Systèmes dynamiques II

Patrice Le Calvez

Cours fondamental II, Master 2

2013-2014

1.1 Introduction

Les décalages de Bernouilli σ_p et σ_q, définis respectivement sur $\{1,\ldots,p\}^{\mathbf{N}}$ et $\{1,\ldots,q\}^{\mathbf{N}}$ ne sont pas topologiquement conjugués si $p \neq q$, il suffit de remarquer par exemple qu'il n'ont pas le même nombre de points fixes. Si on munit chacun des ensembles de la mesure équidistribuée, on obtient deux systèmes dynamiques mesurés et on peut se demander si ces deux systèmes sont conjugués (les points périodiques formant de chaque côté un ensemble dénombrable et donc de mesure nulle peuvent être oubliés). En 1958, Kolmogorov a prouvé que ce n'était pas le cas, en utilisant la notion d'*entropie*, inspirée des travaux de Shanon sur la théorie de l'information, notion améliorée ensuite par Sinai en 1959. On va s'intéresser ici aux principales propriétés de l'entropie.

1.2 Entropie d'une partition

On suppose dans cette section que l'ensemble X est muni d'une σ-algèbre $\mathcal{B}$ et d'une mesure de probabilité μ sur $\mathcal{B}$. Une *partition mesurable* de X est une famille finie $\mathcal{P} = (P_i)_{i\in I}$ d'éléments de $\mathcal{B}$ tels que

- $X = \bigcup_{i\in I} P_i$;
- $\mu(P_i \cap P_{i'}) = 0$ si $i \neq i'$..

Pour tout ensemble $P \in \mathcal{B}$ on écrira $P \in \mathcal{P}$, s'il existe $i \in I$ tel que $P = P_i$.

On dira qu'une partition $\mathcal{Q}$ est *plus fine* que $\mathcal{P}$ ou (*raffine* $\mathcal{P}$) et on écrira $\mathcal{P} \preceq \mathcal{Q}$, si pour tout $Q \in \mathcal{Q}$ il existe $P \in \mathcal{P}$ tel que $\mu(Q \setminus P) = 0$. On obtient un préordre sur l'ensemble des partitions mesurables et on notera $\sim$ la relation d'équivalence associée. Remarquons que deux partitions $\mathcal{P}$ et $\mathcal{Q}$ sont équivalentes si et seulement si, pour tout $Q \in \mathcal{Q}$ il existe $P \in \mathcal{P}$ tel que $\mu(Q\Delta P) = 0$. Remarquons également que toute partition mesurable Q est équivalente à une partition mesurable $\mathcal{P} = (P_i)_{i\in I}$, où les P_i sont disjoints deux à deux (on aurait pu, dans ce qui suit, se restreindre à ces partitions). Remarquons également qu'il existe une bijection naturelle entre ces partitions (au sens fort) et les sous-algèbres finies de $\mathcal{B}$. Enfin, notons que toute partition raffine la partition $\{X\}$ constituée de l'unique élément X.

Si $(\mathcal{P}^j)_{1\leq j\leq n}$ est une famille finie de partitions mesurables de X, on peut définir la partition mesurable $\bigvee_{j=1}^{n} \mathcal{P}^j$ dont les éléments sont les $\bigcap_{1\leq j\leq n} P^j$, où $P^j \in \mathcal{P}^j$. Elle raffine chaque $\mathcal{P}^j$.

Considérons maintenant la fonction

$$\begin{aligned} \phi\ :&\ [0,1] \to [0,+\infty[\\ x \neq 0 &\mapsto -x\ln x\ . \\ 0 &\mapsto 0 \end{aligned}$$

Elle est continue, concave, s'annule en 0 et 1 et atteint son maximum $1/e$ en $1/e$.

Définition. L'*entropie* d'une partition $\mathcal{P} = (P_i)_{i\in I}$ est

$$H(\mathcal{P}) = \sum_{i\in I} \phi(\mu(P_i)) = \sum_{i\in I} -\mu(P_i)\ln\mu(P_i).$$

C'est la moyenne de la *fonction information* $I_{\mathcal{P}}$ définie presque partout par l'égalité $I_{\mathcal{P}}(x) = -\ln \mu(P_i)$ si $x \in P_i$. Ceci rend naturelle la définition suivante de l'entropie conditionnelle d'une partition par rapport à une autre partition.

Définition. Si $\mathcal{P} = (P_i)_{i\in I}$ et $\mathcal{Q} = (Q_j)_{j\in J}$ sont deux partitions mesurables, l'*entropie de* $\mathcal{P}$ *relativement à* $\mathcal{Q}$ est

$$H(\mathcal{P}|\mathcal{Q}) = \sum_{i\in I, j\in J'} \mu(Q_j)\phi\left(\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}\right) = \sum_{i\in I, j\in J'} -\mu(P_i\cap Q_j)\ln\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}$$

où $J' = \{j \in J \mid \mu(Q_j) \neq 0\}$.

Remarque On peut vérifier que $H(\mathcal{P}|\mathcal{Q}) = 0$ si et seulement si $P \preceq Q$, en particulier on a $H(\mathcal{P}|\mathcal{P}) = 0$. On peut vérifier également que $H(\mathcal{P}, \{X\}) = H(\mathcal{P})$.

Énonçons les principales propriétés de l'entropie.

PROPOSITION 1.2.1 : *Supposons que* $\mathcal{P} = (P_i)_{i\in I}$, $\mathcal{Q} = (Q_j)_{j\in J}$ *et* $\mathcal{S} = (S_k)_{k\in K}$ *sont des partitions mesurables. On a alors les propriétés suivantes :*

i) $H(\mathcal{P}) \leq \ln \sharp I$ *avec égalité si et seulement si tous les* $\mu(P_i)$ *sont égaux ;*

ii) $H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S})$ *et* $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P})$ *;*

iii) $H(\mathcal{P}\vee\mathcal{Q}|\mathcal{S}) = H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) + H(\mathcal{Q}|\mathcal{S})$ *et* $H(\mathcal{P}\vee\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q})$ *;*

iv) *si* $Q \preceq P$, *alors* $H(\mathcal{Q}|\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S})$ *et* $H(\mathcal{Q}) \leq H(\mathcal{P})$ *;*

v) *si* $S \preceq Q$, *alors* $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S})$ *;*

vi) $H(\mathcal{P}) \leq H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q})$ *;*

vii) $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S}) + H(\mathcal{S}|\mathcal{Q})$.

Démonstration. L'assertion **i)** est une conséquence de la concavité stricte de la fonction ln. Posons $I' = \{i \in I \mid \mu(P_i) > 0\}$. On a

$$H(\mathcal{P}) = \sum_{i\in I'} -\mu(P_i)\ln\mu(P_i) = \sum_{i\in I'}\mu(P_i)\ln\frac{1}{\mu(P_i)} \leq \ln\left(\sum_{i\in I'}\mu(P_i)\frac{1}{\mu(P_i)}\right) = \ln(\sharp I') \leq \ln(\sharp I).$$

L'égalité a lieu quand chaque inégalité est une égalité, c'est-à-dire quand on a $\mu(P_i) = \dfrac{1}{\sharp I}$, pour tout $i \in I$.

L'assertion **ii)** est une conséquence de la concavité de ϕ. Fixons $i \in I$ et $k \in K' = \{k \in K \mid \mu(S_k) \neq 0\}$, et posons $J_k = \{j \in J \mid \mu(Q_j \cap S_k) \neq 0\}$. On a

$$\begin{aligned}\sum_{j\in J_k}\mu(Q_j\cap S_k)\phi\left(\frac{\mu(P_i\cap Q_j\cap S_k)}{\mu(Q_j\cap S_k)}\right) &\leq \mu(S_k)\phi\left(\sum_{j\in J_k}\frac{\mu(Q_j\cap S_k)}{\mu(S_k)}\frac{\mu(P_i\cap Q_j\cap S_k)}{\mu(Q_j\cap S_k)}\right)\\ &= \mu(S_k)\phi\left(\frac{\mu(P_i\cap S_k)}{\mu(S_k)}\right).\end{aligned}$$

En sommant sur i et k on obtient

$$H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S}).$$

Appliquant ce qui précède à $S = \{X\}$, on obtient

$$H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}).$$

Pour établir **iii)**, écrivons

$$\begin{aligned}
&H(\mathcal{P} \vee \mathcal{Q}|\mathcal{S}) \\
&= \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln \left(\frac{\mu(P_i \cap Q_j \cap S_k)}{\mu(S_k)} \right) \\
&= \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln \left(\frac{\mu(P_i \cap Q_j \cap S_k)}{\mu(Q_j \cap S_k)} \right) + \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln \left(\frac{\mu(Q_j \cap S_k)}{\mu(S_k)} \right) \\
&= H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{Q}|\mathcal{S}).
\end{aligned}$$

Si on applique cette dernière égalité à $S = \{X\}$, on trouve

$$H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}).$$

Pour obtenir **iv)**, il suffit de remarquer que si $\mathcal{Q} \preceq \mathcal{P}$, alors $\mathcal{P} \vee \mathcal{Q} = \mathcal{P}$, et on a donc

$$H(\mathcal{P}|\mathcal{S}) = H(\mathcal{P} \vee \mathcal{Q}|\mathcal{S}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{Q}|\mathcal{S}) \geq H(\mathcal{Q}|\mathcal{S}).$$

Dans le cas où $S = \{X\}$, on obtient

$$H(\mathcal{P}) \geq H(\mathcal{Q}).$$

Pour obtenir **v)**, il suffit de remarquer que si $\mathcal{S} \preceq \mathcal{Q}$, alors $\mathcal{Q} \vee \mathcal{S} = \mathcal{Q}$, et on a donc

$$H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) \leq H(\mathcal{P}|\mathcal{S}).$$

L'assertion **vi)** se déduit de

$$H(\mathcal{P}) \leq H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}),$$

et l'assertion **vii)** de

$$H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P} \vee \mathcal{S}|\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{S}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S}) + H(\mathcal{S}|\mathcal{Q}).$$

□

On en déduit immédiatement :

Corollaire 1.2.2 : *On obtient une distance D sur l'ensemble des classes d'équivalence de la relation $\sim$, appelée la* distance de Rokhlin*, en posant :*

$$D(\mathcal{P}, \mathcal{Q}) = H(\mathcal{P}, \mathcal{Q}) + H(\mathcal{Q}, \mathcal{P}),$$

pour toutes partitions mesurables $\mathcal{P}$ et $\mathcal{Q}$.

Concluons cette section par un résultat qui sera très utile plus tard. Pour toute suite croissante $(\mathcal{P}_m)_{m\geq 0}$ de partitions mesurables de X, notons $\bigvee_{m=0}^{+\infty} \mathcal{P}_m$ la σ-algèbre engendrée par

$\bigcup_{m\geq 0}\mathcal{P}_m$ et $\overline{\bigvee_{m=0}^{+\infty}\mathcal{P}_m}$ la σ-algèbre formée des éléments $A\in\mathcal{B}$ tels qu'il existe $B\in\bigvee_{m=0}^{+\infty}\mathcal{P}_m$ vérifiant $\mu(A\Delta B)=0$. On dira que la famille $(\mathcal{P}_m)_{m\geq 0}$ est *génératrice* si $\overline{\bigvee_{m=0}^{+\infty}\mathcal{P}_m}=\mathcal{B}$.

PROPOSITION 1.2.3 : *Si $(\mathcal{P}_m)_{m\geq 0}$ est une suite croissante génératrice de X, alors pour toute partition mesurable $\mathcal{P}$ on a $\lim_{m\to+\infty}H(\mathcal{P},\mathcal{P}_m)=0$.*

Démonstration. Commençons par quelques lemmes.

LEMME 1.2.4 : *Pour tout $A\in\mathcal{B}$, il existe un entier $m\geq 0$ et un ensemble $B\in\mathcal{B}$, réunion d'éléments de $\mathcal{P}_m$, tels que $\mu(A\Delta B)<\varepsilon$.*

Démonstration. Il faut montrer que l'ensemble $\mathcal{C}$ des éléments $A\in\mathcal{B}$ qui vérifie la condition du lemme est une σ-algèbre qui contient tous les $\mathcal{P}_m$ et qui vérifie la condition suivante :

- si $A\in\mathcal{C}$ et $A'\in\mathcal{B}$ vérifient $\mu(A\Delta A')=0$, alors $A'\in\mathcal{C}$.

Le fait que $\mathcal{C}$ contient chaque $\mathcal{P}_m$ est évident, la fait que la condition précédente est vérifiée se déduit immédiatement de la relation ensembliste suivante :

$$A'\Delta B\subset(A'\Delta A)\cup(A\Delta B).$$

Prouvons maintenant que $\mathcal{C}$ est une σ-algèbre. Cet ensemble contient X et est stable par passage au complémentaire puisque $(X\setminus A)\Delta(X\setminus B)=A\Delta B$. Il reste à montrer qu'il est stable par union dénombrable. Soit $(A_r)_{r\geq 1}$ une famille d'éléments de $\mathcal{B}$ et $\varepsilon>0$. On peut trouver r_0 tel que

$$\mu\left(\bigcup_{r\geq 1}A_r\setminus\bigcup_{1\leq r\leq r_0}A_r\right)\leq\frac{\varepsilon}{2}.$$

Pour tout $r\in\{1,\dots,r_0\}$, on peut trouver un entier $m_r\geq 0$ et un ensemble B_r, réunion d'éléments de $\mathcal{P}_{m_r}$, tel que $\mu(A_r\Delta B_r)<\varepsilon/2r_0$. L'ensemble $\bigcup_{1\leq r\leq r_0}B_r$ est une union d'éléments de $\mathcal{P}_m$, où $m=\sup_{1\leq r\leq r_0}m_r$. Remarquons maintenant que

$$\left(\bigcup_{r\geq 1}A_r\right)\Delta\left(\bigcup_{1\leq r\leq r_0}B_r\right)\subset\left(\bigcup_{r\geq 1}A_r\setminus\bigcup_{1\leq r\leq r_0}A_r\right)\cup\left(\bigcup_{1\leq r\leq r_0}A_r\Delta B_r\right),$$

ce qui implique que

$$\mu\left(\left(\bigcup_{r\geq 1}A_r\right)\Delta\left(\bigcup_{1\leq r\leq r_0}B_r\right)\right)\leq\varepsilon.$$

□

LEMME 1.2.5 : *Pour tout $r\geq 1$ et tout $\varepsilon>0$, il existe $\eta>0$ tel que pour toutes partitions mesurables $\mathcal{P}=(P_i)_{i\in I}$, $\mathcal{Q}=(Q_i)_{i\in I}$ vérifiant $\sharp I=r$ et $\mu(P_i\Delta Q_i)\leq\eta$, on a $H(\mathcal{P}|\mathcal{Q})\leq\varepsilon$.*

Démonstration. Choisissons η assez petit pour que ϕ soit croissante sur $[0,\eta]$ et décroissante sur $[1-r\eta,1]$. Écrivons

$$H(\mathcal{P}|\mathcal{Q})=\sum_{i\neq j}\mu(Q_j)\phi\left(\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}\right)+\sum_i\mu(Q_i)\phi\left(\frac{\mu(P_i\cap Q_i)}{\mu(Q_i)}\right).$$

Si $i \neq j$, alors $\mu(P_i \cap Q_j) \leq \mu(P_i \Delta Q_i) \leq \eta$. Ainsi, on a

$$\begin{aligned}
\mu(Q_j)\phi\left(\frac{\mu(P_i \cap Q_j)}{\mu(Q_j)}\right) &= -\mu(P_i \cap Q_j)\ln\left(\frac{\mu(P_i \cap Q_j)}{\mu(Q_j)}\right) \\
&= -\mu(P_i \cap Q_j)\ln\mu(P_i \cap Q_j) + \mu(P_i \cap Q_j)\ln\mu(Q_j) \\
&\leq -\mu(P_i \cap Q_j)\ln\mu(P_i \cap Q_j) \\
&\leq -\eta\ln(\eta) = \phi(\eta).
\end{aligned}$$

Remarquons maintenant que

$$1 - \sum_i \mu(P_i \cap Q_i) = \sum_i \mu(P_i) - \mu(P_i \cap Q_i) = \sum_i \mu(P_i \setminus P_i \cap Q_i) \leq r\eta.$$

Le fait que ϕ est concave implique que

$$\sum_i \mu(Q_i)\phi\left(\frac{\mu(P_i \cap Q_i)}{\mu(Q_i)}\right) \leq \phi\left(\sum_i \mu(P_i \cap Q_i)\right) \leq \phi(1 - r\eta).$$

On a donc

$$H(\mathcal{P}|\mathcal{Q}) \leq r(r-1)\phi(\eta) + \phi(1 - r\eta).$$

Si η est suffisamment petit, alors le terme à droite est inférieur ou égal à ε. □

Démonstration de la proposition 1.2.3. Soit $\mathcal{P} = (P_i)_{1 \leq i \leq r}$ une partition mesurable. Donnons nous $\varepsilon > 0$. Commençons par considèrer le réel η defini par le lemme 1.2.5. Fixons ensuite δ suffisamment petit pour que l'on ait

$$(1 + r(r-1))\delta \leq \eta$$

et

$$(r-1)(1 + r(r-1))\delta \leq \eta,$$

(la seconde condition implique bien sûr la première si $r > 1$).

On peut trouver un entier $m_0 \geq 0$ et pour tout $i \in \{1 \ldots r\}$, un ensemble B_i, union d'éléments de $\mathcal{P}_{m_0}$, tel que $\mu(P_i \Delta B_i) < \delta$. On va construire une partition $\mathcal{Q} = (Q_i)_{1 \leq i \leq r}$, où chaque Q_i est une union d'éléments de $\mathcal{P}_{m_0}$, tel que $\mu(P_i \Delta Q_i) \leq \eta$. Puisqu'on a $\mathcal{Q} \preceq \mathcal{P}_m$, pour tout $m \geq m_0$, on en déduit que $H(\mathcal{P}, \mathcal{P}_m) \leq H(\mathcal{P}, \mathcal{Q}) \leq \varepsilon$.

Remarquons que si $i \neq j$, alors à un ensemble de mesure nulle près on a

$$B_i \cap B_j \subset (P_i \Delta B_i) \cup (P_j \Delta B_j),$$

ce qui implique que

$$\mu(B_i \cap B_j) \leq 2\delta.$$

On en déduit que l'ensemble

$$B = \cup_{i \neq j} B_i \cap B_j$$

vérifie

$$\mu(B) \leq r(r-1)\delta.$$

Considérons la partition $\mathcal{Q} = (Q_i)_{1 \leq i \leq r}$ où

$$\begin{cases} Q_i = B_i \setminus B & \text{si } i < r, \\ Q_r = X \setminus \bigcup_{i<r} B_i & , \end{cases}$$

et montrons que $\mu(P_i \Delta Q_i) \leq \eta$ pour tout $i \in \{1, \ldots, r\}$. Si $i < r$, alors

$$P_i \Delta Q_i \subset (P_i \Delta B_i) \cup (B_i \Delta Q_i) \subset (P_i \Delta B_i) \cup B,$$

ce qui implique que

$$\mu(P_i \Delta Q_i) \leq \delta + r(r-1)\delta \leq \eta.$$

Le fait que $\mathcal{P}$ et $\mathcal{Q}$ sont des partitions implique que

$$P_r \Delta Q_r \subset \bigcup_{1 \leq i < r} P_i \Delta Q_i,$$

et donc que

$$\mu(P_r \Delta Q_r) \leq (r-1)(1 + r(r-1)\delta) \leq \eta.$$

□

1.3 Entropie d'un système dynamique

On se donne maintenant un système dynamique mesuré $(X, \mathcal{B}, \mu, T)$, où μ est une mesure de probabilité. Remarquons que pour toute partition mesurable $\mathcal{P} = (P_i)_{i \in I}$, on obtient une autre partition mesurable $T^{-1}(\mathcal{P}) = (T^{-1}(P_i))_{i \in I}$. Remarquons également que $h(T^{-1}(\mathcal{P})) = h(\mathcal{P})$ et plus généralement que $h(T^{-1}(\mathcal{P})|T^{-1}(\mathcal{Q})) = h(\mathcal{P}|\mathcal{Q})$ si $\mathcal{Q}$ est une autre partition mesurable.

PROPOSITION 1.3.1 : *Si $\mathcal{P}$ est une partition mesurable de X, alors la suite*

$$\frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)$$

est convergente. Sa limite $h(T, \mathcal{P})$ est l'entropie de T relativement à $\mathcal{P}$, *elle vérifie :*

$$0 \leq h(T, \mathcal{P}) \leq H(\mathcal{P}).$$

Démonstration. La suite $(u_n)_{n \geq 1}$, où $u_n = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)$, est sous-additive. En effet, on a

$$\begin{aligned}
u_{n+m} &= H\left(\bigvee_{i=0}^{n+m-1} T^{-i}(\mathcal{P})\right) \\
&= H\left(\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) \vee T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right) \\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right) \\
&= u_n + u_m.
\end{aligned}$$

Ceci implique que la suite $(u_n/n)_{n \geq 1}$ est convergente et que sa limite est comprise entre 0 et u_1. En effet, fixons $p \geq 1$. Pour tout $n \geq p$, notons $n = kp + r$ la division euclidienne de n par p et remarquons que

$$\frac{u_n}{n} \leq \frac{u_{kp} + u_r}{n} \leq \frac{u_{kp}}{kp} + \frac{u_r}{n} \leq \frac{u_p}{p} + \frac{u_r}{n}.$$

Ceci implique que

$$\limsup_{n\to+\infty} \frac{u_n}{n} \le \frac{u_p}{p} \quad (\le u_1),$$

et donc que

$$\limsup_{n\to+\infty} \frac{u_n}{n} \le \inf_{p\ge 1} \frac{u_p}{p} \le \liminf_{n\to+\infty} \frac{u_n}{n}.$$

□

L'entropie $h(T,\mathcal{P})$ peut être définie de la façon équivalente suivante :

PROPOSITION 1.3.2 : *Si $\mathcal{P}$ est une partition mesurable de X, alors la suite*

$$H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right)$$

est décroissante et converge vers $h(T,\mathcal{P})$.

Démonstration. La décroissance de cette suite est une conséquence de l'assertion **v)** de la proposition 1.2.1. Notons l sa limite. Remarquons que

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \mathcal{P} \vee \left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right),$$

ce qui implique

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &= H\left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right)\\
&= H\left(\bigvee_{i=0}^{n-2} T^{-i}(\mathcal{P})\right) + H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right)\\
&= H(\mathcal{P}) + \sum_{k=1}^{n-1} H\left(\mathcal{P}|\bigvee_{i=1}^{k} T^{-i}(\mathcal{P})\right).
\end{aligned}$$

Par le théorème de Césaro, on obtient

$$h(T,\mathcal{P}) = \lim_{n\to+\infty} H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right).$$

□

Énonçons les propriétés principales de l'entropie relativement aux partitions.

PROPOSITION 1.3.3 : *Soient $\mathcal{P}$ et $\mathcal{Q}$ des partitions mesurables. On a alors les propriétés suivantes :*

i) $h(T,\mathcal{P}) \le h(T,\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q})$ *;*

ii) $h(T,\mathcal{Q}) \le h(T,\mathcal{P})$ *si* $\mathcal{Q} \preceq \mathcal{P}$ *;*

iii) $h(T,\bigvee_{i=0}^{m} T^{-i}(\mathcal{P})) = h(T,\mathcal{P})$ *pour tout* $m \ge 0$ *;*

iv) $h(T^m,\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})) = mh(T,\mathcal{P})$ *pour tout* $m \ge 1$ *;*

v) $h(T,T^{-1}(\mathcal{P})) = h(T,\mathcal{P})$ *;*

v) $h(T^{-1},\mathcal{P}) = h(T,\mathcal{P})$ *si* T *est inversible.*

Démonstration. Pour prouver **i)**, remarquons que

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) | \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + \sum_{i=0}^{n-1} H\left(T^{-i}(\mathcal{P}) | \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + \sum_{i=0}^{n-1} H(T^{-i}(\mathcal{P}) | T^{-i}(\mathcal{Q}))\\
&= H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + nH(\mathcal{P}|\mathcal{Q}).
\end{aligned}$$

Il reste à diviser par n et à faire tendre n vers l'infini pour obtenir **i)**.

On en déduit **ii)** puisque $H(\mathcal{Q}|\mathcal{P}) = 0$ si $\mathcal{Q} \preceq \mathcal{P}$. On peut également déduire **ii)** des relations

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q}) \preceq \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}).$$

Pour obtenir **iii)** posons $\mathcal{S} = \bigvee_{i=0}^{m} T^{-i}(\mathcal{P})$ et remarquons qu'on a

$$\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{P}) \sim \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{S}),$$

ce qui implique

$$\begin{aligned}
h(T,\mathcal{P}) &= \lim_{n\to+\infty} \frac{1}{n+m} H(\left(\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{P})\right)\\
&= \lim_{n\to+\infty} \frac{n}{n+m}\,\frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{S})\right) = h(T,\mathcal{S}).
\end{aligned}$$

Pour établir **iv)**, remarquons qu'on a

$$\bigvee_{i=0}^{n-1} T^{-im}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right) = \bigvee_{i=0}^{nm-1} T^{-i}(\mathcal{P}),$$

ce qui implique

$$\begin{aligned}
h(T^m, \bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})) &= \lim_{n\to+\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-im}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right)\\
&= \lim_{n\to+\infty} \frac{m}{nm} H\left(\bigvee_{i=0}^{nm-1} T^{-i}(\mathcal{P})\right) = mh(T,\mathcal{P}).
\end{aligned}$$

L'assertion **v)** se déduit de

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}(T^{-1}(\mathcal{P}))\right) = H\left(T^{-1}\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)\right) = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right),$$

et l'assertion **vi)** de

$$H\left(\bigvee_{i=0}^{n-1} T^{i}(\mathcal{P})\right) = H\left(T^{n}\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)\right) = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right).$$

□

On définit alors l'entropie métrique $h(T) \in [0, +\infty]$ comme le supremum des entropies relativement à toutes les partitions :

$$h(T) = \sup\{h(T, \mathcal{P}) \ , \ \mathcal{P} \text{ est une partition mesurable de } X\}.$$

PROPOSITION 1.3.4 : *On a les propriétés suivantes :*

i) $h(T^n) = nh(T)$, *pour tout* $n \geq 0$ *;*

ii) *si* T *est inversible, alors* $h(T^k) = |k|h(T)$, *pour tout* $k \in \mathbf{Z}$.

Démonstration. L'assertion **i)** est évidemment vraie si $n = 0$. En effet, pour toute partition mesurable $\mathcal{P}$ et tout $n \geq 1$, on a $\bigvee_{i=0}^{n-1} \mathrm{Id}_X^{-i}(\mathcal{P}) \sim \mathcal{P}$, ce qui implique que

$$h(\mathrm{Id}_X, \mathcal{P}) = \lim_{n\to+\infty} \frac{1}{n} H(\mathcal{P}) = 0.$$

Supposons donc $n \geq 1$ et remarquons que pour toute partition mesurable $\mathcal{P}$, on a

$$h(T^n, \mathcal{P}) \leq h\left(T^n, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) = nh(T, \mathcal{P}),$$

ce qui nous donne

$$h(T^n, \mathcal{P}) \leq nh_\mu(T), \ \ nh(T, \mathcal{P}) \leq h_\mu(T^n)$$

puis

$$h(T^n) \leq nh_\mu(T), \ \ nh(T) \leq h_\mu(T^n).$$

Pour obtenir **ii)**, il suffit de remarquer que pour toute partition mesurable $\mathcal{P}$, on a $h(T^n, \mathcal{P}) = h(T^{-n}, \mathcal{P})$.

□

Nous allons prouver maintenant que l'entropie est un invariant de conjugaison. Si l'on veut spécifier la mesure μ, on peut écrire $H_\mu(\mathcal{P})$ pour l'entropie d'une partition mesurable $\mathcal{P}$; si $T : X \to X$ est mesurable et préserve μ, on peut écrire $h_\mu(T, \mathcal{P})$ pour l'entropie relativement à $\mathcal{P}$ et $h_\mu(T)$ pour l'entropie.

PROPOSITION 1.3.5 : *Soient* $(X, \mathcal{B}, \mu, T)$ *et* $(Y, \mathcal{C}, \nu, S)$ *deux systèmes dynamiques mesurés, tels que* $\mu(X) = \nu(Y) = 1$. *Supposons qu'il existe une application mesurable* $R : X \to Y$ *telle que* $R \circ T = S \circ R$ *et* $R_*(\mu) = \nu$. *Alors* $h_\nu(S) \leq h_\mu(T)$. *De plus il y a égalité si* R *est inversible et* R^{-1} *est mesurable.*

Démonstration. Pour toute partition mesurable $\mathcal{P}$ de Y, et tout entier $n \geq 0$, on a

$$\bigvee_{i=0}^{n-1} T^{-i}(R^{-1}(\mathcal{P})) = R^{-1}\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{P})\right),$$

ce qui implique

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(R^{-1}(\mathcal{P}))\right) = H_\nu\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{P})\right).$$

On en déduit

$$h_\nu(S,\mathcal{P}) = h_\mu(T, R^{-1}(\mathcal{P})) \le h_\mu(T),$$

et donc $h_\nu(S) \le h_\mu(T)$. On a bien sûr égalité si R est inversible et R^{-1} est mesurable. □

L'inégalité peut bien sûr être stricte. L'exemple le plus simple est le cas où ν est la mesurable de Dirac en un point fixe $y_0 \in Y$ de S et R l'application constante égale à y_0. Dans ce cas on a $h_\nu(S) = 0$. En fait, $h_\nu(S)$ est le supremum des entropies (pour T) parmi les partitions mesurables qui sont images d'une partition de Y. Dans l'exemple précédent il n'y a qu'une seule partition de ce type, la partition triviale $\{X\}$. Notons également que dans la seconde assertion, l'inversibilité n'est nécessaire que presque partout : il existe deux ensembles mesurables $X' \subset X$ et $Y' \subset Y$ tels que $\mu(X') = \nu(Y') = 1$ et tels que $R|_{X'}$ est une bijection de X' sur Y' dont l'inverse est mesurable.

1.4 Partitions génératrices

Nous verrons dans cette courte section comment, grâce à Kolomogorov et Sinai, il suffit parfois d'une partition pour définir l'entropie.

Proposition 1.4.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Si $(\mathcal{P}_m)_{m\ge 0}$ est une suite croissante génératrice de partitions mesurables, alors*

$$h(T) = \lim_{m\to+\infty} h(T, \mathcal{P}_m).$$

Démonstration. La suite $(h(T,\mathcal{P}_m))_{m\ge0}$ est croissante et converge donc dans $[0,+\infty]$. Pour toute partition mesurable $\mathcal{P}$ et tout $\varepsilon > 0$, nous avons vu qu'il était possible de trouver un entier m_0 et une partition $\mathcal{Q}$ vérifiant $\mathcal{Q} \preceq \mathcal{P}_{m_0}$ et $H(\mathcal{P}|\mathcal{Q}) \le \varepsilon$. On en déduit

$$h(T,\mathcal{P}) \le h(T,\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}) \le h(T,\mathcal{P}_m) + \varepsilon \le \lim_{m\to+\infty} h(T,\mathcal{P}_m) + \varepsilon,$$

puis

$$h(T) \le \lim_{m\to+\infty} h(T,\mathcal{P}_m).$$

□

Définition. Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. On dira qu'une partition $\mathcal{P}$ est *fortement génératrice* si la suite $\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)_{n\ge1}$ est génératrice. Dans le cas où T est inversible, on dira que $\mathcal{P}$ est*génératrice* si la suite $\left(\bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right)_{n\ge1}$ est génératrice.

Théorème 1.4.2 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$ et $\mathcal{P}$ une partition mesurable. Si $\mathcal{P}$ est fortement génératrice, alors $h_\mu(T) = h(T,\mathcal{P})$. On a la même égalité si T est inversible et si $\mathcal{P}$ est génératrice.*

Démonstration. Nous allons utiliser la proposition 1.4.1 ainsi que l'assertion **iii)** de la proposition 1.3.3. Écrivons

$$h(T) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) = \lim_{n\to+\infty} h(T,\mathcal{P}) = h(T,\mathcal{P}).$$

Dans le cas où T est inversible et $\mathcal{P}$ est un générateur, écrivons

$$h(T) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=0}^{2n-2} T^{-i}(\mathcal{P})\right) = h(T,\mathcal{P}).$$

□

1.5 Exemples

Nous allons calculer des entropies à l'aide notamment du théorème 1.4.2.

Le décalage de Bernouilli

Considérons le décalage σ sur $\{1,\dots,r\}^{\mathbf{N}}$. Remarquons que la (vraie) partition $\mathcal{P} = (P_i)_{1\le i\le r}$ où $P_i = \{x = (x_n)_{n\ge 0} \mid x_0 = i\}$ est fortement génératrice car $\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})$ est la partition en les p^n cylindres de la forme $C^0_{(w_0,\dots,w_{n-1})}$. Pour toute mesure invariante μ, on a donc l'égalité

$$h_\mu(\sigma) = h_\mu(\sigma,\mathcal{P}).$$

Si μ est une mesure produit définie par une famille $(p_i)_{1\le i\le r}$ de réels positifs telle que $\sum_{i=1}^r p_i = 1$, alors, pour tout $n \ge 0$ on a

$$\begin{aligned} H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})) &= \sum_{i_0,\dots,i_{n-1}} -p_{i_0}\dots p_{i_{n-1}} \ln(p_{i_0}\dots p_{i_{n-1}}) \\ &= n\sum_{i=1}^{r} p_i \ln p_i. \end{aligned}$$

On en déduit que

$$h_\mu(\sigma) = \sum_{i=1}^{r} -p_i \ln p_i = H(\mathcal{P}).$$

Remarques

1. On a $h_\mu(\sigma) \le \ln r$, avec égalité uniquement dans le cas de la mesure équidistribuée (voir également exercice **)

2. Si $r \ne s$, les décalages sur $\{1,\dots,r\}^N$ et $\{1,\dots,s\}^N$, ne sont pas conjugués, considérés comme systèmes dynamiques mesurés avec la mesure équidistribuée, car ils n'ont pas la même entropie.

Remarques

1. Ce qui précède peut être fait sur le décalage bilatéral. Écrivons $P_i = \{x = (x_k)_{k\in\mathbf{Z}} \,|\, x_0 = i\}$. La partition $\mathcal{P} = (P_i)_{1\le i\le r}$ étant génératrice (mais pas fortement génératrice) on a, pour toute mesure invariante μ, l'égalité

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}).$$

Pour les mesures produits, on trouvera les mêmes formules que dans le cas unilatéral

2. Un difficile résultat d'Ornstein exprime que deux décalages bilatéraux, considérés comme systèmes dynamiques mesurés avec chacun une mesure de type produit, sont conjugués si et seulement s'ils ont même entropie (le résultat est faux pour les décalages unilatéraux).

Endomorphismes linéaires de $\mathbf{T}^1$.

L'application $F\ :\ \widehat{x} \to p\widehat{x}$ sur $\mathbf{T}^1$, où $|p| \ge 2$, préserve la mesure de Haar. Nous allons calculer l'entropie du système dynamique mesuré obtenu. On peut se limiter au cas où $p \ge 2$, puisque $h(F^2) = 2h(F)$ et que $F^2(\widehat{x}) = p^2\widehat{x}$. Or nous avons vu que le système est alors conjugué au décalage de Bernouilli sur $\{1, \dots, p\}$ avec la mesure équidistribuée. On a donc généralement l'égalité $h(F) = \ln|p|$. On aurait également pu utiliser directement la partition fortement génératrice $\mathcal{P} = (P_i)_{0\le i\le p-1}$, où P_i est la projection dans $\mathbf{T}^1$ de $[i/p, (i+1)/p[$.

Rotations du cercle.

La rotation $T_{\widehat{a}}\ :\ \widehat{x} \to \widehat{x} + \widehat{a}$ préserve la mesure de Haar sur $\mathbf{T}^1$. Si $\widehat{a} \in \mathbf{Q}/\mathbf{Z}$, il existe $q \ge 1$ tel que $T_{\widehat{a}}^q = \mathrm{Id}_{\mathbf{T}^1}$ et donc on a $h(T_{\widehat{a}}) = 0$. Si $\alpha \notin \mathbf{Q}/\mathbf{Z}$, on peut vérifier que la partition $\mathcal{P} = (P_0, P_1)$ est fortement génératrice, où P_i est la projection dans $\mathbf{T}^1$ de $[i/2, (i+1)/2[$. Ceci implique que $h_\mu(T_{\widehat{a}}) = 0$. En effet $T^{-1}(\mathcal{P})$ est également fortement génératrice, et donc :

$$h(T) = h(T, \mathcal{P}) = \lim_{n\to+\infty} H(\mathcal{P} | \bigvee_{i=1}^{n} T^{-i}(\mathcal{P})) = 0.$$

Remarquons que nous avons le résultat plus général suivant :

Proposition 1.5.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré inversible tel que $\mu(X) = 1$. S'il existe une partition fortement génératrice, alors $h(T) = 0$.*

Nous allons associer à toute application continue $T : X \to X$ définie sur un espace métrique compact X un nombre $h(T) \in [0, +\infty]$, invariant par conjugaison, qui mesure le désordre de la dynamique. Cet invariant, *l'entropie topologique*, a été introduit par Adler, Konheim et McAndrew en 1965, par analogie avec l'entropie métrique de Kolmogorov et Sinai.

2.1 Entropie relative par rapport à un recouvrement

Rappelons qu'un recouvrement ouvert d'un espace topologique X est une famille $\mathcal{U} = (U_i)_{i\in I}$ de parties ouvertes telle que $X = \bigcup_{i\in I} U_i$. Si $U \subset X$ est ouvert, on écrira $U \in \mathcal{U}$ s'il existe $i \in I$ tel que $U = U_i$. On dira que le recouvrement $\mathcal{V} = (V_j)_{j\in J}$ est *plus fin* que $\mathcal{U}$ si pour tout $j \in J$, il existe $i \in I$ tel que $V_j \subset U_i$, on écrira alors $\mathcal{U} \preceq \mathcal{V}$. Un cas particulier est la cas où $\mathcal{V}$ est un *sous-recouvrement* de $\mathcal{U}$: pour tout $j \in J$, il existe $i \in I$ tel que $V_j = U_i$. La relation $\preceq$ est un pré-ordre et on notera $\sim$ la relation d'équivalence associée : $\mathcal{U} \sim \mathcal{V}$ si $\mathcal{U} \preceq \mathcal{V}$ et $\mathcal{V} \preceq \mathcal{U}$.

Si $(\mathcal{U}^j)_{1\le j\le n}$ est une famille finie de recouvrements ouverts de X, on peut définir le recouvrement $\bigvee_{j=1}^n \mathcal{U}^j$ dont les éléments sont les $\bigcap_{1\le j\le n} U^j$ où $U^j \in \mathcal{U}^j$. Il est plus fin que tous les $\mathcal{U}^j$. Remarquons que si $(\mathcal{V}^j)_{1\le j\le n}$ est une famille finie de recouvrements ouverts de X et si $\mathcal{U}^j \preceq \mathcal{V}^j$ pour tout $j \in \{1, \dots n\}$, alors $\bigvee_{j=1}^n \mathcal{U}^j \preceq \bigvee_{j=1}^n \mathcal{V}^j$.

Si $H : Y \to X$ est une application continue entre deux espaces topologiques et si $\mathcal{U} = (U_i)_{i\in I}$ est un recouvrement ouvert de X, on peut définir le recouvrement $H^{-1}(\mathcal{U}) = (H^{-1}(U_i))_{i\in I}$ de Y. On a bien évidemment :

- $H^{-1}\left(\bigvee_{j=1}^n \mathcal{U}^j\right) = \bigvee_{j=1}^n H^{-1}(\mathcal{U}^j)$;
- $\mathcal{U} \preceq \mathcal{V} \Longrightarrow H^{-1}(\mathcal{U}) \preceq H^{-1}(\mathcal{V})$.

Si X est compact, tout recouvrement ouvert $\mathcal{U}$ admet un sous-recouvrement fini. On notera alors $N(\mathcal{U})$ le plus petit cardinal des sous-recouvrements finis. Remarquons que :

- $N(\mathcal{U}) = 1$ si et seulement si $X \in \mathcal{U}$;
- $\mathcal{U} \preceq \mathcal{V} \Longrightarrow N(\mathcal{U}) \le N(\mathcal{V})$;
- $\mathcal{U} \sim \mathcal{V} \Longrightarrow N(\mathcal{U}) = N(\mathcal{V})$;
- $N\left(\bigvee_{j=1}^n \mathcal{U}^j\right) \le \prod_{j=1}^n N(\mathcal{U}^j)$.

De plus, si $H : Y \to X$ est une application continue entre deux espaces topologiques compacts, alors

- $N(T^{-1}(\mathcal{U})) \le N(\mathcal{U})$;
- $N(T^{-1}(\mathcal{U})) = N(\mathcal{U})$ si T est surjective.

On supposera dorénavant que X est compact et que $T : M \to M$ est continue. Énonçons maintenant le résultat fondamental suivant :

Proposition 2.1.1 : *Pour tout recouvrement ouvert $\mathcal{U}$ de X, la suite*

$$\frac{1}{n} \ln\left(N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right)\right)$$

converge. Sa limite $h(T,\mathcal{U})$ est l'entropie topologique de T relativement à $\mathcal{U}$, *elle vérifie :*

$$0 \leq h(T,\mathcal{U}) \leq \ln(N(\mathcal{U})).$$

Démonstration.. Remarquons que la suite $(u_n)_{n\geq 1}$, où $u_n = N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right)$, est sous-multiplicative, elle vérifie $u_{n+m} \leq u_n u_m$. En effet,

$$\begin{aligned} u_{n+m} &= N\left(\bigvee_{i=0}^{n+m-1} T^{-i}(\mathcal{U})\right) \\ &= N\left(\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) \vee T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{U})\right)\right) \\ &\leq N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) N\left(T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{U})\right)\right) \\ &\leq u_n u_m. \end{aligned}$$

Posant $v_n = \ln(u_n)$, on obtient une suite $(v_n)_{n\geq 1}$ qui est donc sous-additive et positive. Comme on l'a vu dans la preuve de la proposition 1.3.1., ceci implique que la suite $(\frac{v_n}{n})_{n\geq 1}$ converge et que

$$0 \leq \lim_{n\to+\infty} \frac{v_n}{n} \leq v_1.$$

□

Énonçons maintenant les propriétés principales de cette entropie relative.

PROPOSITION 2.1.2 : *On a les résultats suivants :*

i) $\mathcal{U} \preceq \mathcal{V} \Longrightarrow h(T,\mathcal{U}) \leq h(T,\mathcal{V})$;

ii) $h(T,\mathcal{U}) = h(T, \bigvee_{i=0}^{m} T^{-i}(\mathcal{U}))$ *pour tout* $m \geq 0$;

iii) $h(T^m, \bigvee_{i=0}^{m-1} T^{-i}(\mathcal{U})) = m\, h(T,\mathcal{U})$ *pour tout* $m \geq 1$;

iv) $h(T^{-1},\mathcal{U}) = h(T,\mathcal{U})$ *si T est un homéomorphisme ;*

v) $h(T,\mathcal{U}) = h(T, T^{-1}(\mathcal{U})) = h(T, T(\mathcal{U}))$ *si T est un homéomorphisme.*

Démonstration. L'assertion **i)** découle immédiatement de l'inégalité

$$N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) \leq N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{V})\right),$$

conséquence de

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}) \preceq \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{V}),$$

Pour montrer **ii)**, posons $\mathcal{V} = \bigvee_{i=0}^{m} T^{-i}(\mathcal{U})$ et remarquons d'abord que

$$\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{U}) \sim \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{V}),$$

ce qui implique que

$$h(T,\mathcal{U}) = \lim_{n\to+\infty} \frac{1}{n+m} \ln\left(N\left(\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{U})\right)\right)$$
$$= \lim_{n\to+\infty} \frac{n}{n+m}\,\frac{1}{n} \ln\left(N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{V})\right)\right) = h(T,\mathcal{V}).$$

L'assertion **iii)** se déduit de l'égalité

$$\bigvee_{i=0}^{n-1} T^{-im}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{U})\right) = \bigvee_{i=0}^{nm-1} T^{-i}(\mathcal{U}).$$

Pour montrer **iv)** remarquons que

$$N\left(\bigvee_{i=0}^{n-1} T^{i}(\mathcal{U})\right) = N\left(T^n\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right)\right) = N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right),$$

et pour montrer **v)**, que

$$N\left(\bigvee_{i=0}^{n-1} T^{-i}(T^{-1}(\mathcal{U}))\right) = N\left(T^{-1}\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right)\right) = N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right).$$

□

2.2 Entropie topologique

L'entropie topologique $h(T) \in [0,+\infty]$ est le supremum des entropies de recouvrement :

$$h(T) = \sup\{h(T,\mathcal{U}),\ \mathcal{U} \text{ recouvrement ouvert de } X\}.$$

Énonçons les propriétés principales :

PROPOSITION 2.2.1 : *On a les propriétés suivantes :*

i) $h(\mathrm{Id}_X) = 0$ *;*

ii) $h(T^n) = nh(T)$*, pour tout* $n \geq 0$ *;*

iii) *si* T *est un homéomorphisme, alors* $h(T^k) = |k|h(T)$*, pour tout* $k \in \mathbf{Z}$ *;*

iv) *si* $S : Y \to Y$ *est un facteur de* T*, alors* $h(S) \leq h(T)$ *;*

v) *si* $S : Y \to Y$ *est conjugué à* T*, alors* $h(S) = h(T)$ *;*

vi) *si* $Y \subset X$ *est fermé et positivement invariant, alors* $h(T|_Y) \leq h(T)$*.*

Démonstration. Si $\mathcal{U}$ est un recouvrement ouvert de X, alors pour tout entier $n \geq 1$, on a $\bigvee_{i=0}^{n-1} \mathrm{Id}_X^{-i}(\mathcal{U}) \sim \mathcal{U}$, ce qui bien sûr implique **i)**.

Pour montrer **ii)** dans le cas où $n \geq 1$ (dans le cas où $n = 0$ ce n'est rien d'autre que **i)**) remarquons que pour tout recouvrement ouvert $\mathcal{U}$, on a

$$h(T^n,\mathcal{U}) \leq h\left(T^n, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) = nh(T,\mathcal{U}),$$

ce qui implique
$$h(T^n,\mathcal{U}) \leq nh(T),\ \ nh(T,\mathcal{U}) \leq h(T^n),$$
et donc
$$h(T^n) \leq nh(T),\ \ nh(T) \leq h(T^n)$$
en passant aux supremums.

L'assertion **iii)** se déduit immédiatement de l'égalité $h(T,\mathcal{U}) = h(T^{-1},\mathcal{U})$.

Pour prouver **iv)**, considérons une semi-conjugaison $H\ : X \to Y$ entre T et S. Pour tout recouvrement ouvert $\mathcal{U}$ de Y, et tout entier $n \geq 1$, on a
$$H^{-1}\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{U})\right) = \bigvee_{i=0}^{n-1} T^{-i}(H^{-1}(\mathcal{U}))$$
ce qui implique que
$$N\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{U})\right) = N\left(\bigvee_{i=0}^{n-1} T^{-i}(H^{-1}(\mathcal{U}))\right)$$
puisque H est surjective. On en déduit que
$$h(S,\mathcal{U}) = h(T, H^{-1}(\mathcal{U})) \leq h(T)$$
puis en passant au supremum que $h(S) \leq H(T)$. On en déduit alors immédiatement **iv)**.

Pour prouver **vi)** considérons un recouvrement ouvert $\mathcal{V} = (V_i)_{i\in I}$ de Y. Pour tout $i \in I$, on peut choisir une partie ouverte U_i de X telle que $V_i = Y \cap U_i$. Si on ajoute $X \setminus Y$ à cette famille, on obtient un recouvrement ouvert $\mathcal{U}$ de X. Remarquons maintenant que
$$N\left(\bigvee_{i=0}^{n-1} (T|_Y)^{-i}(\mathcal{V})\right) \leq N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right).$$
En effet, puisque $T(Y) \subset Y$, on sait qu'aucun des ensembles $T^{-i}(X \setminus Y)$ ne rencontre Y. Ceci implique que pour tout sous-recouvrement fini $\mathcal{U}'$ de $\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})$, tout élément non vide de $Y \cap U'$, $U' \in \mathcal{U}'$, appartient à $\bigvee_{i=0}^{n-1} (T|_Y)^{-i}(\mathcal{V})$ On en déduit que
$$h(T|_Y, \mathcal{V}) \leq h(T,\mathcal{U}) \leq h(T)$$
puis en passant au supremum que $h(T|_Y) \leq h(T)$. □

2.3 Recouvrement générateur

Il est bien évidemment difficile de calculer l'entropie topologique à partir de la définition abstraite précédente. Nous verrons dans ce paragraphe qu'on peut se restreindre à certains recouvrements. Nous dirons qu'une famille $(\mathcal{U}^\alpha)_{\alpha\in A}$ de recouvrements ouverts est une *famille génératrice* si, pour tout recouvrement ouvert $\mathcal{U}$, il existe $\alpha \in A$ tel que $\mathcal{U} \preceq \mathcal{U}^\alpha$. Nous avons bien évidemment :

Proposition 2.3.1 : *Si $(\mathcal{U}^\alpha)_{\alpha\in A}$ est une famille génératrice de recouvrements ouverts, alors*
$$h(T) = \sup_{\alpha\in A} h(T,\mathcal{U}^a).$$

Un exemple simple de famille génératrice est donné par les recouvrements finis, un autre exemple est donné, dans le cas où X est un espace métrique, par la famille $(\mathcal{U}^{\varepsilon})_{\varepsilon>0}$, où $\mathcal{U}^{\varepsilon} = (B(x,\varepsilon))_{x\in X}$ est le recouvrement par les boules de rayon ε. Le caractère générateur de cette famille n'est rien d'autre que le lemme de recouvrement de Lebesgue : *pour tout recouvrement ouvert $\mathcal{U}$, il existe $\varepsilon > 0$ tel que toute boule $B(x,\varepsilon)$ est contenue dans une partie ouverte $U \in \mathcal{U}$.* Remarquons que la fonction $\varepsilon \mapsto h(T,\mathcal{U}^{\varepsilon})$ est décroissante et donc que

$$h(T) = \sup_{\varepsilon>0} h(T,\mathcal{U}^{\varepsilon}) = \lim_{\varepsilon\to 0^+} h(T,\mathcal{U}^{\varepsilon}).$$

Dans le cas d'un espace métrique, on peut donner une caractérisation des familles génératrices, grâce au lemme de recouvrement de Lebesgue. Définissons le diamètre d'un recouvrement $\mathcal{U}$:

$$\operatorname{diam}(\mathcal{U}) = \sup_{U\in\mathcal{U}} \operatorname{diam}(U),$$

où

$$\operatorname{diam}(U) = \sup_{x\in U,\, y\in U} d(x,y).$$

Supposons maintenant que $(\mathcal{U}_n)_{n\geq 1}$ est une suite croissante de recouvrements ouverts, c'est-à-dire une suite vérifiant $\mathcal{U}_n \preceq \mathcal{U}_{n+1}$, pour tout $n \geq 0$. Cette suite est génératrice si et seulement si $\lim\limits_{n\to+\infty} \operatorname{diam}(\mathcal{U}_n) = 0$. Dans ce cas, on a

$$h(T) = \lim_{n\to+\infty} h(T,\mathcal{U}_n).$$

Dans de nombreux exemples, une suite génératrice peut être définie à partir d'un seul recouvrement. Plus précisément, on dira qu'un recouvrement ouvert U de X est un *recouvrement générateur* si la suite $\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right)_{n\geq 1}$ est génératrice.

Proposition 2.3.2 : *Si $\mathcal{U}$ est un recouvrement générateur de X, alors*

$$h(T) = h(T,\mathcal{U}) < +\infty.$$

Démonstration. On a

$$h(T) = \lim_{n\to+\infty} h(T, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})) = h(T,\mathcal{U}).$$

□

La remarque précédente sur la famille $(\mathcal{U}^{\varepsilon})_{\varepsilon>0}$ va nous permettre de donner, dans le cas d'un espace métrique, une définition alternative de l'entropie topologique, définition due de façon indépendante à Bowen et Dinaburg. Définissons, pour tout entier $n \geq 1$, la distance suivante d_n sur X :

$$d_n(x,y) = \max_{0\leq i\leq n-1} d(T^i(x), T^i(y)),$$

et notons $B_n(x,\varepsilon)$ la boule ouverte de rayon ε et de centre x. Soient $\varepsilon > 0$ et $n \geq 1$. On dira qu'un ensemble fini $S \subset X$ est (n,ε)-*séparé* si, pour tous points x et y de S, on a $d_n(x,y) \geq \varepsilon$. De même, on dira qu'un ensemble fini $R \subset X$ est (n,ε)-*couvrant* si, pour tout $x \in X$, il existe $y \in R$ tel que $d_n(x,y) < \varepsilon$. On notera alors $s(n,\varepsilon)$ le plus grand cardinal des ensembles (n,ε)-séparés et $r(n,\varepsilon)$ le plus petit cardinal des ensembles (n,ε)-couvrants. Enfin, on écrira $N(n,\varepsilon) = N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}^{\varepsilon})\right)$. Ces nombres sont naturellement liés, comme l'exprime le résultat suivant :

Proposition 2.3.3 : *On a*

$$N(n,\varepsilon) \leq r(n,\varepsilon) \leq s(n,\varepsilon) \leq N(n,\frac{\varepsilon}{2}).$$

Démonstration. Pour montrer la première inégalité, choisissons un ensemble R de cardinal $r(n,\varepsilon)$ qui est (n,ε)-couvrant. Les boules $B_n(x,\varepsilon)$, $x \in R$, recouvrent X et appartiennent toutes à $\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}^{\varepsilon})$. Pour montrer la seconde inégalité, remarquons qu'un ensemble (n,ε)-séparé S de cardinal maximal $s(n,\varepsilon)$ est nécessairement (n,ε)-couvrant. Enfin, pour montrer la dernière égalité, remarquons qu'une partie $U \in \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}^{\frac{\varepsilon}{2}})$ contient au plus un élément de S.

□

Corollaire 2.3.4 : *On a*

$$\begin{aligned} h(T) &= \lim_{\varepsilon\to 0}\lim_{n\to+\infty}\frac{1}{n}\ln N(n,\varepsilon) \\ &= \lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\ln r(n,\varepsilon) \\ &= \lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\ln s(n,\varepsilon) \\ &= \lim_{\varepsilon\to 0}\liminf_{n\to+\infty}\frac{1}{n}\ln r(n,\varepsilon) \\ &= \lim_{\varepsilon\to 0}\liminf_{n\to+\infty}\frac{1}{n}\ln s(n,\varepsilon). \end{aligned}$$

Démonstration. L'égalité

$$h(T) = \lim_{\varepsilon\to 0}\lim_{n\to+\infty}\frac{1}{n}\ln N(n,\varepsilon)$$

est une conséquence, vue plus haut, du caractère générateur de la famille $(\mathcal{U}^{\varepsilon})_{\varepsilon>0}$. L'égalité analogue, où l'on remplace $N(n,\varepsilon)$ par $r(n,\varepsilon)$ ou $s(n,\varepsilon)$ n'est pas nécessairement vraie, car rien ne dit que les suites $\left(\frac{1}{n}\ln s(n,\varepsilon)\right)_{n\geq 1}$ et $\left(\frac{1}{n}\ln r(n,\varepsilon)\right)_{n\geq 1}$ convergent. Cependant, la proposition précédente permet d'obtenir les quatre dernières égalités.

□

Corollaire 2.3.5 : *Si $T : X \to X$ est lipschitzienne de rapport 1, c'est-à-dire si $d(T(x),T(y)) \leq d(x,y)$ pour tous x, y dans X, alors $h(T) = 0$.*

Démonstration. Il suffit de remarquer que la distance d_n coïncide avec la distance d, et que les entiers $r(n,\varepsilon)$ et $s(n,\varepsilon)$ sont donc indépendants de n. □

2.4. Exemples

i) Le décalage de Bernouilli unilatéral :

Notons

$$\sigma \ : A^{\mathbf{N}} \to A^{\mathbf{N}},$$
$$(x_n)_{n\geq 0} \mapsto (x_{n+1})_{n\geq 0}$$

le décalage de Bernouilli, où A est un alphabet fini de cardinal $p \geq 2$. Le recouvrement $\mathcal{U} = (U_a)_{a\in A}$ formé des p cylindres $U_a = \{x = (x_n)_{n\geq 0} \in A^{\mathbf{N}} \,|\, x_0 = a\}$ est générateur. En effet, le recouvrement $\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{U})$ est formé des cylindres à base $\{0, \ldots, n-1\}$. On a p^n cylindres disjoints dont les diamètres tendent vers 0 quand $n \to +\infty$. Ainsi, on a

$$h(T) = h(T,\mathcal{U}) = \lim_{n\to+\infty} \frac{1}{n} \ln \left(N \left(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{U}) \right) \right) = \lim_{n\to+\infty} \frac{1}{n} \ln p^n = \ln p.$$

ii) Le décalage de Bernouilli bilatéral :

Notons

$$\sigma \ : A^{\mathbf{Z}} \to A^{\mathbf{Z}},$$
$$(x_k)_{k\in\mathbf{Z}} \mapsto (x_{k+1})_{k\in\mathbf{Z}}$$

le décalage de Bernouilli, où A est un alphabet fini de cardinal $p \geq 2$. On pourrait montrer qu'il n'y pas de recouvrement générateur (au sens donné plus haut). Cependant si on considère le recouvrement $\mathcal{U} = (U_a)_{a\in A}$ formé des p cylindres $U_a = \{x \in A^{\mathbf{Z}} \,|\, x_0 = a\}$, on peut remarquer que la famille $\left(\bigvee_{i=-n+1}^{n-1} \sigma^{-i}(\mathcal{U})\right)_{n\geq 0}$ est génératrice. Ainsi

$$h(\sigma) = \lim_{n\to+\infty} h\left(\sigma, \bigvee_{i=-n+1}^{n-1} \sigma^{-i}(\mathcal{U})\right) = \lim_{n\to+\infty} h\left(\sigma, \bigvee_{i=0}^{2n-2} \sigma^{-i}(\mathcal{U})\right) = h(\sigma,\mathcal{U}) = \ln p.$$

iii) Endomorphismes linéaires de $\mathbf{T}^1$.

On considère l'application $T \ : x \to px$ sur le cercle $\mathbf{T}^1 = \mathbf{R}/\mathbf{Z}$, où $|p| \geq 2$. On va montrer que

$$h(T) = \ln |p|.$$

Puisque $h(T^2) = 2h(T)$ et puisque $T^2(x) = p^2 x$, il suffit d'étudier le cas où $p \geq 2$. Puisque T est un facteur du décalage unilatéral sur $\{1, \ldots, p\}^N$, on sait déjà que $h(T) \leq \ln p$. Pour prouver que $h(T) = \ln p$ il suffit de trouver un recouvrement $\mathcal{U}$ tel que $h(T,\mathcal{U}) \geq \ln p$. Considérons un recouvrement $\mathcal{U}$ par des intervalles ouverts I de diamètre constant $\delta < \dfrac{1}{p+1}$. Pour tout $I \in \mathcal{U}$, l'ensemble $T^{-1}(I)$ est la réunion de p intervalles de longueur $\frac{\delta}{p}$ separés par des intervalles de longueur $\frac{1-\delta}{p}$. Puisque $\delta < \frac{1-\delta}{p}$ on sait que pour tout intervalle $I' \in \mathcal{U}$ l'ensemble $I' \cap T^{-1}(I)$ est un intervalle (éventuellement vide) de longueur $\leq \frac{\delta}{p}$. Ainsi on a $\mathrm{diam}(\mathcal{U} \vee T^{-1}(\mathcal{U})) \leq \frac{\delta}{p}$. Le même argument nous dit que $\mathcal{U} \vee T^{-1}(\mathcal{U}) \vee T^{-2}(\mathcal{U})$ est un recouvrement par des intervalles de longueur $\leq \frac{\delta}{p^2}$ et plus généralement que $\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})$ est un recouvrement par intervalles et que

$$\mathrm{diam} \left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}) \right) \leq \frac{\delta}{p^{n-1}}.$$

On en déduit que $\mathcal{U}$ est un recouvrement générateur et que $h(T) = h(T,\mathcal{U})$. Remarquons maintenant que

$$N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) \geq \frac{p^{n-1}}{\delta} \geq p^n,$$

et donc que $h(T,\mathcal{U}) \geq \ln p$.

iv) Homéomorphismes du cercle.

On va montrer que l'entropie d'un homéomorphisme T de $\mathbf{T}^1$ est nulle. Remarquons d'abord que si $\mathcal{U}$ et $\mathcal{V}$ sont des recouvrements par des intervalles ouverts de longueur $\leq 1/2$, il en est de même de $\mathcal{U}\vee\mathcal{V}$ et que $N(\mathcal{U}\vee\mathcal{V}) \leq N(\mathcal{U})+N(\mathcal{V})$. En effet, si $\mathcal{U}'$ et $\mathcal{V}'$ sont des sous-recouvrements finis de $\mathcal{U}$ et $\mathcal{V}$ de cardinaux respectifs $N(\mathcal{U})$ et $N(\mathcal{V})$, une extrémité gauche d'un intervalle $I \in \mathcal{U}'$ est associée à un unique intervalle, il y a donc $N(\mathcal{U})$ extrémités gauches. De même il y a $N(\mathcal{V})$ extrémités gauches d'intervalles $J \in \mathcal{V}'$. Remarquons maintenant que si $I \in \mathcal{U}'$ et $J \in \mathcal{V}'$ ont une intersection non vide, cette intersection est un intervalle et que l'extrémité gauche de cet intervalle est soit l'extrémité gauche de I, soit l'extrémité gauche de J. Ainsi $\mathcal{U}' \vee \mathcal{V}'$ est un recouvrement par intervalles, en nombre au plus égal à $N(\mathcal{U}) + N(\mathcal{V})$.

Fixons $\delta \in (0, \frac{1}{2})$ tel que $d(T^{-1}(x), T^{-1}(y)) \leq \frac{1}{2}$ si $d(x,y) \leq \delta$. Puisque la famille des recouvrements par intervalles de longueur $\leq \delta$ est génératrice, il suffit de prouver que $h(T,\mathcal{U}) = 0$ pour un tel recouvrement. Mais nous savons que $\mathcal{U}$ et $T^{-1}(\mathcal{U})$ sont des recouvrements par intervalles de longueur $\leq \frac{1}{2}$. On en déduit que $\mathcal{U} \vee T^{-1}(\mathcal{U})$ est un recouvrement par intervalles de longueur $\leq \delta$ et que

$$N(\mathcal{U} \vee T^{-1}(\mathcal{U})) \leq N(\mathcal{U}) + N(T^{-1}(\mathcal{U})) = 2N(\mathcal{U}).$$

Une simple récurrence permet d'établir que $\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})$ est un recouvrement par intervalles de longueur $\leq \delta$ et que

$$N\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U})\right) \leq nN(\mathcal{U}).$$

Ceci implique, bien sûr que $h(T,\mathcal{U}) = 0$.

2.5 Entropie topologique et entropie métrique

Si $T : X \to X$ est une application continue définie sur un espace métrique compact, on sait que l'ensemble $\mathcal{M}_T$ des mesures boréliennes de probabilité invariantes est une partie compacte (pour la topologie faible*) convexe et non vide. On peut définir l'entropie $h_\mu(T)$ de toute mesure $\mu \in \mathcal{M}_T$. Nous allons nous intéresser dans cette section au lien entre l'entropie topologique et les entropies métriques des mesures invariantes. Le résultat principal, appelé usuellement *principe variationnel pour l'entropie*, a été prouvé par Goodwyn (une des inégalités) et Goodman (égalité), nous allons donner une preuve due à Misiurewicz :

Théorème 2.5.1 : *Si $T : X \to X$ est une application continue définie sur un espace métrique compact, alors*

$$h(T) = \sup_{\mu \in \mathcal{M}_T} h_\mu(T).$$

Démonstration. Nous allons commencer par prouver l'inégalité $h_\mu(T) \leq h(T)$ pour toute mesure $\mu \in \mathcal{M}_T$. Nous utiliserons la régularité de μ.

Lemme 2.5.2 : *Une mesure borélienne de probabilité sur un espace métrique est régulière : pour tout borélien A et tout $\varepsilon > 0$, il existe une partie fermée F et une partie ouverte U telle que $F \subset A \subset U$ et $\mu(U \setminus F) < \varepsilon$.*

Démonstration. Il suffit de prouver que l'ensemble $\mathcal{C}$ des boréliens qui vérifient la condition du lemme est une σ-algèbre qui contient les parties fermées. L'ensemble $\mathcal{C}$ est stable par passage au complémentaire et contient X, pour montrer que c'est une σ-algèbre, il reste à prouver qu'il est stable par réunion dénombrable. Soit $(A_m)_{m\geq 0}$ une suite dans $\mathcal{C}$. Fixons $\varepsilon > 0$ et choisissons pour tout $m \geq 0$ une partie fermée F_m et une partie ouverte U_m telles que $F_m \subset A_m \subset U_m$ et $\mu(U_m \setminus F_m) < \varepsilon/2^{m+1}$. Remarquons que

$$\bigcup_{m\geq 0} F_m \subset \bigcup_{m\geq 0} A_m \subset \bigcup_{m\geq 0} U_m$$

et que

$$\mu\left(\bigcup_{m\geq 0} U_m \setminus \bigcup_{m\geq 0} F_m\right) \leq \mu\left(\bigcup_{m\geq 0} (U_m \setminus F_m)\right) \leq \sum_{m=0}^{+\infty} \mu(U_m \setminus F_m) < \sum_{m=0}^{+\infty} \frac{\varepsilon}{2^{m+1}} = \varepsilon.$$

Ceci implique qu'il existe $m_0 \geq 0$ tel que

$$\mu\left(\bigcup_{m\geq 0} U_m \setminus \bigcup_{0\leq m\leq m_0} F_m\right) < \varepsilon.$$

On en déduit que $A \in \mathcal{C}$ car $U = \bigcup_{m\geq 0} U_m$ est ouvert et $\bigcup_{0\leq m\leq m_0} F_m$ est fermé.

Nous devons montrer maintenant que toute partie fermée F appartient à $\mathcal{C}$. Remarquons que $F = \bigcap_{m\geq 1} U_m$, où

$$U_m = \left\{x \in X \mid d(x, F) < \frac{1}{m}\right\}$$

est ouvert. On en déduit que

$$\lim_{m\to+\infty} \mu(U_m \setminus F) = \mu\left(\bigcap_{m\geq 1} U_m \setminus F\right) = 0.$$

□

Prouvons maintenant que $h_\mu(T) \leq h(T)$ si $\mu \in \mathcal{M}_T$. Nous allons en fait montrer que

$$h_\mu(T) \leq 1 + \ln 2 + h(T).$$

En appliquant cette formule à chaque itéré T^m, $m \geq 1$, on obtiendra

$$mh_\mu(T) = h_\mu(T^m) \leq 1 + \ln 2 + h(T^m) = 1 + \ln 2 + mh(T).$$

Il restera alors à diviser par m et à faire tendre m vers $+\infty$ pour obtenir

$$h_\mu(T) \leq h(T).$$

Choisissons une partition mesurable $\mathcal{P} = (P_i)_{1\leq i\leq r}$. Fixons $\varepsilon > 0$. Pour tout $i \in \{1, \ldots, r\}$ on peut trouver une partie fermée $Q_i \subset P_i$ telle que $\mu(P_i \setminus Q_i) < \varepsilon$. Posons

$$Q_0 = X \setminus \bigcup_{1\leq i\leq r} Q_i, \;\; P_0 = \emptyset.$$

Nous avons deux partitions mesurables $\mathcal{Q} = (Q_i)_{0\leq i\leq r}$ et $\mathcal{P}' = (P_i)_{0\leq i\leq r}$ telles que $\mu(P_i \Delta Q_i) \leq r\varepsilon$ pour tout $i \in \{1, \ldots, r\}$ (puisque $\mu(P_0 \Delta Q_0) = \mu(Q_0) \leq r\varepsilon$). On en déduit que si ε est assez petit, alors

$$H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}'|\mathcal{Q}) \leq 1.$$

Ceci implique que

$$h_\mu(T, \mathcal{P}) \leq h_\mu(T, \mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}) \leq h_\mu(T, \mathcal{Q}) + 1.$$

Pour tout $i \in \{1, \ldots, r\}$ on définit maintenant

$$U_i = Q_0 \cup Q_i = X \setminus \left(\bigcup_{1\leq j\leq r, j\neq i} Q_j \right).$$

On obtient un recouvrement ouvert $\mathcal{U} = (U_i)_{1\leq i\leq r}$. Chaque élément

$$\bigcap_{0\leq k<n} T^{-k}(U_{i_k}) = \bigcap_{0\leq k<n} T^{-k}(Q_0 \cup Q_{i_k})$$

du recouvrement $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{U})$ est la réunion de 2^n éléments de la partition $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{Q})$ (certains éventuellement vides). Si R est le nombre d'éléments non vides de la partition $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{Q})$, on en déduit que

$$H_\mu \left(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{Q}) \right) \leq \ln R \leq \ln \left(2^n N \left(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{U}) \right) \right).$$

En divisant par n et en faisant tendre n vers $+\infty$, on obtient

$$h_\mu(T, \mathcal{Q}) \leq \ln 2 + h(T, \mathcal{U}) \leq \ln 2 + h(T),$$

puis

$$h_\mu(T, \mathcal{P}) \leq 1 + \ln 2 + h(T).$$

Il reste à passer au supremum pour obtenir l'inégalité cherchée.

$$h_\mu(T) \leq 1 + \ln 2 + h(T).$$

□

Nous allons maintenant montrer l'inégalité inverse

$$h(T) \leq \sup_{\mu\in\mathcal{M}_T} h_\mu(T),$$

et pour cela commencer par quelques lemmes.

Lemme 2.5.3 : *Soit X un espace métrique compact et $(\mu_m)_{m\geq 0}$ une suite de mesures boréliennes de probabilité qui converge vers μ pour la topologie faible*. Alors, pour tout borélien A tel que $\mu(\partial A) = 0$, on a*

$$\lim_{m\to+\infty} \mu_m(A) = \mu(A).$$

Démonstration. Définissons une suite de fonctions continues $(f_k)_{k\geq 1}$ sur X, en posant

$$f_k \ : x \mapsto \max(1 - kd(x, A), 0).$$

La suite $(f_k)_{k\geq 1}$ est décroissante et converge vers la fonction caractéristique $\chi_{\overline{A}}$. Pour tout $k \geq 1$, on a

$$\limsup_{m\to+\infty} \mu_m(A) \leq \limsup_{m\to+\infty} \int f_k \, d\mu_m = \lim_{m\to+\infty} \int f_k \, d\mu_m = \int f_k \, d\mu.$$

Ceci implique que

$$\limsup_{m\to+\infty} \mu_m(A) \leq \inf_{k\geq 1} \int f_k \, d\mu = \lim_{k\to+\infty} \int f_k \, d\mu = \mu(\overline{A}).$$

Le même raisonnement appliqué au complémentaire de A nous dit que

$$\liminf_{m\to+\infty} \mu_m(A) \geq \mu(\mathrm{Int}(A)).$$

Puisque, par hypothèse on a $\mu(\overline{A}) = \mu(\mathrm{Int}(A))$, on peut conclure. □

Lemme 2.5.4 : *Soit X un espace métrique compact et μ une mesure borélienne de probabilité. Pour tout $\varepsilon > 0$, il existe une partition borélienne $\mathcal{P} = (P_i)_{i\in I}$ telle que pour tout $i \in I$, on a* $\mathrm{diam}(P_i) \leq \varepsilon$ *et* $\mu(\partial P_i) = 0$.

Démonstration. Pour tout $x \in X$ il existe $\varepsilon_x \in]0, \varepsilon/2[$ tel que

$$\mu\left(\{x' \in X \mid d(x, x') = \varepsilon_x\}\right) = 0.$$

ce qui implique que $\mu\left(\partial B(x, \varepsilon_x)\right) = 0$. Considérons un sous-recouvrement fini $(U_i)_{1\leq i\leq r}$ du recouvrement $(B(x, \varepsilon_x))_{x\in X}$ et définissons une partition mesurable $(P_i)_{1\leq i\leq r}$ en posant

$$P_i = U_i \setminus \bigcup_{1\leq i'<i} U_{i'}.$$

Chaque U_i a un diamètre inférieur à ε et sa frontière est de mesure nulle puisqu'elle est incluse dans $\bigcup_{1\leq i\leq r} \partial U_i$. □

Lemme 2.5.5 : *Pour tout $\varepsilon > 0$, il existe $\mu \in \mathcal{M}_T$ tel que*

$$h_\mu(T) \geq \limsup_{n\to+\infty} \frac{1}{n} \ln\left(s(n, \varepsilon)\right).$$

Démonstration. Pour tout $n \geq 1$, choisissons un ensemble (n, ε)-séparé S_n de cardinal $s(n, \varepsilon)$. Considérons ensuite les mesures

$$\nu_n = \frac{1}{s(n, \varepsilon)} \sum_{x\in S_n} \delta_x$$

et

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} T_*^i(\nu_n).$$

On peut trouver une suite strictement croissante $(n_k)_{k\geq 0}$ dans $\mathbf{N}$ telle que, d'une part, on ait

$$\lim_{k\to+\infty} \frac{1}{n_k} \ln\left(s(n_k, \varepsilon)\right) = \limsup_{n\to+\infty} \frac{1}{n} \ln\left(s(n, \varepsilon)\right)$$

et telle que, d'autre part, la suite $(\mu_{n_k})_{k\geq 0}$ converge pour la topologie faible* vers une mesure de probabilité μ. Cette mesure μ est alors invariante. En effet, pour toute fonction continue $f\ : X \to \mathbf{R}$, on a

$$\int (f\circ T - f)\, d\mu = \lim_{k\to+\infty} \int (f\circ T - f)\, d\mu_{n_k} = \lim_{k\to+\infty} \frac{1}{n_k} \int (f\circ T^{n_k} - f)\, d\nu_{n_k} = 0,$$

puisque

$$\left| \int (f\circ T^{n_k} - f)\, d\nu_{n_k} \right| \leq 2 \max_{x\in X} |f(x)|,$$

(nous venons de refaire l'argument de la preuve du théorème de Krylov-Bogolioubov). Nous allons montrer que μ satisfait la conclusion du lemme.

Fixons une partition borélienne $\mathcal{P} = (P_i)_{i\in I}$ telle que pour tout $i \in I$, on ait $\operatorname{diam}(P_i) \leq \varepsilon$ et $\mu(\partial P_i) = 0$. On va montrer que

$$h_\mu(T, \mathcal{P}) \geq \limsup_{n\to+\infty} \frac{1}{n} \ln\left(s(n,\varepsilon)\right),$$

ce qui prouvera le lemme puisque $h_\mu(T) \geq h_\mu(T,\mathcal{P})$. Définissons la suite de partitions $(\mathcal{P}^n)_{n\geq 1}$ où $\mathcal{P}^n = \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})$. Puisque S_n est (n,ε)-séparé, chaque élément de $\mathcal{P}^n$ contient au plus un point de S_n, ce qui implique que

$$H_{\nu_n}(\mathcal{P}^n) = \ln s(n,\varepsilon).$$

Sous-lemme 2.5.6 : *Pour toute partition borélienne $\mathcal{Q} = (Q_j)_{j\in J}$ et tous entiers $q \leq n$, on a*

$$qH_{\nu_n}(\mathcal{Q}^n) \leq nH_{\mu_n}(\mathcal{Q}^q) + 2q^2 \ln(\sharp J),$$

où $\mathcal{Q}^m = \bigvee_{i=0}^{m-1} T^{-i}(\mathcal{Q})$.

Démonstration. Fixons $q \geq 1$. Pour tout $r < q$ on a

$$\mathcal{Q}^n = \left(\bigvee_{j=0}^{j_r-1} T^{-jq-r}(\mathcal{Q}^q) \right) \vee \left(\bigvee_{i=0}^{r-1} T^{-i}(\mathcal{Q}) \right) \vee \left(\bigvee_{i=qj_r+r}^{n-1} T^{-i}(\mathcal{Q}) \right),$$

où j_r est l'entier tel que

$$n-1-q < r + qj_r - 1 \leq n-1.$$

On obtient

$$\begin{aligned} H_{\nu_n}(\mathcal{Q}^n) &\leq \sum_{j=0}^{j_r-1} H_{\nu_n}(T^{-jq-r}(\mathcal{Q}^q)) + \sum_{i=0}^{r-1} H_{\nu_n}(T^{-i}(\mathcal{Q})) + \sum_{i=qj_r+r}^{n-1} H_{\nu_n}(T^{-i}(\mathcal{Q})) \\ &\leq \sum_{j=0}^{j_r-1} H_{\nu_n}(T^{-jq-r}(\mathcal{Q}^q)) + 2q\ln(\sharp J). \end{aligned}$$

En sommant sur r, on obtient

$$qH_{\nu_n}(\mathcal{Q}^n) \leq \sum_{i=0}^{n-1} H_{\nu_n}(T^{-i}(\mathcal{Q}^q)) + 2q^2 \ln(\sharp J).$$

Pour obtenir le résultat, on va maintenant utiliser la concavité de la fonction $\phi\ : t \mapsto -t\ln t$. En effet, pour toute partition borélienne $\mathcal{S} = (S_k)_{j\in K}$, on a

$$\begin{aligned}
H_{\mu_n}(\mathcal{S}) &= \sum_{k\in K} \phi(\mu_n(S_k)) \\
&= \sum_{k\in K} \phi\left(\frac{1}{n}\sum_{i=0}^{n-1} T_*^i(\nu_n)(S_k)\right) \\
&\geq \sum_{k\in K} \frac{1}{n}\sum_{i=0}^{n-1} \phi\left(T_*^i(\nu_n)(S_k)\right) \\
&= \frac{1}{n}\sum_{i=0}^{n-1} H_{T_*^i\nu_n}(\mathcal{S}), \\
&= \frac{1}{n}\sum_{i=0}^{n-1} H_{\nu_n}(T^{-i}(\mathcal{S})).
\end{aligned}$$

□

Si on applique le sous-lemme à la partition $\mathcal{P}_{n_k}$ et qu'on divise par qn_k, on obtient :

$$\frac{1}{n_k}\ln\left(s(n_k,\varepsilon)\right) = \frac{1}{n_k} H_{\nu_{n_k}}(\mathcal{P}^{n_k}) \leq \frac{1}{q} H_{\mu_{n_k}}(\mathcal{P}^q) + \frac{2q}{n_k}\ln(\sharp I).$$

Puisque la frontière de chaque élément de $\mathcal{P}^q$ est de mesure nulle, on obtient, en faisant tendre k vers $+\infty$:

$$\limsup_{n\to+\infty} \frac{1}{n}\ln\left(s(n,\varepsilon)\right) \leq \frac{1}{q} H_\mu(\mathcal{P}^q),$$

Faisons tendre maintenant q vers $+\infty$, pour obtenir

$$\limsup_{n\to+\infty} \frac{1}{n}\ln\left(s(n,\varepsilon)\right) \leq h_\mu(T,\mathcal{P}) \leq h_\mu(T).$$

□

La preuve de la seconde partie du théorème, c'est-à-dire l'inégalité

$$\sup_{\mu\in\mathcal{M}_T} h_\mu(T) \geq h(\mu)$$

découle alors immédiatement de l'égalité

$$h(T) = \lim_{\varepsilon\to 0}\limsup_{n\to+\infty} \frac{1}{n}\ln s(n,\varepsilon).$$

□

On peut se demander si l'entropie topologique est atteinte par une mesure, c'est-à-dire s'il existe une mesure $\mu \in \mathcal{M}_T$ telle que $h_\mu(T) = h(T)$. Ceci serait vrai, si l'application $\mu \mapsto h_\mu(T)$ était continue (ou même semi-continue supérieurement) pour la topologie faible* puisque $\mathcal{M}_T$ est compact. Malheureusement cette application n'est généralement pas continue et il se peut que l'entropie topologique ne soit pas atteinte. Cependant, comme nous allons le voir, c'est le cas dès que l'application T est *expansive*, c'est-à-dire s'il existe ε_0 (constante d'expansivité) tel que pour tous points distincts x et y, il existe $n \geq 0$ tel que $d(T^n(x), T^n(y)) \geq \varepsilon_0$.

Proposition 2.5.7 : *Soit $T : X \to X$ une application expansive définie sur un espace métrique compact X. Il existe $\mu \in \mathcal{M}_T$ tel que*

$$h_\mu(T) = h(T).$$

Démonstration. On vérifie facilement que le recouvrement $\mathcal{U}^{\varepsilon_1}$ par les boules ouvertes de rayon ε_1 est générateur, si $2\varepsilon_1$ est une constante d'expansivité. On en déduit que

$$h(T) = h(T, \mathcal{U}^{\varepsilon_1}) = \lim_{n\to+\infty} \frac{1}{n} \ln(N(n, \varepsilon_1))$$

où

$$N(n, \varepsilon) = N\left(\bigvee_{i=0}^{n-1} \mathcal{U}^{\varepsilon}.\right)$$

On a vu précédemment que $N(n, \varepsilon) \le s(n, \varepsilon)$ et donc que

$$h(T) \le \limsup_{n\to+\infty} \frac{1}{n} \ln(s(n, \varepsilon_1)),$$

et on vient juste de voir qu'il existe une mesure $\mu \in \mathcal{M}_T$ telle que

$$h_\mu(T) \ge \limsup_{n\to+\infty} \frac{1}{n} \ln(s(n, \varepsilon_1))$$

□

En fait dans le cas d'un système expansif, l'application $\mu \mapsto h_\mu(T)$ est semi-continue supérieurement. Les cas les plus simples de systèmes dynamiques expansifs sont donnés par le décalage de Bernouilli unilatéral $\sigma : A^{\mathbf{N}} \to A^{\mathbf{N}}$, où A est un alphabet fini de cardinal $p \ge 2$, et les endomorphismes du cercle $T : x \to px$, où $|p| \ge 2$. Dans le premier cas la mesure équidistribuée, c'est-à-dire celle pour laquelle tout cylindre

$$C(m, a_0, a_1, \ldots, a_m) = \{(x_n)_{n\ge 0} \,|\, x_i = a_i \text{ si } i \le m\}$$

est de mesure $\frac{1}{p^{m+1}}$, a une entropie égale à $\ln p = h(\sigma)$; dans le second cas la mesure de Lebesgue a une entropie égale à $\ln |p| = h(T)$. On peut se demander également si l'entropie topologique peut-être atteinte par plusieurs mesures. La réponse est évidement oui (pensons au cas où l'entropie est nulle et où il y a plusieurs mesures invariantes, une rotation d'angle rationnel par exemple) ou plus simplement, même dans le cas d'une entropie strictement positive, prenons la réunion disjointe de deux systèmes dynamiques. Nous verrons cependant que pour les systèmes "hyperboliques" il n'y a généralement qu'une seule mesure invariante d'entropie maximale. Nous allons commencer par un modèle-type de ces systèmes, les sous-décalages de type fini.

3.1 Entropie topologique des sous-décalages de type fini

On s'intéressera aux sous-décalages bilatéraux, on pourrait définir de même les sous-décalages unilatéraux. On se donne une matrice carrée $A = (A_{i,j})_{i,j}$ d'ordre p à coefficients égaux à 0 ou 1. L'ensemble

$$X_A = \{\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in \{1,\ldots,p\}^{\mathbf{Z}} \mid A_{i_k,i_{k+1}} = 1 \text{ pour tout } k \in \mathbf{Z}\}$$

est fermé et invariant par le décalage $\sigma : (i_k)_{k\in\mathbf{Z}} \to (i_{k+1})_{k\in\mathbf{Z}}$. La restriction $\sigma_A = \sigma|_{X_A}$ est un *sous-décalage de type fini.*

On supposera qu'il existe toujours un 1 sur chaque ligne et sur chaque colonne, ce qui implique que pour tout $i \in \{1,\ldots,p\}$ et pour tout $k_0 \in \mathbf{Z}$, il existe $\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in X_A$ tel que $i_{k_0} = i$. On appellera *mot* de *longueur* n toute suite $w = (i_k)_{0\le k\le n}$ dans $\{1,\ldots,p\}$. On dira que ce mot *joint* i_0 à i_n. On dira que le mot $w = (i_k)_{0\le k\le n}$ est *admissible* s'il peut être étendu à une suite $\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in X_A$, c'est-à-dire si $A_{i_k,i_{k+1}} = 1$ pour tout $k \in \{0,\ldots,n-1\}$. Les cylindres qui engendrent la topologie de X_A sont les *cylindres admissibles* qui s'écrivent

$$C_w^{k_0} = \{\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in X_A \,|\, i_{k+k_0} = w_k \text{ pour tout } k \in \{0,\ldots,n\}\},$$

où $w = (w_k)_{0\le k\le n}$ est un mot admissible et $k_0 \in \mathbf{Z}$, la *base* d'un tel cylindre est $\{k_0,\ldots,n+k_0\}$.

Nous allons chercher à calculer l'entropie topologique de σ_A. Commençons par le lemme suivant :

LEMME 3.1.1 : *Pour tous i, j dans $\{1,\ldots,p\}$, le nombre de mots admissibles de longueur $n \ge 1$ qui joignent i à j est égale à $A_{i,j}^n$ le coefficient (i,j) de A^n.*

Démonstration. En effet,

$$A_{i,j}^n = \sum_{1\le i_1\le p,\ldots,\,1\le i_{n-1}\le p} A_{i,i_1}A_{i_1,i_2}\ldots A_{i_{n-1},j}.$$

Remarquons maintenant que $A_{i,i_1}A_{i_1,i_2}\ldots A_{i_{n-1},j}$ vaut 1 si $(i,i_1,\ldots,i_{n-1},j)$ est admissible et 0 sinon. □

PROPOSITION 3.1.2 : *L'application σ_A est positivement transitive si, pour tous i, j dans $\{1,\ldots,p\}$, il existe $n \ge 1$ tel que $A_{i,j}^n > 0$. Dans ce cas, on dira que A est* irréductible.

Démonstration. Considérons les cylindres $C_i^0 = \{\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in X_A \,|\, i_0 = i\}$. Si σ_A est positivement transitive, alors, pour tous i, j dans $\{1,\ldots,p\}$, il existe $n \ge 1$ tel que $C_i^0 \cap \sigma_A^{-n}(C_j^0) \ne \emptyset$. Ceci signifie qu'il existe un mot admissible de longueur n qui joint i à j. Par le lemme précédent, cela signifie également qu'il existe $n \ge 1$ tel que $A_{i,j}^n > 0$.

Réciproquement, supposons que pour tous i, j dans $\{1,\ldots,p\}$, il existe $n \ge 1$ tel que $A_{i,j}^n \ne 0$. Soient $C_w^{k_0}$ et $C_{w'}^{k_0'}$ deux cylindres. On veut montrer qu'il existe $k \ge 0$ tel que $C_w^{k_0} \cap \sigma_A^{-k}(C_{w'}^{k_0'}) \ne \emptyset$. Quitte à prendre des cylindres plus petits, on peut supposer que la base commune de $C_w^{k_0}$ et $C_{w'}^{k_0'}$ est $\{-N,\ldots,N\}$. Il existe un mot admissible w'' de longueur $n \ge 1$ qui joint la dernière coordonnée de w à la première coordonnée de w'. Ceci implique que le mot $ww''w'$ construit par assemblage est admissible. Ainsi $C_w^{k_0} \cap \sigma_A^{-2N-n}(C_{w'}^{k_0'}) \neq \emptyset$. $\square$

Avant de donner la formule de l'entropie topologique de σ_A rappelons la définition du rayon spectral.

PROPOSITION 3.1.3 : *Soit E un espace vectoriel réel ou complexe de dimension p et $\mathcal{L}(E)$ l'espace des endomorphismes de E. Si $\| \ \|$ est une norme sur $L(E)$, alors pour tout $L \in \mathcal{L}(E)$, la suite $(\|L^n\|^{\frac{1}{n}})_{n\geq 0}$ converge et sa limite $\rho(L)$ ne dépend pas de $\| \ \|$. C'est le* rayon spectral *de L, il est égal au plus grand module des valeurs propres (complexes) de L. De plus, pour tout $\varepsilon > 0$, il existe une norme $\| \ \|$ sur E telle que $\|L\| \leq \rho(L) + \varepsilon$ où $\| \ \|$ est la norme d'opérateur associée :*

$$\|L\| = \max_{\|v\|=1} \|L(v)\|.$$

Démonstration. Les normes sur $\mathcal{L}(E)$ étant toutes équivalentes, les quantités

$$\liminf_{n\to+\infty} \|L^n\|^{\frac{1}{n}}, \ \ \limsup_{n\to+\infty} \|L^n\|^{\frac{1}{n}}$$

ne dépendent pas de la norme $\| \ \|$.

Commençons par prouver le résultat dans le cas où E est un espace vectoriel sur $\mathbf{C}$. Choisissons une valeur propre $\lambda \in \mathbf{C}$ de module maximal et remarquons que $\liminf_{n\to+\infty} \|L^n\|^{\frac{1}{n}} \geq |\lambda|$ si $\| \ \|$ est une norme d'opérateur. Pour montrer que $\limsup_{n\to+\infty} \|L^n\|^{\frac{1}{n}} \leq |\lambda|$ il suffit de trouver, pour tout $\varepsilon > 0$, une norme $\| \ \|$ sur E telle que $\|L\| \leq |\lambda| + \varepsilon$ pour la norme d'opérateur associée. On sait qu'il existe une base $(v_i)_{1\leq i\leq p}$ dans laquelle la matrice de L, notée A, est triangulaire supérieure. Quitte à choisir $\delta > 0$ petit et à remplacer la base $(v_i)_{1\leq i\leq p}$ par la base $(\delta^{i-1}v_i)_{1\leq i\leq p}$, on peut supposer que les coefficients non diagonaux de A ont tous un module inférieur à ε/p. Considérons la norme

$$\|v\| = \max_{1\leq i\leq p} |x_i|,$$

en notant x_i les coordonnées de v dans la base $(v_i)_{1\leq i\leq p}$ et remarquons que $\|L(v)\| \leq (|\lambda| + \varepsilon)\|v\|$.

Pour étudier le cas où E est un espace vectoriel sur $\mathbf{R}$ on considère l'extension

$$L_{\mathbf{C}} \ : u + iv \mapsto L(u) + iL(v)$$

au complexifié $E_{\mathbf{C}} \sim E + iE$. Chaque norme $\| \ \|$ sur $E_{\mathbf{C}}$ induit par restriction une norme sur E et on a $\|u + iv\| \leq \|u\| + \|v\|$. En considérant les normes d'opérateurs associées, on obtient $\|L^n\| \leq \|L_{\mathbf{C}}^n\|$, ce qui implique que $\limsup_{n\to+\infty} \|L^n\|^{\frac{1}{n}} \leq |\lambda|$. L'inégalité inverse est évidente si on peut choisir $\lambda \in \mathbf{R}$. Sinon, on considère un vecteur propre $w = u + iv$ associé à λ. On a

$$|\lambda|^n \|w\| \leq \|L^n(u)\| + \|L^n(v)\|,$$

ce qui implique que $\liminf_{n\to+\infty} \|L^n\|^{\frac{1}{n}} \geq |\lambda|$.

$\square$

PROPOSITION 3.1.4 : *L'entropie topologique de σ_A vérifie*

$$h(\sigma_A) = \ln \rho(A).$$

Démonstration. Considérons le recouvrement ouvert $\mathcal{U}$ de X_A formé par les p cylindres $(C_i^0)_{1\leq i\leq p}$. Comme dans le cas du décalage de Bernouilli σ, on remarque que la famille $\left(\bigvee_{i=-n+1}^{n-1}\sigma_A^{-i}(\mathcal{U})\right)_{n\geq 0}$ est génératrice et donc que

$$h(\sigma_A)=\lim_{n\to+\infty}h\left(\sigma_A,\bigvee_{i=-n+1}^{n-1}\sigma_A^{-i}(\mathcal{U})\right)=\lim_{n\to+\infty}h\left(\sigma_A,\bigvee_{i=0}^{2n-2}\sigma_A^{-i}(\mathcal{U})\right)=h(\sigma_A,\mathcal{U}).$$

Remarquons maintenant que le recouvrement $\bigvee_{i=0}^{n-1}\sigma_A^{-i}(\mathcal{U})$ n'est rien d'autre que le recouvrement par les cylindres admissibles C_w^0 de base $\{0,\dots,n-1\}$. Ils sont bien sûr disjoints deux-à-deux et il y en a exactement

$$N\left(\bigvee_{i=0}^{n-1}\sigma_A^{-i}(\mathcal{U})\right)=\sum_{i,j}A_{i,j}^{n-1}.$$

La formule

$$\|M\|=\sum_{i,j}|M_{i,j}|$$

définit une norme sur l'espace des matrices carrées d'ordre p. Le lemme précédent nous dit que

$$h(\sigma_A)=\lim_{n\to+\infty}\frac{1}{n}\ln\|A^{n-1}\|=\lim_{n\to+\infty}\frac{n-1}{n}\ln\left(\|A^{n-1}\|^{\frac{1}{n-1}}\right)=\ln\rho(A).$$

$\square$

3.2 Mesures de Markov, mesure de Parry

Commençons par rappeler l'énoncé du théorème de Perron-Frobenius.

Proposition 3.2.1 : *Soit A une matrice carrée d'ordre p à coefficients positifs. On suppose qu'il existe $N\geq 1$ tel que les coefficients de A^N sont tous strictement positifs. Alors :*

i) *la matrice A a un vecteur propre u à coefficients strictement positifs et tout vecteur propre à coefficients positifs est un multiple de u ;*

ii) *la valeur propre λ correspond à u est simple et strictement positive et tout autre valeur propre λ' vérifie $|\lambda'|<\lambda$.*

Démonstration. Considérons le cône

$$C=\{x=(x_1,\dots,x_p)\in\mathbf{R}^p\ \mid\ x_1\geq 0,\dots,x_p\geq 0\}.$$

Fixons un hyperplan affine H qui rencontre $\mathrm{Int}(C)$ et dont la partie linéaire ne rencontre C qu'en 0 puis définissons la partie compacte convexe $C_H=C\cap H$. On peut prendre, par exemple,

$$H=\{(x_1,\dots,x_p)\in\mathbf{R}^p\ \mid\ x_1+\ldots+x_p=1\}.$$

Puisque les coefficients de A^N sont strictement positifs, on a $Au\neq 0$ pour tout $u\in C\setminus\{0\}$. Par conséquent A induit une application continue $A_H\ :C_H\to C_H$ définie par

$$\mathbf{R}A_H(x)=A(\mathbf{R}x),$$

et les points fixes de A_H correspondent à des vecteurs propres.

Rappelons que le birapport $[a, b, c, d]$ de quatre points sur une droite affine Δ de $\mathbf{R}^p$ est

$$[a, b, c, d] = \frac{(d-a)(c-b)}{(c-a)(d-b)},$$

pour n'importe quel système affine de coordonnées sur Δ. Ce nombre est également invariant par homographie.

On va définir une "distance" d_H sur $\mathrm{Int}(C_H)$. Si x et x' sont deux points distincts de $\mathrm{Int}(C_H)$, on considère la droite Δ_0 passant par x et x', puis les deux points a et a' de $\Delta_0 \cap \mathrm{Fr}(C_H)$, le point a choisi du côté de x et le point a' du côté de x'. On a donc $[a, a', x, x'] > 1$ et on peut définir $d_H(x, x') = \ln([a, a', x, x'])$. Si on pose $d_H(x, x) = 0$, on obtient une fonction symétrique réflexive et il n'est pas difficile de voir que d_H est continue (on peut également prouver, mais c'est plus difficile, que d_H vérifie l'inégalité triangulaire et définit vraiment une distance).

Le fait que A est à coefficients positifs implique que pour tous x et x' on a

$$d_H(A_H(x), A_H(x')) \leq d_H(x, x').$$

C'est évident si $A_H(x) = A_H(x')$, expliquons pourquoi c'est encore vrai sinon. Notons Δ_1 la droite passant par $A_H(x)$ et $A_H(x')$, puis b et b' les points de $\Delta_1 \cap \mathrm{Fr}(C_H)$, le point b choisi du côté de $A_H(x)$ et le point b' du côté de $A_H(x')$. Remarquons que $A_H(\Delta_0 \cap C_H) \subset \Delta_1 \cap C_H$ et que le point $A_H(a)$ (resp. $A_H(a')$) est situé entre b et $A_H(x)$ (resp. entre b' et $A_H(x')$) et donc que

$$[b, b', A_H(x), A_H(x')] \leq [A_H(a), A_H(a'), A_H(x), A_H(x')]$$

avec égalité si et seulement si $A_H(a) = b$ et $A_H(a') = b'$. Remarquons de plus que A_H induit une homographie entre Δ_0 et Δ_1 et donc que

$$[A_H(a), A_H(a'), A_H(x), A_H(x')] = [a, a', x, x'].$$

Puisque A^N est à coefficients strictement positifs, on sait que $A_H^N(C_H)$ est une partie compacte de $\mathrm{Int}(C_H)$. Les inégalités précédentes sur les birapports impliquent donc que

$$d_H(A_H^N(x), A_H^N(x')) < d_H(x, x'),$$

si $x \neq x'$. En particulier A_H a au plus un point fixe. Pour montrer l'existence d'un tel point, considérons un point u où la fonction $x \mapsto d(x, A_H(x))$ atteint son minimum. Il doit être fixe puisqu'on aurait $d_H(A_H^N(u), A_H^{N+1}(u)) < d_H(u, A_H(u))$ dans le cas contraire, en contradiction avec la propriété d'extremum.

Remarquons maintenant que pour tout voisinage U de u dans C_H, il existe $n \geq 1$ tel que $A_H^n(C_H) \subset U$. En effet, la suite $(A_H^n(C_H))_{n\geq 0}$ étant décroissante, l'ensemble $K = \bigcap_{n\geq 0} A_H^n(C_H)$ est compact et vérifie $A_H(K) = K$. Nous voulons montrer que K se réduit à u. Dans le cas contraire on pourrait trouver $u' \in K$ qui maximise $x \mapsto d_H(x, u)$, mais ce point ne pourrait pas avoir d'antécédent par A_H^N dans K.

(En fait on aurait pu montrer, mais cela aurait demander plus de travail, non seulement que d_H était une distance mais également que A_H^N était contractante, puis utiliser le théorème de point fixe de Picard.)

Puisque $u = A_N(u)$, ce point est à l'intérieur de C_H, et nous avons montré l'assertion **i)**. Nous voulons maintenant prouver **ii)**. Notons alors λ la valeur propre associée à u. De façon similaire, nous savons que A^t a un unique vecteur propre v à coefficients strictement positifs qui vérifie $\sum_{i=1}^{p} v_i u_i = 1$. On peut identifier u à un vecteur et v à une forme linéaire qui vérifient $v \circ A = \lambda^* v$ et $v(u) = 1$. En appliquant la première égalité à u, on trouve $\lambda^* = \lambda$. L'hyperplan $\mathrm{Ker}(v)$ est invariant par A et son intersection avec C réduite à 0. Pour montrer **ii)** il suffit de prouver que $\rho\left(A|_{\mathrm{Ker}(v)}\right) < \lambda$. Considérons l'hyperplan affine $H = u + \mathrm{Ker}(v)$ et remarquons que

$$A_H(u+w) = u + \frac{1}{\lambda} A(w).$$

Il existe $\varepsilon > 0$ tel que $u + w \in C_H$ si $w \in \mathrm{Ker}(v)$ vérifie $\|w\| \leq \varepsilon$. En appliquant ce qui a été dit plus haut au voisinage

$$U = \{u + w\,,\ w \in \mathrm{Ker}(v)\,,\ \|w\| \leq \frac{\varepsilon}{2}\},$$

on sait qu'il existe $n \geq 1$ tel que $A_H^n(C_H) \subset U$. En particulier pour tout $w \in \mathrm{Ker}(v)$,

$$\|w\| \leq \varepsilon \Rightarrow \|A^n(w)\| \leq \frac{1}{2}\varepsilon\lambda^n$$

et donc

$$\rho\left(A|_{\mathrm{Ker}(v)}\right) \leq \left(\frac{1}{2}\right)^{\frac{1}{n}} \lambda.$$

□

Remarques

i) Pour tous i, j in $\{1, \dots, p\}$ on a

$$\lim_{n \to +\infty} \frac{1}{\lambda^n} A_{i,j}^n = u_i v_j.$$

En effet, on a

$$\lim_{n \to +\infty} \frac{1}{\lambda^n} A^n u = u$$

ainsi que

$$\lim_{n \to +\infty} \frac{1}{\lambda^n} A^n w = 0$$

pour tout $w \in \mathrm{Ker}(v)$. On peut remarquer que la matrice $uv^t = (u_i v_j)_{i,j}$ fixe u et annule tout vecteur $w \in \mathrm{Ker}(v)$.

ii) Si les coefficients de A sont formés de 0 et de 1, alors $\lambda > 1$. En effet

$$p\lambda^N > \mathrm{Tr}(A^N) \geq p.$$

iii) Si A est une *matrice stochastique*, c'est-à-dire si $\sum_{j=1}^{p} A_{i,j} = 1$ pour tout $i \in \{1, \dots, p\}$, alors $\lambda = 1$. En effet l'hyperplan

$$H = \{(x_1, \dots, x_p) \in \mathbf{R}^p \ \mid\ x_1 + \ldots + x_p = 1\}$$

est invariant par A^t. Les applications A^t et $(A^t)_H$ coïncident, le point fixe de A_H est un point fixe de A. En fait, pour toute matrice stochastique M, il existe au moins au moins un vecteur $v = (v_1, \dots, v_p)$ à coefficients positifs tels que $\sum_{i=1}^{p} v_i = 1$ et qui vérifie $vM = v$. En effet l'endomorphisme L dont M^t est la trace dans la base canonique, laisse invariante la partie compacte convexe C_H. Fixons $w \in C$ et définissons

$$w_n = \frac{1}{n} \sum_{k=0}^{n-1} L^k(w).$$

Remarquons que

$$\|L(w_n) - w_n\| = \left\| \frac{1}{n}(L^n(w) - w) \right\| \leq \frac{1}{n} \mathrm{diam}(C),$$

ce qui implique que toute valeur d'adhérence v de la suite $(w_n)_{n\geq 1}$ est fixe par L. Le théorème de Perron Frobenius nous dit que ce vecteur est unique dans le cas où il existe une puissance de M dont les coefficients sont strictement positifs.

Définissons maintenant ce qu'est une *mesure de Markov*. Soit $M = (M_{i,j})_{i,j}$ une matrice stochastique et $v = (v_1, \ldots, v_p)$ un vecteur à coefficients positifs tels que $\sum_{i=1}^p v_i = 1$ et qui vérifie $vM = v$. Il existe alors une unique mesure borélienne de probabilité μ telle que pour tout $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^m$ et tout $n_0 \in \mathbf{N}$, on a

$$\mu(C_w^{n_0}) = v_{w_0} \left(\prod_{k=0}^{m-2} M_{w_k, w_{k+1}} \right).$$

En effet si on écrit

$$iw = (i, w_0, \ldots, w_{m-1}), \;\; wi = (w_0, \ldots, w_{m-1}, i),$$

pour $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^m$, et $i \in \{1, \ldots, p\}$, la formule précédente implique que pour tout $n_0 \in \mathbf{N}$, on a

$$\sum_{i=1}^{p} \mu(C_i^{n_0}) = 1$$

et que pour tout $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^{m+1}$ et tout $n_0 \in \mathbf{N}$, on a

$$\mu(C_w^{n_0}) = \sum_{i=1}^{p} \mu(C_{iw}^{n_0-1}) = \sum_{i=1}^{p} \mu(C_{wi}^{n_0})$$

(la première égalité de la ligne précédente est due au fait que $vM = v$, la seconde au fait que M est une matrice stochastique). Le théorème d'extension de Caratheodory nous dit qu'il existe alors une unique mesure borélienne de probabilité μ définie par ces égalités. On a

$$\mu(\sigma^{-1}(C_w^{n_0})) = \mu(C_w^{n_0+1}) = \mu(C_w^{n_0}),$$

ce qui prouve que μ est invariante par σ.

On peut calculer l'entropie métrique d'une mesure de Markov. Rappelons que la partition $\mathcal{P} = (P_i)_{1\leq i\leq r}$ où $P_i = \{x = (x_n)_{n\geq 0} \,|\, x_0 = i\}$ est génératrice et donc que

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}) = \lim_{n\to+\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})).$$

Le calcul donne alors

$$\begin{aligned}
H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})) &= - \sum_{i_0,i_1,\dots,i_{n-1}} \mu(C^0_{i_0,i_1,\dots,i_{n-1}}) \ln \Big(\mu(C^0_{i_0,i_1,\dots,i_{n-1}}) \Big) \\
&= - \sum_{i_0,i_1,\dots,i_{n-1}} v_{i_0} M_{i_0,i_1} \dots M_{i_{n-2},i_{n-1}} \ln(v_{i_0} M_{i_0,i_1} \dots M_{i_{n-2},i_{n-1}}) \\
&= - \sum_{i_0,i_1,\dots,i_{n-1}} v_{i_0} M_{i_0,i_1} \dots M_{i_{n-2},i_{n-1}} (\ln v_{i_0} + \ln M_{i_0,i_1} + \dots + \ln M_{i_{n-2},i_{n-1}}) \\
&= - \sum_{i_0} v_{i_0} \ln v_{i_0} - (n-1) \sum_{i,j} v_i M_{i,j} \ln M_{i,j}
\end{aligned},$$

et donc

$$h_\mu(\sigma) = - \sum_{i,j} v_i M_{i,j} \ln M_{i,j}.$$

Proposition 3.2.2 : *Soit μ la mesure de Markov définie par une matrice stochastique M et un vecteur fixe v. S'il existe une puissance de M dont les coefficients sont strictement positifs, alors la mesure μ est mélangeante.*

Preuve. Il suffit de prouver que

$$\lim_{n \to +\infty} \mu(C^{n_0}_w \cap \sigma^{-n}(C^{n'_0}_{w'})) = \mu(C^{n_0}_w) \mu(C^{n'_0}_{w'})$$

pour tous cylindres $C^{n_0}_w$ et $C^{n'_0}_{w'}$. Supposons que $w = (w_0, \dots, w_{m-1})$ et $w' = (w'_0, \dots, w'_{m'-1})$. Si $n > m - 1 + n_0 - n'_0$, alors

$$\mu(C^{n_0}_w \cap \sigma^{-n}(C^{n'_0}_{w'})) = v_{w_0} \left(\prod_{k=0}^{m-2} M_{w_k,w_{k+1}} \right) M^{n-m+1-n_0+n'_0}_{w_{m-1},w'_0} \left(\prod_{k=0}^{m'-1} M_{w'_k,w'_{k+1}} \right),$$

où on écrit $M^n = (M^n_{i,j})_{i,j}$, et donc

$$\lim_{n \to +\infty} \mu(C^{n_0}_w \cap \sigma^{-n}(C^{n'_0}_{w'})) = v_{w_0} \left(\prod_{k=0}^{m-2} M_{w_k,w_{k+1}} \right) v_{w'_0} \left(\prod_{k=0}^{m'-1} M_{w'_k,w'_{k+1}} \right) = \mu(C^{n_0}_w) \mu(C^{n'_0}_{w'}),$$

car on a vu précédemment que $\lim_{n \to +\infty} M^n_{i,j} = v_j$ □

Supposons que A est une matrice carrée d'ordre p à coefficients égaux à 0 ou 1 et que

$$A_{i,j} = 0 \Rightarrow M_{i,j} = 0.$$

Le support de la mesure μ est inclus dans X_A car tout cylindre non admissible est de mesure nulle, ainsi μ est une mesure invariante de σ_A.

S'il existe $N \geq 1$ tel que les coefficients de A^N sont strictement positifs, il existe alors une valeur propre réelle $\lambda = \rho(A) > 1$ plus grande strictement que le module de tout autre valeur propre. On dit alors que A est *irréductible et apériodique*. Dans ce cas, il existe une mesure de Markov particulière appelée *mesure de Parry*. Soient u un vecteur propre à coefficients positifs (et donc associé à λ) et v un vecteur propre de A^t à coefficients positifs (donc associé également à λ) tels que $\sum_{i=1}^p v_i u_i = 1$. Considérons la matrice M, où

$$M_{i,j} = \frac{A_{i,j} u_j}{\lambda u_i}.$$

C'est une matrice stochastique car

$$\sum_{j=1}^{p} A_{i,j} u_j = \lambda u_i,$$

la mesure de Markov associée est la mesure de Parry μ_A, elle est supportée sur X_A. Le vecteur propre de M^t qui est dans $C \cap H$ est $(v_1 u_1, \ldots, v_p u_p)$. Par exemple si $A_{i,j} = 1$ pour tout (i, j), alors $\lambda = p$ et $u_i = v_i = \frac{1}{\sqrt{p}}$. Dans ce cas la mesure de Parry est la mesure équidistribuée.

La mesure de Parry maximise l'entropie :

Proposition 3.2.3 : *On a*

$$h_{\mu_A}(\sigma_A) = h(\sigma_A) = \ln \rho(A).$$

Démonstration. Rappelons que l'entropie métrique de μ_A vaut

$$\begin{aligned} h_{\mu_A}(\sigma_A) &= -\sum_{i,j} (u_i v_i) M_{i,j} \ln(M_{i,j}) \\ &= -\sum_{i,j} (u_i v_i) \frac{A_{i,j} u_j}{\lambda u_i} \ln\left(\frac{A_{i,j} u_j}{\lambda u_i}\right) \\ &= \sum_{i,j} \frac{A_{i,j} v_i u_j}{\lambda} (-\ln(A_{i,j} u_j) + \ln u_i + \ln \lambda). \end{aligned}$$

Mais

$$-\sum_{i,j} \frac{A_{i,j} v_i u_j}{\lambda} \ln(A_{i,j} u_j) = -\sum_{i,j} \frac{A_{i,j} v_i u_j}{\lambda} \ln(u_j) = -\sum_j v_j u_j \ln u_j,$$

$$\sum_{i,j} \frac{A_{i,j} v_i u_j}{\lambda} \ln u_i = \sum_i v_i u_i \ln u_i$$

et

$$\sum_{i,j} \frac{A_{i,j} v_i u_j}{\lambda} \ln \lambda = \ln \lambda \sum_i v_i u_i = \ln \lambda.$$

□

En fait, c'est la seule mesure qui maximise l'entropie.

Proposition 3.2.4 : *La mesure de Parry μ_A est la seule mesure $\mu \in \mathcal{M}_{\sigma_A}$ telle que $h_\mu(\sigma_A) = h(\sigma_A)$.*

Démonstration. Supposons qu'il existe une mesure $\mu \in \mathcal{M}_{\sigma_A}$ telle que $h_\mu(\sigma_A) = h(\sigma_A)$ et $\mu \neq \mu_A$. Cette mesure n'est pas uniformément continue par rapport à μ_A. Sinon, on pourrait écrire $d\mu = f d\mu_A$, où $f \in L_1(\mu_A)$ est invariante par T. Mais μ_A étant ergodique, f devrait être constante égale à 1. Puisque μ n'est pas uniformément continue par rapport à μ_A, il existe une partie boréliennne B telle que $\mu_A(B) = 0$ et $\mu(B) > 0$. De la même façon qu'on a montré que μ_A et μ étaient régulières, on peut construire une suite $(B_l)_{l\geq 0}$ formée de réunions finies de cylindres, telle que

$$\lim_{l\to+\infty} \mu(B_l) = \mu(B) \text{ et } \lim_{l\to+\infty} \mu_A(B_l) = \mu_A(B) = 0.$$

La partition $\mathcal{P} = (C_i^0)_{1\leq i\leq p}$ étant génératrice, on sait que

$$h_\mu(T) = h_\mu(T,\mathcal{P}) = \lim_{n\to+\infty} \frac{1}{2n-1} H_\mu\left(\bigvee_{i=0}^{2n-2} T^{-i}(\mathcal{P})\right) = \lim_{n\to+\infty} \frac{1}{2n-1} H_\mu\left(\bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right).$$

Remarquons que pour tout cylindre admissible $C_w^{k_0}$ associé à un mot de longueur n, on a

$$\mu_A(C_w^{k_0}) = \left(\prod_{k=0}^{m-1} A_{i_k,i_{k+1}}\right) v_{i_0} u_{i_m} \lambda^{-m} \geq C\lambda^{-m}$$

où $C = \inf_{i,j} v_i u_j > 0$.

Lemme 3.2.5 : *Pour toute famille $(x_i)_{1\leq i\leq p}$ de nombre positifs*

$$\sum_{i=1}^p x_i = a \Rightarrow -\sum_{i=1}^p x_i \ln x_i \leq a \ln\left(\frac{p}{a}\right).$$

Démonstration. La fonction $\phi : t \mapsto -t\ln t$ étant concave sur $]0,+\infty[$, on a

$$\frac{1}{p}\sum_{i=1}^p \phi(x_i) \leq -\left(\frac{\sum_{i=1}^p x_i}{p}\right) \ln\left(\frac{\sum_{i=1}^p x_i}{p}\right) = \frac{a}{p}\ln\left(\frac{p}{a}\right).$$

□

Fixons l et choisissons n assez grand pour que B_l soit réunion d'éléments de $\mathcal{P}_n = \bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})$. On peut écrire

$$\begin{aligned}
H_\mu(\mathcal{P}_n) &= \sum_{P\in\mathcal{P}_n, P\subset B_l} \phi(\mu(P)) + \sum_{P\in\mathcal{P}_n, P\not\subset B_l} \phi(\mu(P)) \\
&\leq \mu(B_l)\ln\left(\frac{\sharp(\{P\in\mathcal{P}_n, P\subset B_l\})}{\mu(B_l)}\right) + (1-\mu(B_l))\ln\left(\frac{\sharp(\{P\in\mathcal{P}_n, P\not\subset B_l\})}{1-\mu(B_l)}\right) \\
&\leq \mu(B_l)\ln\left(\frac{\lambda^{2n-1}\mu_A(B_l)}{C\mu(B_l)}\right) + (1-\mu(B_l))\ln\left(\frac{\lambda^{2n-1}(1-\mu_A(B_l))}{C(1-\mu(B_l))}\right) \\
&= \mu(B_l)\ln\left(\frac{\mu_A(B_l)}{\mu(B_l)}\right) + (1-\mu(B_l))\ln\left(\frac{1-\mu_A(B_l)}{1-\mu(B_l)}\right) + (2n-1)\ln\lambda - \ln C
\end{aligned}$$

Remarquons maintenant que cette quantité est inférieure à $2n\ln\lambda$ si l est assez grand. On a une contradiction puisque

$$h_\mu(T) = \inf_{n\geq 1} \frac{1}{2n-1} H_\mu\left(\bigvee_{i=0}^{2n-2} T^{-i}(\mathcal{P})\right) = \inf_{n\geq 1} \frac{1}{2n-1} H_\mu\left(\bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right).$$

$\square$

3.3 Croissance des orbites périodiques

On se donne là-encore une matrice carrée $A = (A_{i,j})_{i,j}$ d'ordre p à coefficients égaux à 0 ou 1, et on veut étudier les points périodiques de σ_A.

LEMME 3.3.1 : *Pour tout $n \geq 1$, on a*

$$\sharp\mathrm{Fix}(\sigma_A^n) = \mathrm{Tr}(A^n).$$

Démonstration. Remarquons que les points fixes de σ_A^n sont naturellement associés aux mots admissibles de longueur n dont les extrémités sont égales. Ainsi, d'après le lemme 2.1.1, on a

$$\sharp\mathrm{Fix}(\sigma_A^n) = \sum_{1 \leq i \leq p} A_{i,i}^n = \mathrm{Tr}(A^n).$$

$\square$

Dans le cas où A est irréductible et apériodique, on en déduit que $h(\sigma_A)$ indique le taux de croissance des orbites périodiques :

PROPOSITION 3.3.2 : *Si A est irréductible et apériodique, alors*

$$\lim_{n \to +\infty} \frac{1}{n} \ln\left(\sharp\mathrm{Fix}(\sigma_A^n)\right) = h(\sigma_A) = \ln \rho(A),$$

Démonstration. En effet, on a

$$\sharp\mathrm{Fix}(\sigma_A^n) = \mathrm{Tr}(A^n) = \sum_{1 \leq i \leq p} \lambda_i^n,$$

où $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres de A. Ceci implique que

$$\lim_{n \to +\infty} \frac{\sharp\mathrm{Fix}(\sigma_A^n)}{\lambda^n} = 1.$$

$\square$

Le résultat qui suit nous dit que la mesure de Parry μ_A est la limite, pour la topologie faible*, des mesures équidistribuées sur les orbites périodiques.

PROPOSITION 3.3.3 : *On a $\mu_A = \lim_{n \to +\infty} \mu_n$, où* $\mu_n = \dfrac{1}{\sharp(\mathrm{Fix}(\sigma_A^n))} \displaystyle\sum_{\mathbf{i} \in \mathrm{Fix}(\sigma_A^n)} \delta_{\mathbf{i}}$.

Démonstration. On doit prouver que pour tout cylindre $C_w^{k_0}$, on a

$$\lim_{n \to +\infty} \mu_n(C_w^{k_0}) = \mu_A(C_w^{k_0}).$$

Écrivons $P_n = \mathrm{Fix}(\sigma_A^n)$ et rappelons que $\sharp P_n = \mathrm{Tr}(A^n)$. Fixons $w = (i_0, \ldots, i_m) \in \{1, \ldots, p\}^{m+1}$. Pour tout $n > m$, on a

$$\begin{aligned}\mu_n(C_w^{k_0}) &= \frac{\sharp(P_n \cap C_w^{k_0})}{\mathrm{Tr}(A^n)} \\ &= \frac{1}{\mathrm{Tr}(A^n)} \left(\prod_{k=0}^{m-1} A_{i_k, i_{k+1}} \right) A_{i_m, i_0}^{n-m}.\end{aligned}$$

On a vu que

$$\lim_{n \to +\infty} \frac{1}{\lambda^n} \mathrm{Tr}(A^n) = 1$$

et

$$\lim_{n \to +\infty} \frac{1}{\lambda^n} A_{i_m, i_0}^n = u_{i_m} v_{i_0}.$$

Ainsi, on a

$$\begin{aligned}\lim_{n \to +\infty} \mu_n(C_w^{k_0}) &= \left(\prod_{k=0}^{m-1} A_{i_k, i_{k+1}} \right) \frac{u_{i_m} v_{i_0}}{\lambda^m} \\ &= \left(\prod_{k=0}^{m-1} M_{i_k, i_{k+1}} \right) v_{i_0} u_{i_0} = \mu_A(C_w^{k_0}).\end{aligned}$$

□

4.1 Endomorphismes hyperboliques.

Définition: Soit $(E, \| \ \|_0)$ un espace de Banach. Un endomorphisme $u : E \to E$ est *hyperbolique* s'il existe une décomposition $E = E^s \oplus E^u$ en sous-espaces fermés, une norme $\| \ \|$ équivalente à $\| \ \|_0$ et un réel $\lambda \in (0,1)$ tels que :

i) $u|_{E^s}$ est un endomorphisme de E^s (i.e. $u(E^s) \subset E^s$);

ii) $u|_{E^u}$ est un automorphisme de E^u (en particulier $u(E^u) = E^u$);

iii) $\|u|_{E^s}\| \leq \lambda$ et $\|(u|_{E^u})^{-1}\| \leq \lambda$;

iv) pour tout $x \in E$, on a $\|x\| = \max(\|x^s\|, \|x^u\|)$, où $x = x^s + x^u$ est la décomposition de x dans $E^s \oplus E^u$.

On dit que $E = E^s \oplus E^u$ est une *décomposition hyperbolique*, que $\| \ \|$ est une *norme adaptée* et que u est *λ-hyperbolique.*

Le résultat qui suit nous dit que la décomposition est unique :

Proposition 4.1.1 : *Si $u : E \to E$ est hyperbolique, les sous-espaces E^s et E^u sont caractérisés comme suit :*

- *l'espace E^s est l'ensemble des points $x \in E$ tels que $\lim_{n\to+\infty} u^n(x) = 0$;*
- *l'espace E^u est l'ensemble des points x tels qu'il existe une suite $(x_n)_{n\geq 0}$ vérifiant $\lim_{n\to+\infty} x_n = 0$ et $u^n(x_n) = x$.*

Démonstration. Soit $\| \ \|$ une norme adaptée. Alors, pour tout $x = x^s + x^u \in E = E^s \oplus E^u$, on a $u^n(x) = u^n(x^s) + u^n(x^u)$, où $\|u^n(x^s)\| \leq \lambda^n \|x^s\|$ et $\|u^n(x^u)\| \geq \lambda^{-n}\|x^u\|$. Ceci implique que

$$\lim_{n\to+\infty} \|u^n(x)\| = \begin{cases} 0 & \text{si } x_u = 0, \\ +\infty & \text{si } x_u \neq 0; \end{cases}$$

en d'autres termes

$$\lim_{n\to+\infty} \|u^n(x)\| = \begin{cases} 0 & \text{si } x \in E_s, \\ +\infty & \text{si } x \notin E_s. \end{cases}$$

Ces conditions caractérisent E_s car elles ne dépendent pas de la norme mais de sa classe d'équivalence..

Remarquons maintenant que pour tout $x \in E^u$, on a $\lim_{n\to+\infty} x_n = 0$, où $x_n = ((u|_{E^u})^{-1})^n(x)$. Réciproquement, soit $(x_n)_{n\geq 0}$ une suite telle que $\lim_{n\to+\infty} x_n = 0$ et $u^n(x_n) = x$. Si on écrit $x_n = x_n^s + x_n^u$ et $x = x^s + x^u$, on a $u^n(x_n^s) = x^s$ et $u^n(x_n^u) = x^u$. La norme adaptée $\| \ \|$ étant équivalente à la norme intitiale, on en déduit que $\lim_{n\to+\infty} x_n^s = 0$, ce que implique que $x^s = 0$ car $\|x^s\| \leq \lambda^n \|x_n^s\|$.

Dans le cas où u est un automorphisme, u^{-1} est lui-même hyperbolique et on a

$$\lim_{n\to+\infty} \|u^{-n}(x)\| = \begin{cases} 0 & \text{si } x \in E_u, \\ +\infty & \text{si } x \notin E_u. \end{cases}$$

Remarque : Un endomorphisme d'un espace vectoriel de dimension finie E est hyperbolique si et seulement si les valeurs propres de u ont un module différent de 1. En effet, dans le cas où u est λ-hyperbolique et où $E = E^s \oplus E^u$ est la décomposition hyperbolique, toute valeur propre de u est soit une valeur propre de $u|_{E^s}$ (et donc de module $\leq \lambda$) soit une valeur propre de $u|_{E^u}$ (et donc de module $\geq \lambda^{-1}$). Réciproquement, si les valeurs propres de u ont un module différent de 1, on peut trouver une décomposition $E = E^s \oplus E^u$ en sous-espaces invariants tels que les valeurs propres de $u|_{E^s}$ ont un module < 1 et les valeurs propres de $u|_{E^s}$ un module > 1. Choisissons $\lambda \in]0, 1[$ tel que les valeurs propres de u n'appartiennent pas à $[\lambda, \lambda^{-1}]$. Les rayons spectraux de $u|_{E^s}$ et $(u|_{E^u})^{-1}$ étant strictement inférieurs à λ, on peut trouver une norme $\| \ \|_s$ sur E^s et une norme $\| \ \|_u$ sur E^u tels que les normes d'opérateurs de $u|_{E^s}$ de $(u|_{E^u})^{-1}$ sont inférieurs à λ. On obtient une norme adaptée en posant $\|x\| = \max(\|x^s\|_s, \|x^u\|_u)$.

Dans le cas plus général où E est un espace de Banach, on considère le complexifié $E_{\mathbf{C}}$ et le *spectre* $\mathrm{sp}(u)$ de u, formé des nombres complexes λ tels que $u_{\mathbf{C}} - \lambda \mathrm{Id}_{E_{\mathbf{C}}}$ n'est pas un automorphisme. Cet ensemble est compact. Dire que u est hyperbolique signifie que le spectre ne contient pas de nombre complexe de module 1.

4.2 Automorphismes hyperboliques de $\mathbf{T}^r$.

Fixons une norme $\| \ \|$ sur $\mathbf{R}^r$ et considérons la distance associée d sur $\mathbf{T}^r = \mathbf{R}^r/\mathbf{T}^r$. C'est-à-dire, posons

$$d(\widehat{x}, \widehat{y}) = \inf_{\pi(x)=\widehat{x},\, \pi(y)=\widehat{y}} \|x - y\|,$$

où

$$\begin{aligned} \pi : \mathbf{R}^r &\to \mathbf{T}^r, \\ x &\mapsto x + \mathbf{Z}^r \end{aligned}$$

est la projection. Rappelons qu'il existe $\varepsilon > 0$, tel que pour tout $\widehat{x} \in \mathbf{T}^r$, on a

$$\pi^{-1}(B(\widehat{x}, \varepsilon)) = \bigsqcup_{\pi(x)=\widehat{x}} B(x, \varepsilon)$$

et tel que chaque application $\pi|_{B(x,\varepsilon)}$ est une isométrie de $B(x, \varepsilon)$ sur $B(\widehat{x}, \varepsilon)$.

Soit A un automorphisme hyperbolique de $\mathbf{R}^r$ dont la matrice dans la base canonique est à coefficients entiers et tel que $\det(A) = \pm 1$. On note $\widehat{A}$ l'automorphisme de $\mathbf{T}^r$ relevé par A et $\mathbf{R}^r = E^s \oplus E^u$ la décomposition hyperbolique de A.

Proposition 4.2.1 : *Pour tout $\widehat{x} \in \mathbf{T}^r$ les ensembles*

$$W^s(\widehat{x}) = \{\widehat{y} \in \mathbf{T}^r \ | \ \lim_{k \to +\infty} d(\widehat{A}^k(\widehat{y}), \widehat{A}^k(\widehat{x})) = 0\}$$

et

$$W^u(\widehat{x}) = \{\widehat{y} \in \mathbf{T}^r \ | \ \lim_{k \to -\infty} d(\widehat{A}^k(\widehat{y}), \widehat{A}^k(\widehat{x})) = 0\}$$

sont les projections dans $\mathbf{T}^r$ des espaces affines $x + E^s$ et $x + E^u$, où $x \in \pi^{-1}(\{\widehat{x}\})$.

Démonstration. Pour tout $y \in x + E^s$, on a

$$\lim_{n \to +\infty} \|A^n(y) - A^n(x)\| = \lim_{n \to +\infty} \|A^n(y - x)\| = 0,$$

ce qui implique que

$$\lim_{n \to +\infty} d(\widehat{A}^n(\widehat{y}), \widehat{A}^n(\widehat{x})) = 0,$$

car $d(\widehat{x}', \widehat{y}') \leq \|x' - y'\|$, pour tout x' et y' dans $\mathbf{R}^r$. Ainsi, on a $\pi(x + E^s) \subset W^s(\widehat{x})$.

Fixons $\delta \in (0, \varepsilon)$ tel que

$$\|x - y\| < \delta \Rightarrow \|A(x) - A(y)\| < \varepsilon,$$

pour tout x et y dans $\mathbf{R}^r$. Pour tout $\widehat{y} \in W^s(\widehat{x})$, il existe $N \geq 0$ tel que $d(\widehat{A}^n(\widehat{y}), \widehat{A}^n(\widehat{x})) < \delta$ pour $n \geq N$. Par définition de ε, pour tout $n \geq N$, il existe un unique $k_n \in \mathbf{Z}^r$ tel que $\|A^n(y + k_n) - A^n(x)\| < \varepsilon$ et on a $\|A^n(y + k_n) - A^n(x)\| = d(\widehat{A}^n(\widehat{y}), \widehat{A}^n(\widehat{x})) < \delta$. Observons que $\|A^{n+1}(y + k_n) - A^{n+1}(x)\| < \varepsilon$. Ceci implique que $k_{n+1} = k_n$. Ainsi le point $y + k_N$ appartient à $x + E^s$, ce qui implique que $\widehat{y} \in \pi(x + E^s)$. La preuve est similaire pour $W^u(\widehat{x})$. $\square$

Exemple Supposons que la matrice de A dans la base canonique de $\mathbf{R}^2$ soit

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Les valeurs propres de A sont $\dfrac{3+\sqrt{5}}{2}$ et $\dfrac{3-\sqrt{5}}{2}$, l'équation de la droite E^s est $x_2 = \dfrac{-\sqrt{5}-1}{2}x_1$, celle de la droite E^u est $x_2 = \dfrac{\sqrt{5}-1}{2}x_1$. Les pentes étant irrationnelles, les droites $x + E^s$ et $x + E^u$ se projettent chacune en une partie dense de $\mathbf{T}^2$.

Proposition 4.2.2 : *Pour tout $\widehat{x}$ et $\widehat{y}$ dans $\mathbf{T}^r$, les ensembles $W^s(\widehat{x})$ et $W^u(\widehat{y})$ ont une intersection non vide.*

Démonstration. Les espaces affines $x + E^s$ et $y + E^u$ ont une intersection non vide. $\square$

On peut déduire deux corollaires du résultat précédent (le second pouvant également se déduire d'arguments de théorie ergodique).

Corollaire 4.2.3 : *Pour tout $\widehat{x} \in \mathbf{T}^r$, les ensembles $W^s(\widehat{x})$ et $W^u(\widehat{x})$ sont denses.*

Démonstration. Les ensembles $W^s(\widehat{x})$ et $W^u(\widehat{x})$ se déduisant de $W^s(0)$ et $W^u(0)$ par translation, il suffit de prouver que $W^s(0)$ et $W^u(0)$ sont denses. On va montrer que leur adhérence contient l'ensemble des points périodiques qui, on le sait est dense (c'est $\mathbf{Q}^2/\mathbf{Z}^2$). Puisque 0 est fixe, les ensembles $W^s(0)$ et $W^u(0)$ sont invariants. Si $\widehat{x}$ est un point périodique de $\widehat{A}$ de période q, on peut choisir $\widehat{y} \in W^u(0) \cap W^s(\widehat{x})$. On sait que $\lim_{n\to+\infty} \widehat{A}^{nq}(\widehat{y}) = \widehat{x}$ et donc que $\widehat{x} \in \overline{W^u(0)}$. On démontre de la même façon que $\widehat{x} \in \overline{W^s(0)}$. $\square$

On sait que $\widehat{A}$ préserve la mesure de Haar et que la mesure de Haar est mélangeante, puisque les valeurs propres de A ne sont pas racines de l'unité. La mesure de Haar est donc ergodique, et on en déduit que $\widehat{A}$ est transitif. Donnons une preuve de ce dernier résultat qui ne fait pas appel à la théorie ergodique.

Corollaire 4.2.4 : *L'automorphisme $\widehat{A}$ est positivement transitif.*

Démonstration. Soit U et V deux parties ouvertes non vides. Choisissons un point périodique $\widehat{x} \in U$ et notons q sa période, de même choisissons un point périodique $\widehat{x}' \in U'$ et notons q' sa période. On peut trouver $\widehat{y} \in W^u(\widehat{x}) \cap W^s(\widehat{x}')$ et on sait que $\lim_{k\to-\infty} \widehat{A}^{kq}(\widehat{y}) = \widehat{x}$ et $\lim_{k\to+\infty} \widehat{A}^{kq'}(\widehat{y}) = \widehat{x}'$. On en déduit qu'il existe $n \geq 0$ et $n' \geq 0$ tels que $\widehat{A}^{-n}(\widehat{y}) \in U$ et $\widehat{A}^{n'}(\widehat{y}) \in V$. Ainsi, on a $U \cap \widehat{A}^{-n-n'}(V) \neq \emptyset$. $\square$

4.3 Perturbation des endomorphismes hyperboliques.

Commençons par la version paramètrée du théorème de point fixe de Picard.

THÉORÈME 4.3.1 : *Soit* $(X,)$ *et* (Y,d) *deux espaces métriques,* Y *étant complet. Soit* Φ : $X \times Y \to Y$ *une application continue vérifiant la propriété suivante : il existe* $\lambda \in (0,1)$ *tel que pour tout* x *dans* X *et pour tous* y, y' *dans* Y*, on a*

$$d(\Phi(x,y),\Phi(x,y')) \leq \lambda d(y,y').$$

Alors, pour tout $x \in X$*, il existe une unique solution* $y = \theta(x)$ *de l'équation* $\Phi(x,y) = y$ *et la fonction* $\theta : X \to Y$ *est continue.*

Proof. L'existence d'une unique solution $y = \theta(x)$ de l'équation $\Phi(x,y) = y$ est une conséquence du théorème du point fixe. Pour montrer que θ est continue, considérons x et x' dans X et observons que

$$\begin{aligned} d(\theta(x),\theta(x')) &= d(\Phi(x,\theta(x)),\Phi(x',\theta(x'))) \\ &\leq d(\Phi(x,\theta(x)),\Phi(x',\theta(x))) + d(\Phi(x',\theta(x)),\Phi(x',\theta(x'))) \\ &\leq d(\Phi(x,\theta(x)),\Phi(x',\theta(x))) + \lambda d(\theta(x),\theta(x')), \end{aligned}$$

ce qui implique que

$$d(\theta(x),\theta(x')) \leq \frac{1}{1-\lambda} d(\Phi(x,\theta(x)),\Phi(x',\theta(x))).$$

Fixons $x \in X$. L'application Φ étant continue en $(x,\theta(x))$, on en déduit que

$$\lim_{x' \to x} d(\theta(x),\theta(x')) = 0.$$

□

On en déduit le résultat important suivant :

PROPOSITION 4.3.2 : *Soit* $T : E \to E$ *un endomorphisme* λ*-hyperbolique d'un espace de Banach* E*. Écrivons* $E = E^s \oplus E^u$ *pour la décomposition hyperbolique et considérons une norme adaptée* $\| \ \|$*. Si* $\varphi : E \to E$ *est une application lipschitzienne de rapport* $\varepsilon < \varepsilon_0 = 1 - \lambda$*, alors* $f = T + \varphi$ *a un unique point fixe* $x \in E$*, et on a*

$$\|x\| < \frac{\|\varphi(0)\|}{\varepsilon_0 - \varepsilon}.$$

Démonstration.. Identifions E et $E^s \times E^u$ et écrivons

$$\begin{aligned} T(x^s,x^u) &= (T^s(x^s),T^u(x^u)), \\ \varphi(x^s,x^u) &= (\varphi^s(x^s,x^u),\varphi^u(x^s,x^u)), \\ f(x^s,x^u) &= (f^s(x^s,x^u),f^u(x^s,x^u)). \end{aligned}$$

Définissons maintenant

$$\overline{f} \ : (x^s,x^u) \mapsto (f^s(x^s,x^u), x^u + (T^u)^{-1}(x^u - f^u(x^s,x^u)).$$

Remarquons que $f = (f^s, f^u)$ et $\overline{f} = (f^s, \overline{f}^u)$ ont les mêmes points fixes. Remarquons également que $\overline{f}$ est lipschitzienne de rapport $(\varepsilon + \lambda)$. En effet,

$$\begin{aligned}\|f^s(x^s, x^u) - f^s(y^s, y^u)\| &\leq \|\varphi^s(x^s, x^u) - \varphi^s(y^s, y^u)\| + \|T^s(x^s) - T^s(y^s)\| \\ &\leq (\varepsilon + \lambda)\|(x^s, x^u) - (y^s, y^u)\|,\end{aligned}$$

et

$$\begin{aligned}\|\overline{f}^u(x^s, x^u) - \overline{f}^u(y^s, y^u)\| &= \|x^u - y^u + (T^u)^{-1}(x^u - y^u - f^u(x^s, x^u) + f^u(y^s, y^u))\| \\ &= \|(T^u)^{-1}(T^u(x^u) - T^u(y^u) - f^u(x^s, x^u) + f^u(y^s, y^u) + x^u - y^u)\| \\ &= \|(T^u)^{-1}(\varphi^u(y^s, y^u) - \varphi^u(x^s, x^u) + x^u - y^u)\| \\ &\leq \lambda(\varepsilon\|(x^s, x^u) - (y^s, y^u)\| + \|x^u - y^u\|) \\ &\leq \lambda(\varepsilon + 1)\|(x^s, x^u) - (y^s, y^u)\| \\ &\leq (\varepsilon + \lambda)\|(x^s, x^u) - (y^s, y^u)\|.\end{aligned}$$

Cela implique que $\overline{f}$ a un unique point fixe x. Remarquons également que

$$\|x\| - \|\overline{f}(0)\| \leq \|x - \overline{f}(0)\| = \|\overline{f}(x) - \overline{f}(0)\| \leq (\varepsilon + \lambda)\|x - 0\| = (\varepsilon + \lambda)\|x\|,$$

et donc que

$$\|x\| \leq \frac{1}{1 - \lambda - \varepsilon}\|\overline{f}(0)\| = \frac{1}{\varepsilon_0 - \varepsilon}\|\overline{f}(0)\| \leq \frac{1}{\varepsilon_0 - \varepsilon}\|\varphi(0)\|.$$

□

On va en déduire un résultat fondamental sur les endomorphismes hyperboliques.

Définition. Soit $F : X \to X$ une application continue sur un espace métrique (X, d) et $\alpha > 0$. On dit qu'une suite $(x_i)_{i\in\mathbf{Z}}$ est une *α-pseudo-orbite* si $d(f(x_i), x_{i+1}) < \alpha$ pour tout $i \in \mathbf{Z}$.

Proposition 4.3.3 (lemme de poursuite) : *Soit $T : E \to E$ un endomorphisme λ-hyperbolique sur un espace de Banach E, où $E = E^s \oplus E^u$ est la décomposition hyperbolique et $\| \ \|$ une norme adaptée. Soit $\varphi : E \to E$ une application lipschitzienne de rapport $\varepsilon < \varepsilon_0 = 1-\lambda$. Posons $f = T + \varphi$. Pour toute α-pseudo-orbite $(x_i)_{i\in\mathbf{Z}}$ de f, il existe une unique orbite $(y_i)_{i\in\mathbf{Z}}$ de f telle que*

$$\sup_{i\in\mathbf{Z}} \|y_i - x_i\| < +\infty,$$

et on a

$$\sup_{i\in\mathbf{Z}} \|y_i - x_i\| \leq \frac{\alpha}{\varepsilon_0 - \varepsilon}.$$

Démonstration. Notons $\mathcal{E}$ l'espace des suites $z = (z_i)_{i\in\mathbf{Z}}$ bornées de E, muni de la norme

$$\|z\| = \sup_{i\in\mathbf{Z}} \|z_i\|.$$

C'est un espace de Banach qui se décompose sous la forme $\mathcal{E} = \mathcal{E}^s + \mathcal{E}^u$, où $\mathcal{E}^s$ (resp. $\mathcal{E}^u$) est l'espace des suites bornées de E^s (resp. E^u). Remarquons que

$$\begin{aligned}\mathcal{T} \ &: \mathcal{E} \to \mathcal{E} \\ (z_i)_{i\in\mathbf{Z}} &\mapsto (T(z_{i-1}))_{i\in\mathbf{Z}}\end{aligned}$$

est λ-hyperbolique avec une décomposition hyperbolique $\mathcal{E} = \mathcal{E}^s + \mathcal{E}^u$. Remarquons que la fonction

$$\begin{aligned} \mathcal{F} \ &: \mathcal{E} \to \mathcal{E} \\ (z_i)_{i\in\mathbf{Z}} &\mapsto (\varphi(x_{i-1}+z_{i-1}) + T(x_{i-1}) - x_i)_{i\in\mathbf{Z}} \end{aligned}$$

est bien définie car la suite

$$(\varphi(x_{i-1}+z_{i-1}) + T(x_{i-1}) - x_i)_{i\in\mathbf{Z}} = (\varphi(x_{i-1}+z_{i-1}) - \varphi(x_{i-1}) + f(x_{i-1}) - x_i)_{i\in\mathbf{Z}}$$

est bornée, et qu'elle est lipschitzienne de rapport ε. On en déduit que $\mathcal{T} + \mathcal{F}$ a un unique point fixe $z = (z_i)_{i\in\mathbf{Z}}$ et que

$$\|z\| \leq \frac{\|\mathcal{F}(0)\|}{\varepsilon_0 - \varepsilon} \leq \frac{\alpha}{\varepsilon_0 - \varepsilon}.$$

Posons $y_i = z_i + x_i$. On déduit de l'équation

$$z_i = \varphi(x_{i-1}+z_{i-1}) + T(x_{i-1}) - x_i + T(z_{i-1}),$$

que $(y_i)_{i\in\mathbf{Z}}$est une orbite de f vérifiant

$$\sup_{i\in\mathbf{Z}} \|y_i - x_i\| \leq \frac{\alpha}{\varepsilon_0 - \varepsilon}.$$

L'unicité de $(z_i)_{i\in\mathbf{Z}}$ nous dit que c'est la seule orbite telle que

$$\sup_{i\in\mathbf{Z}} \|y_i - x_i\| < +\infty.$$

□

4.4 Perturbation des automorphismes hyperboliques, résultats de conjugaison.

Dans le cas d'un automorphisme, on peut être plus précis et obtenir le résultat de conjugaison suivant :

Proposition 4.4.1 : *Soit $T : E \to E$ un automorphisme λ-hyperbolique d'un espace de Banach. Écrivons $E = E^s \oplus E^u$ pour la décomposition hyperbolique et considérons une norme adaptée $\| \ \|$. Si $\varphi : E \to E$ est une application bornée et lipschitzienne de rapport $\varepsilon < \varepsilon_1 = \min(\|T^{-1}\|^{-1}, 1-\lambda)$, alors il existe un unique homéomorphisme h de E tel que $h - \mathrm{Id}_E$ soit borné et vérifie $h \circ T \circ h^{-1} = T + \varphi$.*

Démonstration. La proposition se déduira des trois lemmes suivants

Lemme 4.4.2 : *L'application $f = T + \varphi$ est un homéomorphisme.*

Démonstration. Remarquons que

$$f(x) = y \Leftrightarrow T(x) + \varphi(x) = y \Leftrightarrow x = T^{-1}(y) - T^{-1} \circ \varphi(x).$$

La fonction

$$\Phi \ : (x,y) \mapsto T^{-1}(y) - T^{-1} \circ \varphi(x)$$

est continue et chaque application partielle $x \mapsto \Phi(x, y)$ est lipschitzienne de rapport $\|T^{-1}\|\varepsilon$. Par hypothèse, on a $\|T^{-1}\|\varepsilon < 1$, et on peut donc appliquer la version paramètrée du théorème de point fixe. Il existe une fonction continue $\theta\ : E \to E$ telle que

$$f(x) = y \Leftrightarrow x = \theta(y).$$

Ceci signifie que f est un homéomorphisme. □

Lemme 4.4.3 : *Si φ et ψ sont deux applications bornées et lipschitziennes de rapport $\varepsilon < \varepsilon_1$, il existe une unique application continue bornée $\beta\ : E \to E$ telle que*

$$(T + \varphi) \circ (\mathrm{Id}_E + \beta) = (\mathrm{Id}_E + \beta) \circ (T + \psi).$$

Démonstration. L'application $g = T + \psi$ étant un homéomorphisme, l'équation

$$(T + \varphi) \circ (\mathrm{Id}_E + \beta) = (\mathrm{Id}_E + \beta) \circ (T + \psi)$$

peut s'écrire

$$(T + \varphi) \circ (\mathrm{Id}_E + \beta) \circ g^{-1} = \mathrm{Id}_E + \beta,$$

ou encore

$$T \circ \beta \circ g^{-1} + \varphi \circ (\mathrm{Id}_E + \beta) \circ g^{-1} + T \circ g^{-1} - \mathrm{Id}_E = \beta.$$

Notons $\mathcal{E} = C(E, E)$ l'espace des fonctions continues bornées $\beta\ : E \to E$ muni de la norme $\|\ \|$, où $\|\beta\| = \sup_{x \in E} \|\beta(x)\|$, puis définissons deux applications $\mathcal{T}\ : \mathcal{E} \to \mathcal{E}$ et $\mathcal{F}\ : \mathcal{E} \to \mathcal{E}$ de la façon suivante :

$$\begin{aligned}
\mathcal{T}(\beta) &= T \circ \beta \circ g^{-1} \\
\mathcal{F}(\beta) &= \varphi \circ (\mathrm{Id}_E + \beta) \circ g^{-1} + T \circ g^{-1} - \mathrm{Id}_E. \\
&= \varphi \circ (\mathrm{Id}_E + \beta) \circ g^{-1} - \psi \circ g^{-1}
\end{aligned}$$

Remarquons que $\mathcal{T}$ est λ-hyperbolique avec une décomposition hyperbolique $\mathcal{E} = \mathcal{E}^s \oplus \mathcal{E}^u$, où $\mathcal{E}^s = C(E, E^s)$ et $\mathcal{E}^u = C(E, E^u)$, et que $\mathcal{F}$ est lipschitzienne de rapport ε. Par hypothèse, on a $\varepsilon < \varepsilon_1 \leq 1 - \lambda$. Ceci implique que $\mathcal{T} + \mathcal{F}$ a un unique point fixe. □

Lemme 4.4.4 : *Sous les hypothèses du lemme 4.4.3, l'application $h = \mathrm{Id}_E + \beta$ est un homéomorphisme.*

Démonstration. Rappelons que $f = T + \varphi$ et $g = T + \psi$. Le lemme précédent nous dit qu'il existe β et β' dans $\mathcal{E}$ tels que

$$f \circ (\mathrm{Id}_E + \beta) = (\mathrm{Id}_E + \beta) \circ g$$

et

$$(\mathrm{Id}_E + \beta') \circ f = g \circ (\mathrm{Id}_E + \beta').$$

Ainsi, on a

$$f \circ (\mathrm{Id}_E + \beta) \circ (\mathrm{Id}_E + \beta') = (\mathrm{Id}_E + \beta) \circ (\mathrm{Id}_E + \beta') \circ f$$

et

$$g \circ (\mathrm{Id}_E + \beta') \circ (\mathrm{Id}_E + \beta) = (\mathrm{Id}_E + \beta') \circ (\mathrm{Id}_E + \beta) \circ g.$$

L'application

$$(\mathrm{Id}_E + \beta) \circ (\mathrm{Id}_E + \beta') - \mathrm{Id}_E = \beta' + \beta \circ (\mathrm{Id}_E + \beta')$$

étant bornée et continue, le lemme 4.4.3 appliqué au couple (f, f) nous dit que

$$(\mathrm{Id}_E + \beta) \circ (\mathrm{Id}_E + \beta') = \mathrm{Id}_E.$$

De façon analogue, on a

$$(\mathrm{Id}_E + \beta') \circ (\mathrm{Id}_E + \beta) = \mathrm{Id}_E.$$

Ainsi $\mathrm{Id}_E + \beta$ est un homéomorphisme et on a $(\mathrm{Id}_E + \beta)^{-1} = \mathrm{Id}_E + \beta'$. □

4.5 Applications aux homéomorphismes de $\mathbf{T}^r$.

Rappelons que toute application continue $F : \mathbf{T}^r \to \mathbf{T}^r$ est homotope à un unique endomorphisme linéaire $\widehat{F}_* : \mathbf{T}^r \to \mathbf{T}^r$ dont on notera F_* le relèvement linéaire.

Proposition 4.5.1 : *Soit $F : \mathbf{T}^r \to \mathbf{T}^r$ un homéomorphisme tel que F_* est un automorphisme hyperbolique de $\mathbf{R}^r$. Alors $\widehat{F}_*$ est un facteur F. Plus précisément, il existe une unique application continue $H : \mathbf{T}^r \to \mathbf{T}^r$ homotope à l'identité telle que $H \circ F = \widehat{F}_* \circ H$.*

Démontsration. Considérons la décomposition hyperbolique $\mathbf{R}^r = E^s \oplus E^u$ de F_*. Fixons un relèvement $f : \mathbf{R}^r \to \mathbf{R}^r$ de F. On veut trouver une application continue $\beta : \mathbf{R}^r \to \mathbf{R}^r$ invariantes par les translations entières telle que

$$(\mathrm{Id}_{\mathbf{R}^r} + \beta) \circ f = F_* \circ (\mathrm{Id}_{\mathbf{R}^r} + \beta).$$

On peut écrire cette équation sous la forme

$$\beta = F_* \circ f^{-1} + F_* \circ \beta \circ f^{-1} - \mathrm{Id}_{\mathbf{R}^r}.$$

Notons $\mathcal{E}$ l'espace des fonctions $\beta : \mathbf{R}^r \to \mathbf{R}^r$ invariantes par les translations entières. Remarquons que

$$\begin{aligned} \mathcal{T} &: \mathcal{E} \to \mathcal{E} \\ &\beta \mapsto F_* \circ \beta \circ f^{-1} \end{aligned}$$

est hyperbolique avec une décomposition $\mathcal{E} = \mathcal{E}^s \oplus \mathcal{E}^u$, où $\mathcal{E}^s$ (resp. $\mathcal{E}^u$) est l'espace des fonctions $\beta : \mathbf{R}^r \to E^s$ (resp. $\beta : \mathbf{R}^r \to E^u$) invariantes par les translations entières. Puisque la fonction

$$\begin{aligned} \mathcal{F} &: \mathcal{E} \to \mathcal{E} \\ &\beta \mapsto F_* \circ f^{-1} - \mathrm{Id}_{\mathbf{R}^r} \end{aligned}$$

est constante, on déduit de la proposition 4.3.2 qu'il existe une unique application $\beta \in \mathcal{E}$ telle que

$$\beta = F_* \circ f^{-1} + F_* \circ \beta \circ f^{-1} - \mathrm{Id}_{\mathbf{R}^r},$$

c'est-à-dire telle que

$$(\mathrm{Id}_{\mathbf{R}^r} + \beta) \circ f = F_* \circ (\mathrm{Id}_{\mathbf{R}^r} + \beta).$$

L'application $h = \mathrm{Id}_{\mathbf{R}^r} + \beta$ relève une application continue $H : \mathbf{T}^r \to \mathbf{T}^r$ telle que $H \circ F = \widehat{F}_* \circ H$. Il reste à prouver que H est surjective. Ceci peut-être prouvé par des arguments de topologie algébrique, c'est une conséquence du fait que $\deg(H) \neq 0$. On a une preuve très simple dans le cas où $r = 2$. L'ensemble $H(\mathbf{T}^2)$ est compact et invariant par $\widehat{F}_*$, car $H \circ F = \widehat{F}_* \circ H$. L'ensemble des points périodiques de $\widehat{F}_*$ étant dense, il suffit de prouver que $H(\mathbf{T}^2)$ contient tout point périodique $\widehat{x}$. Ce sera le cas si $H(\mathbf{T}^2)$ rencontre $W^s(\widehat{x})$. Fixons un antécédent $x \in \mathbf{R}^2$ de $\widehat{x}$ par la projection de revêtement π et remarquons que $\pi^{-1}(H(\mathbf{T}^2))$ rencontre $W^s(x) = x + E^s$. En effet $\pi^{-1}(H(\mathbf{T}^2)) = h(\mathbf{R}^2)$ est connexe et invariant par les translations entières. □

La proposition suivante nous dit que tout automorphisme linéaire hyperbolique $\widehat{A}$ de $\mathbf{T}^r$ est C^1-*structurellement stable* : si F est un difféomorphisme proche de $\widehat{A}$ en C^1-topologie, alors F est conjugué à A. En d'autres termes, il existe $\varepsilon > 0$ tel que F est conjugué à $\widehat{A}$ si

$$\begin{aligned}
\sup_{\widehat{x}\in\mathbf{T}^r} d(F(\widehat{x}), \widehat{A}(\widehat{x})) &\leq \varepsilon \\
\sup_{\widehat{x}\in\mathbf{T}^r} d(F^{-1}(\widehat{x}), \widehat{A}^{-1}(\widehat{x})) &\leq \varepsilon \\
\sup_{\widehat{x}\in\mathbf{T}^r} \|DF(\widehat{x}) - A\| &\leq \varepsilon \\
\sup_{\widehat{x}\in\mathbf{T}^r} \|DF^{-1}(\widehat{x}) - A^{-1}\| &\leq \varepsilon.
\end{aligned}$$

En fait, on a résultat plus précis :

PROPOSITION 4.5.2 : *Soit A un automorphisme linéaire λ-hyperbolique. Notons $E = E^s \oplus E^u$ la décomposition hyperbolique de A et considérons une norme adaptée $\| \ \|$. Soit $F \ : \mathbf{T}^r \to \mathbf{T}^r$ une application de classe C^1-telle que*

$$\begin{aligned}
&\sup_{\widehat{x}\in\mathbf{T}^r} d(F(\widehat{x}), \widehat{A}(\widehat{x})) \leq \varepsilon_2 = \frac{1}{4} \min_{k\in\mathbf{Z}^r\setminus\{0\}} \|k\|, \\
&\sup_{\widehat{x}\in\mathbf{T}^r} \|DF(\widehat{x}) - A\| \leq \varepsilon_1 = \min(\|A^{-1}\|^{-1}, 1-\lambda).
\end{aligned}$$

Alors F est un difféomorphisme de classe C^1 et il existe un unique homéomorphisme homotope à l'identité $H \ : \mathbf{T}^r \to \mathbf{T}^r$ tel que $H \circ F = \widehat{A} \circ H$.

Démonstration. On sait que la condition

$$\sup_{\widehat{x}\in\mathbf{T}^r} d(F(\widehat{x}), \widehat{A}(\widehat{x})) \leq \varepsilon_2$$

implique que $F_* = A$. Fixons un relèvement $f \ : \mathbf{T}^r \to \mathbf{T}^r$ de F. On peut écrire $f = A + \psi$, où $\psi \ : \mathbf{R}^r \to \mathbf{R}^r$ est une application de classe C^1 invariante par les translations entières. Commençons par prouver que f (et F) sont des difféomorphismes. Pour cela, énonçons la version différentiable du théorème de point fixe avec paramètre.

LEMME 4.5.3 : *Soit E et F deux espaces vectoriels normés, F étant un espace de Banach. Soit $\Phi : E \times F \to F$ une application de classe C^p, $p \geq 1$, telle que*

$$\sup_{x,y} \|D_2\Phi(x, y)\| = \lambda < 1.$$

Alors, pour tout $x \in E$, il existe une unique solution $y = \theta(x)$ de l'équation $\Phi(x, y) = y$. De plus, l'application θ est de classe C^p et on a

$$D\theta(x) = (\mathrm{Id}_F - D_2\Phi(x, \theta(x)))^{-1} \circ D_1\Phi(x, \theta(x)).$$

Démonstration. Chaque application
$$y \mapsto \Phi(x, y)$$
est lipschitzienne de rapport λ. On peut donc appliquer le théorème du point fixe avec paramètre et obtenir une fonction $\theta : E \to F$, où $y = \theta(x)$ est l'unique solution de l'équation $\Phi(x, y) = y$. Pour prouver que θ est de classe C^p, il suffit d'appliquer le théorème des fonctions implicites à l'équation
$$\Phi(x, y) - y = 0$$
en un point $(x_0, \theta(x_0))$. On peut résoudre localement l'équation (car $\mathrm{Id}_F - D_2\Phi(x_0, \theta(x_0))$ est inversible) et obtenir une fonction $\widetilde{\theta}$ de classe C^p définie sur un voisinage de x_0. Mais cette fonction n'est rien d'autre que la restriction de θ à ce voisinage. □

LEMME 4.5.4 : *L'application $f = A + \psi$ est un difféomorphisme de classe C^1.*

Démonstration. Remarquons que
$$f(x) = y \Leftrightarrow A(x) + \psi(x) = y \Leftrightarrow x = A^{-1}(y) - A^{-1} \circ \psi(x).$$
La fonction
$$\Phi \ : (x, y) \mapsto A^{-1}(y) - A^{-1} \circ \psi(x)$$
est de classe C^1 et chaque application partielle $x \mapsto \Phi(x, y)$ est lipschitzienne de rapport $\|A^{-1}\|\varepsilon$ puisque
$$\sup_{x,y} \|D_2\Phi(x, y)\| = \|A^{-1}\|\varepsilon < 1.$$
Utilisant le lemme 4.5.3, on obtient une fonction $\theta \ : E \to E$ de classe C^1 telle que
$$f(x) = y \Leftrightarrow x = \theta(y).$$
Ceci signifie que f est un difféomorphisme de classe C^1. □

La proposition 4.4.1 nous dit qu'il existe un unique homéomorphisme $h \ : \mathbf{R}^r \to \mathbf{R}^r$ tel que $h \circ f = A \circ h$ et tel que $h - \mathrm{Id}_{\mathbf{R}^r}$ soit borné, la proposition 4.5.1 qu'il existe une unique application continue $h \ : \mathbf{R}^r \to \mathbf{R}^r$ telle que $h \circ f = A \circ h$ et telle que soit $h - \mathrm{Id}_{\mathbf{R}^r}$ soit $\mathbf{Z}^r$-périodique. On en déduit que $h = h'$ relève un homéomorphisme H de $\mathbf{T}^2$ homotope à l'identité. □

4.6 Le théorème d'Hartman-Grobman.

Le théorème d'Hartman-Grobman est un résultat de conjugaison locale :

THÉORÈME 4.6.1 : *Soit $f : U \to \mathbf{R}^r$ une application de classe C^1 définie sur une partie ouverte U de $\mathbf{R}^r$ ayant un point fixe x_0 tel que la différentielle $T = Df(x_0)$ est un automorphisme hyperbolique. Alors il existe un voisinage $V \subset U$ de x_0, un voisinage W de 0 et un homéomorphisme $h \ : V \to W$ tel que tout pour tout $x \in V$, le point $f(x)$ appartient à V si et seulement si $T(h(x))$ appartient à W et dans ce cas, on a $f(x) = h^{-1}(T(h(x)))$.*

Démonstration. Quitte à conjuguer f par une translation, on peut toujours supposer que $x_0 = 0$. Le théorème est une conséquence du théorème 4.4.1 et du lemme d'extension suivant, si on choisit ε assez petit.

Lemme 4.6.2 : *Pour tout $\varepsilon > 0$, on peut construire une fonction $\widetilde{f} : \mathbf{R}^r \to \mathbf{R}^r$ telle que :*

i) *$\widetilde{f}$ et f coïncident sur un voisinage de 0;*

ii) *$\widetilde{f} - T$ est bornée et lipschitzienne de rapport ε.*

Proof. Considérons la fonction $\eta : \mathbf{R} \to \mathbf{R}$ où

$$\eta(x) = \begin{cases} 0 & \text{si } x \leq 1, \\ 1 & \text{si } x \geq 2, \\ x - 1 & \text{si } x \in [1, 2] \end{cases}$$

en remarquant qu'elle est lipschitzienne de rapport 1. Fixons α_0 tel que $B(0, 2\alpha_0) \subset U$. Pour tout $\varepsilon > 0$, on peut trouver $\alpha < \alpha_0$ tel que pour tout $x \in B(0, 2\alpha)$ on a

$$\|Df(x) - T\| \leq \frac{\varepsilon}{4},$$

ce qui implique que pour tous x et y dans $B(0, 2\alpha)$ on a

$$\|f(x) - T(x) - f(y) + T(y)\| \leq \frac{\varepsilon}{4}\|x - y\|.$$

En posant $y = 0$, on obtient

$$\|f(x) - T(x)\| \leq \frac{\varepsilon}{4}\|x\|.$$

Definissons

$$\widetilde{f}(x) = \left(1 - \eta\left(\frac{\|x\|}{\alpha}\right)\right) f(x) + \eta\left(\frac{\|x\|}{\alpha}\right) T(x).$$

Cette formule a un sens car $1 - \eta\left(\frac{\|x\|}{\alpha}\right)$ s'annule si $\|x\| \geq 2\alpha$. Posons

$$\varphi(x) = \widetilde{f}(x) - T(x) = \left(1 - \eta\left(\frac{\|x\|}{\alpha}\right)\right)(f(x) - T(x)).$$

L'application φ est bornée et s'annule hors de $B(0, 2\alpha)$. Prouvons qu'elle est lipschitzienne de rapport ε. Fixons x et y dans $\mathbf{R}^r$. Si $\|x\| \geq 2\alpha$ et $\|y\| \geq 2\alpha$, on a

$$\|\varphi(x) - \varphi(y)\| = \|0 - 0\| = 0.$$

Si $\|x\| \leq 2\alpha$ et $\|y\| \geq 2\alpha$, on a

$$\varphi(x) - \varphi(y) = \varphi(x) = \left(\eta\left(\frac{\|y\|}{\alpha}\right) - \eta\left(\frac{\|x\|}{\alpha}\right)\right)(f(x) - T(x))$$

et donc

$$\|\varphi(x) - \varphi(y)\| = \left|\eta\left(\frac{\|y\|}{\alpha}\right) - \eta\left(\frac{\|x\|}{\alpha}\right)\right| \|f(x) - T(x)\| \leq \frac{\|x - y\|}{\alpha}\frac{2\varepsilon\alpha}{4} \leq \varepsilon\|x - y\|.$$

Si $\|x\| \leq 2\alpha$ et $\|y\| \leq 2\alpha$, on a

$$\varphi(x) - \varphi(y) = \left(1 - \eta\left(\frac{\|x\|}{\alpha}\right)\right)(f(x) - f(y) - T(x) + T(y)) + \left(\eta\left(\frac{\|y\|}{\alpha}\right) - \eta\left(\frac{\|x\|}{\alpha}\right)\right)(f(y) - T(y))$$

et donc

$$\|\varphi(x) - \varphi(y)\| \leq \frac{\varepsilon}{4}\|x - y\| + \frac{\|x - y\|}{\alpha}\frac{2\varepsilon\alpha}{4} \leq \varepsilon\|x - y\|.$$

□

Remarque. On peut légèrement modifier le raisonnement en choisissant une fonction lisse η vérifiant

$$\begin{cases} \eta(x) = 0 & \text{si } x \leq 1, \\ \eta(x) = 1 & \text{si } x \geq 2, \\ 0 \leq \eta'(x) \leq 2 & \text{si } x \in [1, 2] \end{cases}$$

et une norme euclidienne $\| \ \|$. On en déduit que si f est de classe C^p, il en est de même de φ.

4.7 Le théorème de la variété stable.

On va énoncer un autre théorème local, qui est complémentaire au théorème d'Hartman-Grobman. Commençons par l'énoncé suivant :

PROPOSITION 4.7.1 : *Soit $T : E \to E$ un endomorphisme λ-hyperbolique d'un espace de Banach E, où $E = E^s \oplus E^u$ est la décomposition hyperbolique et $\| \ \|$ une norme adaptée. Soit $\varphi : E \to E$ une fonction bornée et lipschitzienne de rapport $\varepsilon < \varepsilon_0 = 1 - \lambda$ vérifiant $\varphi(0) = 0$. Si on note $f = T + \varphi$, alors l'ensemble $W^s(0)$ des points $x \in E$ tels que la suite $(f^n(x))_{n\geq 0}$ est bornée est la graphe d'une fonction $\psi : E^s \to E^u$ qui est lipschitzienne de rapport $(\lambda + \varepsilon)$ (on identifie ici E et $E^s \times E^u$). De plus, f est lipschitzienne de rapport $\lambda + \varepsilon$ sur ce graphe et on a donc*

$$W^s(0) = \{x \in E \mid \lim_{n \to +\infty} f^n(x) = 0.\}$$

Démonstration. Fixons $r \in]\lambda + \varepsilon, 1]$ puis définissons l'ensemble $\mathcal{E}$ des suites $x = (x_n)_{n\geq 0}$ dans E telles que la suite $(r^{-n}x_n)_{n\geq 0}$ est bornée. C'est un espace de Banach quand on le munit de la norme

$$\|x\| = \sup_{n\geq 0} r^{-n}\|x_n\|.$$

On a une décomposition $\mathcal{E} = E^s \oplus \mathcal{E}^s \oplus \mathcal{E}^u$, où $\mathcal{E}^s$ est l'espace des suites $x = (x_n)_{n\geq 1}$ dans E^s telles que la suite $(r^{-n}x_n)_{n\geq 1}$ est bornée et $\mathcal{E}^u$ est l'espace des suites $x = (x_n)_{n\geq 0}$ dans E^u telles que la suite $(r^{-n}x_n)_{n\geq 0}$ est bornée.

A chaque suite $(x_n)_{n\geq 0} = (x_n^s + x_n^u)_{n\geq 0}$ on peut associer deux suites $(y_n^s)_{n\geq 1}$ et $(y_n^u)_{n\geq 0}$ en posant

$$\begin{aligned} y_{n+1}^s &= f^s(x_n^s, x_n^u) = T^s(x_n^s) + \varphi^s(x_n^s, x_n^u) \\ y_n^u &= x_n^u + (T^u)^{-1}(x_{n+1}^u - f^u(x_n^s, x_n^u)) = (T^u)^{-1}(x_{n+1}^u - \varphi^u(x_n^s, x_n^u)) \end{aligned}$$

Remarquons que si $(\widetilde{y}_n^s)_{n\geq 1}$ et $(\widetilde{y}_n^u)_{n\geq 0}$ sont associées à une autre suite $(\widetilde{x}_n)_{n\geq 0} = (\widetilde{x}_n^s + \widetilde{x}_n^u)_{n\geq 0}$, alors

$$\begin{aligned} \|y_{n+1}^s - \widetilde{y}_{n+1}^s\| &\leq (\lambda + \varepsilon)\|x_n - \widetilde{x}_n\| \\ \|y_n^u - \widetilde{y}_n^u\| &\leq \lambda\|x_{n+1} - \widetilde{x}_{n+1}\| + \lambda\varepsilon\|x_n - \widetilde{x}_n\|. \end{aligned}$$

Comme les deux suites associées à la suite nulle sont les suites nulles, on en déduit que $(y_n^s)_{n\geq 1} \in \mathcal{E}^s$ et $(y_n^u)_{n\geq 0} \in \mathcal{E}^u$. On a donc défini une application

$$\Phi : \mathcal{E} \sim E^s \times \mathcal{E}^s \times \mathcal{E}^u \to \mathcal{E}^s \times \mathcal{E}^u.$$

Cette application est lipschitzienne de rapport $\dfrac{\lambda + \varepsilon}{r}$. En effet,

$$r^{-n-1}\|y_{n+1}^s - \widetilde{y}_{n+1}^s\| \leq r^{-n-1}(\lambda + \varepsilon)\|x_n - \widetilde{x}_n\| \leq \frac{\lambda + \varepsilon}{r}\|x - \widetilde{x}\|$$

et

$$r^{-n}\|y_n^u - \widetilde{y}_n^u\| \leq \lambda\varepsilon r^{-n}\|x_n - \widetilde{x}_n\| + \lambda r^{-n}\|x_{n+1} - \widetilde{x}_{n+1}\| \leq (\lambda\varepsilon + \lambda r)\|x - \widetilde{x}\| \leq \frac{\lambda+\varepsilon}{r}\|x - \widetilde{x}\|.$$

Puisque $\dfrac{\lambda+\varepsilon}{r} < 1$, on peut appliquer la version à paramètre du théorème de point fixe (théorème 4.3.1) : pour tout $x_0^s \in E^s$, il existe $\theta(x_0^s) \in \mathcal{E}^s \times \mathcal{E}^u$, uniquement défini et dépendant continûment de x_0^s, tel que $\theta(x_0^s) = \Phi(x_0^s, \theta(x_0^s))$. En écrivant

$$\theta(x_0^s) = ((\theta_n^s(x_0^s))_{n\geq 1}, (\theta_n^u(x_0^s))_{n\geq 0}),$$

on obtient

$$x_{n+1}^s = f^s(x_n^s, x_n^u), \;\; x_{n+1}^u = f^u(x_n^s, x_n^u).$$

En d'autres termes, le théorème nous dit que pour tout x_0^s, il existe un unique point $x_0^u = \theta_0^u(x_0^s) \in E^u$ tel que la suite $(f^n(x_0^s, x_0^u))_{n\geq 0}$ appartient à $\mathcal{E}$ et de plus la fonction $\psi = \theta_0^u$ est continue. Si on prend $r = 1$, on en déduit que l'ensemble des points x dont l'orbite positive est bornée est le graphe de ψ. De plus la suite $(r^{-n}f^n(x))_{n\geq 0}$ est bornée pour tout $r > \lambda+\varepsilon$, ce qui implique que $\lim_{n\to+\infty} f^n(x) = 0$ si la suite $(f^n(x))_{n\geq 0}$ est bornée. Prouvons maintenant que θ est lipschitzienne de rapport $(\lambda+\varepsilon)$, ce qui impliquera une propriété analogue pour ψ. Si on pose $r = 1$, on obtient

$$\begin{aligned}\|\theta(x_0^s) - \theta(\widetilde{x}_0^s)\| &= \|\Phi(x_0^s, \theta(x_0^s)) - \Phi(\widetilde{x}_0^s, \theta(\widetilde{x}_0^s))\| \\ &\leq (\lambda+\varepsilon)\max(\|x_0^s - \widetilde{x}_0^s\|, \|\theta(x_0^s) - \theta(\widetilde{x}_0^s)\|\end{aligned}$$

ce qui implique que

$$\|\theta(x_0^s) - \theta(\widetilde{x}_0^s)\| \leq (\lambda+\varepsilon)\|x_0^s - \widetilde{x}_0^s\|.$$

Il reste à prouver que la restriction de f au graphe de ψ est lipschitzienne de rapport $(\lambda+\varepsilon)$. Pour cela, notons que

$$\begin{aligned}\|f(x_0^s, \psi(x_0^s)) - f(\widetilde{x}_0^s, \psi(\widetilde{x}_0^s))\| &= \|f^s(x_0^s, \psi(x_0^s)) - f^s(\widetilde{x}_0^s, \psi(\widetilde{x}_0^s))\| \\ &\leq \lambda\|x_0^s - \widetilde{x}_0^s\| + \varepsilon\|(x_0^s, \psi(x_0^s)) - (\widetilde{x}_0^s, \psi(\widetilde{x}_0^s))\| \\ &\leq (\lambda+\varepsilon)\|(x_0^s, \psi(x_0^s)) - (\widetilde{x}_0^s, \psi(\widetilde{x}_0^s))\|.\end{aligned}$$

□

On peut améliorer le résultat

PROPOSITION 4.7.2 : *En gardant les hypothèses de la proposition 4.7.1 et en supposant de plus que φ est de classe C^p et que $D\varphi(0) = 0$, on peut montrer que ψ également de classe C^p et vérifie $D\psi(0) = 0$.*

Démonstration. Il suffit d'appliquer la version différentiable du théorème de point fixe avec paramètre (Lemme 4.5.3). Il faut donc prouver que la fonction Φ introduite dans la preuve précédente est de classe C^p. Elle vérifiera nécessairement $\|D\Phi(x)\| \leq \dfrac{\lambda+\varepsilon}{r}$ puisqu'elle est lipschitzienne de rapport $\frac{\lambda+\varepsilon}{r}$. Il n'est pas difficile de voir qu'il suffit de prouver que

$$\begin{aligned}\mathcal{F} \;&: \mathcal{E} \to \mathcal{E} \\ (x_n)_{n\geq 0} &\mapsto (f(x_n))_{n\geq 0}\end{aligned}$$

est de classe C^p. On doit nécessairement supposer que $r < 1$ (sinon, c'est faux). On va se limiter à prouver qu'elle est de classe C^1 et que la différentielle en $x = (x_n)_{n\geq 0}$ est la fonction

$$\begin{aligned} \mathcal{L} \ &: \mathcal{E} \to \mathcal{E} \\ (h_n)_{n\geq 0} &\mapsto (Df(x_n).h_n)_{n\geq 0} \end{aligned}.$$

Notons que cette application est bien définie car $\lim_{n\to+\infty} x_n = 0$, ce qui implique que la suite $(Df(x_n))_{n\geq 0}$ est bornée. Écrivons

$$f(x_n + h_n) - f(x_n) - Df(x_n).h_n = \int_0^1 \left(Df(x_n + th_n) - Df(x_n)\right).h_n \, dt.$$

Supposons que $\|h\| \leq 1$ et fixons $\varepsilon > 0$. Il existe un entier $N \geq 0$ tel que pour tout $n \geq N$ et tout $t \in [0,1]$, on a $\|Df(x_n + th_n) - Df(x_n)\| \leq \varepsilon$. De plus, il existe $\eta \in (0,1)$ tel que si $\|h\| \leq \eta$, alors pour tout $n < N$ et tout $t \in [0,1]$, on a $\|Df(x_n + th_n) - Df(x_n)\| \leq \varepsilon$. Ainsi, si $\|h\| \leq \eta$, alors

$$\|f(x+h) - f(x) - \mathcal{L}.h\| \leq \varepsilon \|h\|,$$

ce qui signifie que f est différentiable en x et que $Df(x) = \mathcal{L}$.

On peut calculer $D\psi(0)$ directement. On peut utiliser également l'argument suivant. Puisque

$$\psi(f^s(x^s, \psi(x^s)) = f^u(x^s, \psi(x^s))$$

et

$$\begin{aligned} \psi(x^s) &= D\psi(0).\, x^s + o(x^s) \\ f^s(x^s, \psi(x^s)) &= T^s.\, x^s + o(x^s) \\ f^u(x^s, \psi(x^s)) &= T^u.\, \psi(x^s) + o(x^s) = T^u \circ D\psi(0).\, x^s + o(x^s) \end{aligned},$$

on sait que

$$D\psi(0) \circ T^s.\, x^s = T^u \circ D\psi(0).\, x^s + o(x^s),$$

et donc que

$$D\psi(0) \circ T^s = T^u \circ D\psi(0),$$

ce qui implique

$$D\psi(0) = (T^u)^{-n} \circ D\psi(0) \circ (T^s)^n,$$

for every $n \geq 0$. En faisant tendre n vers $+\infty$, on obtient $D\psi(0) = 0$. □

Énonçons maintenant le théorème de la variété stable locale :

THÉORÈME 4.7.3 : *Soit $f : U \to \mathbf{R}^r$ une application de classe C^p, $p \geq 1$, définie au voisinage d'un point fixe x_0. On suppose que $Df(x_0)$ est λ-hyperbolique, $\lambda \in (0,1)$, on note $\mathbf{R}^r = E^s \oplus E^u$ la décomposition hyperbolique et on considère une norme adaptée $\| \ \|$. Fixons $\lambda' \in]\lambda, 1[$. Alors, il existe $\delta > 0$ tel que l'ensemble $W^s_\delta(x_0)$ des points $x \in U$ vérifiant $\|f^n(x) - x_0\| \leq \delta$ pour tout $n \geq 0$ est le graphe d'une fonction de classe C^p*

$$\psi : x_0 + \left(E^s \cap \overline{B(0,\delta)}\right) \to E^u \cap \overline{B(0,\delta)}$$

vérifiant

$$\psi(x_0) = 0, \ \ D\psi(x_0) = 0.$$

De plus, pour tout $x_0 \in W^s_\delta(x_0)$ et tout $n \geq 0$, on a

$$\|f^n(x) - x_0\| \leq \lambda'^n \|x - x_0\|.$$

Démonstration. On peut toujours supposer que $x_0 = 0$. D'après la remarque qui suit le lemme 4.6.2, on sait qu'il existe une application $\widetilde{f} : \mathbf{R}^n \to \mathbf{R}^n$ de classe C^p qui coïncide avec f au voisinage de x_0, tel que $\widetilde{f} - Df(0)$ est lipschtizienne de rapport $(\lambda' - \lambda)$. Soit $\widetilde{\psi} : E^s \to E^u$ l'application donnée par la proposition 4.7.1. Si $\delta > 0$ est assez petit, f et $\widetilde{f}$ coïncident sur la boule $\overline{B(0,\delta)}$ de E. Définissons maintenant ψ comme la restriction de $\widetilde{\psi}$ à la boule $\overline{B(0,\delta)}$ de E^s. □

Remarques. L'ensemble $W^s_\delta(x_0)$ est appelée la *variété stable locale* de x_0. Dans le cas où les valeurs propres de $Df(x_0)$ ont toutes un module < 1, alors $W^s_\delta(x_0) = \overline{B(x_0,\delta)}$. Dans ce cas, on dit que x_0 est un *puits* ou un *point fixe attractif* : tout point d'un voisinage de x_0 est attiré par ce point. Dans le cas où $Df(x_0)$ est un automorphisme hyperbolique, alors f est un difféomorphisme local et on peut définir de façon similaire la *variété instable* $W^u_\delta(x_0)$ de x_0. Les ensembles $W^s_\delta(x_0)$ et $W^u_\delta(x_0)$ sont des sous-variétés de dimensions complémentaires (respectivement de dimensions $\dim(E^s)$ et $\dim(E^u)$). Elles sont tangentes à E^s et à E^u en x_0 et ne s'intersectent qu'en ce point. Dans le cas où les valeurs propres de $Df(x_0)$ ont toutes un module > 1, alors $W^u_\delta(x_0) = \overline{B(x_0,\delta)}$. Dans ce cas, on dit que x_0 est une *source* ou un *point fixe répulsif*.

Nous allons conclure cette section avec le théorème de la variété stable globale :

Théorème 4.7.4 : *Soit $f : M \to M$ un difféomorphisme de classe C^p, $p \geq 1$, sur une variété M qui fixe x_0. Supposons que $T_{x_0}f$ est hyperbolique et que $T_{x_0}M = E^s \oplus E^u$ est la décomposition hyperbolique. L'ensemble*

$$W^s(x_0) = \{x \in M \mid \lim_{n \to +\infty} f^n(x) = x_0\}$$

est l'image d'une immersion injective $\theta^s : E^s \to \mathbf{R}^r$ de classe C^p telle que $\theta^s(0) = x_0$ et telle que $T_0\theta^s$ est l'inclusion $i^s : E^s \to T_{x_0}M$. On l'appelle la variété stable *de x_0. L'ensemble*

$$W^u(x_0) = \{x \in M \mid \lim_{n \to +\infty} f^{-n}(x) = x_0\}$$

est l'image d'une immersion injective $\theta^u : E^u \to \mathbf{R}^r$ de classe C^p telle que $\theta^u(0) = x_0$ et telle que $T_0\theta^u$ est l'inclusion $i^u : E^u \to T_{x_0}M$. On l'appelle la variété instable *de x_0..*

Démonstration. On suppose que $T_{x_0}f$ est λ-hyperbolique et on considère une norme $\| \ \|$ adaptée. On munit M d'une structure riemmannienne et on note $\Phi : T_{x_0}M \to M$ l'application exponentielle. C'est un difféomorphisme entre un voisinage de 0 dans $T_{x_0}M$ et un voisinage de x_0 dans M tel que $T_0\Phi = \mathrm{Id}_{T_{x_0}M}$. Ainsi, il existe un difféomorphisme $\widetilde{f} : T_{x_0}M \to T_{x_0}M$ de classe C^p qui coïncide avec $\Phi^{-1} \circ f \circ \Phi$ sur une boule $\overline{B(0,\delta)}$, et telle que $\widetilde{f} - T_{x_0}f$ est lipschtizienne de rapport $\lambda' - \lambda$. Soit $\widetilde{\psi} : E^s \to E^u$ l'application introduite précédemment. Pour tout $x \in E^s$, il existe $n \geq 0$ tel que $\|\widetilde{f}^n(x,\psi(x))\| \leq \delta$. Posons $\theta(x) = f^{-n}(\Phi(\widetilde{f}^n(x,\psi(x))))$ et remarquons que $\theta(x)$ ne dépend pas de l'entier n. Remarquons également que θ est de classe C^p, est une immersion injective et que son image est exactement $W^s(x_0)$. On fait la même chose pour la variété instable.

$\square$

Remarques.

i) Dans le cas où x_0 est un puits (resp. une source), l'ensemble $W^s(x_0)$ (resp. $W^u(x_0)$) est une partie ouverte connexe appelée *bassin d'attraction* (resp. *basin de répulsion*) de x_0.

ii) Les variétés $W^s(x_0)$ et $W^u(x_0)$ peuvent s'intersecter en un point différent de x_0, comme on l'a a vu dans le cas des automorphismes hyperboliques de $\mathbf{T}^r$.

ii) Soit f un difféomorphisme de classe C^1 d'une variété M et x_0 un point périodique de f, de période q. On dira que x_0 est *hyperbolique* si $T_{x_0}f^q \ : T_{x_0}M \to T_{x_0}M$ est hyperbolique. On peut appliquer le théorème de la variété stable à f^q. Les ensembles $W^s(x_0)$ et $W^u(x_0)$ vérifient les propriétés suivantes

$$W^s(x_0) = \{x \in M \ | \ \lim_{n\to+\infty} d(f^n(x), f^n(x_0)) = 0\},$$
$$W^u(x_0) = \{x \in M \ | \ \lim_{n\to+\infty} d(f^{-n}(x), f^{-n}(x_0)) = 0\}.$$

4.8 Ensembles hyperboliques

Soit f un difféomorphisme de classe C^1 sur une variété lisse M munie d'une métrique riemmanienne. On dit qu'une partie compacte invariante X est *hyperbolique* s'il existe pour tout point $x \in X$ une décomposition $E(x) = E^s(x) \oplus E^u(x)$, un réel $\mu < 1$ et une constante $C > 0$ tels que

- pour tout $v \in E^s(x)$ et tout $n \geq 0$, on a $\|Tf^n(x).v\| \leq C\mu^n\|v\|$;
- pour tout $v \in E^u(x)$ et tout $n \geq 0$, on a $\|Tf^{-n}(x).v\| \leq C\mu^n\|v\|$.

Il n'est pas difficile de voir que cette condition est indépendante de la structure choisie et que les fonctions $x \mapsto E^s(x)$ et $x \mapsto E^u(x)$ sont continues. En particulier la dimension des espaces stables et instables est localement constante (mais pas nécessairement constante). L'exemple le plus simple est celui d'une orbite périodique hyperbolique, mais nous avons vu également les automorphismes hyperboliques du tore, où le tore tout entier est hyperbolique. Nous en verrons d'autres plus loin. Il n'est pas difficile de voir qu'il existe une borne inférieure uniforme à l'angle formé par $E^s(x)$ et $E^u(x)$. On peut également montrer que $f|_X$ est expansif. Le résultat qui suit permet de définir les variétés stables et instables d'un point de X.

THÉORÈME 4.8.1 : *Soit $f \ : M \to M$ un difféomorphisme de classe C^1 et X une partie compacte invariante hyperbolique. On suppose que chaque ensemble $E^s(x)$ est de dimension r_s et chaque ensemble $E^u(x)$ de dimension r_u. Choisissons $\delta > 0$ assez petit et définissons, pour tout $x \in X$, les ensembles*

$$W^s_{\mathrm{loc}}(x) = \{y \in M \,|\, n \geq 0 \Rightarrow d(f^n(y), f^n(x)) \leq \delta\}$$

et

$$W^u_{\mathrm{loc}}(x) = \{y \in M \,|\, n \geq 0 \Rightarrow d(f^{-n}(y), f^{-n}(x)) \leq \delta\}.$$

i) *Chaque ensemble $W^s(x)$ (resp. $W^u(x)$) est l'image d'un plongement de classe C^1 de la boule euclidienne unité B_{r^s} de $\mathbf{R}^{r^s}$ (resp. B_{r^u} de $\mathbf{R}^{r^u}$) valant x en 0, qui dépend continûment de x pour la C^1-topologie (i.e. pour la convergence uniforme du plongement et de l'application dérivée).*

ii) *Pour tout $\mu' \in]\mu, 1[$, il existe une constante $C > 0$ tel que*

$$y \in W^s_{\text{loc}}(x) \Rightarrow d(f^n(y), f^n(x)) \leq C\mu'^n \text{ si } n \geq 0$$

et

$$y \in W^u_{\text{loc}}(x) \Rightarrow d(f^{-n}(y), f^{-n}(x)) \leq C\mu'^n \text{ si } n \geq 0.$$

iii) *Pour tout $x \in X$ les ensembles*

$$W^s(x) = \{y \in M \mid \lim_{n\to+\infty} d(f^n(y), f^n(x)) = 0\}$$

et

$$W^u_{\text{loc}}(x) = \{y \in M \mid \lim_{n\to+\infty} d(f^{-n}(y), f^{-n}(x)) = 0\}$$

sont des plongements de classe C^1 de $\mathbf{R}$ et on a

$$W^s(x) = \bigcup_{n\geq 0} f^{-n}(W^s_{\text{loc}}(f^n(x))), \;\; W^u(x) = \bigcup_{n\geq 0} f^n(W^s_{\text{loc}}(f^{-n}(x)))$$

On va étudier dans cette section des systèmes dynamiques qui sont facteurs de décalages de type fini. Ce n'est pas toujours le système dynamique qui apparaît comme facteur, mais souvent la partie dynamique la plus intéressante, comme par exemple la restriction à l'ensemble des points non errants. La situation se présente généralement ainsi. On se donne un système dynamique topologique $T : X \to X$. On écrit $X = \bigcup_{i \in I} R_i$ comme réunion finie de parties fermées dont les intérieurs sont mutuellement disjoints. Une suite $\mathbf{i} = (i_n)_{n \geq 0} \in I^{\mathbf{N}}$ *code* l'orbite d'un point $x \in X$, si $T^n(x) \in R_{i_n}$ pour tout $n \geq 0$. Une "bonne" décomposition est une décomposition pour laquelle, pour tout $\mathbf{i} = (i_n)_{n \geq 0} \in I^{\mathbf{N}}$ l'ensemble $H(\mathbf{i}) = \bigcap_{n \geq 0} T^{-n}(R_{i_n})$ contient au plus un point. Dans ce cas, l'ensemble Z des suites $\mathbf{i}$ pour lesquelles $H(\mathbf{i})$ est non vide est invariant par le décalage σ, et il est fermé si X est compact. Si on écrit $H(\mathbf{i}) = \{h(\mathbf{i})\}$, on obtient une semi-conjugaison $h : Z \to X$ de $\sigma|_Z$ à T. Remarquons que si $T : X \to X$ est expansif et si X est compact, alors toute décomposition formée de parties fermées de diamètre suffisamment petit est une bonne décomposition. Une bonne décomposition est intéressante si la dynamique sur Z peut être décrite précisément. Nous verrons plusieurs cas où Z est l'ensemble

$$X_A = \{\mathbf{i} = (i_k)_{k \in \mathbf{N}} \in \{1, \dots, p\}^{\mathbf{Z}} \mid A_{i_k, i_{k+1}} = 1 \text{ pour tout } k \in \mathbf{Z}\}.$$

définie par une matrice carrée $A = (A_{i,j})_{i,j}$ d'ordre p à coefficients égaux à 0 ou 1 et où T est donc un facteur d'un sous-décalage de type fini. On a dans ce cas une *décomposition* (ou par abus de langage *partition*) *de Markov*. Un exemple-type est la partition de Markov $\mathbf{T}^1 = \bigcup_{0 \leq i < p} \left([\frac{i}{p}, \frac{i+1}{p}] + \mathbf{Z} \right)$ qui est associée $T : x \mapsto px$. Ici, c'est le décalage sur $\{1 \dots, p\}^{\mathbf{N}}$ qui code les orbites de T. Bien sûr on peut définir de même des partitions de Markov pour des systèmes inversibles avec codage dans $I^{\mathbf{Z}}$.

5.1 L'application logistique

La *famille logistique* $(f_\lambda)_{\lambda > 0}$ est définie ainsi :

$$\begin{aligned} f_\lambda \ : \mathbf{R} &\to \mathbf{R} \\ x &\mapsto \lambda x(1-x). \end{aligned}$$

Commençons par énoncer certaines propriétés simples. Remarquons d'abord que tout point a au plus deux antécédents, que 0 est un point fixe de f_λ et que 1 est envoyé sur ce point fixe. Notons aussi que $[0, 1]$ est positivement invariant si et seulement si $0 \leq \lambda \leq 4$. Dans le cas où $\lambda \geq 1$, remarquons que si $x < 0$, alors la suite $(f_\lambda^n(x))_{n \geq 0}$ est décroissante et vérifie $\lim\limits_{n \to +\infty} f_\lambda^n(x) = -\infty$. Cette dernière propriété est encore vraie si $x > 1$ car $f_\lambda(x) < 0$. L'ensemble des points d'orbites bornées s'écrit donc

$$X_\lambda = \bigcap_{n \geq 0} f_\lambda^{-n}([0, 1]).$$

C'est une partie compacte incluse dans $[0, 1]$, égale à cet ensemble si $\lambda \in [1, 4]$ et non connexe si $\lambda > 4$. Tout point $x \notin X_\lambda$ vérifie $\lim\limits_{n \to +\infty} f_\lambda^n(x) = -\infty$.

Le cas le plus intéressant est le cas où $\lambda < 4$, il y a une très grande richesse de dynamiques possibles, c'est une famille étudiée par grand nombre d'auteurs. Dans le cas, où $\lambda = 4$, on peut montrer que $f_\lambda|_{[0,1]}$ est un facteur du décalage de Bernouilli sur $\{0, 1\}^{\mathbf{N}}$ à l'aide de la partition de Markov $\{[0,1/2],[1/2,1]\}$. Le cas $\lambda > 4$ est encore plus simple car on peut montrer, à l'aide de la même partition de Markov, que $f_\lambda|_{X_\lambda}$ est conjuguée au décalage. Nous allons en donner une preuve dans le cas où $l > 2+\sqrt{5}$. Dans cette situation, on aura $|f'_\lambda(x)| > 1$ pour tout $x \in X_\lambda$.

PROPOSITION 5.1.1 : *Supposons que $\lambda > 2+\sqrt{5}$. Pour tout $\mathbf{i} = (i_n)_{n\geq 0}$ il existe un unique point $x = h(\mathbf{i}) \in X_\lambda$ tel que pour tout entier $n \geq 0$ on a :*

$$\begin{cases} 0 \leq f_\lambda^n(x) \leq \frac{1}{2} & si\ i_n = 0, \\ \frac{1}{2} \leq f_\lambda^n(x) \leq 1 & si\ i_n = 1. \end{cases}$$

L'application h est un homéomorphisme de $\{0,1\}^{\mathbf{N}}$ sur $X_\lambda = \bigcap_{n\geq 0} f_\lambda^{-n}([0,1])$ qui induit une conjugaison entre le décalage de Bernouilli σ et $f_\lambda|_{X_\lambda}$.

Démonstration.. Les solutions de l'équation $f(x) = 1$ sont

$$\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{\lambda}}.$$

On en déduit que $f^{-1}([0,1]) = \Delta_0 \cup \Delta_1$ où

$$\Delta_0 = \left[0, \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\right], \ \ \Delta_1 = \left[\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\lambda}}, 1\right].$$

Remarquons maintenant que le minimum m de $|f'|$ sur $\Delta_0 \cup \Delta_1$ est donné par

$$m = f'\left(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\right) = -f'\left(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\right).$$

Ainsi,

$$m = 2\lambda\sqrt{\frac{1}{4} - \frac{1}{\lambda}} = \sqrt{\lambda^2 - 4\lambda} > \sqrt{(2+\sqrt{5})^2 - 4(2+\sqrt{5})} = 1.$$

Fixons $\mathbf{i} = (i_n)_{n\geq 0}$. On en déduit que, pour tout $N \geq 0$, l'ensemble $\bigcap_{0\leq n\leq N} f^{-n}(\Delta_{i_n})$ est un intervalle de longueur $\leq m^{-N}$. Ceci implique que $\bigcap_{n\geq 0} f^{-n}(\Delta_{i_n})$ est réduit à un point $h(\mathbf{i})$. L'application h est évidement injective et induit une bijection entre $\{0,1\}^N$ et X_λ (qui est un ensemble de Cantor). Elle induit une conjugaison entre σ et $f_\lambda|_{X_\lambda}$. □

Remarque. On a une conclusion analogue si $\lambda > 4$ mais la preuve demande un outil plus sophistiqué : la *dérivée schwarzienne.*

5.2 Le fer à cheval de Smale

On va décrire un exemple géométrique dû à S. Smale. Soit $R = [a,b] \times [c,d]$ un rectangle de $\mathbf{R}^2$ et $f : R \to f(R)$ un difféomorphisme tel que :

- $f(R) \cap R$ est la réunion de deux rectangles verticaux $R_0 = [a_0, b_0] \times [c,d]$, $R_1 = [a_1, b_1] \times [c,d]$ où $a < a_0 < b_0 < a_1 < b_1 < b$;
- $f^{-1}(R) \cap R$ est la réunion de deux rectangles horizontaux $f^{-1}(R_0) = [a,b] \times [c_0, d_0]$, $f^{-1}(R_1) = [a,b] \times [c_1, d_1]$ où $c < c_0 < d_0 < c_1 < d_1 < d$;
- il existe $\lambda < 1 < \mu$ tels que $f|_{f^{-1}(R_0)}$ est une application affine dont la partie linéaire est $(x,y) \mapsto (\lambda x, \mu y)$ et $f|_{f^{-1}(R_1)}$ une application affine dont la partie linéaire est $(x,y) \mapsto (-\lambda x, -\mu y)$.

Si on veut construire un difféomorphisme global, on écrit D_0 pour le demi-disque en dessous de son diamètre $[a, b] \times \{c\}$ et D_1 pour le demi-disque au dessus de son diamètre $[a, b] \times \{d\}$. On note D le disque topologique $D = D_0 \cup D_1 \cup R$ et D' l'adhérence du complémentaire de D dans la sphère de Riemann $S = \mathbf{R}^2 \sqcup \{\infty\}$. On suppose que f est un difféomorphisme défini sur S qui vérifie les propriétés supplémentaires suivantes :

- $f(D_0) \subset \mathrm{Int}(D_0)$ et $\bigcap_{n\geq 0} f^n(D_0) = \{z_0\}$, où z_0 est un point fixe de type puits ;
- $f(D) \subset \mathrm{Int}(D)$ and $\bigcap_{n\geq 0} f^{-n}(D') = \{\infty\}$ où ∞ est un point fixe de type source.

Sous ses hypothèses remarquons que $f(D_1) \subset D_0$. Ainsi les points $z \in D$ qui ne sont pas attirés par z_0 sont les points $z \in D$ dont l'orbite positive reste dans R, c'est-à-dire l'ensemble $\bigcap_{k\geq 0} f^{-k}(R)$. De même, l'ensemble des points $z \in R$ qui n'appartiennent pas au bassin répulsif de ∞ est $\bigcap_{k\leq 0} f^{-k}(R)$.

Remarquons que pour tout $n \geq 1$, l'ensemble $\bigcap_{0\leq k\leq n} f^{-k}(R)$ est la réunion de 2^n rectangles horizontaux de largeur $\mu^{-n}(d - c)$ et, par conséquent que $\bigcap_{k\geq 0} f^{-k}(R)$ est un ensemble de Cantor de segments horizontaux. De même, pour tout $n \geq 1$, l'ensemble $\bigcap_{0\leq k\leq n} f^k(R)$ est la réunion de 2^n rectangles verticaux de largeur $\lambda^n(b - a)$ et $\bigcap_{k\geq 0} f^k(R)$ un ensemble de Cantor de segments verticaux. En conclusion, l'ensemble invariant maximal contenu dans R, à savoir $X = \bigcap_{k\in\mathbf{Z}} f^{-k}(R)$, est un ensemble de Cantor. Remarquons que $f|_X$ est conjugué au décalage bilatéral σ sur $\{0, 1\}^{\mathbf{Z}}$ par l'homéomorphisme

$$h : (i_k)_{k\in\mathbf{Z}} \mapsto \bigcap_{k\in\mathbf{Z}} f^{-k}(R_{i_k}).$$

La décomposition $\{R_0 \cap X, R_1 \cap X\}$ est une partition de Markov de $f|_X$.

On en déduit alors immédiatement les résultats suivants :

- L'application $f|_X$ est topologiquement transitive.
- Les points périodiques sont denses dans X ainsi que dans $\Omega(f) = X \cup \{z_0\} \cup \{\infty\}$.
- Pour tout point $z = (x, y) \in X$ le segment $W^s_{\mathrm{loc}}(z) = [a, b] \times \{y\}$ est formé de points z' vérifiant $\lim_{k\to+\infty} d(f^k(z'), f^k(z)) = 0$ et le segment $W^u_{\mathrm{loc}}(z) = \{x\} \times [c, d]$ formé de points z' vérifiant $\lim_{k\to-\infty} d(f^k(z'), f^k(z)) = 0$.
- Il y a quatre points fixes de f ; un puits en z_0, une source en ∞, deux points de type selle dans X.
- Il y a $2 + 2^n$ points fixes de f^n. Excepté z_0 et ∞, tous les autres points sont de type selle et on a $W^s(z) = \bigcup_{k\geq 0} f^{-k}(W^s_{\mathrm{loc}}(z))$ et $W^u(z) = \bigcup_{k\geq 0} f^k(W^u_{\mathrm{loc}}(z))$.
- Pour tout point périodique de type selle z, les variétés $W^s(z)$ et $W^u(z)$ ont une intersection transverse en une infinité de points z. Si z' est un autre point périodique de type selle, il en est de même de $W^s(z)$ et $W^u(z')$.
- L'entropie topologique de f vaut $\ln 2$, c'est l'entropie $h_\mu(f)$ d'une unique mesure de probabilité, celle-ci est la limite, pour la topologie faible*, des mesures équidistribuées sur les orbites périodiques.
- L'ensemble X est hyperbolique de même que $\Omega(f)$ (puisqu'on ajoute à X un puits et une source).

5.3 L'application de Hénon

Nous allons voir maintenant un exemple explicite où apparaît un fer à cheval de Smale. La famille de Hénon $(f_{b,c})_{b\in]0,1],c\in\mathbf{R}}$ est une famille de difféomorphismes du plan définie ainsi :

$$\begin{aligned} f_{b,c} \ : \mathbf{R}^2 &\to \mathbf{R}^2 \\ (x,y) &\mapsto (x^2+c-by, x). \end{aligned}$$

Remarquons que la réciproque de $f_{b,c}$ est

$$\begin{aligned} f_{b,c}^{-1} \ : \mathbf{R}^2 &\to \mathbf{R}^2 \\ (x,y) &\mapsto (y, \frac{y^2-x+c}{b}). \end{aligned}$$

Cette famille est également extrêmement riche du point de vue dynamique. Là-encore nous allons nous intéresser à une situation simple.

Remarquons d'abord que le jacobien de $f = f_{b,c}$ est constant égal à b. Ceci implique que f préserve l'orientation. Dans le cas où $b = 1$, le difféomorphisme f préserve la mesure de Lebesgue ; dans le cas où $b < 1$, au contraire f "diminue les aires". Remarquons que $(x_k, y_k)_{k\in\mathbf{Z}}$ est une orbite de f si et seulement si

$$x_k^2 + c = x_{k+1} + bx_{k-1}$$

et

$$y_k = x_{k-1}.$$

En particulier, les orbites bornées de f correspondent aux orbites bornées $(x_k)_{k\in\mathbf{Z}}$ définies par la première relation de récurrence.

Proposition 5.3.1 : *Les conditions suivantes sont équivalentes :*

- $c > \frac{1}{4}(1+b)^2$ *;*
- f *n'a pas de point fixe ;*
- *pour tout point* $z \in \mathbf{R}^2$*, on a* $\lim_{k\to\pm\infty} \|f^k(z)\| = +\infty$.

Démonstration.. L'existence d'un point fixe est équivalent à l'existence d'une solution réelle à l'équation $(*)$:

$$x^2 + c = (1+b)x,$$

et donc à l'inégalité $c \leq \frac{1}{4}(1+b)^2$.

Pour montrer la proposition, on doit prouver qu'en l'absence de solution réelle à l'équation $(*)$, alors pour toute suite $(x_k)_{k\in\mathbf{Z}}$ telle que

$$x_k^2 + c = x_{k+1} + bx_{k-1},$$

on a $\lim_{k\to\pm+\infty} |x_k| = +\infty$. Remarquons que

$$(1+b)|x_k| < x_k^2 + c = x_{k+1} + bx_{k-1},$$

et donc que

$$x_k \leq |x_k| < \max(x_{k-1}, x_{k+1}).$$

Si la suite $(x_k)_{k\in\mathbf{Z}}$ est décroissante (resp. croissante) à partir d'un certain rang, on a nécessairement $\lim_{k\to+\infty} x_k = -\infty$ (resp. $\lim_{k\to+\infty} x_k = +\infty$) puisqu'il n'y a pas de solution réelle à l'équation $(*)$. Si la suite $(x_k)_{k\in\mathbf{Z}}$ n'est pas décroissante, il existe k_0 tel que $x_{k_0-1} \leq x_{k_0}$. On en déduit alors que la suite $(x_k)_{k\geq k_0}$ est croissante. On montre de la même façon que $\lim_{k\to-\infty} |x_k| = +\infty$. □

Remarque. Un résultat de Brouwer affirme qu'un homéomorphisme du plan f qui préserve l'orientation et qui n'a pas de point fixe, n'a que des points errants. On en déduit en particulier que pour tout $z \in \mathbf{R}^2$, on a $\lim_{k\to\pm\infty} \|f^k(z)\| = +\infty$.

Proposition 5.3.2 : *Supposons que $c \leq \frac{1}{4}(1+b)^2$ et posons*

$$M = \frac{1}{2}\left(1 + b + \sqrt{(1+b)^2 - 4c}\right).$$

Fixons $z \in \mathbf{R}^2$. Trois cas sont possibles :
- *$f^k(z) \in [-M, M]^2$ pour tout $k \in \mathbf{Z}$;*
- *$\lim_{k\to+\infty} \|f^k(z)\| = +\infty$;*
- *$\lim_{k\to-\infty} \|f^k(z)\| = +\infty$.*

Démonstration. Remarquons que si $|x_{k_0}| > M$, alors on a

$$x_{k_0} \leq |x_{k_0}| < \max(x_{k_0-1}, x_{k_0+1}) \leq \max(|x_{k_0-1}|, |x_{k_0+1}|),$$

car

$$(1+b)|x_{k_0}| < x_{k_0}^2 + c = x_{k_0+1} + bx_{k_0-1}.$$

Comme précédemment, on en déduit que la suite $(|x_k|)_{k\geq k_0}$ est strictement croissante et vérifie $\lim_{k\to+\infty} |x_k| = +\infty$ ou alors que la suite $(|x_k|)_{k\leq k_0}$ est strictement décroissante et vérifie $\lim_{k\to-\infty} |x_k| = +\infty$. □

En particulier, si $c \leq 0$, l'ensemble des points d'orbites bornées est la partie compacte invariante

$$X = \bigcap_{k\in\mathbf{Z}} f^{-k}([-M, M])^2.$$

Cette partie est non vide puisqu'elle contient un point fixe, et tout point $z \notin X$ vérifie

$$\lim_{k\to+\infty} \|f^k(z)\| = +\infty \text{ ou } \lim_{k\to-\infty} \|f^k(z)\| = +\infty.$$

Nous allons voir que si c est suffisamment petit, la dynamique sur X se décrit par une partition de Markov

Proposition 5.3.3 : *Supposons que $c \leq -10$. Alors $f|_X$ est conjuguée au décalage de Bernouilli σ sur $\{-1, +1\}^{\mathbf{Z}}$ par la partition de Markov $\{X \cap [-M, M] \times [-M, 0], X \cap [-M, M] \times [0, M]\}$. De plus, X est hyperbolique et tous les points périodiques sont de type selle.*

Démonstration. On considère l'espace métrique complet $(\mathcal{B}, d)$ des suites $\mathbf{x} = (x_k)_{k\in\mathbf{Z}}$ dans $[-M, M]$ muni de la distance

$$d(\mathbf{x}, \mathbf{x}') = \sup_{k\in\mathbf{Z}} |x_k - x'_k|.$$

Remarquons que pour toute suite $\mathbf{x} = (x_k)_{k\in\mathbf{Z}}$ dans $\mathcal{B}$, on a

$$x_{k+1} + bx_{k-1} - c \leq (1+b)M - c = M^2$$

et

$$\begin{aligned} x_{k+1} + bx_{k-1} - c &\geq -M(1+b) - c \\ &\geq -2M - c \\ &= -(1+b) - \sqrt{(1+b)^2 - 4c} - c \\ &\geq -2 - \sqrt{4-4c} - c \\ &\geq -2 - \sqrt{44} + 10 = \alpha^2 > 1. \end{aligned}$$

En particulier, pour toute orbite bornée $(x_k, y_k)_{k\in\mathbf{Z}}$ de f, on a $|x_k| \geq \alpha$ et $|y_k| \geq \alpha$ pour tout $k \in \mathbf{Z}$.

Pour tout $\boldsymbol{\varepsilon} = (\varepsilon_k)_{k\in\mathbf{Z}} \in \{-1,1\}^{\mathbf{Z}}$, on définit une application

$$\begin{aligned} \Phi_{\boldsymbol{\varepsilon}} \ : \mathcal{B} &\to \mathcal{B} \\ \mathbf{x} &\mapsto \mathbf{y} \end{aligned}$$

où

$$y_k = \varepsilon_k \sqrt{x_{k+1} + bx_{k-1} - c}.$$

C'est une contraction puisque

$$\begin{aligned} \left| \varepsilon_k \sqrt{x_{k+1} + bx_{k-1} - c} - \varepsilon_k \sqrt{x'_{k+1} + bx'_{k-1} - c} \right| &= \left| \frac{x_{k+1} + bx_{k-1} - x'_{k+1} - b'x_{k-1}}{\sqrt{x_{k+1} + bx_{k-1} - c} + \sqrt{x'_{k+1} + bx'_{k-1} - c}} \right| \\ &\leq \frac{2\|\mathbf{x} - \mathbf{x}'\|}{2\alpha} \end{aligned}$$

pour tous $\mathbf{x}$ et $\mathbf{x}'$ dans $\mathcal{B}$. Ainsi, $\Phi_{\boldsymbol{\varepsilon}}$ a un unique point fixe.

Posons $R_{-1} = [-M, 0] \times [-M, M]$ et $R_1 = [0, M] \times [-M, M]$. On a montré que pour tout $\boldsymbol{\varepsilon} = (\varepsilon_k)_{k\in\mathbf{Z}}$, l'ensemble $\bigcap_{k\in\mathbf{Z}} f^{-k}(R_{\varepsilon_k})$ est réduit à un unique point $h(\boldsymbol{\varepsilon})$, et que l'orbite d'un tel point ne rencontre jamais la bande $[-1,1] \times [-M, M]$. De plus, l'image de h est exactement X. Il est facile d'en déduire que h est un homéomorphisme de $\{-1,1\}^{\mathbf{Z}}$ sur X qui conjugue σ à $f|_X$.

Pour prouver que toute orbite périodique est un point selle, munissons $\mathbf{R}^2$ de la norme $(x,y) = \max(|x|, |y|)$ et considérons la matrice jacobienne

$$Df(x,y) = \begin{pmatrix} 2x & -b \\ 1 & 0 \end{pmatrix}.$$

Fixons $(x,y) \in X$ puis un vecteur (u,v) tel que $|u| \geq |v|$. Si $(u', v') = Df(x,y).(u,v)$, alors

$$|u'| \geq (2\alpha - 1)|u| \geq |v'|$$

et donc

$$\|(u', v')\| = |u'| \geq (2\alpha - 1)|u| = (2\alpha - 1)\|(u,v)\|.$$

On a montré que le cône d'équation $|v| \leq |u|$ est envoyé dans son intérieur par toute matrice $Df(x,y)$, $(x,y) \in X$, et que $Df(x,y)$ dilate la longueur d'un vecteur dans ce cône par un coefficient multiplicatif $\geq 2\alpha - 1$. Si z est un point périodique de période q, on sait que $Df^q(z) = Df(f^{q-1}(z)) \circ \ldots \circ Df(z)$ envoie le cône d'équation $|v| \leq |u|$ dans son intérieur et dilate la longueur d'un vecteur dans ce cône par un coefficient multiplicatif $\geq (2\alpha - 1)^q$. Ceci implique que l'une des valeurs propres λ de $Df^q(z)$ vérifie $|\lambda| \geq (2\alpha - 1)^q > 1$. L'autre valeur propre, notée μ, vérifie $|\mu| \leq (2\alpha - 1)^{-q} < 1$ car $|\lambda||\mu| = b^q \leq 1$. On démontrer plus généralement que X est hyperbolique. $\square$

Un bon exercice est de dessiner l'image de $[-M, M]^2$ par f et d'essayer de comprendre ce qui se passe.

5.4 Automorphismes hyperboliques du tore

On va voir comment construire une partition de Markov pour les automorphismes hyperboliques du tore $\mathbf{T}^r$. Commençons par l'exemple classique de l'automorphisme $\widehat{A}$ de $\mathbf{T}^2$, défini par l'automorphisme linéaire A dont la matrice dans la base canonique est

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Les valeurs propres de A sont $\dfrac{3 \pm \sqrt{5}}{2}$, l'espace vectoriel stable est la droite E^s d'équation $y = \dfrac{-1-\sqrt{5}}{2}x$, l'espace vectoriel instable est la droite E^u d'équation $y = \dfrac{-1+\sqrt{5}}{2}x$. Les ensembles stables et instables d'un point $z \in \mathbf{R}^2$ sont les espaces affines correspondant $z + E^s$ et $z + E^u$.

On se donne deux rectangles R_1 et R_2 de $\mathbf{R}^2$ dont les côtés sont formés d'ensembles stables et instables. Les segments stables de R_1 sont inclus dans E^s et $(1,1) + E^s$, les segments instables dans E^u et $(1,0) + E^u$. Les segments stables de R_2 sont inclus dans $(0,1) + E^s$ et $(1,1) + E^s$, les segments instables dans E^u et $(1,1) + E^u$. La projection $\pi : \mathbf{R}^2 \to \mathbf{T}^2$ est injective sur R_2 mais pas sur R_1. Cependant, elle est injective sur la réunion des deux intérieurs. Les images de R_1 et R_2 sont des parties fermées d'intérieur vide qui recouvrent le tore. Les sommets se projettent en quatre points distincts.

Puisque A^{-1} envoie $(0,0)$ sur $(0,0)$, $(1,0)$ sur $(1,-1)$, $(0,1)$ sur $(-1,2)$ et $(1,1)$ sur $(0,1)$, on peut dessiner $A^{-1}(R_1)$ et $A^{-1}(R_2)$. On va construire notre décomposition $(\widehat{S}_i)_{1 \leq i \leq 5}$ en gardant les cinq rectangles naturellement définis par les ensembles $\widehat{R}_i \cap \widehat{A}^{-1}(\widehat{R}_j)$. Chaque $\widehat{S}_i$ est homéomorphe à S_i par π où:

- les côtés de $S_1 = R_1 \cap A^{-1}(R_1)$ sont inclus dans E^s, $(0,1) + E^s$, E^u, $(1,0) + E^u$;
- les côtés de $S_2 = ((0,-1) + R_1) \cap A^{-1}(R_1)$ sont inclus dans E^s, $(0,1) + E^s$, $(0,-1) + E^u$, $(1,-1) + E^u$;
- les côtés de $S_3 = ((-1,0) + R_1) \cap A^{-1}(R_2)$ sont inclus dans $(-1,2) + E^s$, $(0,1) + E^s$, E^u, $(-1,0) + E^u$;
- les côtés de $S_4 = (0,-1) + R_2) \cap A^{-1}(R_1)$ sont inclus dans E^s, $(0,1) + E^s$, $(1,0) + E^u$, $(0,-1) + E^u$;
- les côtés de $S_5 = (-1,0) + R_2) \cap A^{-1}(R_2)$ sont inclus dans $(-1,2) + E^s$, $(0,1) + E^s$, $(0,1) + E^u$, $(-1,0) + E^u$.

Dans le cas où $\widehat{A}(\mathrm{Int}(\widehat{S}_i)) \cap \mathrm{Int}(\widehat{S}_j) \neq \emptyset$, cet ensemble est l'intérieur d'un sous-rectangle $\widehat{\Delta}_{i,j}$ de $\widehat{S}_j$ qui joint les deux côtés stables de $\widehat{S}_j$. Dans ce cas, $\widehat{\Delta}'_{i,j} = \widehat{A}^{-1}(\widehat{\Delta}_{i,j})$ est un sous-rectangle de $\widehat{S}_i$ qui joint les deux côtés instables de $\widehat{S}_i$. On pose $M_{i,j} = 1$ dans ce cas et $M_{i,j} = 0$ sinon. Remarquons que

$$M = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Considèrons le sous-décalage σ_M sur $Z_M \subset \{1, \ldots, 5\}^{\mathbf{Z}}$. Pour toute suite $\mathbf{i} = (i_k)_{k\in\mathbf{Z}} \in Z_M$ il existe un unique point $z = h(\mathbf{i})$ tel que $\bigcap_{k\in\mathbf{Z}} \widehat{A}^{-k}(\Delta'_{i_k,i_{k+1}}) = \{z\}$. L'application h définit une semi-conjugaison entre σ_M et $\widehat{A}$. Remarquons que si z a plus d'un antécédent, son orbite rencontre l'un des côtés de $\widehat{S}_i$. Ceci implique que $z \in W^s(0)$ ou $z \in W^u(0)$. Remarquons également que les antécédents de 0 sont les trois points fixes de σ_M.

Notre décomposition n'est pas vraiment une décomposition de Markov, car $\bigcap_{k\in\mathbf{Z}} \widehat{A}^{-k}(\widehat{S}_i)$ peut contenir plus d'un point (considérer la suite de période 2 formée de 1 et de 2). En fait on peut construire une vraie partition de Markov avec de tels rectangles (mais il en faut plus). Cependant la semi-conjugaison donnée par notre décomposition peut nous permettre de déduire des propriétés dynamiques de $\widehat{A}$.

Remarquons que

$$\left(\frac{3+\sqrt{5}}{2}\right)^n + \left(\frac{3-\sqrt{5}}{2}\right)^n - 2 = |\det(A^n - \mathrm{Id})| = \sharp\mathrm{Fix}(\widehat{A}^n) = \sharp\mathrm{Fix}(\widehat{\sigma}_M^n) - 2 = \mathrm{Tr}(M^n) - 2.$$

On peut vérifier que M is irréductible et apériodique, on sait donc que

$$h(\sigma_M) = \log\rho(M) = \lim_{n\to+\infty} \frac{1}{n}\ln(\sharp\mathrm{Fix}(\sigma_M^n)) = \ln\left(\frac{3+\sqrt{5}}{2}\right),$$

et par conséquent que

$$h(\widehat{A}) \leq h(\sigma_M) = \ln\rho(A).$$

Expliquons pourquoi on a une égalité. Considérons une métrique $\| \ \|$ adaptée à la décomposition hyperbolique de A. Il existe $\varepsilon_0 > 0$ tel que pour tout $z \in \mathbf{R}^2$ l'application π est injective sur toute boule $B(z, \varepsilon_0)$ (qui est un rectangle !) et induit une isométrie locale. Notons $\lambda = \dfrac{3+\sqrt{5}}{2}$. Cela implique que pour tous points $\widehat{z}$ et $\widehat{z}'$ dans $\mathbf{T}^2$

$$\max_{-1\leq i\leq 1} d(\widehat{A}^i(\widehat{z}), \widehat{A}^i(\widehat{z}')) \leq \varepsilon_0 \Rightarrow \max_{i\in\{-1,1\}} d(\widehat{A}^i(\widehat{z}), \widehat{A}^i(\widehat{z}')) \geq \lambda d(\widehat{z}, \widehat{z}').$$

On en déduit immédiatement que $\mathrm{Fix}(\widehat{A}^n)$ est (n, ε_0)-séparé, ce qui implique que

$$h(\widehat{A}) \geq \lim_{n\to+\infty} \frac{1}{n}\ln(\sharp\mathrm{Fix}(\widehat{A}^n)) = \ln\left(\frac{3+\sqrt{5}}{2}\right).$$

L'arugument précédent nous dit que $\widehat{A}$ est expansif : il existe ε_0 tel que pour tous points distincts x and y, il existe $k \in \mathbf{Z}$ tel que $d(T^k(x), T^k(y)) \geq \varepsilon_0$. Cet argument peut être généralisé à tout automorphisme $\widehat{A}$ hyperbolique du tore $\mathbf{T}^r$. Remarquons que pour un tel automorphisme, on a

$$h(\widehat{A}) \geq \lim_{n\to+\infty} \frac{1}{n}\ln(\sharp\mathrm{Fix}(\widehat{A}^n)) = \sum_{1\leq i\leq s} \ln(|\lambda_i|)$$

où $\lambda_1, \ldots, \lambda_s$ sont les valeurs propres de module > 1. Pour montrer l'inégalité inverse, il suffit de construire une partition de Markov de $\mathbf{T}^r = \bigcup_{1\leq i\leq p} \widehat{S}_i$ par des rectangles à "côtés" dans les espaces stables et instables et vérifiant des propriétés similaires à celles qu'on a vu plus haut et tel que le sous-décalage de type fini σ_M sur $Z_M \subset \{1, \ldots, p\}^{\mathbf{Z}}$ défini par la partition de Markov vérifie :

$$\lim_{n\to+\infty} \frac{1}{n} \ln(\sharp\mathrm{Fix}(\sigma_M^n)) = \lim_{n\to+\infty} \frac{1}{n} \ln(\sharp\mathrm{Fix}(\widehat{A}^n))$$

car

$$h(\sigma_M) = \ln \rho(M) = \lim_{n\to+\infty} \frac{1}{n} \ln(\sharp\mathrm{Fix}(\sigma_M^n)).$$

Ceci sera vrai si le nombre d'antécédents d'un point $\widehat{z} \in \mathbf{T}^r$ est borné. Ce sera le cas si les rectangles sont choisis assez petits pour définir une vraie partition de Markov. En fait on a

Proposition 5.4.1 : *Pour tout $\varepsilon > 0$, on peut construire une telle famille avec des rectangles de diamètre $\leq \varepsilon$.*

Ceci implique

Corrolaire 5.4.2 : *Soit $\widehat{A}$ un automorphisme hyperbolique de $\mathbf{T}^r$. Alors*

$$h(\widehat{A}) = \lim_{n\to+\infty} \frac{1}{n} \ln(\sharp\mathrm{Fix}(\widehat{A}^n)) = \sum_{1\leq i\leq s} \ln(|\lambda_i|)$$

où $\lambda_1, \ldots, \lambda_s$ sont les valeurs propres de module > 1.

5.5 Intersections homoclines

On obtient des ensembles hyperboliques non triviaux dès qu'on a un point fixe (ou périodique) de type selle dont les variétés stables et instables ont une intersection transverse. Plus précisément soit x_0 un point périodique de type selle. Un *point d'intersection homocline* x_1 est un point appartenant à $W^s(x_0) \cap W^u(x_0)$ distinct de x_0. Si les espaces tangents à $W^s(x_0)$ et à $W^u(x_0)$ en x_1 sont supplémentaires, on dit que l'intersection est *transverse*. Le théorème qui suit affirme qu'il existe des fers à cheval similaires à celui construit dans l'application de Hénon, dès qu'un point périodique de type selle a une intersection homocline transverse.

Théorème 5.5.1 : *Soit $f : M \to M$ un difféomorphisme de classe C^1 d'une variété M et x_0 un point périodique de type selle admettant un point d'intersection homocline transverse x_1. Pour tout voisinage U de x_0, il existe un entier N et une partie compacte X, invariante par f^N et hyperbolique, contenant un itéré de x_1, qui est de type fer à cheval : la restriction $f^N|_X$ est conjuguée au décalage de Bernouilli sur $\{0,1\}^{\mathbf{Z}}$ via une partition de Markov.*

On en déduit alors immédiatement les corollaires suivants :

Corollaire 5.5.2 : *Soit $f : M \to M$ un difféomorphisme de classe C^1 d'une variété M et x_0 un point périodique de type selle admettant un point d'intersection homocline transverse x_1. Alors x_1 appartient à l'adhérence de l'ensemble des points périodiques de f.*

Corollaire 4.5.3 : *Soit $f : M \to M$ un difféomorphisme de classe C^1 d'une variété compacte M et x_0 un point périodique de type selle admettant un point d'intersection homocline transverse. Alors on a $h(f) > 0$.*

On peut construire sur des variétés de dimension ≥ 3 des difféomorphismes sans point périodique dont l'entropie topologique est strictement positive. Il suffit de prendre un difféomorphisme f d'une variété compacte M dont l'entropie est non nulle (un automorphisme linéaire hyperbolique de $\mathbf{T}^2$ par exemple) puis de faire le produit, sur $M \times \mathbf{T}^1$, de f et d'une rotation d'angle irrationnel. Par une construction due à M. Rees, on peut également trouver un homéomorphisme minimal de $\mathbf{T}^2$ dont l'entropie est non nulle. Cependant, le théorème suivant, dû à A. Katok, nous dit que ce type d'exemple est impossible sur une surface dès que la différentiabilité est assez grande. Rappelons qu'un difféomorphisme f est de classe $C^{1+\varepsilon}$ si f est un difféomorphisme de classe C^1 et si Df est hölderienne de rapport ε (c'est le cas par exemple si f est un difféomorphisme de classe C^2).

Théorème 5.5.4 : *Soit $f : M \to M$ un difféomorphisme de classe $C^{1+\varepsilon}$ d'une surface compacte M. Si $h(f) > 0$, alors il existe un point périodique de type selle admettant un point d'intersection transverse.*

6.1 Cas des automorphismes hyperboliques du tore

Soit $\widehat{A}$ un automorphisme linéaire hyperbolique de $\mathbf{T}^r$ défini par une matrice A. On a vu que l'entropie topologique de $\widehat{A}$ était égal à $\ln|\lambda|$, où λ est le produit des valeurs propres de module > 1. S'il y a s_0 telles valeurs propres, on peut remarquer que $|\lambda|$ est le rayon spectral de l'automorphisme $\Lambda^{s_0}A$ défini sur le produit extérieur $\Lambda^{s_0}(\mathbf{R}^p)$. On peut remarquer également que $|\lambda|$ est strictement plus grand que le rayon spectral de Λ^s si $s \neq s_0$. Or $\Lambda^s(\mathbf{R}^p)$ apparaît naturellement comme le s-ième groupe d'homologie $H_s(\mathbf{T}^r, \mathbf{R})$ et $\Lambda^s A$ comme l'action $\widehat{A}_{*,s} : H_s(\mathbf{T}^r, \mathbf{R}) \to H_s(\mathbf{T}^r, \mathbf{R})$ définie par $\widehat{A}$ en homologie. On a donc le résultat suivant :

Proposition 6.1.1 : *Si $\widehat{A}$ est un automorphisme linéaire hyperbolique de $\mathbf{T}^r$, alors*

$$h(\widehat{A}) = \sup_{0 \le s \le r} \ln \rho(\widehat{A}_{*,s})$$

où $\widehat{A}_{*,s} : H_s(\mathbf{T}^r, \mathbf{R}) \to H_s(\mathbf{T}^r, \mathbf{R})$ est l'action induite par $\widehat{A}$ sur $H_s(\mathbf{T}^r, \mathbf{R})$.

Si $f : \mathbf{T}^r \to \mathbf{T}^r$ est homotope à $\widehat{A}$ on sait, c'est une conséquence du lemme de poursuite, que que $\widehat{A}$ est un facteur de f et donc que $h(f) \ge h(\widehat{A})$. On sait aussi que les morphismes $\widehat{A}_{*,s}$ et $f_{*,s}$ coïncident. On a donc le résultat plus général suivant :

Proposition 6.1.2 : *Si $f : \mathbf{T}^r \to \mathbf{T}^r$ est homotope à un automorphisme linéaire hyperbolique de $\mathbf{T}^r$, alors*

$$h(f) \ge \sup_{0 \le s \le r} \ln \rho(f_{*,s})$$

où $f_{*,s} : H_s(\mathbf{T}^r, \mathbf{R}) \to H_s(\mathbf{T}^r, \mathbf{R})$ est l'action induite par f sur $H_s(\mathbf{T}^r, \mathbf{R})$.

6.2 Conjecture de l'entropie

Au vu du résultat précédent, on peut se demander si, pour toute transformation continue $f : M \to M$ d'une variété compacte M de dimension r, on a $h(f) \ge \sup_{0 \le s \le r} \ln \rho(f_{*,s})$ où $f_{*,s} : H_s(M, \mathbf{R}) \to H_s(M, \mathbf{R})$ est l'action induite par f sur $H_s(M, \mathbf{R})$. Expliquons pourquoi ce n'est pas vrai. Supposons que M est orientable. Dans ce cas, $\dim(H_r(M, \mathbf{R})) = 1$ et $f_{*,r}$ est une homothétie. Le rapport de cette homothétie, noté $\deg(f)$, s'appelle le *degré* de f. C'est en fait un entier et on vérifie aisément que $\deg(f \circ g) = \deg(f) \circ \deg(g)$, en particulier $\deg(f) \in \{-1, 1\}$ si f est un homéomorphisme. L'inégalité $h(f) \ge \ln(|\deg(f)|)$ est clairement vraie si f est un homéomorphisme mais cesse d'être nécessairement vraie sinon. Illustrons cela sur des exemples. L'application $f : z \mapsto z^2$, définit une application de degré 2 de la sphère de Riemann. Tout point de module < 1 (resp. > 1) tend dans le futur vers 0 (resp. ∞). L'ensemble des points non errant $\Omega(f)$ est donc formé de 0 de ∞ et du cercle unité sur lequel f est conjugué à un endomorphisme de degré 2. Dans ce cas on a bien

$$h(f) = h(f_{\Omega(f)}) = \ln 2 = \ln(\deg(f)).$$

L'application $g\ : z \mapsto z^2/2|z|$, est également de degré 2 et de dynamique très simple, il y a deux points fixes 0 et ∞, et toute autre orbite va de ∞ à 0. On a $\Omega(g) = \{0, \infty\}$ et

$$h(g) = 0 < \ln(\deg(g)).$$

La différence entre les deux cas tient au fait que f est de classe C^1 alors que g n'est pas différentiable en 0 et surtout en ∞. En fait, on a le résultat suivant, dû à M. Misiurewicz et F. Przytycki :

Théorème 6.2.1 : *Si f est une transformation de classe C^1 d'une variété compacte lisse orientable M, on a $h(f) \geq \ln(\deg(f))$.*

Là différentielle joue un rôle important, ce qui a amené M. Shub à conjecturer :

Conjecture 6.2.2 : *Si f est une transformation de classe C^1 d'une variété compacte lisse M, on a $h(f) \geq \sup_{0\leq s\leq r} \rho(f_{*,s})$ où $f_{*,s}\ : H_s(M, \mathbf{R}) \to H_s(M, \mathbf{R})$ est l'action induite par f sur $H_s(M, \mathbf{R})$.*

Cettte conjecture est toujours ouverte mais on sait qu'elle est vraie si f est de classe C^∞. C'est une conséquence des travaux de Yomdin qui lie l'entropie topologique à la croissance des volumes sous l'action de la dynamique. Cependant, on a le résultat suivant de A. Manning qui ne nécessite pas de condition différentiable.

Théorème 6.2.3 : *Si f est une transformation continue d'une variété compacte M, on a $h(f) \geq \ln \rho(f_{*,1})$ où $f_{*,1}\ : H_1(M, \mathbf{R}) \to H_1(M, \mathbf{R})$ est l'action induite par f sur $H_1(M, \mathbf{R})$.*

Démonstration. On munit M d'une structure riemmanienne. On note alors $\|\sigma\|$ la longueur d'un chemin σ et d la distance naturellement définie sur M à l'aide de la structure riemannienne. On peut trouver δ_0 tel que toute boule $B(x, \delta_0)$ est dans le domaine d'une carte contractile et tel que pour tous points y et y' dans $B(x, \delta_0)$ il existe une unique géodésique, notée $[y, y']$ qui joint y à y'. On fixe alors $\delta < \delta_0/4$ tel que $d(x, x') < 2\delta \Rightarrow d(f(x), f(x')) < \delta_0$. On note $\mathcal{C}$ l'ensemble des chemins σ qui sont C^1 par morceaux et qui vérifient $\delta/2 \leq \|\sigma\| \leq \delta$. On peut définir le groupe d'homologie $H_1(M, \mathbf{R})$ à l'aide des chemins dans $\mathcal{C}$. Plus précisément, le $\mathbf{R}$-module libre engendré par les éléments de $\mathcal{C}$ est formé des *chaînes* $\sigma = \sum_{i\in I} a_i\sigma_i$, I fini ; les *cycles* sont les chaînes $\sigma = \sum_{i\in I} a_i\sigma_i$ telles que $\partial\sigma = \sum_{i\in I} a_i\partial\sigma_i = 0$; les *bords* sont les cycles qui bordent une 2-chaîne. On sait alors que l'espace des cycles quotienté par l'espace des bords est le $\mathbf{R}$-espace vectoriel de dimension finie $H_1(M, \mathbf{R})$. On peut voir également les bords comme les cycles $\sigma = \sum_{i\in I} a_i\sigma_i$ tels que $\int_\sigma \omega = \sum_{i\in I} a_i \int_{\sigma_i} \omega = 0$ pour toute forme différentielle fermée ω. En particulier, si σ est un cycle et ω une forme fermée, alors $\int_\sigma \omega = \langle [\omega], [\sigma]\rangle$ ne dépend que de la classe de cohomologie $[\omega] \in H^1(M, \mathbf{R})$ et de la classe d'homologie $[\sigma] \in H_1(M, \mathbf{R})$. Pour toute chaîne $\sigma = \sum_{i\in I} a_i\sigma_i$, on peut définir $\|\sigma\| = \sum_{i\in I} |a_i|\|\sigma_i\|$, puis pour toute classe d'homologie $u \in H_1(M, \mathbf{R})$ on peut définir $\|u\| = \inf_{[\sigma]=u} \|\sigma\|$. On obtient ainsi une norme $\|\ \|$ sur $H_1(M, \mathbf{R})$. Pour vérifier que l'on obtient une norme, il suffit de remarquer que pour toute chaîne σ et toute forme différentielle ω, on a

$$\left|\int_\sigma \omega\right| \leq \|\omega\|\|\sigma\|,$$

où

$$\|\omega\| = \max_{x\in M,\, y\in T_xM,\, \|y\|=1} \|\omega_x(y)\|.$$

Rappelons que $r(n, \delta)$ est le cardinal minimum des ensembles (n, δ) couvrant. Nous allons prouver le lemme suivant :

Lemme 6.2.4 : *Pour tout $n \geq 0$ et tout $u \in H_1(M, \mathbf{R})$, on a*

$$\|f_{*,1}^{n-1}(u)\| \leq 16\, r(n, \delta)\|u\|.$$

Expliquons pourquoi le lemme implique le théorème. On déduit du lemme que

$$\ln\left(\|f_{*,1}^{n-1}\|^{\frac{1}{n-1}}\right) \leq \frac{\ln 16}{n-1} + \frac{n}{n-1}\,\frac{\ln(r(n,\delta))}{n},$$

et donc que

$$\ln \rho(f_{*,1}) \leq \liminf_{n \to +\infty} \frac{\ln(r(n,\delta))}{n} \leq h(f).$$

Démonstration du lemme. Nous allons montrer que l'image par f^{n-1} de tout chemin $\sigma \in \mathcal{C}$, est homotope, à extrémités fixées, à un chemin C^1 par morceaux de longueur $\leq 4\delta r(n, \delta)$. Ceci impliquera le lemme. En effet, pour tout $u \in H_1(M, \mathbf{R})$, on peut trouver un cycle $\sigma = \sum_{i \in I} a_i \sigma_i$, où $\sigma_i \in \mathcal{C}$, tel que $\|\sigma\| \leq 2\|u\|$. Pour tout $i \in I$, on peut trouver un chemin σ_i' de longueur $\leq 4\delta r(n, \delta)$ homotope à $f^{n-1}(\sigma_i)$. Ainsi le cycle $\sigma = \sum_{i \in I} a_i \sigma_i'$ représente $f_{*,1}^{n-1}(u)$. On a donc

$$\|f_{*,1}^{n-1}(u)\| \leq \sum_{i \in I} |a_i| \|\sigma_i'\| \leq \left(\sum_{i \in I} |a_i|\right) 4\delta r(n, \delta) \leq 8\|\sigma\| r(n, \delta) \leq 16\|u\| r(n, \delta).$$

Notons d_n la distance

$$d_n(x, x') = \max_{0 \leq i < n} d(f^i(x), f^i(x')),$$

et fixons un ensemble (n, δ)-couvrant R de cardinal $r(n, \delta)$. Si x et x' sont deux points tels que $d_n(x, x') < 2\delta$, les chemins géodésiques $[f^i(x), f^i(x')]$ sont bien définis pour tout $i \in \{0, \ldots, n-1\}$, puisque $d(f^i(x), f^i(x')) \leq d_n(x, x') < 2\delta \leq \delta_0$. De plus, les chemins $f([f^i(x), f^i(x')])$ et $[f^{i+1}(x), f^{i+1}(x')]$ sont homotopes à extrémités fixées, puisqu'ils sont tous les deux inclus dans la boule $B(f^{i+1}(x), \delta_0)$ et donc dans le domaine d'une carte contractile. On en déduit en particulier que $f^{n-1}[x, x']$ et $[f^{n-1}(x), f^{n-1}(x')]$ sont homotopes.

On peut écrire tout chemin $\sigma \in \mathcal{C}$ comme assemblage $\sigma = \prod_{0 \leq k \leq K} \sigma^k$ de lacets σ^k, où σ^k est inclus dans une boule $B_{d_n}(x_k, \delta)$, $x_k \in R$. En particulier $B_{d_n}(x_k, \delta) \cap B_{d_n}(x_{k+1}, \delta) \neq \emptyset$ et $d_n(x_k, x_{k+1}) < 2\delta$. Notons x et x' l'origine et l'extrémité de σ. Les deux chemins σ et $[x, x_0] \prod_{0 \leq k < K} [x_k, x_{k+1}][x_K, x']$ sont inclus dans $B(z, 4\delta)$ et donc dans $B(z, \delta_0)$. Ils sont donc homotopes à extrémités fixées. De plus, on peut toujours supposer que les x_k sont distincts (s'il y a une boucle, on peut l'enlever). Or on sait que $f^{n-1}(\sigma)$ est homotope à

$$\sigma' = [f^{n-1}(x), f^{n-1}(x_0)] \prod_{0 \leq k < K} [f^{n-1}(x_k), f^{n-1}(x_{k+1})][f^{n-1}(x_K), f^{n-1}(x')]$$

et on peut remarquer que

$$\|\sigma'\| \leq (1 + r(n, r))2\delta \leq 4r(n, \delta).$$

□

Corollaire 6.2.5 : *Si f est un homéomorphisme d'une variété compacte M de dimension ≤ 3, on a $h(f) \geq \sup_{0\leq s\leq 3} \ln \rho(f_{*,s})$.*

Démonstration. Il suffit de rappeler que par dualité de Poincaré, $f_{*,2}$ est conjuguée à l'application duale de $f_{*,1}^{-1}$ et que ses valeurs propres de $f_{*,2}$ sont donc les inverses des valeurs propres de $f_{*,1}$ et d'utiliser l'égalité $h(f) = h(f^{-1})$. □

En fait on a une meilleure minoration de l'entropie si on s'intéresse au groupe fondamental plutôt qu'au premier groupe d'homologie. Si G est un groupe de type fini et si φ est un endomorphisme de G on peut définir *l'entropie algébrique* de φ de la façon suivante. On commence par choisir un ensemble générateur symétrique fini Σ, ce qui permet de définir une longueur l_Σ sur G en notant $l_\Sigma(g)$ le nombre de lettres minimales nécessaires pour écrire g comme mot dans l'alphabet Σ. Si on pose

$$L_n(\varphi, \Sigma) = \max_{g\in\Sigma} l_\Sigma(\varphi^n(g)),$$

on peut montrer que la suite $(L_n(\varphi, \Sigma))_{\geq 0}$ est sous-multiplicative et donc que

$$\lim_{n\to+\infty} \frac{1}{n} \ln(L_n(\varphi, \Sigma))$$

existe. On peut démontrer également que cette limite ne dépend pas de Σ, c'est l'entropie algébrique de φ, on la note $h_{\text{alg}}(\varphi)$. Elle vérifie des propriétés analogues à l'entropie (invariance par conjugaison, décroissance par semi-conjugaison, ...). On peut démontrer également que l'entropie ne varie pas par composition par un automorphisme intérieur. Plus précisément si $\iota : g \mapsto g_0^{-1} g g_0$ est un automorphisme intérieur, alors $h_{\text{alg}}(\iota\varphi) = h_{\text{alg}}(\varphi)$. Supposons maintenant que f est une transformation continue d'une variété compacte M. Fixons $x_0 \in M$ et un chemin α joignant x_0 à $f(x_0)$. L'application, qui à un lacet γ basé en x_0 associe le lacet $\alpha^{-1}\gamma\alpha$ induit un endomorphisme $\varphi_{x_0,\alpha}$ du groupe fondamental $\pi_1(M, x_0)$. L'entropie algébrique de $\varphi_{x_0,\alpha}$ ne dépend pas du choix de φ (un autre choix donnerait le même automorphisme composé par un automorphisme intérieur) ni du choix de x_0 (un autre choix donnerait une conjugaison). On peut donc définir une entropie de f sur le groupe fondamental

$$h_{\pi_1}(f) = h_{\text{alg}}(\varphi_{x_0,\alpha}).$$

Le groupe $H_1(M, \mathbf{Z})$ est l'abélianisé de $\pi_1(M, x_0)$ et l'action de $f_{*,1}$ sur $H_1(M, \mathbf{Z})$ est un facteur (via le morphisme de Hurewicz) de $\tau_{x_0,\alpha}$ et a donc une entropie algébrique plus petite. Il n'est pas difficile de voir que cette entropie n'est rien d'autre que $\ln(\rho(f_{*,1}))$. Ainsi, on a $h_{\pi_1}(f) \geq \ln(\rho(f_{*,1}))$ Le résultat suivant de Bowen améliore donc (même s'il est moins effectif) le résultat de Manning :

Théorème 6.2.6 : *Si f est une transformation continue d'une variété compacte M, on a $h(f) \geq h_{\pi_1}(f)$.*

Les deux théorèmes précédents sont particulièrement intéressants en dimension deux. En effet, supposons que O soit une orbite périodique de f (ou une réunion finie de telles orbites). Supposons de plus que φ soit un difféomorphisme. On peut alors éclater chaque point $x \in O$ en le remplaçant par le cercle des demi-droites sur le plan tangent $T_x(M)$. Le difféomorphisme $f|_{M\setminus O}$ s'étend donc en un homéomorphisme $\overline{f}$ d'une variété compacte dont le H_1 et a fortiori le π_1 est bien plus gros que le H_1 ou le π_1 de M. Or il n'est pas difficile de voir que $h(\overline{f}) = h(f)$ car f est un facteur de $\overline{f}$ et $\overline{f}$ induit sur l'image inverse de O un homéomorphisme d'entropie nulle.

Il se peut très bien que pour des raisons algébriques, la structure de certaines orbites périodiques force l'entropie à être non nulle.

L'énoncé original du théorème de Manning est en fait plus fort que le théorème 6.2.3. Il s'applique non seulement pour une variété, comme il est vu plus haut, mais plus généralement pour tout espace métrique compact (X, d) vérifiant les conditions suivantes (et on peut vérifier que nous n'avons utilisé que ces deux propriétés dans la preuve) :

i) pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que pour tout couple (x, y) de points de X vérifiant $d(x, y) < \eta$, il existe un chemin joignant x à y de diamètre inférieur à ε;

ii) il existe $\varepsilon_0 > 0$ tel que tout lacet de diamètre inférieur à ε_0 est contractile.

Nous allons voir dans la section qui suit une application de cette dernière remarque.

6.3 Ensembles invariants isolés

Nous allons conclure par une courte introduction à la théorie de l'indice de Conley discrète (la version continue pour les flots est due à Conley lui-même) et donner une application du théorème de Manning. Commençons par rappeler la notion plus classique d'indice de Poincaré-Lefschetz.

Indice de Poincaré-Lefschetz

Notons S^{r-1} la sphère unité de l'espace euclidien $\mathbf{R}^r$. Soit U un voisinage de 0 dans $\mathbf{R}^r$ et $f : U \to f(U)$ une application continue ayant un point fixe isolé en 0. Si $\varepsilon > 0$ est assez petit, l'application

$$z \mapsto \frac{\varepsilon z - f(\varepsilon z)}{\|\varepsilon z - f(\varepsilon z)\|}$$

est bien définie sur S^{r-1} et son degré ne dépend pas de ε. On l'appelle l'*indice de Poincaré-Lefschetz* du point fixe et on le note $i(f, 0)$. Dans le cas où f est différentiable en 0 et où 1 n'est pas valeur propre de $Df(0)$ (ce qui implique de fait que le point est isolé), alors $i(f, 0) = (-1)^l$, où l est le nombre de valeurs propres réelles supérieurs à 1. Parmi les propriétés de l'indice, on peut rappeler:

- Si f préserve l' orientation, alors $i(f^{-1}, 0) = i(f, 0)$ si r est pair et $i(f^{-1}, 0) = -i(f, 0)$ si r est impair;
- if f renverse l'orientation, alors $i(f^{-1}, 0) = -i(f, 0)$ si r est pair et $i(f^{-1}, 0) = i(f, 0)$ si r est impair.

L'indice étant invariant par conjugaison, on peut définir l'indice d'un point fixe isolé d'une transformation continue f définie sur une variété M. On sait alors que si x_1 est un point fixe isolé d'une transformation f_1 sur M_1 et x_2 un point fixe isolé d'une transformation f_2 sur M_2, alors on a

$$i(f_1 \times f_2, (x_1, x_2)) = i(f_1, x_1)i(f_2, x_2).$$

On peut en fait généraliser la notion d'indice d'un point fixe à celle d'indice d'une partie compacte invariante X s'il existe un voisinage de X qui ne contient aucun point fixe qui n'est pas dans X. Cet indice vérifie les deux propriétés suivantes :

- si $i(f, X_1)$ et $i(f, X_2)$ sont bien définis et si $X_1 \cap X_2 = \emptyset$, alors $i(f, X_1 \cup X_2)$ est bien défini et on a

$$i(f, X_1 \cup X_2) = i(f, X_1) + i(f, X_2);$$

- si $i(f,X)$ est bien défini et si X contient un nombre fini de points fixes de f, alors

$$i(f,X) = \sum_{x \in \mathrm{Fix}(f) \cap X} i(f,x).$$

Si f est une transformation continue d'une variété compacte M, on peut donc définir l'indice $i(f,M)$. Cet indice s'appelle le *nombre de Lefschetz* et se note $\Lambda(f)$. C'est la somme des indices des points fixes si ceux-ci sont en nombre fini. La *formule de Lefschetz* relie ce nombre à l'action de f sur les groupes d'homologie par l'égalité:

$$\Lambda(f) = \sum_{s=0}^{r} (-1)^s \mathrm{tr}(f_{*,s}),$$

où r est la dimension de M.

Par exemple, si f est un homéomorphisme de la sphère S^r, alors $\Lambda(f) = 0$ si r est impair et f préserve l'orientation, ou alors si r est pair et f renverse l'orientation. Par contre $\Lambda(f) = 2$ si r est pair et f préserve l'orientation, ou alors si r est impair et f renverse l'orientation. En effet, on a

$$\Lambda(f) = \mathrm{tr}(f_{*,0}) + (-1)^r \mathrm{tr}(f_{*,r}) = 1 + (-1)^r \deg(f).$$

La notion d'indice d'un point fixe isolé ainsi que la formule de Lefchetz ont été généralisées par Dold à une classe d'espaces topologiques plus générale que les variétés, les ANR (absolutely neighborhood retract) dont nous donnerons un exemple un peu plus loin. Parmi les propriétés de cette indice généralisé, citons le fait que $i(f,x) = 1$ si x est un point fixe attractif au sens suivant : il existe un voisinage compact V de x vérifiant $f(V) \subset V$ et $\bigcap_{n \geq 0} f^n(V) = \{x\}$.

Ensemble invariant isolé, paire filtrante

Soient U et V deux parties ouvertes d'une variété M et $f : U \to V$ un homéomorphisme. Une partie compacte X, invariante par f, est dite *isolée* s'il existe un voisinage compact N de X tel que $\bigcap_{k \in \mathbf{Z}} f^{-k}(N) = X$. On dit que N est un *voisinage isolant* de X. Remarquons que l'indice $i(f,X)$ est bien défini et plus généralement que la suite des indices des itérés $(i(f^n,X))_{n \geq 0}$ est bien définie. Dans le cas où $\bigcap_{k \leq 0} f^{-k}(N) = X$, l'ensemble X est un *attracteur*, dans le cas où $\bigcap_{k \geq 0} f^{-k}(N) = X$, c'est un *répulseur*. Les attracteurs et les répulseurs sont des cas particuliers d'ensembles invariants isolés. Toute orbite périodique hyperbolique est également un ensemble invariant isolé. Parmi les voisinages isolants, une classe importante est formée des *blocs isolants*. Il s'agit des voisinages isolants N tels que $f^{-1}(N) \cap N \cap f(N) \subset \mathrm{int}(N)$. On définit l'ensemble de sortie N^- d'un bloc isolant N par la formule

$$N^- = \{x \in N \mid f(x) \notin \mathrm{int}(N)\}.$$

On appelle alors *paire filtrante* de X toute paire (N,L) formée d'un bloc isolant N de X et d'une partie fermée $L \subset N$ telle que:

- $\overline{N \setminus L}$ est un voisinage isolant de X;
- L est un voisinage de N^- dans N,
- $f(L) \cap \overline{N \setminus L} = \emptyset$.

Si N, L et $\overline{N \setminus L}$ sont des variétés à bord, on dira que la paire est *régulière*. On peut toujours construire des paires régulières. Nous allons le montrer dans le cas particulier où M est l'espace $\mathbf{R}^m$ et où X est connexe, en établissant le résultat plus précis suivant :

Proposition 6.3.1 : *Si X est une partie compacte connexe invariante isolée d'un homéomorphisme $f : U \to V$ entre deux parties ouvertes de $\mathbf{R}^r$, il existe une paire filtrante régulière (N, L) de X qui vérifie l'une des trois propriétés suivantes:*

i) la variété N est connexe, l'ensemble L est vide et X est un attracteur;

ii) la variété N est connexe, la variété L n'est pas vide, il existe une seule composante connexe bornée de $\mathbf{R}^r \setminus L$ qui est incluse dans N, cette composante contient X, qui est un répulseur;

ii) la variété N est connexe, la variété L n'est pas vide, et aucune composante connexe bornée de $\mathbf{R}^r \setminus L$ n'est incluse dans N.

Démonstration. Soit N un voisinage (compact) isolant de X. Nous savons que $\bigcap_{k\in\mathbf{Z}} f^{-k}(N) = X$ et en déduisons donc qu'il existe $n_0 \geq 1$ tel que

$$\bigcap_{|k|\leq n_0} f^{-k}(N) \subset \mathrm{int}(N).$$

Nous pouvons donc construire une suite $(N_n)_{0\leq n<n_0}$ de voisinages compacts de X, telle que $N_0 = N$, telle que $N_n \subset \mathrm{int}(N_{n-1})$, pour tout $n \in \{1, \ldots, n_0\}$, et telle que

$$\bigcap_{|k|\leq n_0} f^{-k}(N) \subset \mathrm{int}(N_{n_0-1}).$$

Posons alors

$$N' = \bigcap_{|k|<n_0} f^{-k}(N_{n_0-1-|k|}).$$

Il s'agit bien sûr d'un voisinage compact isolant. Vérifions que c'est un bloc isolant, c'est-à-dire que

$$f(N') \cap N' \cap f^{-1}(N') \subset \mathrm{int}(N').$$

en effet, d'une part, on a

$$f(N') \cap N' \cap f^{-1}(N') \subset f(N) \cap N \cap f^{-1}(N) \subset \mathrm{int}(N_{n_0-1}),$$

d'autre part, pour tout $k \in \{1, \ldots, n_{0-1}\}$, on a

$$f(N') \cap N' \cap f^{-1}(N') \subset f^k(N_{n_0-k}) \cap f^{-k}(N_{n_0-k}) \subset \mathrm{int}(f^k(N_{n_0-1-k})) \cap \mathrm{int}(f^{-k}(N_{n_0-1-k})).$$

Quitte à remplacer N par N', on peut supposer que N est un bloc isolant. On peut trouver un voisinage N'' de N qui est une variété à bord et qui est suffisament proche de N pour vérifier

$$f(N'') \cap N'' \cap f^{-1}(N'') \subset \mathrm{int}(N'').$$

C'est donc également un bloc isolant. Quitte à remplacer N par N'' on peut supposer que N est une variété à bord. Notons maintenant N''' la composante connexe de N qui contient X. Puisque $\partial N''' \subset \partial N$, on a

$$f(N''') \cap N''' \cap f^{-1}(N''') \cap \partial N''' = \emptyset$$

et donc

$$f(N''') \cap N''' \cap f^{-1}(N''') \subset \mathrm{int}(N''').$$

Ainsi N''' est un bloc isolant de X. Quitte à remplacer N par N''', on peut donc toujours supposer que N est une variété à bord connexe.

Si l'ensemble sortant N^- est vide, alors N est positivement invariant et $\bigcap_{n\geq 0} f^n(N) = X$. Nous pouvons prendre pour L l'ensemble vide, nous sommes dans le premier cas.

Supposons maintenant que N^- est non vide. Les inclusions $N^- \subset N$ et $f(N^-) \cap N \subset \partial N$ impliquent que

$$f(N^-) \cap N \cap f^{-1}(N) = \emptyset.$$

Nous pouvons trouver un voisinage L de N^- dans N tel que

$$f(L) \cap N \cap f^{-1}(N) = \emptyset$$

et tel que L et $N \setminus L$ sont des variétés à bord. Remarquons que (N, L) est une paire filtrante régulière. En effet, nous avons

$$\overline{N \setminus L} \subset N \cap f^{-1}(N)$$

et donc

$$f(L) \cap \overline{N \setminus L} = \emptyset.$$

Notons alors W_i, $1 \leq i \leq m$, les composantes connexes bornées de $\mathbf{R}^r \setminus L$ qui sont incluses dans N. Nous savons que

$$f(\partial W_i) \cap \overline{W}_i \subset f(L) \cap \overline{N \setminus L} = \emptyset,$$

et donc que

$$\overline{W}_i \subset f(W_i) \text{ ou } \overline{W}_i \cap f(\overline{W}_i) = \emptyset.$$

Dans le cas où $\overline{W}_i \subset f(W_i)$, l'ensemble X est nécessairement inclus dans W_i, c'est un répulseur et on a $X = \bigcap_{k\geq 0} f^{-k}(\overline{W}_i)$. Il existe donc au plus une composante vérifiant la première égalité. Pour montrer la proposition, il suffit de montrer que (N, L') est une paire filtrante, où L' est formée de la réunion de L et des composantes W_i telles que $\overline{W}_i \cap f(\overline{W}_i) = \emptyset$. Il faut montrer que si $\overline{W}_i \cap f(\overline{W}_i) = \emptyset$, alors on a

$$X \cap \overline{W}_i = \emptyset \text{ et } f(W_i) \cap N \cap f^{-1}(N) = \emptyset.$$

La première égalité est évidente puisque X est invariant et $\overline{W}_i \cap f(\overline{W}_i) = \emptyset$. L' inclusion $W_i \subset N \setminus N^-$ implique que $f(W_i) \subset \mathrm{int}(N)$ et donc $f(\partial W_i) \subset N$. L'inclusion $\partial W_i \subset L$ implique que

$$f(\partial W_i) \cap f^{-1}(N) \subset f(L) \cap N \cap f^{-1}(N) = \emptyset.$$

Par connexité de $f^{-1}(N)$, on en déduit que $f^{-1}(N) \subset f(W_i)$ ou $f^{-1}(N) \cap f(\overline{W}_i) = \emptyset$. Le premier cas est impossible puisque $S \subset f^{-1}(N)$ et $S \cap f(W_i) = \emptyset$. Nous sommes donc dans le second cas, on a $f(W_i) \cap N \cap f^{-1}(N) = \emptyset$.

□

Réciproquement, on peut montrer que si X est un attracteur (resp. un répulseur), il existe alors une paire régulière (N, L) vérifiant la condition i) (resp. ii)).

Si (N, L) est une paire régulière, alors l'espace quotient N_L où L est identifié à un point est un ANR et on peut donc utiliser la formule de Lefschetz-Dold. Le fait que (N, L) soit une paire filtrante implique que f induit une transformation continue $\overline{f} : N_L \to N_L$ qui fixe $[L]$ et envoie tout point $x \notin L$ sur la projection de $f(x)$ dans N_L. La continuité de $\overline{f}$ provient du fait que $f(x) \in L$ si $x \notin L$ est proche de L. Remarquons à ce propos que $\overline{f}$ est constante égale à $[L]$

au voisinage de $[L]$. En particulier, $[L]$ est un point fixe attractif et on a donc $i(\overline{f},[L]) = 1$. La dynamique de $\overline{f}$ est facile à comprendre : pour tout point $x \in N_L$, soit l'ensemble ω-limite d'un point x est inclus dans X, soit, il existe n tel que $\overline{f}^n(x) = [L]$.

Si (N, L) est une paire régulière, on peut identifier les groupes d'homologie réduite $\widetilde{H}_s(N_L, \mathbf{R})$ avec les groupes d'homologie relative $H_s(N, L, \mathbf{R})$. Rappelons que $\widetilde{H}_s(N_L, \mathbf{R}) \sim H_s(N_L, \mathbf{R})$ si $s > 0$ et que $\widetilde{H}_0(N_L, \mathbf{R})$ est le noyau de l'application $\sum_{\kappa \in \pi_0(N_L)} \lambda_\kappa \kappa \mapsto \sum \lambda_\kappa$. En particulier, on sait que $\widetilde{H}_s(N_L, \mathbf{R}) = 0$ si $s > r$. Appliquons maintenant la formule de Lefschetz Dold à chaque itéré $\overline{f}^n$, $n \geq 0$. Nous obtenons

$$\Lambda(\overline{f}^n) = \sum_{s=0}^{r} (-1)^s \mathrm{tr}(\overline{f}^n_{*,s}),$$

où

$$\Lambda(\overline{f}^n) = i(\overline{f}^n, [L]) + i(\overline{f}^n, X) = 1 + i(f^n, X).$$

On peut donc écrire

$$i(f^n, X) = \sum_{s=0}^{r} (-1)^s \mathrm{tr}(\widetilde{f}^n_{*,s}),$$

où $\widetilde{f}_{*,s}$ est l'action de $\overline{f}$ sur $\widetilde{H}_s(N_L, \mathbf{R})$ ou, de façon équivalente, l'action naturelle de f sur $H_s(N, L, \mathbf{R})$.

Dans le troisième cas, N_L est connexe par arcs, on en déduit que $\widetilde{H}_0(N_L, \mathbf{R}) = 0$. On sait également que $\widetilde{H}_r(N_L, \mathbf{R}) = 0$. Dans le premier cas, N est positivement invariant, l'ensemble N_L est obtenu par ajout à N du point $\{\emptyset\}$ et $\overline{f}$ est l'extension de $f|_N$ qui fixe ce point. On a $\widetilde{H}_0(N_L, \mathbf{R}) \sim \mathbf{R}$ et $\widetilde{f}_{*,0} = \mathrm{Id}$. Dans le deuxième cas, on a $\widetilde{H}_0(N_L, \mathbf{R}) = 0$. Par contre $\widetilde{H}_r(N_L, \mathbf{R}) \sim \mathbf{R}$, on a $\widetilde{f}_{*,r} = \mathrm{Id}$ si f préserve l'orientation et $\widetilde{f}_{*,r} = -\mathrm{Id}$ si f renverse l'orientation.

Illustrons ce qui précède dans le cas où X se réduit au point fixe hyperbolique $x = 0$. Notons E^s et E^u les espaces stables et instables de $Df(0)$ et posons $r_u = \dim E_u$. Supposons que $DT(0)$ soit λ-hyperbolique, où $\lambda < 1$. Considérons une norme adaptée $\| \ \|$ et écrivons $\mathbf{R}^m = E^s \oplus E^u$. Fixons $\lambda' \in]\lambda, 1[$. Nous savons qu'il existe $\varepsilon > 0$ tels que si $\|x_s + x_u\| = \max(\|x_s\|, \|x_u\|) \leq \varepsilon$, alors $\|y^u\| \geq \lambda' \|x^u\|$ et $\|z^s\| \geq \lambda' \|x^s\|$ où on écrit $T(x^s + x^u) = y^s + y^u$ et $T^{-1}(x^s + x^u) = z^s + z^u$. Si $\eta > 0$ est assez petit, on obtient une paire filtrante régulière en posant

$$N = \{x_s + x_u \ | \ \|x_s\| \leq \varepsilon, \|x^u\| \leq \varepsilon\}$$

et

$$L = \{x_s + x_u \ | \ \|x_s\| \leq \varepsilon, \|x^u\| \leq \eta\}.$$

Observons que l'espace N_L à le type d'homotopie de la sphère S^{r_u}, et que $\widetilde{f}_{*,r_u}$ est l'homothétie de rapport $(-1)^r$, où r est le nombre de valeurs propres réelles inférieures à 1. Pour obtenir l'indice de 0 , il faut multiplier ce nombre par $(-1)^{r_u}$. On trouve donc $(-1)^{r'}$ ou r' est le nombre de valeurs propres supérieures à 1.

Étude de $\widetilde{f}_{*,i}$, suite des indices des itérés

Nous allons montrer le résultat suivant

Proposition 6.3.2 : *Si X est connexe et si l'entropie topologique $h(f|_X)$ est nulle, alors la suite $(\mathrm{tr}(\widetilde{f}^n_{*,1}))_{n\geq 1}$ est périodique et prend périodiquement une valeur positive ou nulle. Il en est de même pour la suite $(\mathrm{tr}(\widetilde{f}^n_{*,2}))_{n\geq 1}$ si $m = 3$.*

Nous allons commencer par montrer les deux résultats algébriques suivants:

LEMME 6.3.3 : *Soient E et F deux espaces vectoriels sur $\mathbf{C}$ et $u : E \to E$ et $v : F \to F$ deux endomorphismes. Si $\operatorname{tr}(u^k) = \operatorname{tr}(v^k)$ pour tout $k \geq 1$, alors u et v ont les mêmes valeurs propres non nulles comptées avec leur multiplicité.*

Démonstration. Notons $\{\lambda_1, \ldots, \lambda_r\}$ et$\{\mu_1, \ldots, \mu_s\}$ les valeurs propres non nulles respectives de u et v. On a

$$\sum_{k=1}^{\infty} \frac{\operatorname{tr}(u^k)}{k} Z^k = \sum_{k=1}^{\infty} \frac{\lambda_1^k + \cdots + \lambda_r^k}{k} Z^k = -\sum_{i=1}^{r} \log(1 - \lambda_i Z).$$

On en déduit l'égalité polynomiale

$$\prod_{i=1}^{r}(1 - \lambda_i Z) = \prod_{j=1}^{s}(1 - \mu_j Z),$$

qui implique que $r = s$ et que pour tout j il existe i_j, tel que $\mu_j = \lambda_{i_j}$. □

LEMME 6.3.4 : *Soit E un espace vectoriel sur $\mathbf{C}$ et u un endomorphisme qui peut être représenté par une matrice à coefficients entiers. Si toutes les valeurs propres de u ont un module inférieur ou égal à 1, alors toutes les valeurs propres non nulles sont racines de l'unité.*

Démonstration. On peut écrire le polynôme caractéristique de u sous la forme

$$P_u(X) = X^s \prod_{1 \leq i \leq r} (X - \lambda_i),$$

où $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres non nulles de u. Les coefficients de P_u étant entiers, par hypothèse, on en déduit que $\prod_{1 \leq i \leq r} \lambda_i \in \mathbf{Z}$. On sait également, par hypothèse, que chaque λ_i vérifie $|\lambda_i| \leq 1$. On en déduit que $|\lambda_i| = 1$. On peut donc écrire $\lambda_i = e^{2i\pi\alpha_i}$, où $\alpha_i \in \mathbf{R}^r/\mathbf{Z}^r$. On sait que tout point de $\mathbf{R}^r/\mathbf{Z}^r$ est un point récurrent de la rotation

$$(t_1, \ldots, t_r) \mapsto (t_1 + \alpha_1, \ldots, t_r + \alpha_r).$$

On peut donc trouver un entier $n > 0$ tel que tout nombre $\lambda_i^n = e^{2i\pi n\alpha_i}$ est proche de 1. Ceci implique que $\operatorname{tr}(u^n) = \lambda_1^n + \ldots + \lambda_r^n$ est proche de r. Cette somme, étant entière, doit être égale à r et ceci n'est possible que si chaque λ_i^n est égal à 1. □

Démonstration de la proposition 6.3.2.

Si (N, L) est une paire régulière, on peut munir aisément l'espace N_L d'une métrique vérifiant les deux propriétés énoncées à la fin du paragraphe 6.2, à savoir

i) pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que pour tout couple (x, y) de points de N_L vérifiant $d(x, y) < \eta$, il existe un chemin joignant x à y de diamètre inférieur à ε;

ii) il existe $\varepsilon_0 > 0$ tel que tout lacet de diamètre inférieur à ε_0 est contractile.

Remarquons que $\Omega(\overline{f}) = \{[L]\} \cup X$. Puisque $h(\overline{f}) = h(\overline{f}|_{\Omega(\overline{f})})$ (exercice 6 du chapitre 2) on sait que $h(\overline{f}) = 0$. On en déduit du théorème de Manning que les valeurs propres de $\overline{f}_{*,1}$ ont un module inférieur ou égal à 1. Grâce au lemme 6.3.4 et au fait que $H_1(N_L, \mathbf{R}) = H_1(N_L, \mathbf{Z}) \otimes \mathbf{R}$, on conclût que toutes les valeurs propres non nulles de $\overline{T}_{*,1}$ sont racines de l'unité. Par conséquent, la suite $(\mathrm{tr}(\overline{f}^n_{*,1}))_{n\geq 1}$ est périodique. Si q est une période commune à toutes ces valeurs propres et si n est un multiple de q, on a $\mathrm{tr}(\overline{f}^n_{*,1}) = \mathrm{tr}(\widetilde{f}^n_{*,1}) = -p$, où p est le nombre de valeurs propres non nulles.

Supposons maintenant que $m = 3$. On peut alors construire une paire filtrante régulière (N', L') pour f^{-1} qui vérifie également l'une des conclusions de la proposition 6.3.1. Notons alors $\overline{f^{-1}} : N'_{L'} \to N'_{L'}$ l'application définie sur l'espace quotient. Notons $M = N_L \vee N'_{L'}$ l'espace obtenu à partir de N_L et $N'_{L'}$ en identifiant les points $[L]$ et $[L']$ en un point $*$. Il s'agit d'un espace métrique compact vérifiant les propriétés énoncées à la fin du paragraphe 6.2. De plus $*$ est un point fixe attractif de l'application induite $\overline{g} = \overline{f} \vee \overline{f^{-1}} : M \to M$. La formule de Lefschetz-Dold donne

$$\Lambda(\overline{g}^n) = \mathrm{tr}(\overline{g}^n_{*,0}) - \mathrm{tr}(\overline{g}^n_{*,1}) + \mathrm{tr}(\overline{g}^n_{*,2}) - \mathrm{tr}(\overline{g}^n_{*,3})$$

où

$$\Lambda(\overline{g}^n) = i(\overline{g}^n, *) + i(f^n, X) + i(f^{-n}, X).$$

On sait que chaque espace $H_s(M, \mathbf{R})$, $1 \leq s \leq 3$, s'écrit $H_s(M, \mathbf{R}) = H_s(N_L, \mathbf{R}) \oplus H_s(N'_{L'}, \mathbf{R})$ et que chaque application $\overline{g}_{*,s}$ se décompose $\overline{g}_{*,s} = \overline{f}_{*,s} \oplus \overline{f^{-1}}_{*,s}$.

Supposons d'abord que f préserve l'orientation. On a alors $i(f^n, X) = -i(f^{-n}, X)$ ce qui implique que $\Lambda(\overline{g}^n) = 1$.

Si X n'est ni un attracteur, ni un répulseur, chaque couple (N, L) et (N', L') vérifie la condition iii) de la proposition 6.3.1: on sait que $\mathrm{tr}(\overline{g}^n_{*,0}) = 1$ et que les espaces $H_3(N_L, \mathbf{R})$ et $H_3(N'_{L'}, \mathbf{R})$ sont réduits à 0. On en déduit que

$$-\mathrm{tr}(\overline{g}^n_{*,1}) + \mathrm{tr}(\overline{g}^n_{*,2}) = 0,$$

pour tout $n \geq 0$. Le lemme 6.3.3 nous dit que les valeurs propres non nulles de $\overline{g}_{*,1}$ et de $\overline{g}_{*,2}$, coïncident. Le rayon spectral de $\overline{g}_{*,1}$ étant inférieur ou égal à 1 puisque $h(\overline{g}) = 0$, on a un résultat similaire pour $\overline{g}_{*,2}$ et donc pour $\overline{f}_{*,2}$. Il reste à utiliser le lemme 6.3.4 pour conclure.

Si X est un attracteur de f, c'est un répulseur de f^{-1}. On peut donc supposer que (N, L) vérifie la condition i) de la proposition 6.3.1 et (N', L') la condition ii).

$$\begin{aligned}
1 + i(f^n, X) &= 2 - \mathrm{tr}(\overline{f}^n_{*,1}) + \mathrm{tr}(\overline{f}^n_{*,2}), \\
1 + i(f^{-n}, X) &= 1 - \mathrm{tr}((\overline{f^{-1}}_{*,1})^n) + \mathrm{tr}((\overline{f^{-1}}_{*,2})^n) - 1.
\end{aligned}$$

et

$$1 = \Lambda(\overline{g}^n) = 2 - \mathrm{tr}(\overline{g}^n_{*,1}) + \mathrm{tr}(\overline{g}^n_{*,2}) - 1$$

On en déduit que

$$-\mathrm{tr}(\overline{g}^n_{*,1}) + \mathrm{tr}(\overline{g}^n_{*,2}) = 0.$$

On conclût comme dans le premier cas. Le cas où X est un répulseur se traite de même.

Dans le cas où f renverse l'orientation, on obtient

$$-\mathrm{tr}(\overline{g}^n_{*,1}) + \mathrm{tr}(\overline{g}^n_{*,2}) = 0,$$

pour tout nombre pair n. Or ceci est suffisant pour prouver que les valeurs propres de $\overline{g}_{*,1}$ sont, au signe près, celles de $\overline{g}_{*,2}$. $\square$

Corollaire 6.3.6 : *Soit x un point fixe d'un homéomorphisme local $f : U \to f(U)$ de $\mathbf{R}^r$, $r \leq 3$, qui est un ensemble invariant isolé. Alors la suite $i(f^n, U)_{n\geq 0}$ est périodique. De plus, si $r \leq 2$, et si le point fixe x n'est ni attractif, ni répulsif, la suite prend périodiquement une valeur négative ou nulle.*

Démonstration. Le théorème est évident dans le cas où $r = 1$. De plus, la première partie est également évidente si x est un point fixe attractif ou répulsif. Dans le cas où $r = 2$ et où x n'est ni attractif, ni répulsif, nous trouvons

$$i(f^n, x) = -\mathrm{tr}(\overline{f}^n_{*,1}),$$

pour tout $n \geq 0$. D'après la proposition 6.3.3, nous savons qu'elle est périodique et prend périodiquement une valeur négative ou nulle. Dans le cas où $r = 3$ et où x n'est ni attractif, ni répulsif, nous trouvons

$$i(f^n, x) = -\mathrm{tr}(\overline{f}^n_{*,1}) + \mathrm{tr}(\overline{f}^n_{*,2})$$

et nous savons, toujours grâce à la proposition 6.3.2 que cette suite est périodique. □

Corollaire 6.3.7 : *Il n'existe aucun homéomorphisme minimal sur $S^2 \setminus X$, où X est une partie finie de S^2.*

Démonstration. Raisonnons par l'absurde et supposons qu'un tel homéomorphisme f existe. Nous pouvons étendre f en un homéomorphisme $\widehat{f}$ de la sphère S^2 qui induit une permutation sur X. Nous pouvons trouver un itéré pair $\widehat{f}^{2r}$ qui fixe tous les points de X. Puisque f est minimal, les seules parties fermées invariantes de $\widehat{f}$, hors la sphère S^2, sont incluses dans X. On en déduit, d'une part que les seuls points périodiques de $\widehat{f}^{2r}$ sont les points de X, d'autre part que chacun de ces points est un ensemble invariant isolé. De plus aucun de ces points est un attracteur ou un répulseur. Le corollaire 6.3.6 implique qu'il existe $q \geq 1$ tel que $i(\widehat{f}^{2rq}, x) \leq 0$, pour tout $x \in X$, et donc que

$$\Lambda(\widehat{f}^{2rq}) = \sum_{x\in X} i(\widehat{f}^{2rq}, x) \leq 0.$$

Or on sait que $\Lambda(\widehat{f}^{2rq}) = 2$, car $\widehat{f}^{2rq}$ préserve l'orientation, d'où la contradiction. □

Corollaire 6.3.8 : *Il n'existe pas d'homéomorphisme expansif de S^1.*

Démonstration. Raisonnons par l'absurde et supposons qu'un tel homéomorphisme f existe. Quitte à remplacer f par f^2, on peut supposer que f préserve l'orientation. Le fait que f est expansif signifie que la diagonale

$$\Delta = \{(x, y) \in S^1 \times S^1 \,|\, x = y\}$$

est un ensemble invariant isolé de l'application produit $f \times f$. La surface $S^1 \times S^1 \setminus \Delta$ a deux bouts N et S et se compactifie en une sphère S^2. La restriction de $f \times f$ à $S^1 \times S^1 \setminus \Delta$ se prolonge en un homéomorphisme F de S^2 préservant l'orientation et fixant N et S. Les dynamiques au voisinage de N et au voisinage de S sont conjuguées par l'application $(x, y) \mapsto (y, x)$. Chacun de ces deux points fixes est un ensemble invariant isolé puisque c'est le cas de Δ. Il est facile de voir que ces points ne sont ni attractifs, ni des répulsifs (ceci signifierait que Δ est un attracteur ou un répulseur de $f \times f$ ou encore que f est positivement ou négativement expansif, ce qui est impossible d'après l'exercice 15 du chapitre 2). Le corollaire 6.3.7 implique qu'il existe $q \geq 1$ tel que $i(F^q, N) \leq 0$ et $i(F^q, S) \leq 0$. On sait qu'un homéomorphisme expansif a un nombre fini de point fixes de période donnée (exercice 16 du chapitre 2). En utilisant la formule de Lefschetz, on peut donc écrire

$$\sum_{(x,y)\in \mathrm{Fix}(F^q)\setminus\{N,S\}} i(F^q, (x, y)) = 2 - i(F^q, N) - i(F^q, S) \geq 2.$$

Or, d'une part, on a

$$\sum_{(x,y)\in \mathrm{Fix}(f^q\times f^q)} i(f^q\times f^q, (x, y)) = \sum_{(x,y)\in \mathrm{Fix}(f^q\times f^q)} i(f^q, x)i(f^q, y) = \left(\sum_{x\in \mathrm{Fix}(T^q)} i(f^q, x)\right)^2 = \Lambda(f^q)^2 = 0,$$

d'autre part, on a

$$\sum_{(x,y)\in \mathrm{Fix}(f^q\times f^q)} i(f^q \times f^q, (x, y)) = \sum_{(x,y)\in \mathrm{Fix}(F^q)\setminus\{N,S\}} i(F^q, (x, y)) + \sum_{x\in \mathrm{Fix}(f^q)} i(f^q, x)^2 \geq 2,$$

d'où la contradiction. □

Bien évidemment, le corollaire 6.3.8 donne un résultat bien connu. En effet, nous savons décrire très précisément les homéomorphismes de S^1 et pouvons montrer directement, sans guère de difficulté, qu'aucun d'entre-eux n'est expansif. La preuve précédente permet cependant de mettre en relief le fait suivant : un homéomorphisme d'une variété compacte est expansif si et seulement si la diagonale est une partie invariante isolée de l'application produit définie sur le carré de la variété. Notons à ce propos que Leiwowicz et Hiraide ont indépendamment décrit tous les homéomorphismes des surfaces compactes. Il n'y en a pas sur la sphère S^2; tout homéomorphisme expansif du tore $\mathbf{T}^2$ est conjugué à un automorphisme hyperbolique ; tout homéomorphisme d'une surface de genre $g \geq 2$ est homéomorphisme à un modèle introduit par Thurston, appelé *homéomorphisme pseudo-Anosov*. Le corollaire 6.3.7 est évident si $X = \emptyset$, d'après la formule de Lefschetz, et était connu depuis longtemps si $\sharp X = 1$, conséquence d'un résultat de Brouwer que l'on a mentionné dans l'étude de l'application de Hénon au paragraphe 5.3. Le cas $\sharp X \geq 3$ était dû à Handel qui avait en fait prouvé le résultat suivant : tout homéomorphisme transitif de $S^2 \setminus \sharp X$ a des orbites périodiques de périodes arbitrairement grandes si $\sharp X \geq 3$. Il manquait le cas où $\sharp X = 2$. L'argument d'indice donné plus haut est dû à Le Calvez-Yoccoz. En fait, on peut affiner la seconde partie du corollaire 6.3.6 : si x un point fixe d'un homéomorphisme local $f : U \to f(U)$ de $\mathbf{R}^2$ qui préserve l'orientation et si ce point fixe est un ensemble invariant isolé, qui n'est ni attractif, ni répulsif, alors il existe deux entiers $q \geq 1$ et $r \geq 1$ tel que

$$i(f^n, x) = \begin{cases} 1 & \text{si } n \notin q\mathbf{N} \\ 1 - rq & \text{si } n \in q\mathbf{N} \end{cases}.$$

L'idée d'utiliser l'indice de Conley discret, c'est-à-dire la construction de paires filtrantes pour démontrer la seconde partie du corollaire 6.3.6 et par conséquent du corollaire 6.3.7 est due à Franks. Un résultat important dans l'étude de Franks est le fait que la famille des valeurs propres non nulles des $\widetilde{f}_{*,s}$, et donc les suites $(\mathrm{tr}(\widetilde{f}^n_{*,s})_{n\geq 1}$ ne dépendent pas de la paire filtrante considérée. L'idée d'utiliser l'entropie et le théorème de Manning, c'est-à-dire la proposition 6.3.4, est due à Le Calvez, Ruiz del Portal et Salazar. Il faut noter que la conclusion du corollaire 6.3.6 est fausse si $m \geq 4$. On peut construire un exemple de point fixe d'un homéomorphisme local $f : U \to f(U)$ de $\mathbf{R}^4$, qui est un ensemble invariant isolé, et pour lequel $\widetilde{f}_{*,2}$ a une valeur propre de module strictement plus grand que 1 : les suites $(\mathrm{tr}(\widetilde{f}^n_{*,2}))_{n\geq 1}$ et $i(f^n, x)_{n\geq 1}$

tendent alors vers $+\infty$ à vitesse exponentielle. Pour conclure, le raisonnement aboutissant à la proposition 6.3.2, avec l'astuce d'agglomérer les espaces N_L et $N'_{L'}$ est un peu mystérieux. Il y a une explication plus rationnelle, provenant d'un résultat de dualité dû à Szymczak: si X est une partie compacte connexe invariante isolée d'un homéomorphisme $f : U \to V$ entre deux ouverts de $\mathbf{R}^r$, si (N, L) et (N', L') sont deux paires filtrantes régulières de X respectivement pour f et f^{-1}, alors l'application duale de $\widetilde{f}_{*,s}$ est conjuguée à $\deg(f)\widetilde{f^{-1}}_{*,r-s}$, à partie nilpotente près.

Chapitre 1

Exercice 1

Soit $(X, \mathcal{B}, \mu)$ un espace de probabilités. On dira que deux partitions $\mathcal{P}$ et $\mathcal{Q}$ sont *indépendantes* si, pour tout $P \in \mathcal{P}$ et tout $Q \in \mathcal{Q}$, on a $\mu(P \cap Q) = \mu(P)\mu(Q)$. Montrer que les propriétés suivantes sont équivalentes :

- $\mathcal{P}$ et $\mathcal{Q}$ sont indépendantes ;
- $H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q})$;
- $H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P})$.

Exercice 2

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, où $\mu(X) = 1$. Prouver l'*inégalité de Rokhlin* : pour toutes partitions mesurables $\mathcal{P}$ et $\mathcal{Q}$, on a

$$|h(T, \mathcal{P}) - h(T, \mathcal{Q})| \leq D(\mathcal{P}, \mathcal{Q}).$$

Exercice 3

Soit $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$. On considère le produit $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \otimes \nu, T \times S)$. Montrer que

$$h_{\mu\otimes\nu}(T \times S) = h_\mu(T) + h_\nu(S).$$

Exercice 4

Soit $(X, \mathcal{B})$ un ensemble muni d'une σ-algèbre. On suppose que $T \ : X \to X$ est une transformation mesurable et que μ et ν sont deux mesures de probabilité invariantes.

Montrer que pour toute partition mesurable $\mathcal{P} = (P_i)_{\in I}$, on a :

$$tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) \leq H_{t\mu+(1-t)\nu}(\mathcal{P}) \leq tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) - t\ln t - (1-t)\ln(1-t).$$

En déduire que

$$H_{t\mu+(1-t)\nu}(T, \mathcal{P}) = tH_\mu(T, \mathcal{P}) + (1-t)H_\nu(T, \mathcal{P}),$$

puis que

$$h_{t\mu+(1-t)\nu}(T) = th_\mu(T) + (1-t)h_\nu(T).$$

Exercice 5

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, avec $\mu(X) = 1$. Soit $\mathcal{P}$ une partition mesurable telle que $\mathcal{P} \in \overline{\bigvee_{m=1}^{+\infty} T^{-m}\mathcal{P}}$. Montrer que $h(T, \mathcal{P}) = 0$ (en fait la réciproque est vraie).

Exercice 6

Soit $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$. On suppose que $h_\mu(T) = h_\nu(S)$. Cela implique t-il les systèmes sont conjugués ?

Exercice 7

On note σ le décalage de Bernouilli défini sur $\{1, \dots, r\}^{\mathbf{N}}$. Monter que pour toute mesure borélienne de probabilité invariante par σ, on a $h_\mu(\sigma) \leq \ln r$. Montrer que $h_\mu(\sigma) = \ln r$ si et seulement si μ est la mesure produit équidistribuée.

Exercice 8

On fixe $r \geq 2$ et on considère la transformation

$$\begin{aligned} T \ : \mathbf{T}^1 &\to \mathbf{T}^1 \\ x &\mapsto rx \end{aligned}.$$

Monter que pour toute mesure borélienne de probabilité invariante par T, on a $h_\mu(T) \leq \ln r$. Montrer que $h_\mu(T) = \ln r$ si et seulement si μ est la mesure de Haar.

Chapitre 2

Exercice 1

Montrer que l'entropie topologique d'un homéomorphisme de $[0, 1]$ est nulle.

Exercice 2

Donner un exemple d'homéomorphisme $T \ : X \to X$ d'un espace compact dont l'entropie topologique est infinie.

Exercice 3

Soient $T_1 \ : X_1 \to X_1$ et $T_2 \ : X_2 \to X_2$ deux applications définies sur des espaces métriques compacts. Montrer que l'entropie de l'application produit

$$\begin{aligned} T_1 \times T_2 \ : X_1 \times X_2 &\to X_1 \times X_2 \\ (x_1, x_2) &\mapsto (T_1(x_1), T_2(x_2)) \end{aligned}$$

vérifie $h(T_1 \times T_2) = h(T_1) + h(T_2)$.

Exercice 4

Quelle est l'entropie de $T \ : z \mapsto z^2$ définie sur la sphère de Riemann.

Exercice 5

Soit $T \ : X \to X$ une application continue définie sur un espace métrique compact tel que $\Omega(T)$ est fini. Montrer que $h(T) = 0$.

Exercice 6

Soit $T \ : X \to X$ une application continue définie sur un espace métrique compact. On rappelle que x est non errant si pour tout voisinage U de x, il existe $y \in U$ et $n \geq 0$ tel que $T^n(y) \in U$.

1) Montrer que l'ensemble $\Omega(T)$ des points non errants est fermé et vérifie $T(\Omega(T)) \subset \Omega(T)$.

2) À l'aide du principe variationnel, montrer que $h(T) = h(T|_{\Omega(T)})$.

Exercice 7

Soit $T : X \to X$ une application continue définie sur un espace métrique compact. On suppose qu'il existe une famile finie $(X_i)_{1\le i\le p}$ de parties fermées positivement invariantes (i.e. $T(X_i) \subset X_i$), telles que $X = \bigcup_{1\le i\le p} X_i$. Montrer que $h(T) = \sup_{1\le i\le p} h(T|_{X_i})$.

Exercice 8

On se donne une application continue sur un espace métrique compact $T : X \to X$.

1) On fixe dans cette question μ, ν dans $\mathcal{M}_T$ et $t \in [0,1]$. Montrer que pour toute partition borélienne $\mathcal{P} = (P_i)_{\in I}$, on a :

$$tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) \le H_{t\mu+(1-t)\nu}(\mathcal{P}) \le tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) - t\ln t - (1-t)\ln(1-t).$$

En déduire que

$$H_{t\mu+(1-t)\nu}(T,\mathcal{P}) = tH_\mu(T,\mathcal{P}) + (1-t)H_\nu(T,\mathcal{P}),$$

puis que

$$h_{t\mu+(1-t)\nu}(T) = th_\mu(T) + (1-t)h_\nu(T).$$

2) On suppose qu'il existe une unique mesure $\mu \in \mathcal{M}_T$ telle que $h_\mu(T) = h(T)$. Montrer que μ est ergodique.

3) On suppose que $h(T) = +\infty$. Montrer qu'il existe $\mu \in \mathcal{M}_T$ telle que $h_\mu(T) = h(T)$.

4) Donner un exemple où $h(T) = +\infty$ et où il n'existe aucune mesure ergodique μ telle que $h_\mu(T) = h(T)$.

Exercice 9

Soit $\alpha \in \mathbf{T}^1$. Calculer l'entropie topologique de

$$\begin{aligned} F \ : \mathbf{T}^2 &\to \mathbf{T}^2 \\ (x,y) &\mapsto (x+y, y+\alpha) \end{aligned}.$$

Exercice 10

Soit $T : X \to X$ une application lispshitzienne sur un espace métrique compact. L'entropie peut elle être infinie ? et dans le cas où X est une variété différentiable ?

Exercice 11

Soit $T : X \to X$ une application continue définie sur un espace métrique compact. Rappelons que T est (positivement) expansive s'il existe $\delta > 0$ tels que pour tous x et y, il existe $n \ge 0$ tel que $d(T^n(x), T^n(y)) \ge \delta$.

1) Donner des exemples d'applications qui sont expansives et d'applications qui ne le sont pas.

2) Montrer que si X est compact, la propriété d'expansivité est topologique (elle ne dépend pas de la métrique mais de la topologie).

Exercice 12

Soit $T\;: X \to X$ une application continue définie sur un espace métrique compact. Montrer l'équivalence des propriétés suivantes :

- T est expansive;
- il existe un recouvrement générateur;
- si $\varepsilon > 0$ est assez petit, le recouvrement $\mathcal{U}^{\varepsilon}$ par boules de rayon ε est générateur.

Exercice 13

L'entropie d'une application continue expansive $T\;: X \to X$ sur un espace métrique compact peut-elle être infinie ?

Exercice 14

Soit $T\;: X \to X$ une application continue expansive $T\;: X \to X$ sur un espace métrique compact.

1) Montrer que pour tout $n \geq 1$, l'ensemble $\mathrm{Fix}(T^n)$ est fini.

2) Montrer que $h(T) \geq \limsup_{n \to =\infty} \frac{1}{n} \ln(\sharp \mathrm{Fix}(T^n))$.

Exercice 15

Soit $T\;: X \to X$ un homéomorphisme d'un espace métrique compact. Montrer que si T est (positivement) expansif, alors X est fini.

Exercice 16

Soit $T\;: X \to X$ un homéomorphisme défini sur un espace métrique compact. Rappelons que T est expansif s'il existe $\delta > 0$ tels que pour tous x et y, il existe $k \in \mathbf{Z}$ tel que $d(T^k(x), T^k(y)) \geq \delta$.

1) Donner des exemples d'applications qui sont expansives et d'applications qui ne le sont pas.

2) Montrer que T est expansif si et seulement si la suite $\left(\bigvee_{i=0}^{n-1} T^i(\mathcal{U}_\varepsilon)\right)_{n \geq 0}$ est génératrice si $\varepsilon > 0$ est assez petit, et que $h(T) < +\infty$.

3) Là-encore, montrer que pour tout $n \geq 1$, l'ensemble $\mathrm{Fix}(T^n)$ est fini et que $h(T) \geq \limsup_{n \to =\infty} \frac{1}{n} \ln(\sharp \mathrm{Fix}(T^n))$.

Exercice 17

Montrer qu'il n'y a pas d'homéomorphisme expansif sur $\mathbf{T}^1$.

Exercice 18

Prouver que la mesure de Lebesgue est la seule mesure borélienne de probabilité invariante par $T\;: \widehat{x} \mapsto p\widehat{x}$ sur $\mathbf{T}^1$, telle que $h_\mu(T) = h(T)$, si $p \geq 2$.

Chapitre 3

Exercice 1

Considérons la matrice

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

et le sous-décalage de type fini σ_A défini sur $X_A \subset \{1, 2, 3\}^{\mathbf{Z}}$.

1) Calculer le nombre de points périodiques de période q, pour $q \leq 3$. Quel est le nombre de points fixes de σ_A^n, pour $n \geq 1$?

2) Quelle est l'entropie topologique de σ_A ?

3) Peut-on trouver une partie fermée invariante $X \subset X_A$ telle que la restriction $\sigma_A|_X$ soit conjuguée au décalage de Bernouilli sur $\{1, 2\}^{\mathbf{Z}}$?

4) On considère la matrice stochastique suivante

$$M = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Expliquer pourquoi il existe sur $\{1, 2, 3\}^{\mathbf{Z}}$ une unique mesure de Markov associée μ et calculer $\mu(C)$, où

$$C = \{(x_i)_{i \in \mathbf{Z}} \in \{1, 2, 3\}^{\mathbf{Z}} \,|\, x_0 = 1, x_1 = 1, x_2 = 2\}.$$

La mesure μ est-elle supportée sur X_A ? que vaut l'entropie $h_\mu(\sigma_A)$?

5) Soit ν la mesure de Parry de σ_A. Calculer $\nu(C)$ et donner la valeur de l'entropie $h_\nu(\sigma_A)$.

Chapitre 4

Exercice 1

Montrer qu'un automorphisme linéaire du tore $\mathbf{T}^r$ est expansif si et seulement s'il est hyperbolique.

Exercice 2

1) Soit A une matrice carrée d'ordre 2 à coefficients entiers de déterminant 1. Montrer que les conditions suivantes sont équivalentes :

- $\mathrm{Tr}(A)| > 2$;
- A est hyperbolique ;
- $\widehat{A}$ est mélangeante (pour la mesure de Haar).

2) Soit A une matrice carrée d'ordre 2 à coefficients entiers de déterminant -1. Montrer que les conditions suivantes sont équivalentes :

- $|\mathrm{Tr}(A) \neq 0$;
- A est hyperbolique ;
- $\widehat{A}$ est mélangeante (pour la mesure de Haar).

3) Trouver une matrice carrée d'ordre 4 à coefficients entiers de déterminant 1 telle que

- A n'est pas hyperbolique ;

- $\widehat{A}$ est mélangeante (pour la mesure de Haar).

(On commencera par montrer que les racines de $P(X) = X^4 + 2X^3 + X^2 + 2X + 1$ ne sont pas racines de l'unité).

Exercice 3

Donner un argument dynamique au fait que les valeurs propres d'une matrice carrée hyperbolique d'ordre 2 à coefficients entiers de déterminant 1 sont irrationnelles.

Exercice 4

Soit $F\ : \mathbf{T}^r \to \mathbf{T}^r$ un homéomorphisme tel que F_* soit un automorphisme hyperbolique de $\mathbf{R}^r$. On va donner un autre preuve du fait qu'il existe une unique application continue $H\ : \mathbf{T}^r \to \mathbf{T}^r$ homotope à l'identité telle que $H \circ F = \widehat{F}_* \circ H$.

1) Soit f un relèvement de F. Montrer qu'il existe $M > 0$ tel que pour tout $x \in \mathbf{R}^r$, il existe un unique point $y = h(x) \in \mathbf{R}^r$ tel que la suite $(f^k(x) - F_*^k(y))_{k\in\mathbf{Z}}$ est bornée et que de de plus on a $\|f^k(x) - F_*^k(y)\| \leq M$ pour tout $k \in \mathbf{Z}$.

2) Montrer que $h \circ f = F_* \circ h$ et que $h - \mathrm{Id}_{\mathbf{R}^r}$ est invariante par les translations entières.

3) Montrer que h est continue.

4) Conclure

Exercice 5

Soit $\widehat{A}$ un automorphisme hyperbolique de $\mathbf{T}^r$. Prouver le lemme de fermeture suivant : pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que toute δ-pseudo-orbite $(x_k)_{k\in\mathbf{Z}}$ de période q peut être ε-pistée par une orbite périodique $(y_k)_{k\in\mathbf{Z}}$ de période q.

Exercice 6

On notera $\mathrm{GL}(2,\mathbf{Z})$ (resp. $\mathrm{SL}(2,\mathbf{Z})$) le groupe des matrices carrées d'ordre 2 à coefficients entiers de déterminant ± 1 (resp. 1). Par abus de langage on identifiera une matrice de $\mathrm{GL}(2,\mathbf{Z})$ à l'automorphisme linéaire $A\ : \mathbf{R}^2 \to \mathbf{R}^2$ représenté par cette matrice dans la base canonique ; on notera alors $\widehat{A}$ l'automorphisme de $\mathbf{T}^2 = \mathbf{R}^2/\mathbf{Z}^2$ naturellement défini par A. On écrira $\mathrm{GL}_{\mathrm{hyp}}(2,\mathbf{Z})$ pour l'ensemble des automorphismes $A \in \mathrm{GL}(2,\mathbf{Z})$ qui sont hyperboliques et on posera $\mathrm{SL}_{\mathrm{hyp}}(2,\mathbf{Z}) = \mathrm{GL}_{\mathrm{hyp}}(2,\mathbf{Z}) \cap \mathrm{SL}(2,\mathbf{Z})$.

On notera $\mathrm{Homeo}(\mathbf{T}^2)$ le groupe des homéomorphismes de $\mathbf{T}^2$. On écrira alors $\mathrm{Homeo}_+(\mathbf{T}^2)$ pour le sous-groupe formé des homéomorphismes de degré 1 et $\mathrm{Homeo}_0(\mathbf{T}^2)$ pour le sous-groupe formé des homéomorphismes homotopes à l'identité.

Si G est un groupe, on dira que $g' \in G$ est une *racine* de $g \in G$, s'il existe un entier $n \geq 1$ tel que $(g')^n = g$. On rappelle que le *centralisateur* d'un élément $g \in G$ est le sous-groupe $\mathrm{Cent}_G(g)$ des éléments $g' \in G$ tels que $gg' = g'g$.

1.a) On définit l'ensemble Λ_0 (resp. Λ_1) des valeurs propres d'automorphismes $A \in \mathrm{GL}_{\mathrm{hyp}}(2,\mathbf{Z})$ (resp. $A \in \mathrm{SL}_{\mathrm{hyp}}(2,\mathbf{Z})$). Montrer que les ensembles $\Lambda_0 \cup \{-1, 1\}$ et $\Lambda_1 \cup \{-1, 1\}$ sont des parties discrètes de $\mathbf{R}$ puis calculer

$$\lambda_0 = \min(\Lambda_0 \cap]1, +\infty[), \ \ \lambda_1 = \min(\Lambda_1 \cap]1, +\infty[).$$

1.b) En déduire que le centralisateur d'un automorphisme $A \in \mathrm{GL}_{\mathrm{hyp}}(2,\mathbf{Z})$ (resp. $A \in \mathrm{SL}_{\mathrm{hyp}}(2,\mathbf{Z})$) dans $\mathrm{GL}(2,\mathbf{Z})$ (resp. $\mathrm{SL}(2,\mathbf{Z})$) est isomorphe à $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}$.

1.c) On considère les automorphismes linéaires

$$A_0 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Déterminer deux éléments qui engendrent $\mathrm{Cent}_{\mathrm{GL}(2,\mathbf{Z})}(A_0)$. Même question pour $\mathrm{Cent}_{\mathrm{GL}(2,\mathbf{Z})}(A_1)$ et $\mathrm{Cent}_{\mathrm{SL}(2,\mathbf{Z})}(A_1)$.

2.a) On note $\widehat{A}_0$ et $\widehat{A}_1$ les automorphismes de $\mathbf{T}^2$ définis respectivement par A_0 et A_1. Montrer que $\widehat{A}_0$ a un unique point fixe. En déduire que tout élément $G \in \mathrm{Cent}_{\mathrm{Homeo}(\mathbf{T}^2)}(\widehat{A}_1) \cap \mathrm{Homeo}_0(\mathbf{T}^2)$ a un relèvement à $\mathbf{R}^2$ qui commute avec A_0. Montrer que $\mathrm{Cent}_{\mathrm{Homeo}(\mathbf{T}^2)}(\widehat{A}_0) \cap \mathrm{Homeo}_0(\mathbf{T}^2)$ se réduit à l'identité. Prouver un résultat analogue pour $\mathrm{Cent}_{\mathrm{Homeo}(\mathbf{T}^2)}(\widehat{A}_1) \cap \mathrm{Homeo}_0(\mathbf{T}^2)$

2.b) En déduire ce que sont les groupes $\mathrm{Cent}_{Homeo(\mathbf{T}^2)}(\widehat{A}_0)$, $\mathrm{Cent}_{\mathrm{Homeo}(\mathbf{T}^2)}(\widehat{A}_1)$ et $\mathrm{Cent}_{\mathrm{Homeo}(\mathbf{T}^2)}(\widehat{A}_1)$?

2.c) Quelles sont les racines de $\widehat{A}_0$ dans $\mathrm{Homeo}(\mathbf{T}^2)$? Quelles sont les racines de $\widehat{A}_1$ dans $\mathrm{Homeo}(\mathbf{T}^2)$? et dans $\mathrm{Homeo}_+(\mathbf{T}^2)$?

Chapitre 5

Exercice 1

Soit (X, d) un espace métrique compact et $T : X \to X$ une application continue (positivement) expansive. On suppose qu'il existe un sous-décalage de type fini σ_A irréductible apériodique défini sur $X_A \subset \{1, \dots, p\}^{\mathbf{N}}$, une application continue surjective $H : X_A \to X$ telle que $T \circ H = H \circ \sigma_A$ et un entier M tel que, tout élément x de X a au plus M antécédents par H. Montrer que l'entropie topologique de T est égale à l'entropie topologique de σ_A.

Exercice 2

Pour tout $\alpha \in]0, 1[$, on définit $g_\alpha : [0, 1] \to [0, 1]$ par

$$g_\alpha(x) = \begin{cases} \alpha x + \gamma & \text{si } x \in [0, \gamma], \\ \beta(1 - x) & \text{si } x \in [\gamma, 1], \end{cases}$$

où $\beta = 1 + \dfrac{1}{\alpha}$ et $\gamma = \dfrac{1}{1 + \alpha}$.

1) Dessiner le graphe de g_α et expliquer pourquoi $\{[0, \gamma], [\gamma, 1]\}$ est une partition de Markov qui définit une semi-conjugaison $h_\alpha : \{0, 1\}^{\mathbf{N}} \to [0, 1]$ d'un sous-décalage de type fini sur g_α.

2) Calculer l'entropie de g_a.

Exercice 3

On définit $f : [0, 1] \to [0, 1]$ par

$$f(x) = \begin{cases} 2x + \frac{4}{3}x & \text{si } x \in [0, \frac{1}{4}], \\ \frac{4}{3} - \frac{4}{3}x & \text{si } x \in [\frac{1}{4}, 1]. \end{cases}$$

1) Dessiner le graphe de f et prouver que pour tout $n \geq 1$ il existe $a_n \in \mathbf{N}$ non multiple de 3 tel que $f^n(0) = \dfrac{a_n}{3^n}$.

2) Peut-on construire une partition de Markov (finie) ?

Exercice 4
1) Montrer que

$$h : x \mapsto \sin^2\left(\frac{\pi x}{2}\right)$$

définit une conjugaison entre $f\ : [0,1] \to [0,1]$ et $g\ : [0,1] \to [0,1]$ où

$$f(x) = \begin{cases} 2x & \text{si } x \in [0,1/2], \\ 1-2x & \text{si } x \in [1/2,1]. \end{cases}$$

et

$$g(x) = 4x(1-x).$$

2) Constuire une partition de Markov pour g.

3) Soit $\varphi\ : [0,1] \to \mathbf{R}$ Expliquer pourquoi il existe $c \in \mathbf{R}$ tel que presque tout point x (pour la mesure de Lebesgue) vérifie

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(g^i(x)) = c.$$

Calculer ce nombre et prouver que

$$\lim_{n\to+\infty} \frac{1}{\sharp\mathrm{Fix}(g^n)} \sum_{x\in\mathrm{Fix}(g^n)} \varphi(x) = c.$$

Chapitre 6

Exercice 1
Soit f un homéomorphisme de $\mathbf{T}^2$. Montrer que $h(f) \geq \sup_{0\leq s\leq 2} \rho(f_{*,s})$.

Exercice 2
Soit $f : I \to J$ un homéomorphisme entre deux intervalles ouverts de $\mathbf{R}$ ayant un point fixe x. Montrer que si f renverse l'orientation, alors x est un point fixe isolé et $i(f,x) = 1$. Montrer que si f renverse l'orientation, alors x est un point fixe isolé si et seulement si c'est un ensemble invariant isolé et que $i(f,x) \in \{-1,0,1\}$.

Exercice 3
Vérifier la formule de Lefschetz pour les itérés de $\widehat{A}$, où $\widehat{A}$ est l'automorphisme de $\mathbf{T}^2$ associé à l'automorphisme linéaire de $\mathbf{R}^2$ dont la matrice dans la base canonique est

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Exercice 4

Soit X une partie compacte connexe invariante attractive (resp. répulsive) d'un homéomorphisme $f : U \to V$ entre deux parties ouvertes de $\mathbf{R}^r$. Montrer qu'il existe une paire filtrante régulière (N, L) qui vérifie la propriété i) (resp. ii)) de la proposition 6.3.1.

Exercice 5

En gardant les notations du paragraphe 5.2 consacré au fer à cheval, construire une paire filtrante (N, L) naturelle et étudier les endomorphismes $\overline{f}_{*,s}$ associés (définies au paragraphe 6.3).