Systèmes dynamiques : cours fondamental II

Partie I : Entropie topologique et systèmes hyperboliques

§1. Entropie topologique

Nous allons associer à toute application continue $T : X \to X$ définie sur un espace métrique compact 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.

1.1 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 $\mathcal{V} = (V_j)_{i\in I}$ 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 $\mathcal{V} \preceq \mathcal{U}$. 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 $(\mathcal{U}^j)_{1\leq j\leq n}$ 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 $\bigcap_{1\leq j\leq n} U^j$ où $U^j \in \mathcal{U}^j$. Il est plus fin que tous les $\mathcal{U}^j$. Remarquons que si $(\mathcal{V}^j)_{1\leq j\leq n}$ est une famille finie de recouvrements ouverts de X et si $\mathcal{V}^j \preceq \mathcal{U}^j$ pour tout $j \in \{1, \dots n\}$, alors $\bigvee_{j=1}^n \mathcal{V}^j \preceq \bigvee_{j=1}^n \mathcal{U}^j$.

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 :

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

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 $X \in \mathcal{U}$;
- $\mathcal{V} \preceq \mathcal{U} \Longrightarrow N(\mathcal{V}) \geq N(\mathcal{U})$;
- $N\left(\bigvee_{j=1}^n \mathcal{U}^j\right) \leq \prod_{j=1}^n N(\mathcal{U}^j)$.

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

- $N(T^{-1}(\mathcal{U})) \leq N(\mathcal{U})$;
- $N(T^{-1}(\mathcal{U})) = N(\mathcal{U})$ 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 1.1.1 : *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 relative par rapport 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,

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

La suite $(v_n)_{n\geq 1}$, où $v_n = \ln(u_n)$ est donc sous-additive et positive. Ceci implique que la suite $(\frac{v_n}{n})_{n\geq 1}$ converge et que

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

En effet si $n \geq p \geq 1$ sont deux entiers et si $n = kp + r$ est la division euclidienne de n par p on a

$$\frac{v_n}{n} \leq \frac{v_{kp} + v_r}{kp + r} \leq \frac{v_{kp}}{kp} + \frac{v_r}{kp + r} \leq \frac{v_p}{p} + \frac{v_r}{n}.$$

On en déduit

$$\limsup_{n\to+\infty} \frac{v_n}{n} \leq \frac{v_p}{p} \quad (\leq v_1)$$

puis

$$\limsup_{n\to+\infty} \frac{v_n}{n} \leq \inf_{p\geq 1} \frac{v_p}{p} \leq \liminf_{n\to+\infty} \frac{v_n}{n}.$$

□

Énonçons maintenant les propriétés principale de cette entropie relative.

Proposition 1.1.2 : *On a les propriétés suivantes :*

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

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

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

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

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

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

□

1.2 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 1.2.1 : *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 $\mathcal{U}$ of 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 $h(T|_Y) \leq h(T)$. □

1.3 Recouvrement générateur

Il est bien évidemment difficile de calculer l'entropie topologique à partir de la définition précédente. Nous verrons dans ce paragraphe d'autres façons de calculer l'entropie. 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 $\mathcal{U}_\alpha \preceq \mathcal{U}$. Nous avons bien évidemment :

Proposition 1.3.1 : *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 d'ailleurs que la fonction $\varepsilon \mapsto h(T,\mathcal{U}^\varepsilon)$ est croissante 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 $\mathcal{U}$:

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

où

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

Supposons maintenant que $\mathcal{U}_n$ est une suite décroissante de recouvrements ouverts ($\mathcal{U}_{n+1} \preceq \mathcal{U}_n$). Cette suite est génératrice si et seulement si $\lim_{n\to+\infty} \mathrm{diam}(\mathcal{U}_n) = 0$. Dans ce cas, on a

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

Il arrive qu'une telle suite génératrice peut être définie à partir d'un seul recouvrement. 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 1.3.2 : *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. Pour tout $n \geq 0$ définissons la distance suivante d_n sur X :

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

et notons $B_n(x,\varepsilon)$ la boule de rayon ε et de centre x. Soient $\varepsilon > 0$ et $n \geq 0$. 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 $N(n,\varepsilon) = N\left(\bigvee_{i=0}^{n} T^{-i}(\mathcal{U}^\varepsilon)\right)$. Ces nombres sont naturellement liés, comme l'exprime le résultat suivant :

Proposition 1.3.3 : *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 1.3.4 : *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. Cependant, la proposition précédente permet d'obtenir les quatre autres égalités.

□

Corollaire 1.3.5 : *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 in X, alors $h(T) = 0$.*

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

1.4. 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 en p cylindres $\mathcal{U} = (U_a)_{a\in A}$, avec $U_a = \{x \in A^{\mathbf{N}} \mid x_0 = a\}$

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 de diamètre qui tend 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 en p cylindres $\mathcal{U} = (U_a)_{a \in A}$, avec $U_a = \{x \in A^{\mathbf{Z}} \,|\, x_0 = a\}$, on remarque 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 $T^1 = \mathbf{R}/\mathbf{Z}$, 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 $p \geq 2$. On sait déja que $h(T) \leq \ln p$ puisque T est un facteur du décalage unilatéral sur $\{1, \ldots, p\}^N$. 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 tels que

$$\operatorname{diam}(\mathcal{U}) = \delta < \frac{1}{p+1}.$$

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 $\operatorname{diam}(\mathcal{U} \vee T^{-1}(\mathcal{U})) \leq \frac{\delta}{p}$. 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

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

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 $\leq 1/2$, alors on a $N(\mathcal{U} \vee \mathcal{V}) \leq N(\mathcal{U}) + N(\mathcal{V})$. 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 $h(T,\mathcal{U}) = 0$.

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

Nous avons le résultat suivant qui sera prouvé plus tard :

PROPOSITION 1.4.1: *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.

1.5 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 1.5.1 : *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 1.5.2 : *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. 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

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

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 $i \in \{1, \dots, r\}$ 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 $\mathcal{Q} = (Q_i)_{0\leq i\leq r}$ et $\mathcal{P}' = (P_i)_{0\leq i\leq r}$ telles que $\mu(P_i\Delta Q_i) \leq r\varepsilon$ pour tout $i \in \{1, \ldots, r\}$. 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 1.5.3 : *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

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

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

Lemme 1.5.4 : *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 1.5.5 : *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 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 $\mathrm{diam}(P_i) \leq \varepsilon$ 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 $\mathcal{P}^n = \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})$. 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 1.5.6 : *Pour tous entiers $q \leq n$, on a $qH_{\nu_n}(\mathcal{P}^n) \leq nH_{\mu_n}(\mathcal{P}^q) + 2q^2 \ln(\sharp I)$.*

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

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

où J_r est l'entier tel que $n-1-q < r + qJ_r - 1 \leq n-1$ On obtient

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

En sommant sur r, on obtient

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

Puisque $\phi\ : t \mapsto -t\ln t$ est concave sur $[0,1]$, on sait que pour toute partition borélienne $\mathcal{Q} = (Q_j)_{j\in J}$ on a

$$\begin{aligned}
H_{\mu_n}(\mathcal{Q}) &= \sum_{j\in J} \phi(\mu_n(Q_j)) \\
&= \sum_{j\in J} \phi\left(\frac{1}{n} \sum_{i=0}^{n-1} T_*^i(\nu_n)(Q_j) \right) \\
&\geq \sum_{j\in J} \frac{1}{n} \sum_{i=0}^{n-1} \phi\left(T_*^i(\nu_n)(Q_j) \right) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} H_{T_*^i \nu_n}(\mathcal{Q}),
\end{aligned}$$

et on en déduit que

$$\frac{1}{n}\sum_{i=0}^{n-1} H_{\nu_n}(T^{-i}(\mathcal{P}^q) = \frac{1}{n}\sum_{i=0}^{n-1} H_{T_*^i(\nu_n)}(\mathcal{P}^q) \leq H_{\mu_n}(\mathcal{P}^q).$$

□

On a donc prouvé que

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

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 pour la topologie faible* puisque $\mathcal{M}_T$ est compact ou même semi-continue supérieurement. 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 and y, il existe $n \geq 0$ tel que $d(T^n(x), T^n(y)) \geq \varepsilon_0$.

PROPOSITION 1.5.7 : *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 $2e_1$ 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 $N(n,\varepsilon) \leq s(n,\varepsilon)$ et donc que

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

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

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

$\square$

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 $p \geq 2$, et les endomorphismes du cercle $T\ : x \to px$, où $|p| \geq 2$. Dans le premier cas la mesure équidistribuée telle que la mesure d'un cylindre

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

vaut $\frac{1}{p^m}$ 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.

§2. Sous-décalages de type fini

2.1 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

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

est fermé et invariant par le décalage $\sigma\ : (i_k)_{k\in\mathbf{Z}} \to (i_{k+1})_{k\in\mathbf{Z}}$. La restiction $\sigma_A = \sigma|_{X_A}$ 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 $i \in \{1,\dots,p\}$ et pour tout $k_0 \in \mathbf{Z}$, il existe $\bar{i} = (i_k)_{k\in\mathbf{Z}} \in X_A$ tel que $i_{k_0} = i$. Un *mot* de *longueur* n est une suite $w = (i_k)_{0\le k\le n}$ dans $\{1,\dots,p\}$. Ce mot *joint* i_0 à i_n. Nous dirons qu'un mot $w = (i_k)_{0\le k\le n}$ est *admissible* s'il peut être étendu à une suite $\bar{i} = (i_k)_{k\in\mathbf{Z}} \in X_A$, c'est-à-dire si $A_{i_k,i_{k+1}} = 1$ pour tout $k \in \{0,\dots,n-1\}$. Les cylindres qui engendrent la topologie de X_A sont les *cylindres admissibles* qui s'écrivent

$$C_w^{k_0} = \{\bar{i}' = (i'_k)_{k\in\mathbf{Z}} \in X_A \,|\, i'_{k+k_0} = i_k \text{ pour tout } k \in \{0,\dots,n\}\},$$

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

Nous allons chercher à calculer l'entropie topologique de T. Commençons par le lemme suivant

Lemme 2.1.1 : *Pour tous i, j dans $\{1,\dots,p\}$, le nombre de mots admissibles de longueur $n \ge 1$ qui joignent i à j est égale à $A^n_{i,j}$ le coefficient (i,j) de A^n.*

Démonstration En effet,

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

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

Proposition 2.1.2 : *L'application σ_A est positivement transitive si, pour tous i, j dans $\{1,\dots,p\}$, il existe $n \ge 1$ tel que $A^n_{i,j} > 0$. Dans ce cas, on dira que A est* irréducible.

Démonstration Considérons les cylindres $C_i^0 = \{\bar{i} \in X_A \,|\, i_0 = i\}$. Si σ_A est positivement transitive, alors, pour tous i, j dans $\{1,\dots,p\}$, 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 $A^n_{i,j} > 0$.

Réciproquement, supposons que pour tous i, j dans $\{1,\dots,p\}$, il existe $n \ge 1$ tel que $A^n_{i,j} \ne 0$. Soient $C_w^{k_0}$ et $C_{w'}^{k'_0}$ deux cylindres. 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 $\{-N,\dots,N\}$. 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 $C_w^{k_0} \cap \sigma_A^{-2N-n}(C_{w'}^{k'_0}) \ne \emptyset$. □

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

Proposition 2.1.3 : *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 égale au plu 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 $(v_i)_{1\leq i\leq p}$ 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

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

où les x_i sont les coordonnées de v de $(v_i)_{1\leq i\leq p}$ 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 une norme sur E et on a $\limsup_{n\to+\infty} \|L^n\|^{\frac{1}{n}} \leq |\lambda|$. 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 2.1.4 : *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 $\{0,\ldots,n-1\}$. Ils sont bien sûr disjoints deux-à-deux et il y en a exactement

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

L'équivalence des normes dans un espace de dimension finie nous dit qu'il existe $c>0$ et $C>0$ tels que

$$c\|M\| \leq \sum_{i,j}|M_{i,j}| \leq C\|M\|,$$

pour toute matrice carrée d'ordre n. Ainsi, pour tout $n\geq 1$, on a

$$c\|A^{n-1}\| \leq N\left(\bigvee_{i=0}^{n-1}\sigma_A^{-i}(\mathcal{U})\right) \leq C\|A^{n-1}\|,$$

ce qui implique la proposition. □

2.2 Mesures de Markov, mesure de Parry

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

Proposition 2.2.1 : *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 v à coefficients strictement positifs et tout vecteur propre à coefficients positifs est un multiple de v ;*

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

Démonstration Considérons le cône

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

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 $C_H = C\cap H$, par exemple prenons

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

Puisque les coefficients de A^N sont strictement positifs, on a $Av\neq 0$ pour tout $v\in C\setminus\{0\}$. 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.

Si x et x' sont deux points distincts de $\mathrm{Int}(C_H)$, on considère la droite Δ passant par x et x' puis les deux points a et b de $\Delta \cap \mathrm{Fr}(C_H)$, le point a choisi du côté de x et le point b du côté de x'. On a donc $[a,b,x,x'] > 1$ et on peut définir $d_H(x,x') = \ln([a,b,x,x'])$ puis poser par extension $d_H(x,x) = 0$. Il n'est pas difficile de voir que d_H est continue.

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' la droite passant par $A_H(x)$ et $A_H(x')$, puis a' et b' les points de $\Delta' \cap \mathrm{Fr}(C_H)$, le point a' choisi du côté de x et le point b' du côté de x'. Remarquons que $A_H(\Delta \cap C_H) \subset \Delta' \cap C_H$ et que le point a est situé entre a et x (resp. entre b et x') et donc que

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

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

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

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

Remarquons maintenant que pour tout voisinage U de v 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 v. Dans le cas contraire on pourrait trouver $v' \in K$ qui maximise $x \mapsto d_H(x,v)$, 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, que d_H était une distance et que A_H^N était contractante, puis utiliser le théorème de point fixe de Picard.)

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

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

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

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

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

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

et donc

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

□

Remarques

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

$$\lim_{n\to+\infty} \frac{1}{\lambda^n} A^n_{i,j} = v_i f_j.$$

En effet, on a

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

et

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

pour tout $w \in \text{Ker}(f)$.

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

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

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

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

est invariant par H. Les applications A et A_H coïncident, le point fixe de A_H est un point fixe de A.

Le décalage de Bernouilli admet beaucoup de mesures invariantes, nous allons nous intéresser à certaines d'enre-elles les *mesures de Markov*

Soit $M = (M_{i,j})_{i,j}$ une matrice stochastique d'ordre p. On a vu plus haut que

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

était invariant par M et que la partie compacte convexe C_H était positivement invariante. Puisque $M|_H$ est affine, on montre facilement que H a au moins un point fixe v dans C_H (et on sait de plus que ce point est unique s'il existe une puissance de M a coefficients strictement positifs). On vérifie facilement que le vecteur $f = (1, \ldots, 1)$ est fixe par M^t.

Proposition 2.2.2 : *Il existe une mesure borélienne de probabilité μ telle que pour tout $w = (i_0, \ldots, i_m) \in \{1, \ldots, p\}^{m+1}$ et tout $k_0 \in \mathbf{Z}$, on a*

$$\mu(C_w^{k_0}) = \left(\prod_{k=0}^{m-1} M_{i_k, i_{k+1}} \right) v_{i_m}.$$

Démonstration Pour $w = (i_0, \ldots, i_m) \in \{1, \ldots, p\}^{m+1}$, et $i \in \{1, \ldots, p\}$ écrivons

$$iw = (i, i_0, \ldots, i_n), \;\; wi = (i_0, \ldots, i_n, i).$$

Puisque $v \in H$, on sait que pour tout $k_0 \in \mathbf{Z}$, on a

$$\sum_{i=1}^{p} \mu(C_i^{k_0}) = 1.$$

On peut remarquer également que pour tout $w = (i_0, \ldots, i_m) \in \{1, \ldots, p\}^{m+1}$ et tout $k_0 \in \mathbf{Z}$, on a

$$\mu(C_w^{k_0}) = \sum_{i=1}^{p} \mu(C_{iw}^{k_0-1}) = \sum_{i=1}^{p} \mu(C_{wi}^{k_0}).$$

La première égalité est une conséquence du fait que M est une matrice stochastique, la seconde du fait que v est un vecteur propre de M associée à la valeur propre 1. Grâce au théorème d'extension de Caratheodory, on sait qu'il existe une unique mesure borélienne de probabilité définie par ces relations. Cette mesure est bien évidemment invariante par le décalage. Le coefficient $M_{i,j}$ représente la probabilité d'être dans C_i^0 au temps $k-1$ sous la dynamique si on est dans C_j^0 au temps k.

□

Proposition 2.2.3 : *Dans le cas où il existe une puissance de M dont les coefficients sont strictement positifs, la mesure μ est mélangeante.*

Démonstration Il suffit de prouver que

$$\lim_{n \to +\infty} \mu(C_w^{k_0} \cap \sigma^{-n}(C_{w'}^{k_0'})) = \mu(C_w^{k_0}) \mu(C_{w'}^{k_0'})$$

pour tous cylindres $C_w^{k_0}$ and $C_{w'}^{k_0'}$. Supposons que $w = (i_0, \ldots, i_m)$ et $w' = (i_0', \ldots, i_{m'}')$. Si $n > m + k_0 - k_0'$, alors

$$\begin{aligned}
\mu(C_w^{k_0} \cap \sigma^{-n}(C_{w'}^{k_0'})) &= \left(\prod_{k=0}^{m-1} M_{i_k, i_{k+1}} \right) M_{i_m, i_0'}^{n-m-k_0+k_0'} \left(\prod_{k=0}^{m'-1} M_{i_k', i_{k+1}'} \right) v_{i_{m'}'} \\
&= \mu(C_w^{k_0}) \mu(C_{w'}^{k_0'}),
\end{aligned}$$

puisque, comme on l'a vu dans le remarque qui suit la proposition 2.2.1, on sait que

$$\lim_{n \to \infty} M_{i_m, i_0'}^n = v_{i_m}.$$

□

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 v un vecteur propre à coefficients positifs (et donc associé à λ) et f un vecteur propre de A^* à coefficients positifs (donc associé également à λ) tels que $\sum_{i=1}^{p} f_i v_i = 1$. Considérons la matrice M, où

$$M_{i,j} = \frac{A_{i,j} f_i}{\lambda f_j}.$$

C'est une matrice stochastique car

$$\sum_{i=1}^{j} A_{i,j} f_i = \lambda f_j,$$

la mesure de Markov associée est la mesure de Parry μ_A, elle est supportée sur X_A. Le vecteur propre de M qui est dans $C \cap H$ est $(f_1 v_1, \ldots, f_p v_p)$. Par exemple si $A_{i,j} = 1$ pour tout (i,j), alors $\lambda = p$ et $v_i = f_i = \frac{1}{\sqrt{p}}$. Dans ce cas la mesure de Parry est la mesure équidistribuée.

La mesure de Parry maximise l'entropie :

PROPOSITION 2.2.4 : *On a*

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

Démonstration Commençons par rappeler la valeur de l'entropie métrique d'une mesure de Markov

LEMME 2.2.5 : *Si μ est une mesure de Markov définie par une matrice stochastique $M = (M_{i,j})_{i,j}$ et un vecteur fixe $v = (v_1, \ldots, v_p)$ à coefficients positifs, alors*

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

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

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

Mais

$$\begin{aligned}
H_\mu(\mathcal{P} | \bigvee_{i=1}^{n} \sigma^{-i}(\mathcal{P})) &= - \sum_{i_0, i_1, \ldots, i_n} \mu(C^0_{i_0, i_1, \ldots, i_n}) \ln \left(\frac{\mu(C^0_{i_0, i_1, \ldots, i_n})}{\mu(C^1_{i_1, \ldots, i_n})} \right) \\
&= - \sum_{i_0, i_1, \ldots, i_n} \mu(C^0_{i_0, i_1, \ldots, i_n}) \ln(M_{i_0, i_1}) \\
&= - \sum_{i_0, i_1} \mu(C^0_{i_0, i_1}) \ln(M_{i_0, i_1}) \\
&= - \sum_{i,j} M_{i,j} \ln(M_{i,j}) v_j
\end{aligned}$$

□

Appliquons le lemme à la mesure de Parry, on trouve

$$\begin{aligned} h_\mu(\sigma_A) &= -\sum_{i,j} \frac{A_{i,j} f_i}{\lambda f_j} \ln\left(\frac{A_{i,j} f_i}{\lambda f_j}\right) v_j f_j \\ &= \sum_{i,j} (-\ln(A_{i,j} f_i) + \ln f_j + \ln \lambda) \frac{A_{i,j} f_i v_j}{\lambda}. \end{aligned}$$

Mais

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

$$\sum_{i,j} \ln f_j \frac{A_{i,j} f_i v_j}{\lambda} = \sum_j \ln f_j \, f_j v_j$$

et

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

□

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

PROPOSITION 2.2.6 : *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 un borélien B tel 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 $\lim_{l\to+\infty} \mu(B_l) = \mu(B)$ et $\lim_{m\to+\infty} \mu_A(B_l) = \mu_A(B) = 0$. La partition $\mathcal{P} = (C_i^0)_{1\leq i\leq p}$ étant génératrice, on sait que

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

Remarquons que pour tout cylindre admissible $C_w^{k_0}$ on a

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

où $C = \inf_{i,j} f_i v_j > 0$.

LEMME 2.2.7 : *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 and 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

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

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

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

□

2.3 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 2.3.1 : *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

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

□

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 2.3.2 : *Si A is 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

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

□

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

Proposition 2.3.3 : *On a $\mu_A = \lim_{n\to+\infty} \mu_n$, où $\mu_n = \dfrac{1}{\sharp(\mathrm{Fix}(\sigma_A^n))} \displaystyle\sum_{\bar{i}\in \mathrm{Fix}(\sigma_A^n)} \delta_{\bar{i}}$.*

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

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

Ainsi, on a

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

□

§ 3. Partitions de Markov

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 est un facteur mais une partie de celui-ci, parfois la partie la plus intéressante : 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 $\bar{i} = (i_n)_{n\geq 0} \in I^{\mathbf{N}}$ *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 $\bar{i} = (i_n)_{n\geq 0} \in I^{\mathbf{N}}$ l'ensemble $H(\bar{i}) = \bigcap_{n\geq 0} T^{-n}(R_{i_n})$ contient au plus un point. Dans ce cas, l'ensemble Z des suites $\bar{i}$ pour lesquelles $H(\bar{i})$ est non vide est invariant par le décalage σ, et il est fermé si X est compact. Si on écrit $H(\bar{i}) = \{h(\bar{i})\}$, 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

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

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 $I^{\mathbf{Z}}$.

3.1 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 $\lambda \leq 4$. Dans le cas où $\lambda \geq 1$, remarquons que la suite $(f_\lambda^n(x))_{n\geq 0}$ est décroissante et que $\lim\limits_{n\to+\infty} f_\lambda^n(x) = -\infty$. Cette dernière propriété est encore vraie si $x > 1$ car $f_\lambda(x) < 0$. L'ensemble des points d'orbites bornées est 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 $\lambda \leq 4$ 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 $\lambda \in]0,4[$, il y a une très grande richesse des 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 tous $x \in X_\lambda$.

Proposition 3.1.1: *Supposons que* $\lambda > 2+\sqrt{5}$. *Pour tout* $\bar{i} = (i_n)_{n\geq 0}$ *il existe un unique point* $x = h(\bar{i}) \in X_\lambda$ *tel que pour tout point* $n \geq 0$ *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 $\bar{i} = (i_n)_{n\geq 0}$. 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 $h(\bar{i})$. L'application h est évidement injective et induit une bijecetion entre $\{0,1\}^N$ et X_λ (qui est un ensemble de Cantor). Elle induit une conjugaison entre σ et $f_\lambda|_{X_\lambda}$. □

On a une conclusion analogue si $\lambda > 4$ mais la preuve demande des outils plus sophistiqués, comme la dérivée schwarzienne.

3.2 Le fer à cheval de Smale

On va décrire un exemple géométrique du 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 du diamètre $[a,b] \times \{c\}$ et D_1 pour le demi-disque au dessus du 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 $z \in \mathbf{R}$ 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 lignes horizontales. 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 lignes horizontales. 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 on $\{0,1\}^{\mathbf{Z}}$ 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 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.

3.3 L'application de Hénon

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 3.3.1: *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 solutions réelles 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, on a nécessairement $\lim_{k\to+\infty} x_k = -\infty$ 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 puis, pour les mêmes raisons que précédemment, que $\lim_{k\to+\infty} x_k = +\infty$. On montre de la même façon que $\lim_{k\to-\infty} |x_k| = +\infty$. □

En fait, 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 3.3.2: *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 $|x_k| > M$, alors on a

$$x_k \leq |x_k| < \max(x_{k-1}, x_{k+1}) \leq \max(|x_{k-1}|, |x_{k+1}|),$$

car

$$(1+b)|x_k| < x_k^2 + c = x_{k+1} + bx_{k-1},$$

Il ne reste donc plus qu'à utiliser l'argument précédent. □

En particulier, si $c \leq 0$, l'ensemble des points d'orbites bornées est la partie compacte invariante non vide (puisqu'elle contient un point fixe) $X = \bigcap_{k\in\mathbf{Z}} f^{-k}([-M, M])^2$, tout point $z \notin X$ vérifie $\lim_{k\to+\infty} \|f^k(z)\| = +\infty$ 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 3.3.3: *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, tous les points périodiques sont de type selle.*

Démonstration. On considère l'espace métrique complet $(\mathcal{B}, d)$ des suites $\overline{x} = (x_k)_{k\in\mathbf{Z}}$ dans $[-M, M]$ muni de la distance

$$d(\overline{x}, \overline{x}') = \sup_{k\in\mathbf{Z}} |x_k - x'_k|.$$

Remarquons que pour toute suite $\overline{x} = (x_k)_{k\in\mathbf{Z}}$ dans $\mathcal{B}$, on a

$$x_{k+1} + bx_{k-1} - c \leq (1+b)M - c = M^2$$

et

$$\begin{aligned} x_{k+1} + bx_{k-1} - c &\geq -M(1+b) - c \\ &\geq -2M - c \\ &\geq -(1+b) - \sqrt{(1+b)^2 - 4c} - c \\ &\geq -2 - \sqrt{4 - 4c} - c \\ &\geq -2 - \sqrt{44} + 10 = \alpha^2 > 1. \end{aligned}$$

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 $\overline{\varepsilon} = (\varepsilon_k)_{k\in\mathbf{Z}}$, on définit une application

$$\begin{aligned}\Phi_{\overline{\varepsilon}}\ :\ &\mathcal{B} \to \mathcal{B}\\ &\overline{x} \mapsto \overline{y}\end{aligned}$$

où

$$y_k = \varepsilon_k \sqrt{x_{k+1} + b x_{k-1} - c}.$$

C'est une contraction puisque

$$\begin{aligned}\left|\varepsilon_k \sqrt{x_{k+1} + b x_{k-1} - c} - \varepsilon_k \sqrt{x'_{k+1} + b x'_{k-1} - c}\right| &= \left|\frac{x_{k+1} + b x_{k-1} - x'_{k+1} - b' x_{k-1}}{\sqrt{x_{k+1} + b x_{k-1} - c} + \sqrt{x'_{k+1} + b x'_{k-1} - c}}\right|\\ &\leq \frac{2\|\overline{x} - \overline{x}'\|}{2\alpha}\end{aligned}$$

pour tous $\overline{x}$ et $\overline{x}'$ dans $\mathcal{B}$. Ainsi, $\Phi_{\overline{\varepsilon}}$ a un unique point fixe.

Posons $R_- = [-M, M] \times [-M, 0]$ et $R_+ = [-M, M] \times [0, M]$. On a montré que pour tout $\overline{\varepsilon} = (\varepsilon_k)_{k\in\mathbf{Z}}$, l'ensemble $\bigcap_{k\in\mathbf{Z}} f^{-k}(R_{\varepsilon_k})$ est réduit à un unique point $h(\overline{e})$, et que l'orbite d'un tel point ne rencontre jamais la bande $\mathbf{R} \times [-1, 1]$. De plus, l'image de h est exactement X. Il est facile d'en déduire que h est un homéomorphisme de $\{-1, 1\}^Z$ 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

$$Df(x, y) = \begin{pmatrix} 2x & -b \\ 0 & 1 \end{pmatrix}.$$

Fixons $(x, y) \in X$ puis un vecteur (u, v) tel que $|u| \geq |v|$. Si $(u', v') = Df(x, y).(u, v)$, alors

$$|u'| \geq (2\alpha - 1)|u| > |u| \geq |v| = |v'|$$

et donc

$$\|(u', v')\| = |u'| \geq (2\alpha - 1)|u| = (2\alpha - 1)\|(u, v)\|.$$

On a montré que le cône $|v| \leq |u|$ 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 en déduit que $Df^q(z) = Df(f^{q-1}(z)) \circ \ldots \circ Df(z)$ envoie le cône $|v| \leq |u|$ 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 $|\lambda||\mu| = b^q \leq 1$. $\square$

Un bon exercice est de dessiner l'image de $[-M, M]^2$ par f et d'essayer de comprendre ce qui se passe.

3.4 Automorphismes hyperboliques du tore

On va construire une partition de Markov pour les endomorphismes hyperboliques du tore $\mathbf{T}^r$. Commençons par l'exemple classique $\widehat{A}$ of $\mathbf{T}^2$, 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'ensemble 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 regadant 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 $\bar{\imath} = (i_k)_{k\in\mathbf{Z}} \in Z_M$ il existe un unique point $z = h(\bar{\imath})$ 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 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 tesl 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

$$h(\sigma_M) = \log\rho(M) = \lim_{n\to+\infty} \frac{1}{n}\log(\sharp\mathrm{Fix}(\sigma_M^n)) = \log\left(\frac{3+\sqrt{5}}{2}\right),$$

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 Si on pose $\lambda = \dfrac{3+\sqrt{5}}{2}$, ceci implique que pour tous points $\widehat{z}$ and $\widehat{z}'$ in $\mathbf{T}^2$

$$\max_{-1\leq i\leq 1} d(\widehat{A}^i(\widehat{z}), \widehat{A}^i(\widehat{z}')) \leq \varepsilon_0 \Rightarrow \min_{i\in\{-1,1\}} d(\widehat{A}^i(\widehat{z}), \widehat{A}^i(\widehat{z}')) \geq \lambda d(\widehat{z}, \widehat{z}').$$

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 $d(T^k(x), T^k(y)) \geq \varepsilon_0$ et 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 les rectangles sont choisis assez petit pour définir une vraie partition de Markov. En fait on a

Proposition 3.4.1: *Pour tout $\varepsilon > 0$, on peut construire une telle famille avec des rectangles de diamètre $\leq \varepsilon$.*

Ceci implique

Corrolaire 3.4.2: *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.

§4. Entropie et homologie

4.1 Cas 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 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 4.1.1: *Si $\widehat{A}$ est un automorphisme linéaire hyperbolique de $\mathbf{T}^r$, alors*

$$h(\widehat{A}) = \sup_{0\leq s\leq r} \rho(\widehat{A}_{*,s})$$

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 $T\ : \mathbf{T}^r \to \mathbf{T}^r$ est homotope à $\widehat{A}$ on sait, c'est une conséquence du lemme de poursuite, que que $\widehat{A}$ est un facteur à T et donc que $h(T) \geq h(A)$. On sait aussi que les morphismes $\widehat{A}_{*,s}$ et $T_{*,s}$ coïncident. On a donc le résultat plus général suivant :

Proposition 4.1.2: *Si $T\ : \mathbf{T}^r \to \mathbf{T}^r$ est homotope à un automorphisme linéaire hyperbolique de $\mathbf{T}^r$, alors*

$$h(T) \geq \sup_{0\leq s\leq r} \rho(T_{*,s})$$

où $T_{*,s}\ : H_s(\mathbf{T}^r, \mathbf{R}) \to H_s(\mathbf{T}^r, \mathbf{R})$ est l'action induite par T sur $H_s(\mathbf{T}^r, \mathbf{R})$.

4.2 Conjecture de l'entropie

Au vu du résultat précédent, on peut se demander si, pour toute application continue $T\ : M \to M$ sur une variété compacte M de dimension r, on a $h(T) \geq \sup_{0\leq s\leq r} \rho(T_{*,s})$ où $T_{*,s}\ : H_s(M, \mathbf{R}) \to H_s(M, \mathbf{R})$ est l'action induite par T 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 $T_{*,r}$ est une homothétie. Le rapport de cette homothétie, noté $\deg(T)$, s'appelle le *degré* de T. C'est en fait un entier et on vérifie aisément que $\deg(T \circ S) = \deg(T) \circ \deg(S)$, en particulier $\deg(T) \in \{-1, 1\}$ si T est un homéomorphisme. L'inégalité $h(T) \geq \ln(\deg(T))$ est clairement vraie si T est un homéomorphisme mais cesse d'être nécessairement vraie sinon. Expliquons pourquoi. L'application $T\ : z \mapsto z^2$, $t > 0$, 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 errants $\Omega(T)$ est donc formé de 0 de ∞ et du cercle unité sur lequel T_1 est conjugué à un endomorphisme de degré 2. Dans ce cas on a bien $h(T) = h(T_{\Omega(T)}) = \ln 2 = \ln(\deg(T_1))$.

L'application $S\ : z \mapsto z^2/2|z|$, 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 $\Omega(T) = \{0, \infty\}$ et $h(S) = 0$. La différence entre les deux cas est le fait que T est de classe C^1 alors que S n'est pas différentiable en 0 et surtout en ∞. En fait, on a le résultat suivant, du à Misiurewicz et Przytycki :

Théorème 4.2.1: *Si $T : M \to M$ est une application de classe C^1 sur une variété compacte lisse orientable M, on a $h(T) \geq \ln(\deg(T))$.*

Là différentielle joue un rôle important, ce qui a amené Shub à conjecturer :

Conjecture 4.2.2: *Si $T : M \to M$ est une application de classe C^1 sur une variété compacte lisse M, on a $h(T) \geq \sup_{0 \leq s \leq r} \rho(T_{*,s})$ où $T_{*,s} : H_s(M, \mathbf{R}) \to H_s(M, \mathbf{R})$ est l'action induite par T 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 Manning qui ne nécessite pas de condition différentiable.

Théorème 4.2.3: *Si $T : M \to M$ est une application continue sur une variété compacte M, on a $h(T) \geq \rho(T_{*,1})$ où $T_{*,1} : H_1(M, \mathbf{R}) \to H_1(M, \mathbf{R})$ est l'action induite par T sur $H_1(M, \mathbf{R})$.*

On a également le corollaire suivant, conséquence de la dualité de Poincaré :

Corollaire 4.2.4: *Si $T : M \to M$ est un homéomorphisme sur une variété compacte M de dimension ≤ 3, on a $h(T) \geq \sup_{0 \leq s \leq 3} \rho(T_{*,s})$.*

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 τ est un endomorphisme de G on peut définir *l'entropie algébrique* de τ de la façon suivante : on choisit un ensemble générateur symétrique fini Σ, on peut alors 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

$$L_n(\tau, \Sigma) = \max_{g \in \Sigma} l_\Sigma(\tau^n(g)),$$

on peut montrer que la suite $(L_n(\tau, \Sigma))_{\geq 0}$ est sous-multiplicative et donc que

$$\lim_{n \to +\infty} \frac{1}{n} \ln(L_n(\tau, \Sigma))$$

existe. On peut démontrer également que cette limite ne dépend pas de Σ, c'est l'entropie algébrique de τ, on la note $h_{\mathrm{alg}}(\tau)$. 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 $h_{\mathrm{alg}}(\iota\tau) = h_{\mathrm{alg}}(\tau)$. Supposons maintenant que $T : M \to M$ est une application continue sur une variété compacte. Fixons $x_0 \in M$ et un chemin α joignant x_0 à $T(x_0)$. L'application, qui à un lacet γ basé en x_0 associe le lacet $\alpha^{-1}\gamma\alpha$ induit un endomorphisme $\tau_{x_0,\alpha}$ du groupe fondamental $\pi_1(M, x_0)$. L'entropie algébrique de $\tau_{x_0,\alpha}$ ne dépend pas du choix de α (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

$$h_{\pi_1}(T) = h_{\mathrm{alg}}(\tau_{x_0,\alpha}).$$

Le groupe $H_1(M, \mathbf{Z})$ est l'abélianisé de $\pi_1(M, x_0)$ et l'action de $T_{*,1}$ 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 $\ln(\rho(T_{*,1}))$. Ainsi, on a $h_{\pi_1}(T) \geq \ln(\rho(T_{*,1}))$ Le résultat suivant de Bowen améliore donc (même s'il est moins effectif) le résultat de Manning :

THÉORÈME 4.2.5: *Si $T : M \to M$ est une application continue sur une variété compacte M, on a $h(T) \geq h_{\pi_1}(T)$.*

Les deux théorèmes précédents sont particulièrement intéressants en dimension deux. En effet, supposons que O soit une orbite périodique de T (ou une réunion finie de telles orbites). Supposons de plus que T 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 $\overline{T}$ 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 $h(\overline{T}) = h(T)$ car T est un facteur de $\overline{T}$ et $\overline{T}$ 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.

§5. Systèmes dynamiques hyperboliques

5.1 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 $E(x) = E^s(x) \oplus E^u(x)$ et un réel $\mu < 1$ et une constante $C > 0$ tel 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 en avons vu d'autres : le fer à cheval de Smale mais plus généralement l'ensemble des points non errants du difféomorphisme de S^2 apparaissant dans la construction le fer à cheval (en effet, on ajoute un puits et une source), l'ensemble X introduit précédemment dans l'étude de la famille de Hénon, le tore $\mathbf{T}^r$ lui-même pour un automorphisme hyperbolique du tore. 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 5.1.1: *Soit $f : M \to M$ un difféomorphisme de classe C^1 et X une partie compacte invariante hyperbolique. Choisissons $\delta > 0$ assez petit et définissons, pour tout $x \in X$, les ensembles*

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

et

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

i) *Chaque ensemble $W^s(x)$ (resp. $W^u(x)$) est l'image d'un plongement de classe C^1 de $[-1, 1]$ 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*

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

et

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

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

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

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

$$W^s(x) = \bigcup f^{-n}(W^s_{\mathrm{loc}}(f^n(x))), \ \ W^u(x) = \bigcup f^n(W^s_{\mathrm{loc}}(f^{-n}(x)))$$

5.1 Points d'intersection homoclines, présence de fers à cheval

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 5.2.1: *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 transverse en un point 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 5.2.1: *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 transverse en un point x_1. Alors x_1 appartient à l'adhérence de l'ensemble des points périodiques de f*

Corollaire 5.2.2: *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 transverse. Alors on a $h(f) > 0$.*

En dimension ≥ 3, il existe des difféomorphisme de variétés sans points périodiques. 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, du 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 5.2.4: *Soit $f : M \to M$ un difféomorphisme de classe $C^{1+\varepsilon}$ d'une surface compacte M. Si $h(f) > 0$, il existe alors un point périodique de type selle admettant un point d'intersection transverse.*

§6. Exercices

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 $T(\Omega(T)) = \Omega(T)$.

2) À l'aide du principe variationnel, montrer que $h(T) = h(T|_{\Omega(T)})$.

Exercice 7

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 :

$$tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) \leq H_{t\mu+(1-t)\nu}(\mathcal{P}) \leq tH_\mu(\mathcal{P}) + (1-t)H_\nu(\mathcal{P}) - t\ln t - (1-t)\ln(1-t).$$

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

$$h_{t\mu+(1-t)\nu}(T) = th_\mu(T) + (1-t)h_\nu(T).$$

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 8

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 and y, il existe $n \geq 0$ tel que $d(T^n(x), T^n(y)) \geq \delta$.

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 le topologie).

Exercice 9

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 10

L'entropie d'une application continue expansive $T : X \to X$ sur un espace métrique compact peut-elle être infinie ?

Exercice 11

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 12

Soit $T : X \to X$ un homéomorphisme d'un espace métrique compact. Montrer que si T est (positivement) expansive, alors X est fini.

Exercice 13

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 and 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 14

Montrer qu'il n'y a pas d'homéomorphisme expansif sur $\mathbf{T}^1$.

Exercice 15

Montrer qu'un automorphisme linéaire du tore $\mathbf{T}^r$ est expansif si et seulement s'il est hyperbolique.

Exercice 16

Prouver que la mesure de Lebesgue est la seule mesure borélienne de probabiltié invariante par $T\ : \widehat{x} \mapsto p\widehat{x}$ sur $\mathbf{T}^1$, telle que $h_\mu(T) = h(T)$, si $p \geq 2$.

Exercice 17

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 $\beta = 1 + \dfrac{1}{a}$ et $\gamma = \dfrac{1}{1+a}$.

1) Dessiner le graphe de $g_\alpha$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 18

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 19

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 presuqe toute 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

$$\lim_{n\to+\infty} \frac{1}{\sharp \mathrm{Fix}(g^n)} \sum_{x\in \mathrm{Fix}(g^n)} \varphi(x) = c.$$

Exercice 20

Soit f un homéomorphisme de $\mathbf{T}^2$. Montrer que $h(T) \geq \sup_{0\leq s\leq 2} \rho(f_{*,s})$.