\int_0^{t+u}\phi(F^s x)\,ds = \bar\phi(x).$$

On a donc : $\bar\phi = \mathbb{E}_\mu(\phi|\mathcal{I}(F))$ modulo μ. □

6. Encore quelques applications du théorème de Birkhoff

6.1. Proposition. *Soit (X, T, F) une dynamique métrique compacte.*

Elle est uniquement ergodique si et seulement si, pour toute fonction $\phi \in C(X)$ et tous $x, y \in X$:

$$\lim_{n\to\infty} \frac{1}{n}|S_n\phi(x) - S_n\phi(y)| = 0. \tag{19}$$

6.2. Exercice. Soit (X, T, F) une dynamique topologique métrique compacte. Si tous les points de X sont génériques pour une même mesure $\mu \in \mathbb{P}(F)$ alors F est uniquement ergodique.

Démonstration. Le théorème ergodique uniforme montre que l'unique ergodicité implique (19). Réciproquement, le théorème de Birkhoff implique que si $\mu, \nu \in \mathbb{P}(F)$ et $\phi \in C(X)$, alors

$$\mu(\phi) = \int_X \lim_n \frac{1}{n} S_n^F \phi \, d\mu = \int_X \lim_n \frac{1}{n} S_n^F \phi \, d\mu = \nu(\phi)$$

la deuxième égalité découlant de l'hypothèse (19). Donc $\mu = \nu$. □

6.3. Théorème. *Soit (Y, T, G) une dynamique compacte métrique uniquement ergodique avec $\{\nu\} = \mathbb{P}(G)$.*

Soit (X, T, F) une extension selon $\alpha : Y \to K$ continue à valeurs dans un groupe métrique compact. Soit m_K la mesure de Haar normalisée sur K.

Les propriétés suivantes sont équivalentes :

(1) la mesure $\nu \times m_K$ est ergodique ;

(2) F est uniquement ergodique.

Démonstration. On considère les deux implications.

(2) $\Longrightarrow$ (1) : l'extension F préserve la mesure produit $\mu \times m_K$. Si F est uniquement ergodique, alors cette mesure produit est l'unique mesure de probabilité invariante et elle est donc ergodique.

(1) $\Longrightarrow$ (2) : On considère l'ensemble $\mathfrak{G}$ des points génériques pour $\nu \times m_K$. Òn a vu que $\nu \times m_K(\mathfrak{G}) = 1$. Pour chaque $g \in K$, on considère l'autoconjugaison $h : X \times K \to X \times K$, $(x, k) \mapsto (x, kg)$. Comme h est continue, x générique pour $\nu \times m_K$ implique que $h(x)$ est générique pour $h_*(\mu \times m_K) = \mu \times m_k$: $\mathfrak{G}$ est donc invariant par toute translation à droite. Il s'ensuit que $\mathfrak{G} = \mathfrak{H} \times K$ où $\mathfrak{H} = \pi(\mathfrak{G})$ vérifie $\nu(\mathfrak{H}) = 1$.

Prenons maintenant $\mu_0 \in \mathbb{P}(F)$ et considérons l'ensemble de ses points génériques $\mathfrak{G}_0$. La projection $\pi_*(\mu_0)$ est une mesure de probabilité G-invariante c'est donc ν. En particulier, $\nu(\pi(\mathfrak{G}_0)) = 1$. Il existe donc $y_0 \in \pi(\mathfrak{G}_0)) \cap \pi(\mathfrak{G})$. Mais ceci implique que $\mathfrak{G}_0$ intersecte $\{y_0\} \times K \subset \mathfrak{G}$. Donc μ_0 et $\nu \times m_K$ partage un point générique ; elles sont donc égales. □

**** application ****

Chapitre 7

Théorème ergodique sous-additif de Kingman

Ce chapitre introduit une généralisation du théorème de Birkhoff précieuse pour étudier l'itération d'un difféomorphisme $f : M \to M$ préservant une mesure de probabilité μ sur M. On aimerait alors comprendre l'asymptotique de sa différentielle $Df : TM \to TM$ et l'existence ainsi que les valeurs de limites du type :

$$\lim_{n\to\infty} \frac{1}{n} \log \|Df^n.v\|$$

pour μ-p.t. $x \in M$ et $v \in T_xM \setminus O$. Une réponse détaillée est fournie par le théorème ergodique multiplicatif d'Oseledets. Ce chapitre est consacré au théorème de Kingman, un résultat plus simple mais déjà non trivial dans cette direction.

1. Définitions et exemples

Soit $(X, \mathcal{B}, \mu, \mathbb{N}, F)$ une dynamique probabiliste en temps discret.

1.1. Définition. Un *processus sous-additif* pour la dynamique probabiliste $(X, \mathcal{B}, \mu, \mathbb{N}, f)$ est une suite ϕ de fonctions $\phi_n : X \to \mathbb{R}$ telles que, pour μ-p.t. $x \in X$,

$$\forall m, n \geq 1 \quad \phi_{m+n}(x) \leq \phi_m(x) + \phi_n(F^m x).$$

La *constante* de ce processus est :

$$\gamma(\phi) := \inf_{n\geq 1} \frac{1}{n} \int_X \phi_n \, d\mu \in [-\infty, \infty)$$

Le processus est respectivement*intégrable, négatif ou uniformément borné* si on a pour tout $n \geq 1$ respectivement : $\phi_n \in L^1(\mu)$, $\phi_n \leq 0$ ou $\|\phi_n/n\|_{L^\infty} \leq C$ pour un certain $C < \infty$.

1.2. Remarque. Un processus est additif si l'inégalité précédente est une égalité. Il est alors de la forme $(S_n^f \phi_1)_{n\geq 1}$: c'est le cas traité par le théorème de Birkhoff.

1.3. Remarque. La somme $\phi + \psi$ de deux processus sous-additifs est encore un processus sous-additif. On verra que : $\gamma(\phi + \psi) = \gamma(\phi) + \gamma(\psi)$.

1.4. Remarque. La constante d'un processus sous-additif négatif ou bornée par une constante C est elle-même négative ou bornée par cette même constante. Mais, même si le processus est intégrable, la constante n'est pas nécessairement finie. Il suffit de considérer la suite de fonctions constantes $\phi_n := -n^2$.

Si le processus est additif avec $\phi_1 \in L^1(\mu)$, la constante est simplement $\gamma(\phi) = \mu(\phi)$.

Donnons deux exemples définis à partir d'un cocycle linéaire.

1.5. Exemple. *Soit M une variété riemannienne de dimension $d \geq 1$*[1]*. Soit $(M, \mathcal{B}_M, \mu, \mathbb{N}, f)$ une dynamique différentiable préservant une mesure borélienne $\mu \in \mathbb{P}(M)$. On pose, pour tout $x \in \mathbb{R}^d$,*

$$\phi_n(x) := \log \|Df^n_x\|_x = \log \|Df_{f^{n-1}x} Df_{f^{n-2}x} \dots Df_x\|_x.$$

où $\|A\|_x := \sup\{\|A.v\|_{T_{f(x)}M} : v \in T_xM, \|v\|_{T_xM} = 1\}$ est la norme d'opérateur induite par la norme sur les espaces tangents $T_xM, T_{f(x)}M$ définie par la structure riemannienne[2]

Les formules $D(f \circ g) = D_{g(x)}f \circ Dg_x$ et $\|B \circ A\|_x \leq \|B\|_{f(x)}\|A\|_x$ montrent que que cela définit un processus sous-additif. Ce processus peut-être intégrable ou non.

1.6. Exemple. *Soient $(\Sigma_d, \mathcal{B}, \mu, \mathbb{Z}, \sigma_d)$ une dynamique probabiliste sur le décalage à d symboles et $A : \mathcal{A}_d \to GL_d(\mathbb{R})$. Le cocycle :*

$$A^n(x) := A(x_{n-1})A(x_{n-2}) \dots A(x_0)$$

définit un processus sous-additif en posant : $\phi_n(x) := \log \|A^n(x)\|$ définit un processus sous-additif borné donc intégrable. Si μ est une mesure de Bernoulli (i.e., $\mu = P^{\otimes \mathbb{Z}}$ où P est une probabilité sur $\mathcal{A}_d$), on parle de "produit aléatoire de matrices". On peut alors voir $A^n(x)$ comme une marche aléatoire sur le groupe de matrices engendré par $\{A(i) : i \in \mathcal{A}_d\}$.

2. Théorème ergodic de Kingman

Nous allons démontrer la version suivante du théorème de Kingman :

2.1. Théorème. *Soit $(X, \mathcal{B}, \mu, \mathbb{N}, f)$ une dynamique probabiliste et ergodique. Soit $\{\phi_n\}_{n\geq 1}$ un processus sous-additif intégrable de constante supposée finie :*

$$\gamma(\phi) := \inf_{n\geq 1} \frac{1}{n} \int_X \phi_n \, d\mu \neq -\infty.$$

Alors pour μ-p.t. $x \in X$, on a la convergence μ-presque partout vers la constante du processus :

$$\lim_{n\to\infty} \frac{1}{n} \phi_n(x) = \gamma(\phi).$$

2.2. Remarque. Ce théorème admet des versions non-ergodiques, en topologie L^p, semi-intégrables ($\max(\phi(x), 0) \in L^1(\mu)$), lorsque $\gamma(\phi) = -\infty$, ...

2.3. Corollaire (Lemme de Fekete). *Soit $a_1, a_2, \dots$ une suite numérique. Si elle est sous-additive, alors la limite $\lim_n a_n/n$ existe dans $[-\infty, \infty[$ et vaut* $\inf_{n\geq 1} a_n/n$

En particulier,

$$\gamma(\phi + \psi) = \lim_n \frac{1}{n} \int_X \phi_n + \psi_n, d\mu = \gamma(\phi) + \gamma(\psi).$$

Appliqué à la suite $a_n := \mu(\phi_n)$, cela donne :

$$\inf_{n\geq 1} \frac{1}{n} \int_X \phi_n \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int_X \phi_n \, d\mu$$

et donne une autre expression de la limite obtenue dans le théorème de Kingman.

1. On pourra se contenter des cas $M = \mathbb{R}^d$ ou $M = \mathbb{T}^d$ dont les espaces tangents T_xM s'identifient à $\mathbb{R}^d$.

2. Si $M = \mathbb{R}^d$ ou $\mathbb{T}^d$ alors $\|\cdot\|_{T_yM}$ la norme euclidienne de $\mathbb{R}^d$.

Démonstration. Considérons la dynamique triviale (mesure de Dirac en un point fixe) et le processus sous-additif $\phi_n := a_n$. Si $\gamma(\phi) := \inf_{n\geq 1} a_n/n > \infty$ alors on peut appliquer le théorème de Kingman et obtenir

$$\lim_n \frac{a_n}{n} = \gamma(\phi)$$

Si $\gamma(\phi) = -\infty$.... □

2.4. Lemme. *Soit $(X, \mathcal{B}, \mu, \mathbb{N}, f)$ une dynamique probabiliste en temps discret.*

Si $\phi : X \to \mathbb{R}$ est une fonction mesurable telle que $\phi \leq \phi \circ f$ μ-p.p., alors $\phi = \phi \circ f$ μ-p.p.

Démonstration. Remarquons que pour tout $c \in \mathbb{R}$:

$$f^{-1}(\{\phi \leq c\}) = \{\phi \circ f \leq c\} \subset \{\phi \leq c\}$$

Vu l'invariance $\mu(f^{-1}(\{\phi \leq c\}) = \mu(\{\phi \leq c\})$. Les deux ensembles sont donc égaux modulo μ. Ceci étant vrai pour tout c, $\phi = \phi \circ f$. L'invariance de ϕ est démontrée. □

Démonstration. Le processus $\tilde{\phi}_n := \phi_n - S_n^f \phi_1$ est encore sous-additif et intégrable avec :

$$\frac{1}{n} \int \tilde{\phi}_n \, d\mu = \frac{1}{n} \int \phi_n \, d\mu - \int \phi_1 \, d\mu$$

La sous-additivité de ϕ donne : $\phi_n \leq S_n \phi_1$ donc $\tilde{\phi}_n \leq 0$ pour tout $n \geq 1$. Une application du théorème ergodique de Birkhoff donne, presque partout et dans L^1 :

$$\lim_n \frac{1}{n}(\phi_n - \tilde{\phi}_n) = \lim_n \frac{1}{n} S_n^f \phi_1 = \int_X \phi_1.$$

et donc $\gamma(\tilde{\phi}) = \gamma(\phi) - \mu(\phi_1)$.

On peut donc supposer le processus négatif. Un tel processus satisfait l'inégalité suivante. Si on a une collection disjointe de sous-intervalles entiers :

$$[\![j_1, j_1 + k_1 [\![\sqcup [\![j_2, j_2 + k_2 [\![\sqcup \cdots \sqcup [\![j_r, j_r + k_r [\![\subset [\![0, N [\![$$

alors

$$\phi_N \leq \phi_{k_1} \circ f^{j_1} + \phi_{k_2} \circ f^{j_2} + \cdots + \phi_{k_r} \circ f^{j_{k_r}}. \tag{20}$$

On le démontre par une application répétée de l'inégalité de sous-additivité et en notant qu'on peut toujours supprimer certains termes $\phi_k \circ f^j$ puisqu'ils sont négatifs et que l'équation (20) est une majoration.

Majoration. On va montrer : $\limsup_{n\to\infty} \frac{1}{n}\phi_n \leq \gamma$.

Fixons $m \geq 1$. On va majorer ϕ_N pour $N \to \infty$ par une somme de ϕ_m qu'on pourra interpréter comme une somme de Birkhoff. Pour ce faire on écrit, pour tout $n \geq 1$ et chaque $0 \leq k, q < m$:

$$\phi_{mn} \leq \sum_{j=0}^{n-2} \phi_m \circ f^{jm+k}$$

On somme ensuite sur $0 \leq k < m$ et on divise par $m^2 n$ pour obtenir :

$$\frac{1}{mn}\phi_{nm} \leq \frac{1}{mn} \sum_{k=0}^{m-1} \sum_{j=0}^{n-2} \phi_m \circ f^{jm+k} \leq \frac{n-1}{n} \cdot \frac{1}{m(n-1)} \sum_{i=0}^{m(n-1)-1} \frac{\phi_m \circ f^i}{m}.$$

Vu $\phi_m \in L^1(\mu)$ et l'ergodicité, le théorème de Birkhoff donne la convergence suivante μ-p.p. et dans $L^1(\mu)$:

$$\forall m \geq 1 \quad \limsup_{n\to\infty} \frac{1}{mn}\phi_{mn} \leq \frac{1}{m}\int \phi_m \, d\mu$$

Mais, le processus étant négatif et sous-additif, on a pour tout $N \geq 1$, en posant $n = E(N/m)$:

$$\frac{1}{N}\phi_N \leq \frac{1}{N}(\phi_{mn} + \phi_{N-mn}\circ f^{mn}) \leq \frac{mn}{N}\frac{1}{mn}\phi_{mn}$$

et donc

$$\limsup_{N\to\infty} \frac{1}{N}\phi_N \leq \frac{1}{m}\int \phi_m \, d\mu$$

En prenant la borne inférieure sur $m \geq 1$, on obtient la majoration annoncée de la limite μ-p.p.

Minoration. On va montrer : $\underline{\phi}(x) := \liminf_{n\to\infty} \frac{1}{n}\phi_n(x) \geq \gamma$ pour μ-p.t. $x \in X$. Si la constante $\gamma(\phi) = -\infty$, il n'y a rien à démontrer. Supposons désormais que cette constante est finie.

Montrons d'abord que $\underline{\phi} \leq \underline{\phi}\circ f$. La sous-additivité donne :

$$\frac{\phi_{n+1}(x)}{n+1} \leq \frac{1}{n+1}\phi_1(x) + \frac{n}{n+1}\frac{\phi_n(fx)}{n}$$

Il existe une suite d'entiers $n_i \to \infty$ telle que $\frac{\phi_{n_i}(fx)}{n_i} \to \underline{\phi}(fx)$ et que le membre de gauche ci-dessus converge. On a :

$$\underline{\phi}(x) \leq \lim_i \frac{\phi_{n_i+1}(x)}{n_i+1} \leq \lim_i \frac{\phi_{n_i}(fx)}{n_i} = \underline{\phi}(fx).$$

($\lim_n \phi_1(x)/(n+1) = 0$ car ϕ_1 est finie μ-p.p. étant intégrable).

Le lemme 2.4 montre que la fonction $\underline{\phi}$ étant sous-invariante, elle est invariante μ-p.p. L'ergodicité implique que $\underline{\phi}$ est constante modulo μ.

Pour montrer la minoration, on procède par l'absurde, ie, en la supposant en défaut :

$$\underline{\phi} < \gamma - \epsilon \ \mu\text{-p.p. avec } \epsilon > 0$$

(rappelons-nous que $\gamma(\phi) \in \mathbb{R}$).

On procède comme dans la preuve du théorème de Birkhoff donnée précédemment en définissant :

$$N(x) := \inf\{n \geq 1 : \phi_n(x) < (\gamma - \epsilon)n\}$$

Vu $\underline{\phi}(x) < \gamma - \epsilon$ μ-p.p., $N(x) < \infty$ μ-p.p. Soit $\delta > 0$. Par σ-additivité de μ, il existe $M \geq 1$ tel que

$$\mu(\{x : N(x) > M\}) < \delta. \tag{21}$$

On définit une suite d'intervalles $[\![a_i, b_i[\![$ disjoints et inclus dans $\mathbb{N}$ et vérifiant $\phi_k(f^j x) < (\gamma - \epsilon)n$ en posant, par récurrence :

$$b_0 := 0,\ a_i := \min\{j \geq b_{i-1} : N(f^j x) \leq M\} \text{ et } b_i := a_i + N(f^{a_i}x).$$

Soit $r := \sup\{i \geq 1 : b_i \leq n\}$, de sorte que $\bigsqcup_{i=1}^{r} [\![a_i, b_i[\![\subset [\![0, n[\![$. Le complémentaire est constitué d'itérés $j \geq 0$ tel que $N(f^j x) > M$ ou $j \geq n - M$. Donc, pour tout $n \geq 1$,

$$\sum_{i=1}^{r-1} |b_i - a_i| \geq S_n 1_{N \leq M}(x) - M.$$

En se souvenant que $\gamma \leq 0$,

$$\phi_n(x) \leq (\gamma - \epsilon)\,(S_n 1_{N \leq M}(x) - M)$$

et en divisant par $n \geq 1$ et en intégrant par rapport à μ :

$$\int_X \frac{\phi_n(x)}{n}\,d\mu \leq (\gamma - \epsilon)\left(\mu(\{N \leq M\}) - \frac{M}{n}\right) \leq (\gamma - \epsilon)(1 - \delta)$$

Mais en prenant l'infimum sur $n \geq 1$, on trouve $\gamma \leq (\gamma - \epsilon)(1 - \delta)$, une contradiction. $\square$

Chapitre 8

Difféomorphismes du cercle

Ce chapitre est consacré aux bases de la dynamique du cercle préservant l'orientation. Nous introduisons d'abord le nombre de rotation, un invariant défini par Poincaré. On montrera que tout homéomorphisme du cercle préservant l'orientation est une extension topologique d'une rotation. Puis nous démontrerons le théorème de Denjoy selon lequel tout difféomorphisme C^2, préservant l'orientation et de nombre de rotation irrationnel est topologiquement conjugué à une rotation. On donnera les énoncés les plus simples de la théorie d'Arnold-Herman-Yoccoz qui, sous des conditions arithmétiques supplémentaires, produit une conjugaison lisse.

1. Relèvements, linéarisés et degré topologique

On désigne par $(e_1, \ldots, e_d)$ la base canonique de $\mathbb{R}^d$: e_i est le vecteur dont la $i$ème vaut 1 et les autres 0.

1.1. Définition. On appelle *relèvement* d'une application continue $h : \mathbb{T}^d \to \mathbb{T}^d$ toute application continue $H : \mathbb{R}^d \to \mathbb{R}^d$ vérifiant $\pi \circ H = h \circ \pi$ où $\pi : \mathbb{R}^d \to \mathbb{T}^d$ est la projection canonique.

1.2. Lemme. *Toute application continue $h : \mathbb{T}^d \to \mathbb{T}^d$ avec $d \geq 1$ admet un relèvement $H : \mathbb{R}^d \to \mathbb{R}^d$. Si $G : \mathbb{R}^d \to \mathbb{R}^d$ est un autre relèvement alors $G - H$ est une fonction constante et à valeurs dans $\mathbb{Z}^d$.*

Réciproquement, toute fonction continue $H : \mathbb{R}^d \to \mathbb{R}^d$ vérifiant :

$$\forall n \in \mathbb{Z}^d \quad H(\cdot + n) = H + n \tag{22}$$

est le relèvement d'une et une seule fonction continue $h : \mathbb{T}^d \to \mathbb{T}^d$.

Démonstration. Notons d'abord que si F, G sont deux relèvements d'une même application, alors $F - G : \mathbb{R}^d \to \mathbb{R}^d$ est continue et à valeurs dans $\mathbb{Z}^d$ donc nécessairement constante.

Ensuite si $H : \mathbb{R}^d \to \mathbb{R}^d$ est une fonction continue vérifiant (22), on vérifie qu'on peut définir une fonction $h : \mathbb{T}^d \to \mathbb{T}^d$ en posant $h(x) = \pi \circ H(y)$ pour n'importe quel $y \in \pi^{-1}(x)$ puis que cette fonction est continue : si $x_0 = \pi(y_0)$, alors il existe deux voisinages U_0, V_0 de x_0, y_0 tels que la restriction $\pi : V_0 \to U_0$ est un homéomorphisme. On a alors,

$$h : U_0 \to \mathbb{T}^d, \ x \longmapsto \pi \circ H \circ (\pi|U_0)^{-1}(x)$$

En particulier, h est continue en $x_0 \in \mathbb{T}^d$ arbitraire.

Il reste à construire un relèvement. Si $d = 1$ et h de classe C^1, il suffit de choisir $H_0 \in \pi^{-1}(h(0))$ et de poser :

$$\forall t \in \mathbb{R} \ H(t) = H_0 + \int_0^t h'(s)\, ds.$$

On vérifie que H est de classe C^1, puis que $\pi \circ H(0) = h \circ \pi(0)$ et $\frac{d}{dt}(\pi \circ H)(t) = \frac{d}{dt}(h \circ \pi)(t)$ pour tout $t \in \mathbb{R}$.

On a construit un relèvement dans ce cas particulier. Le cas général utilise que $\pi : \mathbb{R}^d \to \mathbb{T}^d$ est un homéomorphisme local et est laissé au lecteur. $\square$

1.3. Définition. Soit $h : \mathbb{T}^d \to \mathbb{R}^d$ et H un relèvement.

L'application linéaire $L_h : \mathbb{R}^d \to \mathbb{R}^d$ définie par $L(e_i) = H(e_i) - H(0)$ pour $i = 1, \ldots, d$ est appelée *linéarisé* de h.

Le *degré topologique* de h est $\deg(h) := \det L_h$

1.4. Exercice. Démontrer les deux lemmes suivants.

1.5. Lemme. *Le linéarisé de $h : \mathbb{T}^d \to \mathbb{T}^d$ ne dépend pas du choix du relèvement H de h.*

1.6. Lemme. *Si $f, g : \mathbb{T}^d \to \mathbb{T}^d$ alors $L_{f \circ g} = L_f \circ L_g$. En particulier :*

(1) si f est un homéomorphisme alors $L_f \in GL_d(\mathbb{Z})$, i.e., est une matrice à coefficients entiers et à déterminant ± 1 ;

(2) si f, g sont conjugués par un homéomorphisme h, alors les linéarisés sont conjugués par une matrice $L_f = L_g L_h (L_g)^{-1}$ et ont même degré topologique.

On considère les homéomorphismes du type suivant. On peut vérifier que ce sont les homéomorphismes homotopes à l'identité (une notion valide en toute dimension).

1.7. Définition. Soit $\mathrm{homeo}_0(\mathbb{T})$ l'ensemble des *homéomorphismes préservant l'orientation* au sens que tout relèvement est strictement croissant.

On dit que $f, g \in \mathrm{homeo}_0(\mathbb{T})$ sont *conjugués dans* $\mathrm{homeo}_0(\mathbb{T})$ s'il existe $h \in \mathrm{homeo}_0(\mathbb{T})$ tel que $h \circ f = g \circ h$.

1.8. Exercice. Démontrer les deux lemmes suivants.

1.9. Lemme. *Pour tout homéomorphisme $h : \mathbb{T} \to \mathbb{T}$ et H n'importe lequel de ses relèvements, les propriétés suivantes sont équivalentes : $h \in \mathrm{homeo}_0(\mathbb{T})$; $H(\cdot + 1) = H + 1$.*

En particulier, $\sup_{x \in \mathbb{R}} |H(x) - x| < \infty$.

1.10. Lemme. *Si $g \circ h = h \circ f$ avec $f, g, h \in \mathrm{homeo}_0(\mathbb{T})$ et si F, G, H sont des relèvements de f, g, h, alors il existe un entier n tel que $(G + n) \circ H = H \circ F$ et $G \circ H = H \circ (F - n)$.*

2. Nombre de rotation

Dans cette section, $f \in \mathrm{homeo}_0(\mathbb{T})$ et F est un relèvement de f.

2.1. Définition. Le *nombre de rotation* du relèvement F est :

$$\rho(F) := \lim_{n \to \infty} \frac{1}{n} F^n(0) \in \mathbb{R}.$$

Le *nombre de rotation* de l'homéomorphisme f du cercle est :

$$\rho(f) := \rho(F) + \mathbb{Z} \in \mathbb{T}.$$

Le nombre de rotation s'exprime de diverses façons intéressantes.

2.2. Définition. La fonction *déplacement* d'un relèvement F de $h \in \mathrm{homeo}_0(\mathbb{T})$ est la fonction $\Delta_F : \mathbb{T} \to \mathbb{R}$ définie par :

$$\Delta_F(x) := F(y) - y \text{ pour n'importe quelle préimage } y \in \pi^{-1}(x).$$

2.3. Remarque. Le déplacement est une fonction bien définie sur $\mathbb{T}$ grâce à (22). Elle dépend du choix du relèvement. On voit qu'elle est continue car π est un homéomorphisme local.

2.4. Lemme. *Le nombre de rotation $\rho(F)$ est bien défini et coïncide avec*

$$\forall x \in \mathbb{R}\ \forall \mu \in \mathbb{P}(f) \quad \rho(F) := \lim_{n\to\infty} \frac{1}{n} F^n(0) = \lim_{n\to\infty} \frac{1}{n} F^n(x) = \int_X \Delta_F \, d\mu.$$

Démonstration. Remarquons que, pour tout $y \in \mathbb{R}$, $S_n^f \Delta_F(\pi(y)) = F^n(y) - y$ et donc :

$$\forall y \in \mathbb{R} \quad \lim_{n\to\infty} \frac{1}{n} F^n(y) = \lim_{n\to\infty} \frac{1}{n} S_n^f(\pi(y))$$

au sens où ces deux limites existent simultanément et sont alors égales.

La dynamique topologique $(\mathbb{T}, f)$ est compacte et métrique donc $\mathbb{P}(f) \neq \emptyset$. Le théorème ergodique ponctuel s'applique à la fonction continue donc intégrable Δ et donne $x_0 \in \mathbb{T}$ tel que la limite suivante existe :

$$\lim_{n\to\infty} \frac{1}{n} S_n^f \Delta_F(x_0).$$

F étant croissante et équivariante par les translations entières, on a les inégalités suivantes. Pour tous $x, y \in \mathbb{R}$, en posant $p := E(x-y)$ (partie entière) de sorte que $y + p \leq x < y + p + 1$, on vérifie :

$$\forall n \in \mathbb{Z} \quad F^n(y+p) = F^n(y) + p \leq F^n(x) < F^n(y+p+1) = F^n(y) + p + 1$$

soit

$$\exists p \in \mathbb{Z}\ \forall n \in \mathbb{Z} \quad 0 \leq F^n(x) - (F^n(y) - p) < 1$$

En particulier pour $x = 0$ et $y = x_0$:

$$\left| \frac{F_n(0) - F_n(x_0)}{n} \right| \leq \frac{|E(x_0)| + 1}{n} \to 0$$

Donc, pour tout $x \in \mathbb{T}$,

$$\lim_{n\to\infty} \frac{1}{n} F^n(x) \text{ existe et vaut } \lim_{n\to\infty} \frac{1}{n} F^n(x_0).$$

En particulier, le nombre de rotation de F est bien défini et vérifie :

$$\forall x \in \mathbb{T} \quad \rho(F) := \lim_{n\to\infty} \frac{1}{n} F^n(0) = \lim_{n\to\infty} \frac{1}{n} F^n(x).$$

Si $\mu \in \mathbb{P}(f)$, le théorème ergodique de Birkhoff implique que, μ-p.p.

$$\mathbb{E}_\mu(\Delta_F | \mathcal{I}(f)) = \lim_{n\to\infty} \frac{1}{n} S_n^f \Delta_F = \rho(F)$$

et donc

$$\int_X \Delta_F \, d\mu = \int_X \mathbb{E}_\mu(\Delta_F | \mathcal{I}(f)) \, d\mu = \rho(F)$$

□

2.5. Lemme. *Le nombre de rotation $\rho(f)$ est bien défini dans $\mathbb{T}$.*

Démonstration du lemme. Le nombre de rotation est bien défini dans $\mathbb{T}$ car tout autre relèvement est de la forme $G = F + q$ avec $q \in \mathbb{Z}$ ce qui, avec (22), donne : $G^n(x) = F^n(x) + nq$ donc $\rho(G) = \rho(F) + q$. $\square$

2.6. Remarque. On peut éviter le recours au théorème ergodique dans la preuve ci-dessus en déduisant la convergence de $F^n(0)/n$ d'une propriété d'additivité approximative :

$$(F^{m+n}(0) - (F^n(0) + F^m(0))| \leq \text{const.}$$

2.7. Lemme. *Si $f \in \mathrm{homeo}_0(\mathbb{T})$ et F est un relèvement de f alors, pour tout $n \in \mathbb{Z}$, $\rho(F^n) = n \cdot \rho(F)$, $\rho(f^n) = n \cdot \rho(f)$. De plus si $f, g \in \mathrm{homeo}_0(\mathbb{T})$ sont conjugués par $h \in \mathrm{homeo}_0(\mathbb{T})$ $\rho(f) = \rho(g)$.*

Démonstration. Les deux premières identités découlent du calcul :

$$\rho(F^n) = \lim_m \frac{1}{m}(F^n)^m(0) = n \lim_m \frac{1}{mn} F^{mn}(0) = n\rho(F)$$

D'après le lemme 1.10, si $f, g, h \in \mathrm{homeo}_0(\mathbb{T})$, il existe des relèvements G, H de g, h tels que $F \circ H = H \circ G$. On en déduit :

$$\rho(F) = \lim_n \frac{F^n(H(0))}{n} = \lim_n \frac{H(G^n(0))}{n} = \lim_n \frac{G^n(0)}{n} = \rho(G)$$

où la dernière égalité découle de ce que $\sup_{x \in \mathbb{R}} |H(x) - x|$ est borné. $\square$

2.8. Proposition. *Soit $f \in \mathrm{homeo}_0(\mathbb{T})$. Son nombre de rotation $\rho(f)$ est rationnel si et seulement s'il admet un point périodique.*

Démonstration. Supposons d'abord qu'il existe $x = f^p(x)$ avec $x \in \mathbb{T}$ et $p \in \mathbb{N}^*$. Par passage au relèvement : $F^p(y) = y + q$ pour $y \in \pi^{-1}(x)$ et $q \in \mathbb{Z}$. Donc $\Delta_{F^p}(x) = q$ et $f^p(x) = x$, soit

$$\rho(F) = \lim_{n \to \infty} \frac{F^{nq}(y) - y}{nq} = \frac{q}{p} \in \mathbb{Q}.$$

Réciproquement, supposons $\rho(F) = q/p$ avec $q \in \mathbb{Z}$ et $p \in \mathbb{N}^*$. On a :

$$q = \rho(F^p) = \int_{\mathbb{T}} \Delta_{F^p} \, d\mu \text{ pour n'importe quelle mesure } \mu \in \mathbb{P}_{\mathrm{erg}}(f).$$

En particulier, Δ_{F^q} doit prendre des valeurs supérieures à p et inférieures à p. Par continuité, il existe $z \in \mathbb{T}$ tel que $\Delta_{F^q}(z) = p$, c'est-à-dire : $F^q(z) = z + p$. On a trouvé un point périodique. $\square$

3. Semiconjugaison à une rotation

3.1. Théorème (Poincaré). *Soit $f \in \mathrm{homeo}_0(\mathbb{T})$ de nombre de rotation α.*

Si $\alpha \notin \mathbb{Q}$, alors f est une extension topologique de la rotation $R_\alpha : \mathbb{T} \to \mathbb{T}$, $x \mapsto x + \alpha \mod \mathbb{Z}$. Plus précisément, il existe $h : \mathbb{T} \to \mathbb{T}$ continue, croissante, de degré topologique 1 telle que $h \circ f = R_\alpha \circ h$.

3.2. Exemple. *Si $\rho(f) \in \mathbb{Q}$, alors f peut ne pas être une extension topologique de la rotation R_α comme le montre l'exemple de l'homéomorphisme $x \mapsto x^2$ pour $0 \leq x \leq 1$.*

3.3. Exercice. Vérifier que l'homéomorphisme cité dans l'exemple ci-dessus n'est pas l'extension topologique d'aucune rotation.

3.4. EXERCICE. Soit (F,μ) une dynamique probabiliste. Montrer que s'il existe un point a tel que $\mu(\{a\}) > 0$, alors ce point est périodique.

DÉMONSTRATION. On fixe $\mu \in \mathbb{P}(f)$. On en construit un relèvement M, une mesure borélienne sur $\mathbb{R}$, en posant :

$$\forall A \in \mathcal{B}_{\mathbb{R}}\ M(A) = \sum_{i\in I} \mu(\pi(A_i))$$

pour toute partition $A = \bigsqcup_{i\in I} A_i$ avec $\operatorname{diam}(A_i) < 1$.

Pour vérifier que $M(A)$ est bien définie, il suffit de vérifier que découper la partition ne change pas la valeur de $M(A)$. On pourra alors comparer deux partitions quelconques en introduisant la partition jointe. Considérons donc :

$$A = \bigsqcup_{i\in I} A_i = \bigsqcup_{(i,j)\in I} B_{i,j}$$

où $A_i = \bigsqcup_j B_{i,j}$. La mesure μ étant additive et $\pi|A_i$ étant un homéomorphisme sur son image, on a bien :

$$\mu(\pi(A_i)) = \sum_j \mu(\pi(B_{i,j})).$$

Soit

$$M(A) = \sum_{i\in I} \mu(\pi(A_i)) = \sum_{(i,j)\in I} \mu(\pi(B_{i,j})).$$

Vérifions que M est une mesure. Clairement, $M(\emptyset) = 0$. Si $A_1, A_2, \ldots$ sont des parties mesurables disjointes,

$$\begin{aligned} M(\bigcup_{i\geq 1} A_i) &= \sum_{n\in\mathbb{Z}} \sum_{i\geq 1} \mu(\pi(A_i \cap [n/2,(n+1)/2[)) \\ &= \sum_{i\geq 1} \sum_{n\geq 1} \mu(\pi(A_i \cap [n/2,(n+1)/2[)) \\ &= \sum_{i\geq 1} M(A_i) \end{aligned}$$

on a donc la σ-additivité. M est une mesure positive sur $\mathbb{R}$.

Notons que $M([x,x+1[) = 1$: les mesures des compacts sont bien finies. M est une mesure de Radon sur $\mathbb{R}$. M n'a pas d'atomes cas μ n'en a pas (voir l'exercice ci-dessus).

Montrons que M est F-invariante. Il suffit de le vérifier sur tout intervalle $[a,b]$ avec $a < b < a+1$. Notons que M n'a pas d'atomes puisque μ n'en a pas. On note que F^{-1} est le relevé de f^{-1}. On calcule :

$$\begin{aligned} M(F^{-1}[a,b]) &= M([F^{-1}a, F^{-1}b]) = \mu([\pi(F^{-1}a), \pi(F^{-1}b)]) \\ &= \mu(f^{-1}\pi([a,b])) = \mu(\pi([a,b])) = M([a,b]) \end{aligned}$$

Construisons la conjugaison à partir de la mesure M. On pose :

$$P : \mathbb{R} \to \mathbb{R}, \quad P(x) = \begin{cases} +M([0,x]) & \text{si } x \geq 0 \\ -M([x,0]) & \text{sinon.} \end{cases}$$

P est manifestement croissante (non-nécessairement strictement croissante).

L'absence d'atomes de la mesure M implique la continuité de P.

Notons que :

$$P(x)-P(y)=\begin{cases} M([x,y]) & \text{si } x\le y\\ -M([y,x]) & \text{sinon.}\end{cases}$$

Il suffit d'établir la première formule. Elle découle d'un examen des cas selon les signes de x et y : $0\le x\le y$, $x\le 0\le y$, $x\le y\le 0$.

On en déduit que $P(x+1)-P(x)=1$ pour tout $x\in\mathbb{R}$. La fonction $P:\mathbb{R}\to\mathbb{R}$ est donc le relèvement d'une application continue $p:\mathbb{T}\to\mathbb{T}$ de degré topologique 1.

Vérifions que $P\circ F=R\circ P$.

$$P(F(x))-P(x)=\begin{cases} M[x,F(x)] & \text{si } x\le F(x)\\ -M[F(x),x] & \text{sinon.}\end{cases} \tag{23}$$

Par continuité et vu l'absence de points fixes pour f, on a ou bien $x<F(x)$ pour tout $x\in\mathbb{R}$, ou bien $x>F(x)$ pour tout $x\in\mathbb{R}$. Supposons qu'on soit dans le premier cas.

Notons que $0<\inf F(x)-x<\sup F(x)-x<\infty$.

Montrons que pour tous $x,y\in\mathbb{R}$, $M[y,F(y)]=M[x,F(x)]$. Quitte à introduire des points intermédiaires, il suffit de le montrer pour $x\le y<F(x)$. Nécessairement on a alors : $x\le y<F(x)\le F(y)$ et donc :

$$M[x,F(x)]-M[y,F(y)]=M[x,y]-M[F(x),F(y)]$$

L'invariance de M par F implique que cette différence est nulle. On peut poser $\beta:=M[x,F(x)]$, indépendant de $x\in\mathbb{R}$.

L'équation (23) donne que $P\circ F=R\circ P$ avec $R:\mathbb{R}\to\mathbb{R}$, $x\mapsto x+\beta$, et $P(x+1)=P(x)+1$.

La relation $P\circ F=R\circ P$ implique $\rho(F)=\rho(R)$, i.e., $\alpha=\beta$.

En passant au quotient par $\pi:\mathbb{R}\to\mathbb{Z}$, on en déduit :

$$p\circ f=r\circ p$$

$p:\mathbb{T}\to\mathbb{T}$ est une semiconjugaison topologique faisant de f une extension topologique de la rotation $r:\mathbb{T}\to\mathbb{T}$ d'angle α. □

4. Conjugaison à la rotation

4.1. Non-conjugaison et intervalles errants. Dans la situation du théorème de Poincaré, on n'a pas nécessairement de conjugaison topologique. La notion suivante nous sera utile :

4.1. Définition. Etant donné f une application continue de $\mathbb{T}$ dans $\mathbb{T}$ ou de $[0,1]$ dans $[0,1]$, un *intervalle errant* est un intervalle de longueur non-nulle dont tous les itérés futurs $f^n(I)$, $n\ge 0$, sont disjoints et qui ne s'accumule pas sur une orbite périodique.

Nous utiliserons cette notion pour $f\in\mathrm{homeo}_+(\mathbb{T})$ sans points périodiques. Dans ce cadre, il suffit d'avoir $f^n(I)\cap I=\emptyset$ pour tout $n\ge 1$ car f est inversible et il n'y a pas d'accumulation vers une orbite périodique car il n'y a pas de telles orbites.

4.2. Lemme. *Soit* $f\in\mathrm{homeo}_+(\mathbb{T})$ *sans points périodiques. Alors les propriétés suivantes sont équivalentes :*

(1) f est topologiquement conjuguée à une rotation irrationnelle ;

(2) f est topologiquement transitive ;

(3) il n'existe pas d'un intervalle errant.

Démonstration. Démontrons un cycle d'implications.

(1) $\Longrightarrow$ (2) : clair.

(2) $\Longrightarrow$ (3) : supposons qu'il existe un intervalle errant I. Soit $x \in \mathbb{T}$. Notons que pour tout n assez grand $f^n(x) \notin I$: ou bien parce que $x \in\notin \bigcup_{n\in\mathbb{Z}} f^n(I)$ ou bien $x \in f^{n_0}(I)$ et donc $f^n(x) \notin I$ pour tout $n > -n_0$. Aucune orbite n'est dense et f n'est donc pas topologiquement transitive.

(3) $\Longrightarrow$ (1) : On applique la construction précédente et on obtient $P\circ F = R\circ P$ où P, R, F sont des relèvements croissants de p, r, f. Si p n'est pas une conjugaison topologique, alors il existe $x \in \mathbb{T}$ tel que $I := \pi^{-1}(x)$ n'est pas un point donc est un intervalle car P est croissant. Le point x n'étant pas périodique pour r, I est disjoint de tous ses itérés. L'application f n'a pas de points périodiques donc I est un intervalle errant. $\square$

4.3. Exercice. Soit $r : \mathbb{T} \to \mathbb{T}$ la rotation d'angle $\alpha \in \mathbb{R}\setminus\mathbb{Q}$. On considère la mesure suivante sur $\mathbb{T}$:

$$\mu := \frac{1}{2+\pi^2/3}\left(m_{\mathbb{T}} + \sum_{n\in\mathbb{Z}} \frac{1}{\max(n^2,1)}\delta_{r^n(0)}\right).$$

On définit $p : \mathbb{T} \to \mathbb{T}$ par

$$\forall t \in [0,1[\ p(\pi(t)) = \pi(x) \iff \mu([0,x[) \leq t \leq \mu([0,x]).$$

et on pose $I_n := p^{-1}(r^n(0))$.

1. Montrer que $\mu \in \mathbb{P}(\mathbb{T})$ et vérifier que $p : \mathbb{T} \to \mathbb{T}$ est continue et croissante de degré topologique égal à 1.

On définit $f : \mathbb{T} \to \mathbb{T}$ par $f : I_n \to I_{n+1}$ est un homéomorphisme croissant et, si $x \notin \bigcup_{n\in\mathbb{Z}} I_n$, $p(f(x)) = r(p(x))$.

2. Montrer que $f \in \mathrm{homeo}_+(\mathbb{T})$, $\rho(f) = \alpha \mod 1$ et f n'est pas topologiquement conjuguée à une rotation irrationnelle.

Le théorème suivant montre qu'un peu de régularité empêche l'apparition d'intervalles errants.

4.4. Définition. On note $\mathrm{Diffeo}_+{}^r(\mathbb{T})$ l'ensemble des $f \in \mathrm{homeo}_+(\mathbb{T})$ dont un relèvement est un difféomorphisme de $\mathbb{R}$ de classe C^r.

4.5. Théorème (Denjoy). *Soit $f \in \mathrm{Diffeo}_+{}^2(\mathbb{T})$ sans points périodiques. Alors f est topologiquement conjuguée à la rotation d'angle $\rho(f)$.*

4.6. Remarque. Un théorème d'Arnold exhibe des difféomorphismes de nombre de rotation irrationnel qui ne peuvent pas être conjugués de façon C^1 à une rotation. Ce théorème s'applique à la famille $f_\alpha : \mathbb{T} \to \mathbb{T}$, $x \mapsto x + \frac{1}{4\pi}\sin(2\pi x) + \alpha$: il existe des valeurs de α telles que $\rho(f_\alpha)$ est irrationnelle mais l'unique mesure de probabilité invariante par f_α est singulière par rapport à la mesure uniforme. Le lecteur intéressé pourra consulter [**5**, I.5].

La preuve du théorème de Denjoy consiste à utiliser la régularité pour montrer qu'il n'existe pas d'intervalle errant.

L'outil technique principal est une estimée de *distorsion* :

4.7. LEMME. *Soit* $f \in \mathrm{Diffeo_+}^2(\mathbb{T})$. *Il existe* $C < \infty$ *tel que, pour tout intervalle* J *de* $\mathbb{T}$ *et tout* $n \geq 1$,

$$\sup_{x,y\in J} \left| \frac{(f^n)'(x)}{(f^n)'(y)} \right| \leq \exp\left(C \sum_{k=0}^{n-1} |f^k(I)| \right).$$

Remarquons que si I est un intervalle errant alors $\sum_{k\geq 0} |f^k(I)| \leq |\mathbb{T}| = 1$.

DÉMONSTRATION. Remarquons que $\log|f'|$ est de classe C^1 donc admet une constante de Lipschitz C, ce qui donne, pour tout $x, y \in J$:

$$\log\left| \frac{(f^n)'(x)}{(f^n)'(y)} \right| = \sum_{k=0}^{n-1} \log\left| \frac{f'(f^k(x))}{f'(f^k(y))} \right| \leq \sum_{k=0}^{n-1} C.|f^k(J)|.$$

Il suffit maintenant de prendre l'exponentielle puis la borne supérieure. □

DÉMONSTRATION DU THÉORÈME. Vu le lemme 4.2, il suffit de montrer que f n'a pas d'intervalle errant. On procède par contradiction, en supposant cette existence.

Le lemme de Zorn montre qu'il existe alors un intervalle errant J qui est maximal pour l'inclusion (l'union d'une collection totalement ordonnée d'intervalle errant est un intervalle errant). On fixe, C étant la constante du lemme 4.7,

K intervalle ouvert contenant J avec $|K| = (1+\delta)|J| \quad (\delta := \exp(-2C))$

On va utiliser la distorsion pour voir que

$$U := \bigcup_{n\geq 0} f^n(K) \subsetneq \mathbb{T}. \tag{24}$$

Les composantes connexes de U sont alors des intervalles (et pas le cercle entier). Prenons la composante connexe V de U contenant K. Si $f^n(V) \cap V \neq \emptyset$ alors $f^n(V)$ composante connexe de $f^n(U) \subset U$ doit être incluse dans V. Mais ceci implique l'existence d'un point périodique, une contradiction. Donc $V \supset K \supsetneq J$ est un intervalle errant qui contredit la maximalité de J.

L'inclusion stricte (24) découlera de la majoration :

$$\forall n \geq 0 \ |f^n(K)| \leq 2|f^n(J)| \tag{25}$$

en remarquant de plus qu'on peut supposer $\sum_{n\geq 0} |f^n(J)| < 1/2$ quitte à remplacer J par un itéré suffisamment grand. La majoration précédente impliquera donc $|U| < |\mathbb{T}| = 1$ et donc (24).

Pour démontrer (25), on procède par récurrence sur $n \geq 0$ en supposant que :

$$\forall 0 \leq i \leq n \quad |f^i(K)| \leq 2|f^i(J)|. \tag{26}$$

Ceci est vrai pour $n = 0$. Si on l'a pour un certain $n \geq 0$, alors le lemme 4.7 implique, en notant K_1, K_2 les deux composantes connexes de $K \setminus J$:

$$|f^{n+1}(K)| \leq |f^{n+1}(J)| + |f^{n+1}(K_1 \cup K_2)| \leq |f^{n+1}(J)| + \sup_{x\in K} |(f^{n+1})'(x)| \cdot |K \setminus J|.$$

Comme J est un intervalle, il existe $y \in J$ tel que $|(f^{n+1})'(y)| \cdot |J| = |f^{n+1}(J)|$.

L'hypothèse de récurrence (26) donne :

$$\sum_{i=0}^{n} |f^i(K)| \leq 2 \sum_{i=0}^{n} |f^i(J)| \leq 2.$$

Le lemme 4.7 s'applique à $f^{n+1}|K$ et donne :

$$\forall x \in K \quad |(f^{n+1})'(x)| \leq e^{2C}|(f^{n+1})'(y)| = e^{2C}\frac{|f^{n+1}(J)|}{|J|}.$$

D'où

$$\begin{aligned}|f^{n+1}(K)| &\leq |f^{n+1}(J)| + e^{2C}\frac{|f^{n+1}(J)|}{|J|}|K \setminus J| \\ &\leq (1+e^{2C}\delta)|f^{n+1}(J)| \leq 2|f^{n+1}(J)|\end{aligned}$$

car on a choisi $\delta \leq e^{-2C}$.

L'hypothèse de récurrence est démontrée au rang $n+1$. La preuve du théorème est complète. $\square$

4.8. Remarque. Denjoy a construit des difféomorphismes $f \in \mathrm{Diffeo}_+{}^1(\mathbb{T})$ de nombre de rotation irrationnel et qui ne sont pas topologiquement conjugués à une rotation.

5. Un énoncé d'Herman-Yoccoz

5.1. Théorème. *Soit $f \in \mathrm{Diffeo}_+{}^\infty(\mathbb{T})$ de nombre de rotation α satisfaisant la condition diophantienne suivante :*

$$\exists C, \nu > 0 \quad \forall (q,p) \in \mathbb{Z} \times \mathbb{N}^* \ \left|\alpha - \frac{q}{p}\right| \geq \frac{C}{q^{2+\nu}}$$

Alors f est C^∞-conjugué à la rotation d'angle α.

Chapitre 9

Application de Gauss et fractions continues

L'application de Gauss est définie par :

$$G : [0,1] \to [0,1],\ x \mapsto \frac{1}{x} - E\left(\frac{1}{x}\right) \text{ si } x \neq 0 \text{ et } 0 \text{ si } x = 0.$$

On a $[0,1] = \bigsqcup_{n\geq 1} I_n$ où $I_n :=]1/(n+1), 1/n[$ (à un ensemble dénombrable près).

Chaque restriction : $G : I_n \to]0,1[$ est un difféomorphisme analytique. Tout point $x \in]0,1[$ a une infinité de préimages, une dans chaque I_n.

1. L'application de Gauss préserve la mesure

1.1. Théorème. *L'application de Gauss préserve la mesure de probabilité* $\mu_G := \frac{1_{[0,1]}}{\log 2}\frac{dx}{1+x}$.

Démonstration. La mesure proposée est bien définie et borélienne. Elle est absolument continue par rapport à la mesure de Lebesgue et donc sans atomes. En particulier, tout ensemble dénombrable est de mesure nulle.

Calculons sa masse totale :

$$\int_0^1 \frac{1}{\log 2}\frac{dx}{1+x} = \frac{\log(2) - \log(1)}{\log 2} = 1$$

Vérifions son invariance en utilisant le théorème de changement de variable dans les intégrales. Pour ce faire, on pose $I_n :=]\frac{1}{n+1}, \frac{1}{n}[$ pour tout $n \geq 1$. On pose $\mathcal{P} := \{I_n : n \geq 1\}$. C'est une partition de $]0,1[$ (à un ensemble dénombrable donc de mesure nulle près). On a, pour chaque $n \geq 1$,

$$G_n : I_n \to]0,1[,\ x \mapsto \frac{1}{x} - n \text{ est un difféomorphisme d'inverse } y \mapsto \frac{1}{y+n}.$$

On note $((G_n)^{-1})'(y) = G'(G_n^{-1}(y)) = (y+n)^2$.

On a donc, pour tout $\phi \geq 0$ mesurable :

$$\begin{aligned}
\int_0^1 \phi \circ G \, d\mu_G &= \frac{1}{\log 2}\sum_{n\geq 1}\int_{I_n} \phi \circ G \frac{dx}{1+x} \\
&= \frac{1}{\log 2}\sum_{n\geq 1}\int_{G(I_n)} \phi(y) \frac{dy}{(1+1/(y+n))G'(G_n^{-1}(y))} \\
&= \frac{1}{\log 2}\int_0^1 \phi(y) \sum_{n\geq 1}\frac{1}{(y+n+1)(y+n)}\, dy
\end{aligned}$$

Mais

$$\sum_{n\geq 1}\frac{1}{(y+n+1)(y+n)} = \sum_{n\geq 1}\frac{1}{y+n} - \frac{1}{y+n+1} = \frac{1}{y+1}.$$

Donc :

$$\int_0^1 \phi \circ G \, d\mu_G = \frac{1}{\log 2} \int_0^1 \phi(y) \, \frac{dy}{y+1}$$

On a bien : $\mu_G(\phi \circ G) = \mu_G(\phi)$ pour tout $\phi \in L^1(\mu)$. On en déduit que la mesure μ_G est bien invariante par G. □

2. Dilatation et distorsion

L'application de Gauss est uniformément dilatante :

2.1. PROPOSITION. *Pour tout* $n \geq 1$ *et tout* $x \in [0,1] \setminus \bigcup_{k=0}^{n-1} G^{-1}(\{0, 1/2, 1/3, 1/4, \dots\}$,

$$|(G^n)'(x)| \leq 4^{-1} 2^n$$

Attention, $G'(1) = 1$.

DÉMONSTRATION. On calcule, pour $x \in I_n$ et donc $G(x) = \frac{1}{x} - n$,

$$(G^2)'(x) = G'(G(x))G'(x) = \frac{1}{(\frac{1}{x} - n)^2} \frac{1}{x^2} = \frac{1}{(1-nx)^2} \geq (n+1)^2 \geq 4$$

□

On aura besoin d'un contrôle de la distorsion, rendu plus délicat car G'' n'est pas bornée.

2.2. LEMME. *Pour* x, y *appartenant au même élément de* $\mathcal{P}^n$ *pour* $n \geq 1$,

$$\log \left| \frac{(G^n)'(x)}{(G^n)'(y)} \right| \leq \sum_{k=0}^{n-1} |y_{k+1} - x_{k+1}| \sup_{u,v \in \mathcal{P}(x_k)} \left| \frac{G''(u)}{G'(u)G'(v)} \right|$$

où $x_k := G^k(x)$, $y_k := G^k(y)$. *En particulier,*

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

où C *ne dépend ni de* $n \geq 1$, *ni de* x, y.

DÉMONSTRATION. La formule $(\log |G'|)' = \frac{G''}{G'}$ implique :

$$\log \left| \frac{G'(x_k)}{G'(y_k)} \right| \leq |y_k - x_k| \cdot \sup_{u \in [x_k, y_k]} \left| \frac{G''(u)}{G'(u)} \right|$$

Pour gagner un facteur $1/G'$, on introduit $vin[x_k, y_k]$ tel que $|y_k - x_k||G'(v)| \cdot |y_{k+1} - x_{k+1}|$:

$$|y_k - x_k| \cdot \sup_{u \in [x_k, y_k]} \left| \frac{G''(u)}{G'(u)} \right| \leq |y_{k+1} - x_{k+1}| \sup_{u,v \in [x_k, y_k]} \left| \frac{G''(u)}{G'(u)G'(v)} \right|$$

Pour démontrer la dernière assertion, on remarque que

$$\sup_{k \geq 1} \sup_{u,v \in \mathcal{P}(x_k)} \left| \frac{G''(u)}{G'(u)G'(v)} \right| \leq \sup_{k \geq 1} \frac{2(k+1)}{k^2} \leq \text{const}$$

D'autre part,

$$\sum_{k=1}^{n} |x_k - y_k| \leq \sum_{k=1}^{n} |x_k - y_k| 4 \cdot 4^{-(n-k)/2} \leq 8$$

L'affirmation est démontrée. □

3. Convergence des martingales

On utilise l'outil fondamental suivant :

3.1. Théorème (Convergence des martingales de Doob). *Soient $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \dots$ des tribus croissantes engendrant la tribu $\mathcal{B}$. Soit m une mesure de probabilité borélienne.*

Pour tout $\phi \in L^1(m)$, les espérances conditionnelles convergent :

$$\lim_{n\to\infty} E_m(\phi|\mathcal{B}_n) = \phi \ m\text{-}p.p \ et \ dans \ L^1(m).$$

On l'applique en prenant pour $\mathcal{B}_n$ la tribu engendrée par la partition mesurable modulo μ (elle exclut les points $1/n$ et leurs préimages) :

$$\mathcal{P}^n := \{A_1 \cap \dots \cap A_n : A_i \in \mathcal{P}\}.$$

En notant $\mathcal{P}^n(x)$ l'élément de $\mathcal{P}^n$ contenant x on a :

$$\operatorname{diam}(\mathcal{P}^n(x)) \leq 2 \cdot 2^{-n/2}.$$

En particulier, si $x \neq y$ n'appartiennent pas à l'ensemble exceptionnel, alors il existe $n \geq 1$, $\mathcal{P}^n(x) \neq \mathcal{P}^n(y)$. Il s'ensuit que $\bigcup_{n\geq 1} \mathcal{P}^n$ engendre la tribu des boréliens.

On a donc, pour tout $A \in \mathcal{B}$, pour μ-p.t. $x \in X$:

$$E_\mu(1_A|\mathcal{P}^n)(x) = \frac{\mu(A \cap \mathcal{P}^n(x))}{\mu(\mathcal{P}^n(x))} \to 1_A(x)$$

4. Ergodicité

4.1. Théorème. *L'application de Gauss est ergodique pour la mesure μ_G.*

Démonstration de l'ergodicité. On se donne un ensemble A invariant modulo μ dont on suppose $\mu(A) > 0$. Vu $\mu \ll m$, on a $m(A) > 0$ où m est la mesure de Lebesgue sur $[0,1]$.

Le théorème de convergence des martigales implique :

$$\text{pour } m\text{-p.t. } x \in A \quad \frac{m(A^c \cap \mathcal{P}^n(x))}{m(\mathcal{P}^n(x))} \to 0$$

On fixe un tel point x_0. En prenant n assez grand on rend la fraction ci-dessus arbitrairement petite.

L'invariance de A, $F^{-n}(A) = A$ modulo μ donc modulo m (car $m \ll \mu$) implique que $F^n(A \cap \mathcal{P}^n(x)) \subset A \cap]0,1[$ et donc $A^c \cap]0,1[\subset F^n(]0,1[\setminus A)$.

On utilise maintenant une estimée de distorsion :

$$m(A^c) \leq \frac{m(A^c \cap [0,1])}{m([0,1])} \leq \sup_{s,t \in \mathcal{P}^n(x)} \left|\frac{(f^n)'(s)}{(f^n)'(t)}\right| \cdot \frac{m(A^c \cap \mathcal{P}^n(x))}{m(\mathcal{P}^n(x))} \leq K \cdot \frac{m(A^c \cap \mathcal{P}^n(x))}{m(\mathcal{P}^n(x))}$$

avec K indépendant de $n \to \infty$ d'après le lemme 2.2. En prenant $n \to \infty$, on en déduit que $m(A^c)$ est arbitrairement petit donc nul. Donc $\mu(A) = 1$. L'ergodicité est démontrée. $\square$

5. Développement en fraction continue

Soit $x \in [0,1] \setminus \mathbb{Q}$. On introduit la fonction $a : [0,1] \to \{1,2,\ldots,\infty\}$, $x \mapsto E(1/x)$. On pose $a_k(x) := a(G^{k-1}x)$: ce sont les *quotients partiels* de x.

On pose également :

$$q_0 := 1,\ q_1 := a_1,\ q_{n+1} := a_{n+1}q_n + q_{n-1}$$
$$p_0 := 0,\ p_1 := 1,\ p_{n+1} := a_{n+1}p_n + p_{n-1}$$

On définit $[a_1,\ldots,a_n]$ $(n \geq 0)$ par récurrence en posant :

$$[\,] := 0 \text{ et } [a_1,\ldots,a_n] := \frac{1}{a_1 + [a_2,\ldots,a_n]}.$$

Remarquons que

$$[a_1,\ldots,a_k + [a_{k+1},\ldots,a_n]] = [a_1,\ldots,a_n]$$

On introduit également les matrices :

$$A_n := \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}, \quad M_n := \begin{pmatrix} q_n & p_n \\ q_{n-1} & p_{n-1} \end{pmatrix} \quad (n \geq 1)$$

Les relations ci-dessus se réécrivent : $M_{n+1} = A_{n+1}M_n$ pour tout $n \geq 1$

5.1. Lemme. *On a :* $x = [a_1,\ldots,a_{n-1},a_n + G^n(x)]$.

Démonstration. On procède par récurrence. On vérifie directement :

$$G^{n-1}(x) = [a_n + G^n(x)]$$

En supposant $G^{k+1}(x) = [a_{k+2},\ldots,a_{n-1},a_n + G^n(x)]$, on calcule :

$$\begin{aligned} G^k(x) = [a_{k+1} + G^{k+1}(x)] &= [a_{k+1} + [a_{k+2},\ldots,a_{n-1},a_n + G^n(x)]] \\ &= [a_{k+1},\ldots,a_{n-1},a_n + G^n(x)]. \end{aligned}$$

Le lemme en découle. □

5.2. Lemme. *On a :* $\frac{p_n}{q_n} = [a_1,\ldots,a_n]$.

Démonstration. On procède par récurrence sur n. Les cas $n = 0,1$ sont triviaux. Supposons que pour tous $a_1,\ldots,a_n > 0$, les convergents définis comme précédemment satisfont

$$\frac{p_n}{q_n} = [b_1,\ldots,b_n]$$

Etant donné $b_{n+1} > 0$ quelconque, on calcule

$$[b_1,\ldots,b_n,b_{n+1}] = [b_1,\ldots,b_n + 1/b_{n+1}]$$

En posant $\tilde{b}_k = b_k$ for $k < n$ et $\tilde{b}_n = b_n + 1/b_{n+1}$, les convergents $\frac{\tilde{p}_n}{\tilde{q}_n}$ définis par la suite $\tilde{b}$ vérifient, par l'hypothèse de récurrence :

$$\frac{\tilde{p}_n}{\tilde{q}_n} = [b_1,\ldots,b_n + 1/b_{n+1}]$$

mais par définition des convergents (et vu $\tilde{p}_k = p_k$, $\tilde{q}_k = q_k$ pour $k < n$) :

$$\tilde{p}_n = (b_n + 1/b_{n+1})p_{n-1} + p_{n-2} \text{ et } \tilde{q}_n = (b_n + 1/b_{n+1})q_{n-1} + q_{n-2}$$

Soit :

$$\tilde{p}_n = p_n + p_{n-1}/b_{n+1} \text{ et } \tilde{q}_n = q_n + q_{n-1}/b_{n+1}$$

D'où :

$$\frac{\tilde{p}_n}{\tilde{q}_n} = \frac{p_n + p_{n-1}/b_{n+1}}{q_n + q_{n-1}/b_{n+1}} = \frac{b_{n+1}p_n + p_{n-1}}{b_{n+1}q_n + q_{n-1}} = \frac{p_{n+1}}{q_{n+1}}$$

par définition des convergents d'ordre $n+1$. Ceci complète la récurrence. $\square$

5.3. Lemme. *On a :*

$$x = \frac{p_n + p_{n-1}G^n(x)}{q_n + q_{n-1}G^n(x)} \text{ et } G^n(x) = -\frac{x \cdot q_n - p_n}{x \cdot q_{n-1} - p_{n-1}}$$

Démonstration. Le lemma précédent donne

$$x = [a_1, \dots, a_n + G^n(x)] = \frac{\tilde{p}_n}{\tilde{q}_n} = \frac{(a_n + G^n(x))p_{n-1} + p_{n-2}}{(a_n + G^n(x))q_{n-1} + q_{n-2}}$$
$$= \frac{p_n + G^n(x)p_{n-1}}{q_n + G^n(x)q_{n-1}}$$

en utilisant la définition de p_n, q_n.

On peut alors résoudre en $G^n(x)$ pour obtenir la deuxième expression annoncée. $\square$

5.4. Lemme. *On a :*

$$(-1)^n = q_n p_{n-1} - q_{n-1}p_n \text{ et } x \cdot q_{n-1} - p_{n-1} = \frac{(-1)^{n+1}}{q_n + q_{n-1}G^n(x)}$$

Démonstration. La première identité résulte du calcul des déterminants de M_n et A_k.

Pour obtenir la seconde idéntité, on utilise l'expression de x fournie par le lemme précédent et la première identité. $\square$

5.5. Remarque. Les convergents sont précisément les meilleures approximations rationnelles de l'irrationnel x au sens suivant, pour tout $(p, q) \in \mathbb{Z} \times \mathbb{N}$ si

$$q \leq q_n \text{ et } \left|x - \frac{p}{q}\right| \leq \left|x - \frac{p_n}{q_n}\right|$$

alors $(p, q) = (p_n, q_n)$.

6. Conclusion

En utilisant la dynamique, on va étudier le comportement de l'erreur $\left|x - \frac{p_n}{q_n}\right|$ quand $n \to \infty$ pour Lebesgue-p.t. $x \in [0, 1]$.

6.1. Théorème (Lévy). *Pour Lebesgue presque tout $x \in [0, 1]$, les convergents approchent x exponentiellement vite :*

$$\lim_{n\to\infty} -\frac{1}{n}\left|x - \frac{p_n}{q_n}\right| = \frac{\pi^2}{12}$$

Démonstration. L'équation **??** donne

$$G^n(x) \cdot G^{n-1}(x) \cdots G(x) = (-1)^n \frac{xq_n - p_n}{xq_{n-1} - p_{n-1}} \cdot \frac{xq_{n-1} - p_{n-1}}{xq_{n-2} - p_{n-2}} \cdots \frac{xq_1 - p_1}{xq_0 - p_0}$$

C'est-à-dire en se rappelant que $xq_0 - p_0 = x$:

$$G^n(x)G^{n-1}(x) \dots G(x) \cdot x = |xq_n - p_n|$$

La dynamique probabiliste (G, μ_G) étant ergodique et la fonction $\log x \in L^1(\mu_G)$, le théorème de Birkhoff donne, pour μ_G-p.t. $x \in [0,1]$,

$$g := \lim_{n\to\infty} \frac{1}{n} \log(x \cdot G(x) \cdot G^2(x) \cdots G^{n-1}(x)) = \frac{1}{\log 2} \int_0^1 \log x \frac{dx}{1+x}$$

Calculons l'intégrale en faisant une intégration par partie puis un développement en série entière :

$$\begin{aligned}
\int_0^1 \frac{\log x}{1+x}\, dx &= 0 - \int_0^1 \frac{1}{x} \log(1+x)\, dx = - \int_0^1 \frac{1}{x} \sum_{k\geq 1} \frac{(-1)^{k+1} x^k}{k}\, dx \\
&= - \sum_{k\geq 1} \int_0^1 \frac{(-1)^{k-1} x^{k-1}}{k}\, dx = \sum_{k\geq 1} \frac{(-1)^k}{k^2} \\
&= \sum_{k\geq 1} \frac{(-1)^k}{k^2} = \sum_{k\geq 1} \frac{1}{4k^2} - \left(\sum_{k\geq 1} \frac{1}{k^2} - \sum_{k\geq 1} \frac{1}{4k^2} \right) \\
&= -\frac{1}{2} \sum_{k\geq 1} \frac{1}{k^2} = -\frac{\pi^2}{12}
\end{aligned}$$

vu $\sum_{k\geq 1} 1/k^2 = \pi^2/6$. On a donc, pour μ_G-p.t. $x \in [0,1]$,

$$\lim_{n\to\infty} \frac{1}{n} \log |xq_n - p_n| = -\frac{\pi^2}{12 \log 2}.$$

Vu le lemme 5.4,

$$x \cdot q_{n-1} - p_{n-1} = \frac{(-1)^{n+1}}{q_n + q_{n-1} G^n(x)}$$

On en déduit (car $0 < q_{n-1} G^n(x) < q_n$) :

$$\frac{1}{2q_n} < |x \cdot q_{n-1} - p_{n-1}| < \frac{1}{q_n}$$

et donc l'existence et la valeur de la limité suivante :

$$\lim_n \frac{1}{n} \log q_n = \lim_n -\frac{1}{n} \log |x \cdot q_{n-1} - p_{n-1}| = \frac{\pi^2}{12 \log 2}.$$

L'estimée de Lévy s'en déduit immédiatement. $\square$

7. Mélanges et théorèmes limites

On peut obtenir des résultats *quantitatifs* :

7.1. Théorème (Lévy, 1929). *On a le mélange exponentiel. Soit $a \in (\mathbb{N}^*)^r$, $b \in (\mathbb{N}^*)s$. On pose : $A := \{x \in [0,1] : a_1(x) \dots a_r(x) = a\}$, $B := \{x \in [0,1] : a_1(x) \dots a_s(x) = b\}$.*

Il existe $c > 0$ et $q < 1$ tels que :

$$\forall n \geq 1\ |\mu_G(A \cap G^{-n-r}(B)) - \mu_G(A)\mu_G(B)| \leq c \cdot q^n \cdot \mu_G(A)\mu_G(B)$$

7.2. Exercice. Déduire du théorème précédent que (G, m) est fortement mélangeante.

Ce genre d'estimées exponentielles permet d'obtenir des théorèmes limites. En voici un exemple qui précise la divergence des quotients partiels vue ci-dessus :

7.3. THÉORÈME (Galambos 1972, Doeblin 1940). *On considère les records successifs :* $R_n(x) := \max\{a_k(x) : 1 \leq k \leq n\}$. *On a, pour tout* $t > 0$ *:*

$$\lim_{n\to\infty} \mu_G(\{x \in [0,1] : \frac{R_n(x)}{n} \geq \frac{t}{\log 2}\}) \to e^{-1/t}.$$

Autrement dit, R_n/n *converge en loi.*

Le résultat précédent est motivé par le cas indépendant :

7.4. EXERCICE. Soit $X_1, X_2, \ldots$ des variables aléatoires i.i.d. dont la loi coïncide avec celle du quotient partiel a. Montrer que, pour tout $t > 0$:

$$\lim_{n\to\infty} \mathbb{P}(\max(X_1, X_2, \ldots, X_n) \geq n\frac{t}{\log 2}) \to e^{-1/t}$$

Chapitre 10

Introduction à l'entropie des dynamiques probabilistes

Dans ce chapitre, $(X, \mathcal{B}, \mu, \mathbb{Z}, f)$ est une dynamique probabiliste standard en temps $\mathbb{Z}$ et toutes les partitions sont des collections finies de parties mesurables dont les intersections sont de mesure nulle et dont l'union est de mesure totale.

1. Entropie statique

1.1. Définition. Etant données une mesure μ et une partition $\mathcal{A}$, l'entropie statique est la quantité :

$$H_\mu(\mathcal{A}) := -\sum_{A\in\mathcal{A}} \mu(A)\log\mu(A)$$

avec la convention $\mu(A)\log\mu(A) = 0$ si $\mu(A) = 0$.

1.2. Lemme. *Si $\mathcal{A}$ est une partition, alors $0 \leq H_\mu(\mathcal{A}) \leq \log\#\mathcal{A}$. De plus : $H_\mu(\mathcal{A}) = 0 \iff \mu(A) = 1$ pour un certain $A \in \mathcal{A}$; $H_\mu(\mathcal{A}) = \log\#\mathcal{A} \iff \mu(A) = 1/\#\mathcal{A}$ pour tout $A \in \mathcal{A}$.*

Démonstration. La positivité de l'entropie statique est évidente. Si l'entropie est nulle alors chaque $\phi(\mu(A)) = 0$ donc $\mu(A) = 0$ ou 1. On a le cas d'égalité annoncé.

Pour la majoration, on utilise la concavité stricte de la fonction $\phi(t) = -t\log t$. En notant $n = \#\mathcal{A}$,

$$\sum_{A\in\mathcal{A}} \frac{1}{n}\phi(\mu(A)) \leq \phi\left(\frac{1}{n}\sum_{A\in\mathcal{A}}\mu(A)\right) = \phi(1/n) = \frac{1}{n}\log n$$

Soit $H_\mu(\mathcal{A}) \leq \log n$. En cas d'égalité, la concavité stricte force $\mu(A)$ à être indépendant de A, i.e., $\mu(A) = 1/n$ pour tout $A \in \mathcal{A}$. On a donc la majoration et son cas d'égalite. □

Le joint de deux partitions est :

$$\mathcal{A} \vee \mathcal{B} := \{A \cap B \neq \emptyset : A \in \mathcal{A},\ B \in \mathcal{B}\}$$

1.3. Définition. Etant données une mesure μ et deux partitions $\mathcal{A}$ et $\mathcal{B}$ l'*entropie statique de $\mathcal{A}$ sachant $\mathcal{B}$* est la quantité :

$$H_\mu(\mathcal{A}|\mathcal{B}) := H_\mu(\mathcal{A} \vee \mathcal{B}) - H_\mu(\mathcal{B}).$$

Remarquons que $H_\mu(\mathcal{A}|\{\emptyset, X\}) = H_\mu(\mathcal{A})$.

1.4. Lemme. *On a : $0 \leq H_\mu(\mathcal{A}|\mathcal{B}) \leq H_\mu(\mathcal{A})$*

Démonstration. On remarque d'abord (les termes où $\mu(B) = 0$ étant pris égaux à zéro) :

$$H_\mu(\mathcal{A}|\mathcal{B}) = \sum_{B \in \mathcal{B}} \mu(B) \cdot H_{\mu|B}(\mathcal{A})$$

En particulier, $H_\mu(\mathcal{A}|\mathcal{B}) \geq 0$. Puis on vérifie que

$$H_\mu(\mathcal{A}|\mathcal{B}) = -\sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} \mu(B) \cdot \frac{\mu(A \cap B)}{\mu(B)} \log \frac{\mu(A \cap B)}{\mu(B)}$$

On utilise la concavité de $\phi(t) = -t \log t$:

$$\sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} \mu(B) \cdot \phi\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \leq \sum_{A \in \mathcal{A}} \phi\left(\sum_{B \in \mathcal{B}} \mu(B) \cdot \frac{\mu(A \cap B)}{\mu(B)}\right) = \sum_{A \in \mathcal{A}} \phi(\mu(A)) = H_\mu(\mathcal{A}).$$

□

On dit que $\mathcal{A}$ est moins fine que $\mathcal{B}$ et on note $\mathcal{A} \preceq \mathcal{B}$ si tout élément de $\mathcal{A}$ est une union d'éléments de $\mathcal{B}$.

1.5. Lemme. *L'entropie est croissante par rapport à la partition :*
1. Si $\mathcal{A}_1 \preceq \mathcal{A}_2$, alors $H_\mu(\mathcal{A}_1) \leq H_\mu(\mathcal{A}_2)$.
2. Si $\mathcal{A}_1 \preceq \mathcal{A}_2$, alors $H_\mu(\mathcal{A}_1|\mathcal{B}) \leq H_\mu(\mathcal{A}_2|\mathcal{B})$.

Démonstration. On se ramène à la définition de l'entropie conditionnelle.

1. On a : $\mathcal{A}_2 = \mathcal{A}_1 \vee \mathcal{A}_2$ et donc $H_\mu(\mathcal{A}_2) = H_\mu(\mathcal{A}_1) + H_\mu(\mathcal{A}_2|\mathcal{A}_1)$ et ce dernier terme est positif.

2. On en déduit :

$$H_\mu(\mathcal{A}_1|\mathcal{B}) = H_\mu(\mathcal{A}_1 \vee \mathcal{B}) - H_\mu(\mathcal{B}) \leq H_\mu(\mathcal{A}_2 \vee \mathcal{B}) - H_\mu(\mathcal{B}) = H_\mu(\mathcal{A}_2|\mathcal{B}).$$

□

1.6. Lemme. *L'entropie est décroissante par rapport au conditionnement :*
Si $\mathcal{B}_1 \preceq \mathcal{B}_2$, alors $H_\mu(\mathcal{A}|\mathcal{B}_2) \leq H_\mu(\mathcal{A}|\mathcal{B}_1)$.

Démonstration. On reprend la preuve du lemme 1.4.
Tout $B \in \mathcal{B}_1$ est une union d'éléments $C \in \mathcal{B}_2$:

$$\sum_{\substack{C \in \mathcal{B}_2 \\ C \subset B}} \frac{\mu(C)}{\mu(B)} \sum_{A \in \mathcal{A}} -\frac{\mu(A \cap C)}{\mu(C)} \log \frac{\mu(A \cap C)}{\mu(C)} = \sum_{A \in \mathcal{A}} \sum_{\substack{C \in \mathcal{B}_2 \\ C \subset B}} \frac{\mu(C)}{\mu(B)} \phi(\frac{\mu(A \cap C)}{\mu(C)})$$

$$\leq \sum_{A \in \mathcal{A}} \phi\left(\sum_{\substack{C \in \mathcal{B}_2 \\ C \subset B}} \frac{\mu(C)}{\mu(B)} \frac{\mu(A \cap C)}{\mu(C)}\right) = \sum_{A \in \mathcal{A}} \phi\left(\frac{\mu(A \cap B)}{\mu(B)}\right)$$

On en déduit :

$$\begin{aligned} H_\mu(\mathcal{A}|\mathcal{B}_2) &= \sum_{B \in \mathcal{B}_1} \mu(B) \sum_{\substack{C \in \mathcal{B}_2 \\ C \subset B}} \frac{\mu(C)}{\mu(B)} \sum_{A \in \mathcal{A}} -\frac{\mu(A \cap C)}{\mu(C)} \log \frac{\mu(A \cap C)}{\mu(C)} \\ &\leq \sum_{B \in \mathcal{B}_1} \mu(B) \sum_{A \in \mathcal{A}} -\frac{\mu(A \cap B)}{\mu(B)} \log \frac{\mu(A \cap B)}{\mu(B)} \\ &= H_\mu(\mathcal{A}|\mathcal{B}_1). \end{aligned}$$

□

1.7. PROPOSITION. *Si la suite de partitions $\mathcal{B}_1 \preceq \mathcal{B}_2 \preceq \mathcal{B}_3 \preceq \dots$ engendre la tribu des boréliens, alors, pour toute partition finie $\mathcal{A}$, on a :*

$$\lim_n H_\mu(\mathcal{A}|\mathcal{B}_n) = 0$$

1.8. LEMME. *Soit $\mathcal{B}_1 \subset \mathcal{B}_2 \subset$ une suite croissante de partitions dont l'union engendre la tribu des boréliens. Alors pour tout $A \in \mathcal{B}$ et $\epsilon > 0$, il existe $n \geq 1$ et un partie $\mathcal{B}$-mesurable B avec $\mu(A\Delta B) < \epsilon$.*

DÉMONSTRATION. Il suffit de vérifier que

$$\mathcal{C} := \{C \in \mathcal{B} : \forall \epsilon > 0\ \exists n \geq 1\ \exists B \in \mathcal{B}_n\ \mu(C\Delta B) < \epsilon\}$$

est une tribu qui contient $\bigcup_n \mathcal{B}_n$. □

DÉMONSTRATION DE LA PROPOSITION. On remarque que

$$H_\mu(\mathcal{A}|\mathcal{B}_n) = \sum_{A\in\mathcal{A}} \int_X \phi(\mathbb{E}_\mu(1_A|\mathcal{B}_n))\, d\mu$$

où $\phi(t) = -t\log t$ et l'espérance conditionnelle par rapport à la partition finie est simplement :

$$\mathbb{E}_\mu(1_A|\mathcal{B}_n) = \frac{\mu(A \cap \mathcal{B}_n(x))}{\mu(\mathcal{B}_n(x))}$$

Le théorème de convergence des martingales donne :

$$\mathbb{E}_\mu(1_A|\mathcal{B}_n)(x) \to 1_A(x) \text{ pour } \mu\text{-presque tout } x$$

Mais $\phi(0) = \phi(1) = 0$ et ϕ est continue. ϕ est également bornée, le théorème de convergence dominée donne donc :

$$\sum_{A\in\mathcal{A}} \int_X \phi(\mathbb{E}_\mu(1_A|\mathcal{B}_n))\, d\mu \to 0$$

□

1.9. LEMME. *Si $f : (X,\mu) \to (Y,\nu)$ préserve la mesure : On a, pour toute partition $\mathcal{B}$ de Y,*

$$H_\mu(f^{-1}(\mathcal{B})) = H_{f_*(\mu)}(\mathcal{B}).$$

2. Entropie de Kolmogorov-Sinai

2.1. DÉFINITION. L'*entropie de Kolmogorov-Sinaï* d'une dynamique probabiliste $(X,\mathcal{B},\mu,f)$ est :

$$h(f,\mu) := \sup_{\mathcal{A} \text{ partition}} h(f,\mu,\mathcal{A}) \text{ avec } h(f,\mu,\mathcal{A}) := \inf_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{A}^n)$$

La borne inférieure ci-dessus est aussi une limite :

2.2. LEMME (Fekete). *Soit $a_1, a_2, \dots$ des réels positifs vérifiant $a_{n+m} \leq a_n + a_m$. Alors :*

$$\lim_n \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}$$

Démonstration. On a évidemment $\liminf_{n\to\infty}(a_n/n) \geq \inf_{n\geq 1}(a_n/n)$. Il suffit donc de majorer la limite supérieure, i.e., de montrer que :

$$\forall m \geq 1\ \forall \epsilon > 0\ \exists N\ \forall n \geq N \quad \frac{a_n}{n} \leq \frac{a_m}{m} + \epsilon.$$

Pour cela on écrit la division euclidienne $n = mq + r$ avec $q, r \in \mathbb{N}$, $0 \leq r < m$ et on calcule :

$$\frac{a_n}{n} \leq \frac{1}{mq+r}\,(q \cdot a_m + a_r) \leq \frac{a_m}{m} + \frac{\max_{0\leq s<m} a_s}{n}$$

Il suffit de prendre $N \geq \epsilon^{-1}\max_{0\leq s<m}(a_s)$. □

2.3. Théorème. *Si (f,μ) et (g,ν) sont des dynamiques probabilistes conjuguées en mesure alors*

$$h(f,\mu) = h(g,\nu).$$

Plus précisémenet si $\pi : (f,\mu) \to (g,\nu)$ est une semiconjugaison en mesure alors :

$$h(f,\mu) \geq h(g,\nu).$$

Démonstration. On remarque :

$$h(f,\mu) \geq \sup_{\mathcal{B}} h(f,\mu,\pi^{-1}(\mathcal{B})) = \sup_{\mathcal{B}} h(g,\nu,\mathcal{B}) = h(g,\nu).$$

vu

$$h(f,\mu,\pi^{-1}(\mathcal{B})) = \inf_{n\geq 1}\frac{1}{n}H_\mu((\pi^{-1}\mathcal{B})^n) = \inf_{n\geq 1}\frac{1}{n}H_\mu(\pi^{-1}(\mathcal{B}^n)) = \inf_{n\geq 1}\frac{1}{n}H_\nu(\mathcal{B}^n)$$

en appliquant le lemme 1.9 à la semiconjugaison. □

2.4. Définition. Une partition $\mathcal{A}$ est génératrice si $\bigvee_{n\in\mathbb{Z}} f^{-n}\mathcal{A} = \mathcal{B}$.

2.5. Proposition. *Sur un espace de probabilité standard, ceci est équivalent à l'injectivité modulo μ du codage associé :*

$$c : X \to \mathcal{A}^{\mathbb{Z}}, \quad x \mapsto (\mathcal{A}(f^n x))_{n\in\mathbb{Z}}$$

Autrement dit, il existe $X' \subset X$ avec $\mu(X') = 1$ tel que : $x, y \in X'$ et $c(x) = c(y)$ implique $x = y$.

2.6. Théorème (Sinai). *Si $\mathcal{A}$ est une partition génératrice, alors*

$$h(f,\mu) = h_\mu(f,\mu,\mathcal{A}).$$

Le théorème de Sinaï découlera immédiatement de la proposition 1.7 et de la proposition suivante.

On pose $\mathcal{B}_a^b := \bigvee_{a\leq k\leq b-1} f^{-k}(\mathcal{B})$.

2.7. Proposition. *Soit $\mathcal{A}, \mathcal{B}$ des partitions et $N \geq 1$. On a :*

$$h(f,\mu,\mathcal{A}) \leq h(f,\mu,\mathcal{B}) + H_\mu(\mathcal{A}|\mathcal{B}_{-N}^N).$$

Démonstration. On a :

$$H_\mu(\mathcal{A}^n) \leq H_\mu(\mathcal{B}^n) + H_\mu(\mathcal{A}^n|\mathcal{B}^n)$$

Il suffit donc montrer que

$$\limsup_n \frac{1}{n}H_\mu(\mathcal{A}^n|\mathcal{B}^n) \leq H_\mu(\mathcal{A}|\mathcal{B}_{-N}^N)$$

On calcule :

$$H_\mu(\mathcal{A}^n|\mathcal{B}^n) \leq \sum_{k=0}^{n-1} H_\mu(f^{-k}\mathcal{A}|\mathcal{B}^n) \leq \sum_{k=0}^{n-1} H_\mu(\mathcal{A}|f^k\mathcal{B}^n) \leq \sum_{k=0}^{n-1} H_\mu(\mathcal{A}|\mathcal{B}_{-k}^{n-k})$$
$$\leq 2N \cdot H_\mu(A) + \sum_{k=N}^{n-N} H_\mu(\mathcal{A}|\mathcal{B}_N^{-N}) \leq 2N \cdot H_\mu(A) + (n-2N)\epsilon$$

On a donc :

$$h(f,\mu,\mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) \leq h(f,\mu,\mathcal{B}) + \epsilon.$$

□

2.8. Exercice. Pour P une probabilité sur $\{0,1,\ldots,d-1\}$ où $d \geq 1$. Soit $(\Sigma_d, \mathcal{B}, P^{\mathbb{Z}}, \sigma_d)$ le schéma de Bernoulli correspondant. On a :

$$h(f, P^{\mathbb{Z}}) = -\sum_{i=0}^{d-1} P(i) \log P(i).$$

On en déduit facilement :

2.9. Exercice. Posons $\mu_t := (t\delta_1 + (1-t)\delta_0)^{\mathbb{Z}}$. Alors le les schémas de Bernoulli (σ_2, μ_t) et (σ_2, μ_s) sont conjugués en mesure si et seulement si $s = t$ ou $s = t - 1$.

Le programme de l'examen s'arrête ici.

3. Ebauche de quelques résultats fondamentaux

3.1. Entropie, dilatation et dimension. Pour les dynamiques différentiables, l'entropie apparaît comme le produit de l'expansion et de la dimension, convenablement définies au niveau des mesures. La formule suivante est le précurseur d'une vaste théorie

3.1. Théorème. *Soit $T_A : \mathbb{T}^d \to \mathbb{T}^d$ un automorphisme du tore défini par $A \in GL(d,\mathbb{Z})$. Soit $\lambda_1, \ldots, \lambda_d$ les valeurs propres de A répétées selon leur multiplicité ($\dim \operatorname{Ker}(A - \lambda I)$). Alors*

$$h(f, m_{\mathbb{T}^d}) = \sum_{i=1}^{d} \log \max(|\lambda_i|, 1).$$

3.2. Entropie et codage.

3.2. Théorème (Shannon-McMillan-Breiman). *Soit (f,μ) une dynamique probabiliste et $\mathcal{A}$ une partition mesurable et finie. Alors, μ-p.p. et dans L^1 :*

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu(\mathcal{A}^n(x)) = h(x)$$

où h est une fonction invariante modulo μ et d'intégrale $h(f,\mu,\mathcal{A})$.

Ce résultat implique qu'il y a de l'ordre de $e^{nh(f,\mu,\mathcal{A})}$ "atomes significatifs" dans $\mathcal{A}^n$ et qu'au moins sur une échelle logarithmique, ils ont tous la même mesure :

3.3. Exercice. Supposons de plus (f,μ) ergodique et $\epsilon > 0$. Pour tout n assez grand, il existe une collection d'éléments A de la partition $\mathcal{A}^n$, chacun de mesure :

$$e^{-n(h(f,\mu,\mathcal{A})+\epsilon)} \leq \mu(A) \leq e^{-n(h(f,\mu,\mathcal{A})-\epsilon)}$$

et dont l'union est de mesure au moins $1-\epsilon$.

3.4. Théorème (Jewett-Krieger). *Soit* (X, f, μ) *une dynamique probabiliste quelconque ergodique et d'entropie finie.*

Il existe une partition génératrice $\mathcal{A}$ *de cardinal* $\leq e^{h(f,\mu)} + 1$.

Toute dynamique probabiliste ergodique est donc conjuguée en mesure avec un décalage sur $e^{h(f,\mu)} + 1$ symboles muni d'une certaine mesure de probabilité borélienne. L'exercice suivant montre qu'on ne peut remplacer $e^{h(f,\mu)}+1$ par $e^{h(f,\mu)}$.

3.5. Exercice. Soit σ_d le décalage sur $d \geq 1$ symboles.

Montrer que pour toute mesure $\mu \in \mathbb{P}(\sigma_d)$, $h(\sigma_d, \mu) \leq \log d$.

Montrer que le cas d'égalité ne se produit que si $\mu = P^{\mathbb{Z}}$ où P est la probabilité uniforme sur $\{0, \ldots, d-1\}$.

3.3. Facteurs de Bernoulli. Toute l'entropie provient de schémas de Bernoulli :

3.6. Théorème (Sinai). *Soit* (X, f, μ) *une dynamique probabiliste quelconque d'entropie finie.*

Un schéma de Bernoulli $(\sigma_d, P^{\mathbb{Z}})$ *est un facteur mesuré de* (f, μ) *si et seulement si* $h(\sigma_d, P^{\mathbb{Z}}) \leq h(f, \mu)$.

3.7. Exercice. Dans la situation du théorème, montrer qu'il existe une partition $\mathcal{A}$ de X telle que : (1) $H_\mu(\mathcal{A}) = h(f, \mu)$; (2) le processus $X_n : X \to \mathcal{A}^{\mathbb{Z}}$ défini par $X_n(x) = A$ si et seulement si $f^n(x) \in A$ est i.i.d. (les X_n sont de même loi et ils sont indépendants).

3.8. Remarque. Le théorème de Sinaï montre qu'une dynamique (f, μ) d'entropie finie est une extension d'un schéma de Bernoulli de même entropie. Il n'est pas vrai qu'on puisse décomposer comme un produit direct d'un schéma de Bernoulli et d'une dynamique d'entropie nulle. Un résultat plus faible a été récemment obtenu par T. Austin.

3.4. Classification des schémas de Bernoulli. Le résultat suivant est difficile dès que $\max(d, e) \geq 3$.

3.9. Théorème (Ornstein). *Deux schémas de Bernoulli* $(\Sigma_d, \mathcal{B}_d, P^{\mathbb{Z}}, \sigma_d)$ *et* $(\Sigma_e, \mathcal{B}_e, Q^{\mathbb{Z}}, \sigma_q)$ *avec* $d, e \geq 1$*,* P *et* Q *des probabilités sur* $\{0, \ldots, d-1\}$ *et* $\{0, \ldots, e-1\}$ *sont conjugués en mesure si et seulement si leurs entropies sont égales.*

La théorie sous-jacente est très profonde et aboutit à une classification en mesure de beaucoup de systèmes qui s'avèrent être conjugués à des schémas de Bernoulli.

Par exemple, si A est sans valeurs propres de module 1, alors $(T_A, m_{\mathbb{T}^d})$ est conjugué en mesure avec tout schéma de Bernoulli d'égale entropie.

3.10. Exercice. Trouver un schéma de Bernoulli conjugué en mesure avec $(T_{\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}}, m_{\mathbb{T}^2})$. Ce schéma est-il unique ? Quel nombre minimal de symboles doit-il avoir ?

Bibliographie

[1] T. Downarowicz, Entropy in dynamical systems, Cambridge University Press, 2011.

[2] A. Katok, B. Hasselblatt, An introduction to the modern theory of dynamical systems, Cambridge University Press, 1993.

[3] A. Ya. Khinchin, Continued fractions, Phenix press, 1964.

[4] U. Krengel, Ergodic theorem, De Gruyter, 1985.

[5] W. De Melo, S. Van Strien, One-dimensional dynamics, Springer, 1993.