Introduction aux systèmes dynamiques

Patrice Le Calvez

Cours fondamental I, Master 2

2021-2022

1.1 Rappels sur les systèmes dynamiques mesurés, mélange faible

Rappelons qu'un système dynamique mesuré est la donnée d'un quadruplet $(X, \mathcal{B}, \mu, T)$, où X est un ensemble, où $\mathcal{B}$ est une σ-algèbre sur X, où μ est une mesure σ-finie sur X et où $T\ : X \to X$ est une application mesurable qui préserve μ. Ceci signifie que

$$\mu(A) = \mu(T^{-1}(A)) \text{ pour tout } A \in \mathcal{B},$$

où de façon équivalente que $T_*(\mu) = \mu$, où $T_*(\mu)$ est la *mesure image* définie ainsi :

$$T_*(\mu)(A) = \mu(T^{-1}(A)) \text{ pour tout } A \in \mathcal{B}.$$

Si T est inversible, on obtient un système dynamique inversible. Très souvent, l'inversibilité n'a lieu que presque partout, ce qui est suffisant pour avoir un système dynamique inversible : il existe une application mesurable $S\ : X \to X$, telle que $S \circ T(x) = T \circ S(x) = x$ pour presque tout point x. De façon équivalente, il existe $Y \in \mathcal{B}$ telle que $\mu(X \setminus Y) = 0$ et telle que T induit une bijection mesurable en restriction à Y.

Parmi les exemples classiques de systèmes dynamiques mesurés, rappelons:

- les rotations $\widehat{x} \mapsto \widehat{x} + \widehat{a}$ sur le tore $\mathbf{T}^r = \mathbf{R}^r/\mathbf{Z}^r$, où $\widehat{a} \in \mathbf{T}^r$ ($\mathcal{B}$ est alors la tribu borélienne et μ la mesure de Haar) ;

- les endomorphismes $\widehat{x} \mapsto p\widehat{x}$ sur le cercle $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, où $p \in \mathbf{Z} \setminus \{0\}$, et plus généralement les endomorphismes linéaires surjectifs de $\mathbf{T}^r$ ($\mathcal{B}$ est encore la tribu borélienne et μ la mesure de Haar) ;

- le décalage de Bernouilli σ défini sur $\{1, \dots, p\}^{\mathbf{N}}$ (où sur $\{1, \dots, p\}^{\mathbf{Z}}$ pour une version inversible), où $\sigma((x_n)_{n\geq 0}) = (x_{n+1})_{n\geq 0}$, muni de l'une de ses nombreuses mesures boréliennes de probabilité invariantes, par exemple muni d'une mesure produit ou d'une mesure de Markoff.

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Une partie $A \in \mathcal{B}$ est *invariante* si $T^{-1}(A) = A$. On obtient alors un nouveau système dynamique mesuré $(A, \mathcal{B}_A, \mu_{|\mathcal{B}_A}, T_{|A})$, un sous-système, en notant $\mathcal{B}_A = \{A \cap B \mid B \in \mathcal{B}\} = \{B \in \mathcal{B} \mid B \subset A\}$. De même, un système dynamique mesuré $(Y, \mathcal{C}, \nu, S)$ est un *facteur* de $(X, \mathcal{B}, \mu, T)$ s'il existe une application mesurable $H\ : X \to Y$ telle que $H \circ T = S \circ H$ presque partout et $H_*(\mu) = \nu$, c'est-à-dire telle que $\nu(A) = \mu(T^{-1}(A))$ pour tout $A \in \mathcal{C}$. Si H est inversible (à ensembles négligeables près), les deux systèmes dynamiques sont *conjugués*.

Rappelons qu'un système dynamique mesuré $(X, \mathcal{B}, \mu, T)$ est *ergodique* si pour toute partie $A \in \mathcal{B}$ qui vérifie $T^{-1}(A) = A$, alors on a $\mu(A) = 0$ ou $\mu(X \setminus A) = 0$, ou de façon équivalente, si toute fonction mesurable $f\ : X \to \mathbf{C}$ invariante par T est constante presque partout. Une propriété plus forte que l'ergodicité, mais définie uniquement si μ est une mesure de probabilité, est la propriété de mélange : le système $(X, \mathcal{B}, \mu, T)$ est *mélangeant* si, pour tous A, B dans $\mathcal{B}$, on a $\lim_{n\to+\infty} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B)$. Autrement dit les évènements "être dans A au temps 0" et "être dans B au temps n" sont asymptotiquement indépendants quand $n \to +\infty$. Le fait que le mélange implique l'ergodicité provient de l'égalité

$$\mu(A)^2 = \lim_{n\to+\infty} \mu(A \cap T^{-n}(A)) = \mu(A),$$

si $T^{-1}(A) = A$. En effet, on doit avoir $\mu(A) = 0$ ou $\mu(A) = 1$.

Introduisons une propriété intermédiaire, définie également dans le cas où μ est une mesure de probabilité : le système $(X, \mathcal{B}, \mu, T)$ est *faiblement mélangeant* si, pour tous A, B dans $\mathcal{B}$, il existe une partie $I \subset \mathbf{N}$ de densité 1 telle que

$$\lim_{n\to+\infty,\, n\in I} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B).$$

Rappelons qu'une partie $I \subset \mathbf{N}$ est de *densité* $\delta \in [0,1]$ si

$$\lim_{n\to+\infty} \frac{\#(\{0,\dots,n-1\} \cap I)}{n} = \delta.$$

On sait que la rotation $\widehat{x} \mapsto \widehat{x}+\widehat{a}$ sur le tore $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ est ergodique si et seulement $\widehat{a} \notin \mathbf{Q}/\mathbf{Z}$. Il n'est pas difficile de voir qu'elle n'est jamais mélangeante ni même faiblement mélangeante (propriété vérifiée plus généralement pour toute rotation sur $\mathbf{T}^r$). Par contre l'endomorphisme $\widehat{x} \mapsto p\widehat{x}$ est mélangeant si $|p| \geq 2$.

Nous caractériserons dans la prochaine section ces propriétés dynamiques en termes spectraux. Nous verrons en particulier la richesse, en termes spectraux, du mélange faible.

1.2 Un peu de théorie spectrale des systèmes dynamiques

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Si $f : X \to \mathbf{C}$ est intégrable, il en est de même de $f \circ T$ et on a $\int f \circ T \, d\mu = \int f \, d\mu$. On obtient ainsi une autre façon de définir l'invariance de la mesure. Plus généralement, si $f \in L^p(\mu)$, $p \geq 1$, alors $f \circ T$ appartient à $L^p(\mu)$ et vérifie $\|f \circ T\|_p = \|f\|_p$. Un cas particulièrement intéressant est celui où $p = 2$. L'opérateur

$$\begin{aligned} U_T \ &: L^2(\mu) \to L^2(\mu) \\ &\qquad f \mapsto f \circ T \end{aligned},$$

appelé *opérateur de Koopman*, est alors une isométrie de $L^2(\mu)$: on a

$$\langle U_T(f), U_T(g) \rangle = \int (f \circ T)(\overline{g} \circ T) \, d\mu = \int f\overline{g} \, d\mu = \langle f, g \rangle.$$

Dans le cas où T est inversible, U_T est aussi inversible, c'est donc un opérateur unitaire. On a $U_T^* = U_T^{-1}$ puisque

$$\langle U_T(f), g \rangle = \langle U_T(f), U_T(U_T^{-1}(g)) \rangle = \langle f, U_T^{-1}(g) \rangle.$$

Les *moyennes de Birkhoff* d'une fonction $f : X \to \mathbf{C}$ sont les moyennes temporelles $\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i$. Énonçons le théorème de Von Neumann qui donne une convergence, au sens quadratique, de ces moyennes.

Théorème 1.2.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Définissons l'espace $L^2_{\mathrm{inv}}(\mu) = \mathrm{Fix}(U_T) = \mathrm{Ker}(\mathrm{Id} - U_T)$ ainsi que la projection orthogonale $p_{\mathrm{inv}} : L^2(\mu) \to L^2_{\mathrm{inv}}(\mu)$. Alors, pour tout $f \in L^2(\mu)$, on a*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} U_T^i(f) = p_{\mathrm{inv}}(f).$$

Preuve. On doit prouver que

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} U_T^i(f) = f,$$

si $f \in L^2_{\mathrm{inv}}(\mu)$, ce qui évident, et

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} U_T^i(f) = 0$$

si $f \in L^2_{\mathrm{inv}}(\mu)^\perp$, ce qui l'est moins.

Commençons par prouver que

$$L^2_{\mathrm{inv}}(\mu)^\perp = \overline{\mathrm{Im}(\mathrm{Id} - U_T)}$$

ou, de façon équivalente, que

$$L^2_{\mathrm{inv}}(\mu) = \mathrm{Im}(\mathrm{Id} - U_T)^\perp.$$

Si $f \in L^2_{\mathrm{inv}}(\mu)$, alors pour tout $g \in L^2(\mu)$, on a

$$\langle f, g - U_T(g)\rangle = \langle f, g\rangle - \langle f, U_T(g)\rangle = \langle f, g\rangle - \langle U_T(f), U_T(g)\rangle = 0,$$

Ainsi, on a $L^2_{\mathrm{inv}}(\mu) \subset \mathrm{Im}(\mathrm{Id} - U_T)^\perp$.

Réciproquement, si f appartient à $\mathrm{Im}(\mathrm{Id} - U_T)^\perp$ alors, pour tout $g \in L^2(\mu)$, on a

$$\langle f - U_T^*(f), g\rangle = \langle f, g - U_T(g)\rangle = 0,$$

où U_T^* est l'adjoint de U_T. Ainsi, on a $U_T^*(f) = f$. On peut alors écrire

$$\begin{aligned}\langle U_T(f) - f, U_T(f) - f\rangle &= \langle U_T(f), U_T(f)\rangle - \langle f, U_T(f)\rangle - \langle U_T(f), f\rangle + \langle f, f\rangle \\ &= \langle f, f\rangle - \langle U_T^*(f), f\rangle - \langle f, U_T^*(f)\rangle + \langle f, f\rangle = 0\end{aligned},$$

ce qui prouve que $U_T(f) = f$. On a prouvé l'inclusion inverse $\mathrm{Im}(\mathrm{Id} - U_T)^\perp \subset L^2_{\mathrm{inv}}(\mu)$.

Supposons donc que $f \in L^2_{\mathrm{inv}}(\mu)^\perp$. Pour tout $\varepsilon > 0$, on peut trouver $g \in L^2(\mu)$ tel que $\|f - (U_T(g) - g)\| \le \varepsilon$. Ceci implique que

$$\frac{1}{n}\left\|\sum_{i=0}^{n-1} U_T^i(f) - U_T^n(g) - g\right\| \le \varepsilon,$$

et donc que

$$\frac{1}{n}\left\|\sum_{i=0}^{n-1} U_T^i(f)\right\| \le \frac{1}{n}(\|U_T^n(g)\| + \|g\|) + \varepsilon = \frac{2}{n}\|g\| + \varepsilon \le 3\varepsilon,$$

si n est grand. □

Lea preuve précédente se généralise à n'importe quelle isométrie U d'un espace de Hilbert $\mathcal{H}$. Pour tout $x \in \mathcal{H}$, on a

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} U^i(x) = p_{\mathrm{inv}}(x)$$

où $p_{\mathrm{inv}} : \mathcal{H} \to \mathrm{Ker}(\mathrm{Id} - \mathrm{U}) = \mathrm{Fix}(\mathrm{U})$ est la projection orthogonale. Dans le cas qui nous concerne, l'espace propre $\mathrm{Fix}(U_T)$ contient toujours les fonctions constantes et est réduit à cet espace, c'est-à-dire est de dimension un, si et seulement si le système est ergodique. Si c'est le cas, le théorème de Von Neumann nous dit que pour tout $f \in L^2(\mu)$, la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f)\right)_{n\geq 0}$ converge dans $L^2(\mu)$ vers $(\int f\, d\mu)\mathbf{1}$. En effet, $p_{\rm inv}(f)$ est une fonction constante puisqu'elle invariante par T. La mesure μ étant finie, f appartient à $L^1(\mu)$ et par l'inégalité de Cauchy-Schwartz, on sait que

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f) - p_{\rm inv}(f)\right\|_1 \leq \left\|\frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f) - p_{\rm inv}(f)\right\|_2.$$

On en déduit que la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f)\right)_{n\geq 0}$ converge dans $L^1(\mu)$ vers $p_{\rm inv}(f)$. On conclut que $p_{\rm inv}(f) = (\int f\, d\mu)\mathbf{1}$ du fait que

$$\int p_{\rm inv}(f)\, d\mu = \lim_{n\to+\infty}\int \frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f)\, d\mu = \int f\, d\mu.$$

On peut alors énoncer :

Proposition 1.2.2 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, tel que $\mu(X) = 1$. Les trois propriétés suivantes sont équivalentes :*

i) *le système est ergodique ;*

ii) *pour tout A et B dans $\mathcal{B}$, on a*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}(B)) = \mu(A)\mu(B).$$

iii) *pour tous f, g dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int f(g\circ T^i)\, d\mu = \left(\int f\, d\mu\right)\left(\int g\, d\mu\right).$$

Preuve. Commençons par prouver que **ii)** implique **i)**. Si **ii)** est vraie, alors pour tout ensemble invariant $A \in \mathcal{B}$, on a

$$\mu(A)^2 = \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}(A)) = \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A) = \mu(A),$$

ce qui implique que $\mu(A) = 0$ ou $\mu(A) = 1$. Le système est donc ergodique.

On déduit **ii)** de **iii)** en appliquant cette dernière propriété aux fonctions caractéristiques χ_A et χ_B.

Il reste à prouver que **i)** implique **iii)**. Supposons donc que **i)** soit vraie. Fixons f et g dans $L^2(\mu)$. D'après le théorème de Von Neumann, on sait que la suite $\frac{1}{n}\sum_{i=0}^{n-1} g\circ T^i$ converge dans $L^2(\mu)$ vers la fonction constante $\int g\, d\mu$. On en déduit que

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int f(g\circ T^i)\, d\mu = \lim_{n\to+\infty}\int f\left(\frac{1}{n}\sum_{i=0}^{n-1} g\circ T^i\right)\, d\mu = \int f(\int g\, d\mu)\, d\mu = \int f\, d\mu\int g\, d\mu.$$

$\square$

De façon analogue, on a:

Proposition 1.2.3 : *Soit $(X,\mathcal{B},\mu,T)$ un système dynamique mesuré tel que $\mu(X)=1$. Les deux propriétés suivantes sont équivalentes : :*

i) *le système est mélangeant;*

ii) *pour tous f, g dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty}\int f(g\circ T^n)\,d\mu=\left(\int f\,d\mu\right)\left(\int g\,d\mu\right).$$

Preuve. Pour montrer **i)** à partir de **ii)**, c'est-à-dire pour montrer que pour tous A, B dans $\mathcal{B}$, on a

$$\lim_{n\to+\infty}\mu(A\cap T^{-n}(B))=\mu(A)\mu(B)$$

il suffit d'appliquer **ii)** aux applications χ_A et χ_B. Supposons maintenant que **i)** est vrai. Fixons $A\in\mathcal{B}$ et considérons l'ensemble des fonctions $g\in L^2(\mu)$ telles que

$$\lim_{n\to+\infty}\int\chi_A(g\circ T^n)\,d\mu=\mu(A)\int g\,d\mu.$$

Il s'agit d'un espace vectoriel fermé. Puisqu'il contient les fonctions caractéristiques, d'après **i)**, il contient les fonctions étagées, et puisque les fonctions étagées sont denses dans $L^2(\mu)$, il est égal à $L^2(\mu)$. Fixons $g\in L^2(\mu)$ et considérons l'ensemble des fonctions $f\in L^2(\mu)$, telles que

$$\lim_{n\to+\infty}\int f(g\circ T^n)\,d\mu=\left(\int f\,d\mu\right)\left(\int g\,d\mu\right).$$

Il s'agit la encore d'un espace vectoriel fermé. On vient de prouver qu'il contient les fonctions caractéristiques, on en déduit comme précédemment qu'il est égal à $L^2(\mu)$. $\square$

Le caractère mélangeant du système se traduit par le fait que pour tout $f\in L^2(\mu)$ et tout $g\in L^2(\mu)$, on a

$$\lim_{n\to+\infty}\langle U_T^n(f),g\rangle=\lim_{n\to+\infty}\int(f\circ T^n)\,\overline{g}\,d\mu=\left(\int f\,d\mu\right)\left(\int\overline{g}\,d\mu\right)=\langle p_{\mathrm{inv}}(f),g\rangle,$$

c'est-à-dire par la convergence faible de la suite $(U_T^n(f))_{n\geq 0}$ vers $(\int f\,d\mu)\mathbf{1}$. Le caractère ergodique se traduit par la convergence faible au sens de Césaro de la suite $(U_T^n(f))_{n\geq 0}$ vers $(\int f\,d\mu)\mathbf{1}$, mais cette convergence faible provient de la convergence forte (c'est-à-dire dans $L^2(\mu)$) de la suite des moyennes.

Considérons maintenant le mélange faible.

Proposition 1.2.4 : *Soit $(X,\mathcal{B},\mu,T)$ un système dynamique mesuré tel que $\mu(X)=1$. Les trois conditions suivantes sont équivalentes :*

i) *le système est faiblement mélangeant;*

ii) *pour tous A, B dans $\mathcal{B}$, on a*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(A\cap T^{-i}(B))-\mu(A)\mu(B)\right|=0\,;$$

iii) *pour tous f, g dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \left| \int f(g\circ T^i)\, d\mu - \left(\int f\, d\mu\right)\left(\int g\, d\mu\right) \right| = 0.$$

Preuve. Le fait que les propriétés ii) et iii) sont équivalentes se prouve comme dans la proposition 1.2.3. Pour montrer que i) et ii) sont équivalentes, il suffit de prouver le lemme suivant

LEMME 1.2.5 : *Soit $(u_n)_{n\geq 0}$ une suite réelle bornée. Alors les conditions suivantes sont équivalentes :*

i) $\displaystyle\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} |u_i| = 0$*;*

ii) *il existe une partie $I\subset \mathbf{N}$ de densité 0 telle que* $\displaystyle\lim_{n\to+\infty,\, n\notin I} u_n = 0.$

Preuve. Pour toute partie $J\subset\mathbf{N}$ et tout $n\geq 1$, on notera $\alpha_J(n) = \#(\{0,\ldots,n-1\}\cap J)$. Commençons par prouver que **i)** implique **ii)**. Pour tout $r\geq 1$, l'ensemble

$$I_r = \{n\geq 0 \,|\, |a_n| \geq \frac{1}{r}\}$$

est de densité 0 puisque

$$\frac{1}{n}\sum_{i=0}^{n-1}|u_i| \geq \frac{1}{r}\frac{\alpha_{I_r}(n)}{n}.$$

On peut donc construire dans $\mathbf{N}$ une suite strictement croissante $(n_r)_{r\geq 0}$ telle que pour tout $n\geq n_r$, on ait

$$\frac{\alpha_{I_{r+1}}(n)}{n} \leq \frac{1}{r+1}.$$

Posons

$$I = \bigcup_{r\geq 0} (I_{r+1}\cap [n_r, n_{r+1}[).$$

Il s'agit d'un ensemble de densité zéro car, pour tout $n\in [n_r, n_{r+1}[$, on a

$$I\cap [0,n[\subset I_{r+1}\cap [0,n[,$$

ce qui implique

$$\frac{\alpha_I(n)}{n} \leq \frac{\alpha_{I_{r+1}}(n)}{n} \leq \frac{1}{r+1}.$$

Remarquons maintenant que si $n\geq n_r$ n'appartient pas à I, alors n n'appartient pas à I_{r+1} et donc on a $|a_n|\leq 1/(r+1)$. On a bien montré que $\displaystyle\lim_{n\to+\infty,\, n\notin I} u_n = 0$.

Montrons maintenant que **ii)** implique **i)**. Soit $K>0$ un majorant de la suite $(|u_n|)_{n\geq 0}$. Pour tout $\varepsilon>0$, il existe $n_0\geq 0$ tel que pour tout $n\geq n_0$, on a $\alpha_I(n)\leq \varepsilon n$ et de plus $|u_n|\leq\varepsilon$ si $n\notin I$. On en déduit que pour tout $n\geq \max(n_0, Kn_0/\varepsilon)$, on a

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}|u_i| &= \frac{1}{n}\left(\sum_{i<n,\, i\in I}|u_i| + \sum_{i<n_0,\, i\notin I}|u_i| + \sum_{n_0\leq i<n,\, i\notin I}|u_i|\right)\\ &\leq \frac{K\alpha_I(n)}{n} + K\frac{n_0}{n} + \varepsilon\\ &\leq (K+2)\varepsilon.\end{aligned}$$

$\square$

Le lemme 1.2.5 admet le corollaire évident suivant :

COROLLAIRE 1.2.6 : *Soit $(u_n)_{n\geq 0}$ une suite réelle bornée. Alors les conditions suivantes sont équivalentes :*

i) $\lim\limits_{n\to+\infty} \frac{1}{n}\sum\limits_{i=0}^{n-1} |u_i| = 0;$

i) $\lim\limits_{n\to+\infty} \frac{1}{n}\sum\limits_{i=0}^{n-1} u_i^2 = 0;$

On en déduit :

COROLLAIRE 1.2.7 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Les trois conditions suivantes sont équivalentes :*

i) *le système est faiblement mélangeant*

ii) *pour tous A, B dans $\mathcal{B}$, on a*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Big(\mu(A\cap T^{-i}(B)) - \mu(A)\mu(B)\Big)^2 = 0\,;$$

ii) *pour tous f, g dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \left|\int f(g\circ T^i)\,d\mu - \left(\int f\,d\mu\right)\left(\int g\,d\mu\right)\right|^2 = 0.$$

On va conclure par cette caractérisation fondamentale des systèmes faiblement mélangeants:

PROPOSITION 1.2.8 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré inversible tel que $\mu(X) = 1$. Le système est faiblement mélangeant si et seulement si 1 est la seule valeur propre de U_T et $\dim(\mathrm{Ker}(\mathrm{Id} - U_T)) = 1$.*

Preuve. Si le système est faiblement mélangeant, il est ergodique et on sait donc que $\dim(\mathrm{Ker}(\mathrm{Id}-U_T)) = 1$. Montrons qu'il n'y a pas d'autre valeur propre que 1. Soit $\lambda \in \mathbf{C}\setminus\{1\}$ et $f \in L^2(\mu)$ tel que $U_T(f) = \lambda f$. Notons $\mathbf{1}$ la fonction constante égale à 1. Puisque

$$\langle f, \mathbf{1}\rangle = \langle U_T(f), U_T(\mathbf{1})\rangle = \lambda\langle f, \mathbf{1}\rangle,$$

on sait que $\langle f, \mathbf{1}\rangle = 0$; puisque

$$\langle f, f\rangle = \langle U_T(f), U_T(f)\rangle = \lambda\overline{\lambda}\langle f, f\rangle,$$

on sait que $f = 0$ si $|\lambda| \neq 1$. Supposons donc que $|\lambda| = 1$. Le fait que T est faiblement mélangeant implique que

$$0 = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |\langle U_T(f), f\rangle - |\langle f, \mathbf{1}\rangle|^2\,| = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |\langle U_T(f), f\rangle| = \langle f, f\rangle,$$

et donc que $f = 0$. Ainsi, si le système est faiblement mélangeant, il n'y a pas de valeur propre autre que 1.

La proposition réciproque sera basée sur le théorème spectral pour les opérateurs unitaires.

THÉORÈME 1.2.9 : *Soit U un opérateur unitaire sur un espace de Hilbert $\mathcal{H}$. Pour tout $x \in \mathcal{H}$, il existe une unique mesure borélienne de probabilité ν_x sur le cercle $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ telle que pour tout $k \in \mathbf{Z}$, on a*

$$\langle U^k(x), x\rangle = \int z^k \, d\nu_x(z).$$

L'adhérence V du sous-espace vectoriel engendré par la famille $(U^k(x))_{k\in\mathbf{Z}}$ est alors invariante par U et $U|_V$ est conjugué à l'opérateur ψ défini sur $L^2(\nu_x)$ qui à $g \in L^2(\nu_x)$ associe l'application $z \mapsto zg(z)$.

Fin de la preuve de la proposition 1.2.8. Supposons que 1 soit la seule valeur propre de U_T et que l'espace propre associé soit de dimension 1. Il faut montrer que pour tout $f \in L^2(\mu)$ et $g \in L^2(\mu)$, on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \langle U_T^k(f), g\rangle - \langle f, \mathbf{1}\rangle \overline{\langle g, \mathbf{1}\rangle} \right| = 0.$$

D'après le lemme 3.3.8 il faut montrer que

$$\lim_{n\to\infty} \langle U_T^n(f), g\rangle = \langle f, \mathbf{1}\rangle \overline{\langle g, \mathbf{1}\rangle}$$

pour n variant dans un ensemble de densité 1. Puisque ceci est évidemment vrai si f ou g est constante, il suffit par linéarité de montrer cette égalité si f et g sont dans l'orthogonal de $\mathrm{Fix}(U_T)$, égalité qui devient

$$\lim_{n\to\infty} \langle U_T^n(f), g\rangle = 0.$$

Par linéarité, il suffit de démonter que pour tout $f \in \mathrm{Fix}(U_T)^\perp$, on a

$$\lim_{n\to\infty} \langle U_T^n(f), f\rangle = 0,$$

pour n variant dans un ensemble de densité 1. En effet, soient f et g dans l'orthogonal de $\mathrm{Fix}(U_T)$. On a

$$\lim_{n\to\infty} \langle U_T^n(f+g), f+g\rangle = \lim_{n\to\infty} \langle U_T^n(f+ig), f+ig\rangle = \lim_{n\to\infty} \langle U_T^n(f), f\rangle = \lim_{n\to\infty} \langle U_T^n(g), g\rangle = 0,$$

pour n variant dans un ensemble de densité 1. On en déduit que

$$\lim_{n\to\infty} \langle U_T^n(g), f\rangle + \lim_{n\to\infty} \langle U_T^n(f), g\rangle = i \lim_{n\to\infty} \langle U_T^n(g), f\rangle - i \lim_{n\to\infty} \langle U_T^n(f), g\rangle = 0,$$

et donc

$$\lim_{n\to\infty} \langle U_T^n(g), f\rangle = \lim_{n\to\infty} \langle U_T^n(f), g\rangle = 0,$$

pour n variant dans cet ensemble.

Toujours d'après le lemme 3.3.8, il suffit donc de montrer que pour tout $f \in \mathrm{Fix}(U_T)^\perp$, on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle U_T^n(f), f\rangle|^2 = 0.$$

Notons ν_f la mesure spectrale associée à f par le théorème 3.5.2. Commençons par montrer qu'elle n'a pas d'atome. En effet, si $\nu_f(\{z_0\}) \neq 0$, l'opérateur ψ qui à $g \in L^2(\nu_f)$ associe l'application $z \mapsto zg(z)$ a z_0 comme valeur propre puisque $\psi(\chi_{\{z_0\}}) = z_0\chi_{\{z_0\}}$ et $\chi_{\{z_0\}} \neq 0$. Par le théorème 3.5.2. ll doit en être de même de la restriction de U_T à l'adhérence V de l'espace engendré par les itérés $U_T^k(f)$, $k \in \mathbf{Z}$. Ceci est impossible puisque cet espace est dans l'orthogonal de $\mathbf{1}$. On veut donc prouver que

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1}\left|\int z^k\, d\nu_f(z)\right|^2 = 0.$$

On écrit

$$\begin{aligned} \frac{1}{n}\sum_{k=0}^{n-1}\left|\int z^k\, d\nu_f(z)\right|^2 &= \frac{1}{n}\sum_{k=0}^{n-1}\left(\int z^k\, d\nu_f(z)\right)\left(\int \overline{z'}^k\, d\nu_f(z')\right) \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int\int (z\overline{z'})^k\, d(\nu_f\times\nu_f)(z,z') \\ &= \int\int \frac{1}{n}\left(\sum_{k=0}^{n-1}(z\overline{z'})^k\right)\, d(\nu_f\times\nu_f)(z,z') \end{aligned}.$$

Or, on sait que si $z \neq z'$, on a

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1}(z\overline{z'})^k = \lim_{n\to+\infty}\frac{1}{n}\left(\frac{1-(z\overline{z'})^n}{1-z\overline{z'}}\right) = 0.$$

Puisque ν_f n'a pas d'atome on en déduit que

$$\lim_{n\to+\infty} \frac{1}{n}\left(\sum_{k=0}^{n-1}(z\overline{z'})^k\right) = 0$$

presque partout (pour $\nu_f \times \nu_f$). Cette suite étant bornée en module par 1, on peut appliquer le théorème de convergence dominée de Lebesgue. □

Le théorème 1.2.8 nous donne une caractérisation spectrale du mélange faible, le système U_T restreint à l'orthogonal des constantes n'a pas de valeur propre : on dit que le système est à *spectre continu.* Le cas opposé est le cas où il existe une base hilbertienne de $L^2(\mu)$ formée de vecteurs propres, on dit alors que le système est à *spectre purement ponctuel.* C'est le cas par exemple des rotations sur $\mathbf{T}$, où même sur $\mathbf{T}^r$ (avec la mesure de Haar), puisque la famille $(e_k)_{k\in\mathbf{Z}^r}$, où

$$e_k(x+\mathbf{Z}^r) = e^{2i\pi\langle k,x\rangle},$$

est formée de vecteurs propres.

On peut également écrire la proposition 1.2.8 sous la forme équivalente suivante :

PROPOSITION 1.2.10 : *Soit $(X,\mathcal{B},\mu,T)$ un système dynamique mesuré inversible ergodique tel que $\mu(X)=1$. Le système est faiblement mélangeant si et seulement s'il ne se factorise sur aucune rotation non triviale de $\mathbf{T}$ (munie d'une mesure invariante).*

Preuve. Si le système est faiblement mélangeant, il ne se factorise sur aucune rotation non triviale de $\mathbf{T}$ puisque ces rotations ne sont pas faiblement mélangeantes. Remarquons également que si T est une rotation d'angle $\alpha \in \mathbf{T}^r$, et si $h : X \to \mathbf{T}$, est la semi-conjugaison, alors les fonctions $e_k \circ h$ sont des vecteurs propres de U_T. Réciproquement supposons que le système ne soit pas faiblement mélangeant. Puisqu'il est ergodique, ceci signifie que U_T admet une valeur propre $\lambda = e^{2i\pi\alpha} \in S^1$ différente de 1. Soit $f : X \to \mathbf{C}$ une fonction propre. On a $f(T(x)) = \lambda f(x)$ pour presque tout $x \in X$. On en déduit que la fonction $|f|$ est invariante par T et donc constante. Puisque f est non nulle, on peut supposer que f est à valeurs dans S^1. La multiplication par λ sur S^1, munie de la mesure $f_*(\mu)$ est un facteur de $(X, \mathcal{B}, \mu, T)$, l'application f étant une semi-conjugaison. Ce sytème dynamique est naturellement conjugué à la rotation d'angle $\alpha + \mathbf{Z}$ sur $\mathbf{T}$. Notons que si α est irrationnel, alors la mesure image est la mesure de Haar. Ce n'est pas le cas nécessairement si α est rationnel. pesons par exemple au cas où T est une rotation non triviale de $\mathbf{Z}/p\mathbf{Z}$, $p \geq 2$. □

1.3 Entropie métrique

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

1.3.1 Entropie d'une partition

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

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

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

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

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

Considérons maintenant la fonction

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

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

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

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

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

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

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

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

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

Énonçons les principales propriétés de l'entropie.

Proposition 1.3.1 : *Supposons que $\mathcal{P} = (P_i)_{i\in I}$, $\mathcal{Q} = (Q_j)_{j\in J}$ et $\mathcal{S} = (S_k)_{k\in K}$ sont des partitions mesurables. On a alors les propriétés suivantes :*

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

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

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

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

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

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

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

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

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

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

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

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

En sommant sur i et k on obtient

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

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

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

Pour établir **iii)**, écrivons

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

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

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

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

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

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

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

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

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

L'assertion **vi)** se déduit de

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

et l'assertion **vii)** de

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

$\square$

On en déduit immédiatement :

COROLLAIRE 1.3.2 : *On obtient une distance D sur l'ensemble des classes d'équivalence de la relation $\sim$, appelée* distance de Rokhlin*, en posant :*

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

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

Concluons cette section par un résultat qui sera très utile plus tard. Pour toute suite croissante $(\mathcal{P}_m)_{m\geq 0}$ de partitions mesurables de X, notons $\bigvee_{m=0}^{+\infty} \mathcal{P}_m$ la σ-algèbre engendrée par $\bigcup_{m\geq 0} \mathcal{P}_m$ et $\overline{\bigvee_{m=0}^{+\infty} \mathcal{P}_m}$ la σ-algèbre formée des éléments $A \in \mathcal{B}$ tels qu'il existe $B \in \bigvee_{m=0}^{+\infty} \mathcal{P}_m$ vérifiant $\mu(A\Delta B) = 0$. On dira que la famille $(\mathcal{P}_m)_{m\geq 0}$ est *génératrice* si $\overline{\bigvee_{m=0}^{+\infty} \mathcal{P}_m} = \mathcal{B}$.

PROPOSITION 1.3.3 : *Si $(\mathcal{P}_m)_{m\geq 0}$ est une suite croissante génératrice de X, alors pour toute partition mesurable $\mathcal{P}$ on a $\lim_{m\to+\infty} H(\mathcal{P}, \mathcal{P}_m) = 0$.*

Démonstration. Commençons par quelques lemmes.

LEMME 1.3.4 : *Pour tout $A \in \mathcal{B}$, il existe un entier $m \geq 0$ et un ensemble $B \in \mathcal{B}$, réunion d'éléments de $\mathcal{P}_m$, tels que $\mu(A\Delta B) < \varepsilon$.*

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

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

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

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

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

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

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

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

ce qui implique que

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

□

LEMME 1.3.5 : *Pour tout $r \geq 1$ et tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour toutes partitions mesurables $\mathcal{P} = (P_i)_{i\in I}$, $\mathcal{Q} = (Q_i)_{i\in I}$ vérifiant $\#I = r$ et $\mu(P_i\Delta Q_i) \leq \eta$, on a $H(\mathcal{P}|\mathcal{Q}) \leq \varepsilon$.*

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

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

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

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

Remarquons maintenant que

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

Le fait que ϕ est concave implique que

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

On a donc

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

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

Démonstration de la proposition 1.3.3. Soit $\mathcal{P} = (P_i)_{1\leq i\leq r}$ une partition mesurable. Donnons nous $\varepsilon > 0$. Commençons par considèrer le réel η defini par le lemme 1.3.5. Fixons ensuite δ suffisamment petit pour que l'on ait

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

et

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

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

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

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

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

ce qui implique que

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

On en déduit que l'ensemble

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

vérifie

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

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

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

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

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

ce qui implique que

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

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

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

et donc que

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

□

1.3.2 Entropie d'un système dynamique

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

Proposition 1.3.6 : *Si $\mathcal{P}$ est une partition mesurable de X, alors la suite*

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

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

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

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

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

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

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

Ceci implique que

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

et donc que

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

□

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

Proposition 1.3.7 : *Si $\mathcal{P}$ est une partition mesurable de X, alors la suite*

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

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

Démonstration. La décroissance de cette suite est une conséquence de l'assertion **v)** de la proposition 1.3.1. Notons l sa limite. Remarquons que

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

ce qui implique

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

Par le théorème de Césaro, on obtient

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

□

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

Proposition 1.3.8 : *Soient $\mathcal{P}$ et $\mathcal{Q}$ des partitions mesurables. On a alors les propriétés suivantes :*

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

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

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

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

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

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

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

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

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

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

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

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

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

ce qui implique

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

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

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

ce qui implique

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

L'assertion **v)** se déduit de

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

et l'assertion **vi)** de

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

□

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

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

PROPOSITION 1.3.9 : *On a les propriétés suivantes :*

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

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

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

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

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

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

ce qui nous donne

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

puis

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

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

□

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

PROPOSITION 1.3.8 : *Soient $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, tels que $\mu(X) = \nu(Y) = 1$. Supposons qu'il existe une application mesurable $R : X \to Y$ telle que $R \circ T = S \circ R$ et $R_*(\mu) = \nu$. Alors $h_\nu(S) \leq h_\mu(T)$. De plus il y a égalité si R est inversible et R^{-1} est mesurable.*

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

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

ce qui implique

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

On en déduit

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

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

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

1.3.3 Partitions génératrices

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

PROPOSITION 1.3.9 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Si $(\mathcal{P}_m)_{m \geq 0}$ est une suite croissante génératrice de partitions mesurables, alors*

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

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

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

puis

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

□

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

THÉORÈME 1.3.10 : *Soit $(X,\mathcal{B},\mu,T)$ un système dynamique mesuré tel que $\mu(X) = 1$ et $\mathcal{P}$ une partition mesurable. Si $\mathcal{P}$ est fortement génératrice, alors $h_\mu(T) = h(T,\mathcal{P})$. On a la même égalité si T est inversible et si $\mathcal{P}$ est génératrice.*

Démonstration. Nous allons utiliser la proposition 4.4.1 ainsi que l'assertion **iii)** de la proposition 4.3.3. Écrivons

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

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

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

□

1.3.4 Exemples

Nous allons calculer des entropies à l'aide notamment du théorème 1.3.10.

Le décalage de Bernouilli

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

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

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

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

On en déduit que

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

Remarques

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

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

Supposons maintenant que μ est une mesure de Markov définie par une matrice stochastique $M = (M_{i,j})_{i,j}$ et par un vecteur fixe $v = (v_1, \dots, v_r)$ à coefficients positifs tels que $\sum_{i=1}^r v_i = 1$. Pour tout $w = (w_0, \dots, w_{m-1}) \in \{1, \dots, r\}^m$ et tout $n_0 \in \mathbf{N}$, on a

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

On sait que

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

Utilisant le fait que M est stochastique et l'égalité $vM = M$, le calcul donne alors

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

et donc

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

Remarques

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

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

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

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

Endomorphismes linéaires de T.

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

Rotations du cercle.

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

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

La dernière égalité provient du fait que $T^{-1}(\mathcal{P})$ est également fortement génératrice, puisque c'est le cas de $\mathcal{P}$.

Remarquons que nous avons le résultat plus général suivant :

PROPOSITION 1.3.11 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré inversible tel que $\mu(X) = 1$. S'il existe une partition fortement génératrice, alors $h(T) = 0$.*

1.3.5 Entropie topologique et entropie métrique

Si $T : X \to X$ est une application continue définie sur un espace topologique compact métrisable, 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 1.3.12 : *Si $T : X \to X$ est une application continue définie sur un espace topologique compact métrisable, 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 1.3.13 : *Une mesure borélienne de probabilité sur un espace topologique compact métrisable est régulière : pour tout borélien A et tout $\varepsilon > 0$, il existe une partie fermée F et une partie ouverte U telle que $F \subset A \subset U$ et $\mu(U \setminus F) < \varepsilon$.*

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

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

et que

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

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

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

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

Nous devons montrer maintenant que toute partie fermée F appartient à $\mathcal{C}$. Considérons une distance d définissant la topologie de X et 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

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

$\square$

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

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

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

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

Ceci implique que

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

Pour tout $i \in \{1,\dots,r\}$ 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). Pour un recouvrement donné $\mathcal{V}$, notons $N(\mathcal{V})$ le plus petit cardinal de tous les sous-recouvrements de $\mathcal{V}$. 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 1.3.14 : *Soit X un espace métrique compact et $(\mu_m)_{m\geq 0}$ une suite de mesures boréliennes de probabilité qui converge vers μ pour la topologie faible*. Alors, pour tout borélien A tel que $\mu(\partial A) = 0$, on a*

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

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

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

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

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

Ceci implique que

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

Le même raisonnement appliqué au complémentaire de A nous dit que

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

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

LEMME 1.3.15 : *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 $\bigcup_{1\leq i\leq r} \partial U_i$. $\square$

PROPOSITION 1.3.16 : *Pour tout $\varepsilon > 0$, il existe $\mu \in \mathcal{M}_T$ tel que*

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

Avant de prouver cette proposition, expliquons pourquoi elle implique la seconde partie du théorème 1.3.12. Puisque

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

on sait que pout tout $\alpha > 0$, il existe $\varepsilon > 0$, tel que $\limsup_{n\to+\infty} \frac{1}{n} \ln s(n,\varepsilon) \geq h(T) - \alpha$ (où $1/\alpha$ si $h(T) = +\infty$) et donc par la proposition 1.3.16, il existe $\mu \in \mathcal{M}_T$ tel que $h_\mu(T) \geq h(T) - \alpha$ (ou $\geq 1/\alpha$ si $h(T) = +\infty$).

Démonstration de la proposition 1.3.16. Pour tout $n \geq 1$, choisissons un ensemble (n, ε)-séparé S_n de cardinal $s(n, \varepsilon)$. Considérons ensuite les mesures

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

et

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

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

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

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

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

puisque

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

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

Lemme 1.3.17 : *Pour toute partition borélienne $\mathcal{P} = (P_i)_{i \in I}$, on a*

$$H_{\mu_n}(\mathcal{P}) \geq \frac{1}{n} \sum_{k=0}^{n-1} H_{\nu_n}(T^{-k}(\mathcal{P})).$$

Démonstration. Puisque la fonction $\phi : t \mapsto -t \ln t$ est concave sur $[0, 1]$, on a:

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

□

Lemme 1.3.18 : *Pour toute partition borélienne $\mathcal{P} = (P_i)_{i\in I}$ et tous entiers $q \le n$, on a*

$$qH_{\nu_n}(\mathcal{P}^n) \le nH_{\mu_n}(\mathcal{P}^q) + 2q^2 \ln(\sharp I),$$

où $\mathcal{P}^n = \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{P})$.

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

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

où j_r est l'entier tel que

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

On obtient

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

En sommant sur r, on obtient

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

Il ne reste plus qu'à utiliser le lemme 1.3.17. $\square$

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

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

ce qui prouvera le lemme puisque $h_\mu(T) \ge h_\mu(T,\mathcal{P})$. Puisque S_n est (n,ε)-séparé, chaque élément de $\mathcal{P}^n$ contient au plus un point de S_n, ce qui implique que

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

Si on applique le lemme 1.3.18 à la partition $\mathcal{P}_{n_k}$ et qu'on divise par qn_k, on obtient :

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

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

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

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

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

□

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

PROPOSITION 1.3.19 : *Soit $T : X \to X$ une application expansive définie sur un espace métrique compact X. Il existe $\mu \in \mathcal{M}_T$ tel que*

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

Démonstration. On vérifie facilement que le recouvrement $\mathcal{U}^{\varepsilon_1}$ par les boules ouvertes de rayon ε_1 est générateur, si $2\varepsilon_1$ est une constante d'expansivité. On effet, pour tout $\eta > 0$, il existe n tel que le diamètre des éléments de $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{U}^\varepsilon)$ est uniformément majoré par η. 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

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

On montre facilement que $N(n, \varepsilon) \leq s(n, \varepsilon)$. En effet, si S est un ensemble (n, ε) séparé de cardinal $s(n, \varepsilon)$, il est (n, ε) couvrant : c'est-à-dire que pour tout $x \in X$, il existe $y \in S$ tel que pour tout $k \in \{0, \dots, n-1\}$ on a $d(T^k(x), T^k(y)) < \varepsilon$. On en déduit que les ensembles $\bigcap_{0\leq k<n} T^{-k}(B(T^k(y), \varepsilon))$, $y \in S$, recouvrent X. On a donc

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

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

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

□

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, c'est-à-dire celle pour laquelle tout cylindre

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

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

2.1. Notations

Soit F un homéomorphisme de $\mathbf{T} = \mathbf{R}/\mathbf{Z}$. Fixons un relèvement f de F à $\mathbf{R}$, c'est-à-dire une application continue $f : \mathbf{R} \to \mathbf{R}$ vérifiant $\pi \circ f = F \circ \pi$, où $\pi : x \mapsto x + \mathbf{Z}$ est la projection de revêtement. L'application $x \mapsto f(x+1) - f(x)$ est à valeurs entières, et constante puisque f est continue : il existe donc $k \in \mathbf{Z}$ tel que $f(x+1) = f(x) + k$, pour tout $x \in \mathbf{R}$. On sait, puisque F est injective, que $f|_{[0,1[}$ est injective et que son image ne contient pas deux points de différence entière. On en déduit que $k = \pm 1$. Dans le cas où $k = 1$, f est un homéomorphisme croissant qui vérifie $f(x+1) = f(x) + 1$, pour tout $x \in \mathbf{R}$; dans le cas où $k = -1$ f est un homéomorphisme décroissant qui vérifie $f(x+1) = f(x) - 1$, pour tout $x \in \mathbf{R}$. Le premier cas a lieu quand F préserve l'orientation (son degré est 1), le second cas a lieu quand F renverse l'orientation (son degré est -1). Nous nous intéresserons principalement au premier cas et noterons alors $\mathrm{Homeo}_+(\mathbf{T})$ le groupe des homéomorphismes de $\mathbf{T}$ qui préservent l'orientation. Nous noterons $D^0(\mathbf{T})$ le groupe de tous les relèvements des éléments de $\mathrm{Homeo}_+(\mathbf{T})$, autrement dit le groupe des homéomorphismes croissants f tels que $f - \mathrm{Id}_{\mathbf{R}}$ est périodique, de période 1. Ce groupe contient bien sûr les translations $T_a : x \mapsto x + a$, où $a \in \mathbf{R}$. On écrira fréquemment $f + a$ au lieu de $T_a \circ f$, si $f \in D^0(\mathbf{T})$. On munira $D^0(\mathbf{T})$ de la distance suivante

$$d(f,g) = \max\left(\max_{x\in\mathbf{R}} |f(x) - g(x)|,\ \ \max_{x\in\mathbf{R}} |f^{-1}(x) - g^{-1}(x)|\right).$$

Proposition 2.1.1 : *On a les résultats suivants :*

i) *l'application* $(f,g) \mapsto f \circ g$ *est continue ;*

ii) *l'application* $f \mapsto f^{-1}$ *est continue ;*

iii) *l'espace* $(D^0(\mathbf{T}), d)$ *est complet.*

Preuve. L'assertion **ii)** est évidente puisque $d(f^{-1}, g^{-1}) = d(f,g)$. Prouvons **i)**. Supposons donc que $\lim_{n\to+\infty} f_n = f$ et $\lim_{n\to+\infty} g_n = g$ et montrons que $\lim_{n\to+\infty} f_n \circ g_n = f \circ g$. Fixons $\varepsilon > 0$. Il existe $N \geq 0$ tel que $d(f_n, f) \leq \varepsilon/2$ pour tout $n \geq N$. L'application f est uniformément continue puisque c'est la somme de $\mathrm{Id}_{\mathbf{R}}$ et d'une application périodique. Il existe donc $\alpha > 0$ tel que

$$|x - y| \leq \alpha \Rightarrow |f(x) - f(y)| \leq \varepsilon/2.$$

Il existe $N' \geq 0$ tel que $d(g_n, g) \leq \alpha$ pour tout $n \geq N'$. On en déduit que pour tout $n \geq \max(N, N')$ on a

$$|f_n \circ g_n(x) - f \circ g(x)| \leq |f_n \circ g_n(x) - f \circ g_n(x)| + |f \circ g_n(x) - f \circ g(x)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Le même argument nous dit que $|g_n^{-1} \circ f_n^{-1}(x) - g^{-1} \circ f^{-1}(x)| \leq \varepsilon$, si n est assez grand, ce qui implique que l'assertion **i)** est vraie. Pour prouver **iii)**, considérons une suite de Cauchy $(f_n)_{n\geq 0}$ pour d. Chacune des suites $(f_n)_{n\geq 0}$ et $(f_n^{-1})_{n\geq 0}$ est une suite de Cauchy pour la distance d', où

$$d'(f,g) = \max_{x\in\mathbf{R}} |f(x) - g(x)|,$$

et converge uniformément vers une application continue. Notons f la limite de la première suite et g la limite de la seconde. Les égalités $f(x+1) = f(x) + 1$ et $g(x+1) = g(x) + 1$ sont également vraies. Il reste à prouver que f et g sont inversibles et que $g = f^{-1}$. Or la preuve de **i)** nous dit en fait que

$$g \circ f = \lim_{n\to+\infty} f_n^{-1} \circ f_n = \mathrm{Id}_{\mathbf{R}}$$

et

$$f \circ g = \lim_{n\to+\infty} f_n \circ f_n^{-1} = \mathrm{Id}_{\mathbf{R}},$$

et donc que la suite $(f_n)_{n\geq 0}$ converge vers f pour la distance d. □

2.2 Nombre de rotation de Poincaré

Énonçons le résultat fondamental de cette section :

THÉORÈME 2.2.1 : *Soit $f \in D^0(\mathbf{T})$. Il existe $\rho \in \mathbf{R}$ tel que pour tout $x \in \mathbf{R}$ et tout $k \in \mathbf{Z}$, on a*

$$-1 < f^k(x) - x - k\rho < 1.$$

En particulier, ρ vérifie

$$\rho = \lim_{k\mapsto\pm\infty} \frac{f^k(x)}{k},$$

on l'appelle le nombre de rotation *de f et on le note $\rho(f)$. En d'autre termes, on a*

$$-1 < f^k(x) - T^k_{\rho(f)}(x) < 1$$

pour tout $x \in \mathbf{R}$ et tout $k \in \mathbf{Z}$.

Preuve du théorème. Commençons par écrire $f(x) = x + \varphi(x)$, où $\varphi\ : \mathbf{R} \to \mathbf{R}$ est périodique, de période 1.

LEMME 2.2.2 : *Pour tous réels x et y, on a $-1 < \varphi(y) - \varphi(x) < 1$.*

Preuve. Soit y' l'unique élément de $[x, x+1[$ tel que $y' - y \in \mathbf{Z}$. Des inégalités

$$x + \varphi(x) \leq y' + \varphi(y') < x + 1 + \varphi(x),$$

on déduit que

$$-1 = x - x - 1 < x - y' \leq \varphi(y') - \varphi(x) < x + 1 - y' \leq x + 1 - x = 1.$$

□

Si on applique le lemme précédent aux applications f^k, $k \geq 1$, on obtient

$$0 \leq M_k - m_k < 1,$$

où

$$m_k = \min_{x\in\mathbf{R}} f^k(x) - x, \ \ M_k = \max_{x\in\mathbf{R}} f^k(x) - x.$$

Pour tous entiers strictement positifs k et k', et pour tout réel x, on a

$$f^{k+k'}(x) - x = f^k(f^{k'}(x)) - f^{k'}(x) + f^{k'}(x) - x$$

et on a donc
$$m_k + m_{k'} \leq f^{k+k'}(x) - x \leq M_k + M_{k'},$$
ce qui implique
$$m_k + m_{k'} \leq m_{k+k'} \leq M_{k+k'} \leq M_k + M_{k'}.$$
Une récurrence évidente nous dit que pour tous entiers strictement positifs k et k', on a $M_{k'k} \leq k'M_k$ et $m_{k'k} \geq km_{k'}$ et donc que
$$\frac{m_{k'}}{k'} \leq \frac{m_{k'k}}{k'k} \leq \frac{M_{k'k}}{k'k} \leq \frac{M_k}{k}.$$
L'ensemble des nombres de la forme $\dfrac{m_k}{k}$ est donc à gauche de l'ensemble des nombres de la forme $\dfrac{M_k}{k}$. Puisque $\dfrac{M_k}{k} - \dfrac{m_k}{k} < \dfrac{1}{k}$, on en déduit que
$$\sup_{k\geq 1} \frac{m_k}{k} = \inf_{k\geq 1} \frac{M_k}{k}.$$
Si on note ρ cette borne commune, on a $m_k \leq k\rho \leq M_k$ pour tout $k \geq 1$. Il existe x_k et y_k tels que $f^k(x_k) - x_k = m_k$ et $f^k(y_k) - y_k = M_k$. Le théorème des valeurs intermédiaires nous dit qu'il existe z_k tel que $f^k(z_k) - z_k = k\rho$. Si on fixe $x \in \mathbf{R}$ et qu'on applique le lemme 2.2.2 à la fonction f^k et aux points x et z_k, on obtient
$$-1 < f^k(x) - x - k\rho < 1.$$
Le théorème 2.2.1 a donc été démontré pour $k \geq 1$. Il est évident si $k = 0$ et il reste à le prouver pour $k < 0$. Appliquons donc l'inégalité obtenue au point $f^k(x)$ pour l'itéré f^{-k}. On obtient
$$-1 < f^{-k}(f^k(x)) - f^k(x) + k\rho < 1,$$
ce qui implique
$$-1 < f^k(x) - x - k\rho < 1.$$
□

Remarque La preuve précédente nous donne les informations suivantes sur le nombre de rotation. Si $p \in \mathbf{Z}$ et $q \geq 1$ sont deux entiers, alors

i) $\rho(f) = p/q$ si et seulement s'il existe $x \in \mathbf{R}$ tel que $f^q(x) = x + p$;

ii) $\rho(f) > p/q$ si et seulement si $f^q(x) > x + p$ pour tout $x \in \mathbf{R}$;

iii) $\rho(f) < p/q$ si et seulement si $f^q(x) < x + p$ pour tout $x \in \mathbf{R}$.

Énonçons maintenant quelques propriétés du nombre de rotation.

PROPOSITION 2.2.3 : **i)** *Le nombre de rotation de T_a est a.*

ii) *Pour tout $f \in D^0(\mathbf{T})$ et tout $p \in \mathbf{Z}$, on a $\rho(f + p) = \rho(f) + p$.*

iii) *Pour tout $f \in D^0(\mathbf{T})$ et tout $q \in \mathbf{Z}$, on a $\rho(f^q) = q\rho(f)$.*

iv) *Si $f \leq g$, alors $\rho(f) \leq \rho(g)$.*

v) *L'application $f \mapsto \rho(f)$ est continue.*

Preuve. Pour obtenir **i)**, il suffit d'écrire

$$\rho(T_a) = \lim_{k\to+\infty} \frac{T_a^k(x)}{k} = \lim_{k\to+\infty} \frac{x+ka}{k} = a.$$

Pour montrer **ii)**, remarquons que

$$(f+p)^k = (T_p \circ f)^k = T_p^k \circ f^k$$

puisque f et T_p commutent. Ainsi, on a

$$\rho(f+p) = \lim_{k\to+\infty} \frac{(f+p)^k(x)}{k} = \lim_{k\to+\infty} \frac{f^k(x)+kp}{k} = \rho(f)+p.$$

Pour établir **iii)**, écrivons

$$\rho(f^q) = \lim_{k\to+\infty} \frac{(f^q)^k(x)}{k} = \lim_{k\to+\infty} \frac{f^{qk}(x)}{k} = q\rho(f).$$

Pour prouver **iv)**, commençons par montrer par récurrence que pour tout $k \geq 1$, on a $f^k \leq g^k$. L'inégalité est supposée vraie pour $k=1$. Si elle est vraie au rang k elle est également varie au rang $k+1$ car pour tout $x \in \mathbf{R}$ on a

$$f^{k+1}(x) = f(f^k(x)) \leq g(f^k(x)) \leq g(g^k(x)) = g^{k+1}(x).$$

On peut donc écrire

$$\rho(f) = \lim_{k\to+\infty} \frac{f^k(x)}{k} \leq \lim_{k\to+\infty} \frac{g^k(x)}{k} = \rho(g).$$

Il reste à montrer **v)**. Commençons par remarquer que les applications $f \mapsto f^k$ sont continues, puisque c'est le cas de l'application $(f,g) \to f \circ g$. Supposons donc que $\lim_{n\to+\infty} f_n = f$. Fixons $\varepsilon > 0$, puis choisissons $k \geq 1$ pour que $3 < k\varepsilon$. Fixons $x \in \mathbf{R}$. Il existe $N \geq 1$ tel que $-1 < f^k(x) - f_n^k(x) < 1$, pour tout $n \geq N$. Des inégalités

$$\begin{aligned} -1 &< f_n^k(x) - x - k\rho(f_n) < 1, \\ -1 &< -f^k(x) + x + k\rho(f) < 1, \\ -1 &< f^k(x) - f_n^k(x) < 1, \end{aligned}$$

on déduit

$$-3 < k(\rho(f) - \rho(f_n)) < 3,$$

ce qui implique

$$-\varepsilon < \rho(f) - \rho(f_n) < \varepsilon.$$

□

Remarques :

1. Remarquons que $\rho(f^{-1}) = -\rho(f)$ et que $\rho(f^q + p) = q\rho(f) + p$, pour tous entiers p et q.

2. L'assertion **ii)** implique que pour tout $F \in \mathrm{Homeo}_+(\mathbf{T})$, la classe $\rho(f) + \mathbf{Z} \in \mathbf{T}$ ne dépend pas du choix du relèvement f de F. C'est le *nombre de rotation* $\rho(F) \in \mathbf{T}$ de F.

3. La formule $\rho(f\circ g)=\rho(f)+\rho(g)$ est fausse en général (voir exercices). Cependant cette formule est vraie si f et g commutent. En effet, des inégalités

$$-1 < f^k\circ g^k(x) - g^k(x) - k\rho(f) < 1,$$
$$-1 < g^k(x) - x - k\rho(g) < 1,$$

on obtient

$$-2 < (f\circ g)^k(x) - x - k(\rho(f)+\rho(g)) < 2,$$

ce qui implique

$$\rho(f\circ g) = \lim_{k\to+\infty} \frac{(f\circ g)^k(x)}{k} = \rho(f)+\rho(g).$$

Exemple : Fixons $\alpha \in (-\frac{1}{2\pi}, \frac{1}{2\pi})$ et définissons la famille, indexée par $t\in\mathbf{R}$, d'applications :

$$f_t\ : x\mapsto x+\alpha\sin(2\pi x)+t.$$

On vérifie que $f_t\in D^0(\mathbf{T})$ et que $t\mapsto f_t$ est continue. La proposition 2.2.3 nous dit alors que l'application $r\ : t\mapsto \rho(f_t)$ est continue, croissante et vérifie $r(t+1)=r(t)+1$, pour tout $t\in\mathbf{R}$. On peut prouver que si $\alpha\neq 0$, alors l' intervalle $r^{-1}(\{a\})$ est réduit à un point si et seulement si $a\notin\mathbf{Q}$ (voir exercices).

2.3 Dynamique des homéomorphismes de nombre de rotation rationnel

Commençons par étudier les homéomorphismes dont le nombre de rotation est nul.

Proposition 2.3.1 : *Soit $f\in D^0(\mathbf{T})$. Alors $\rho(f)=0$ si et seulement f a un point fixe.*

Preuve. Si f a un point fixe x, alors $\rho(f)=\lim_{k\to+\infty}\frac{f^k(x)}{k}=0$. Réciproquement, nous avons vu dans la section précédente que pour tout $f\in D^0(\mathbf{T})$, il existe $x\in\mathbf{R}$ tel que $f(x)-x=\rho(f)$, ce qui implique la proposition. □

Donnons nous $f\in D^0(\mathbf{T})$ tel que $\rho(f)=0$. L'ensemble $\mathrm{Fix}(f)$ est une partie fermée invariante par $T\ : x\mapsto x+1$. Si $]a,b[$ est une composante connexe de $\mathbf{R}\setminus\mathrm{Fix}(f)$, deux cas sont possibles suivant que la fonction $f-\mathrm{Id}_{\mathbf{R}}$ est positive ou négative sur $]a,b[$. Dans le premier cas, pour tout $x\in]a,b[$, la suite $(f^k(x))_{k\in\mathbf{Z}}$ est croissante et vérifie

$$\lim_{k\to-\infty} f^k(x)=a,\quad \lim_{k\to+\infty} f^k(x)=b\,;$$

dans le second cas, elle est décroissante et vérifie

$$\lim_{k\to-\infty} f^k(x)=b,\quad \lim_{k\to+\infty} f^k(x)=a.$$

Remarquons que les points fixes de l'homéomorphisme F relevé par f sont les images par π : $x\mapsto x+\mathbf{Z}$ des points fixes de f. Les ensembles α-limite et ω-limite d'un point $\widehat{x}\notin\mathrm{Fix}(F)$ sont les extrémités, égales si et seulement si F n'a qu'un point fixe, de la composante connexe de $\mathbf{R}\setminus\mathrm{Fix}(F)$ qui contient $\widehat{x}$.

Passons maintenant au cas général. Nous allons voir qu'un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T})$ a un nombre de rotation rationnel si et seulement s'il admet une orbite périodique et que dans ce cas toutes les périodes des orbites périodiques sont égales. Plus précisément :

Proposition 2.3.2 : *Supposons que $F \in \mathrm{Homeo}_+(\mathbf{T})$ a un nombre de rotation rationnel $\rho(F) = p/q + \mathbf{Z}$, où $p \in \mathbf{Z}$ et $q \geq 1$ sont premiers entre-eux. Alors*

i) *F a une orbite périodique de période q ;*

ii) *toutes les orbites périodiques de F sont de période q ;*

iii) *pour tout $\widehat{x} \in \mathbf{T}$, les ensembles $\alpha(\widehat{x})$ et $\omega(\widehat{x})$ sont des orbites périodiques.*

Preuve. Soit $f \in D^0(\mathbf{T})$ le relèvement de F tel que $\rho(f) = p/q$. Puisque $\rho(f^q - p) = 0$, on sait que $f^q - p$ a un point fixe. On en déduit que F^q a un point fixe et que tout point fixe de F^q est la projection d'un point fixe de $f^q - p$. Si $x + \mathbf{Z}$ est un point périodique de F de période q', il existe $p' \in \mathbf{Z}$ tel que $f^{q'}(x) = x + p'$. Remarquons que $f^{qq'}(x) - qp' = x$, ce qui implique que $0 = \rho(f^{qq'} - qp') = q'p - qp'$. Ainsi, il existe $r \geq 1$ tel que $q' = rq$ et $p' = rp$. On en déduit que x est un point périodique de $f^q - p$ de période r. C'est donc un point fixe et on a $r = 1$. On vient d'établir **i)** et **ii)**. Pour prouver **iii)**, nous utilisons le fait que pour tout $\widehat{x} \in \mathbf{T}$, il existe un point fixe $\widehat{y}$ de F^q tel que $\lim_{k\to+\infty} F^{kq}(\widehat{x}) = \widehat{y}$. C'est un point périodique de F, de période q, et son orbite est égale à $\omega(\widehat{x})$. On montre de façon similaire que $\alpha(\widehat{x})$ est une orbite périodique de F.$\square$

2.4 Dynamique des homéomorphismes de nombre de rotation irrationnel

Le résultat principal est le résultat de semi-conjugaison suivant, dû à Poincaré :

Théoreme 2.4.1 : *Soit $F \in \mathrm{Homeo}_+(\mathbf{T})$. Si $\rho(F) \notin \mathbf{Q}/\mathbf{Z}$, alors, il existe une surjection continue $H : \mathbf{T} \to \mathbf{T}$ de degré 1, relevée par une application croissante $h : \mathbf{R} \to \mathbf{R}$, telle que $H \circ F = R_{\rho(F)} \circ H$, où $R_{\rho(F)} : \widehat{x} \mapsto \widehat{x} + \rho(F)$.*

Preuve. Fixons un relèvement $f \in D^0(\mathbf{T})$ de F et écrivons $\rho(f) = \rho$. Fixons également $x \in \mathbf{R}$. L'application

$$\begin{aligned} c\ :\ & \mathbf{Z}^2 \to \mathbf{R} \\ & (p,q) \mapsto q\rho - p \end{aligned}$$

est injective, puisque $\rho \notin \mathbf{Q}$, et son image est dense dans $\mathbf{R}$. L'application

$$\begin{aligned} c'\ :\ & \mathbf{Z}^2 \to \mathbf{R} \\ & (p,q) \mapsto f^q(x) - p \end{aligned}$$

est également injective. En effet

$$\begin{aligned} f^q(x) - p = f^{q'}(x) - p' &\Rightarrow f^q(x) - f^{q'}(x) = p - p' \\ &\Rightarrow f^{q-q'}(f^{q'}(x)) - f^{q'}(x) = p - p' \\ &\Rightarrow q = q' \text{ et } p = p', \end{aligned}$$

car $\rho \notin \mathbf{Q}$. Notons Z l'image de c et Z' celle de c'. Les ensembles Z et Z' sont invariants par $T : x \mapsto x + 1$, l'ensemble Z invariant par T_ρ et l'ensemble Z' invariant par f.

Lemme 2.4.2 : *L'application $h = c \circ c'^{-1}$ est une bijection strictement croissante de Z' sur Z telle que $h \circ T = T \circ h$ et $h \circ f = T_\rho \circ h$.*

Preuve. L'application h est bien sûr bijective et les égalités $h \circ T = T \circ h$ et $h \circ f = T_\rho \circ h$ sont évidentes. Pour montrer que h est croissante, il suffit de montrer que c'est le cas de h^{-1}. Supposons que l'on ait $q\rho - p < q'\rho - p'$, c'est-à-dire que $(q - q')\rho < p - p'$.

Dans le cas où $q = q'$, on en déduit que $p > p'$ puis que $f^q(x) - p < f^{q'}(x) - p'$.

Dans le cas où $q > q'$, alors on a $\rho < \frac{p-p'}{q-q'}$ ce qui implique que $f^{q-q'}(f^{q'}(x)) - f^{q'}(x) < p - p'$.

Dans le cas où $q < q'$, alors on a $\rho > \frac{p-p'}{q-q'}$ ce qui implique que $f^{q-q'}(f^{q'}(x)) - f^{q'}(x) > p - p'$.

□

L'image de h étant dense dans $\mathbf{R}$, on peut étendre h en une fonction définie sur $\mathbf{R}$ en posant

$$h(x) = \sup_{y \leq x,\, y \in Z'} h(y) = \inf_{y \geq x,\, y \in Z'} h(y).$$

On obtient une application croissante, dont l'image est dense. Elle est donc continue et surjective. Elle vérifie

$$h(x+1) = \sup_{y \leq x+1,\, y \in Z'} h(y) = \sup_{y \leq x,\, y \in Z'} h(y) + 1 = h(x) + 1$$

et

$$h(f(x)) = \sup_{y \leq f(x),\, y \in Z'} h(y) = \sup_{y \leq x,\, y \in Z'} h(f(y)) = \sup_{y \leq x,\, y \in Z'} h(y) + \rho = h(x) + \rho.$$

La première égalité nous dit que h relève une application $H\ : \mathbf{T} \to \mathbf{T}$ de degré 1 ; la seconde nous dit que $h \circ f = T_\rho \circ h$ et donc que $H \circ F = R_{\rho(F)} \circ H$. □.

Remarques :

1. Remarquons que la préimage $h^{-1}(\{x\})$ de tout point $x \in \mathbf{R}$ est un intervalle compact I. L'ensemble des points x tels que $h^{-1}(\{x\})$ n'est pas réduit à un point est une partie au plus dénombrable invariante par $T_{\rho(f)}$. On a des résultats similaires pour la préimage $H^{-1}(\{\widehat{x}\})$ d'un point $\widehat{x} \in \mathbf{T}$.

2. Si l'orbite de $x + \mathbf{Z}$ pour F est dense, alors Z' est dense dans $\mathbf{R}$. Ceci implique que h est un homéomorphisme Par conséquent H est un homéomorphisme et F est conjugué à $R_{\rho(F)}$. En d'autres termes, puisqu'un homéomorphisme de nombre de rotation rationnel n'est jamais transitif, nous en déduisons que pour un homéomorphisme du cercle F préservant l'orientation, les propriétés suivantes sont équivalentes :

- F est transitif ;
- F est minimal ;
- F est conjugué à une rotation d'angle irrationnel.

Continuons notre étude et expliquons plus précisément la dynamique des homéomorphismes $F \in \mathrm{Homeo}_+(\mathbf{T})$ de nombre de rotation irrationnel, c'est-à-dire des homéomorphismes sans point périodique.

Théorème 2.4.3 : *Supposons que $F \in \mathrm{Homeo}_+(\mathbf{T})$ n'a pas de point périodique. Il existe alors une partie fermée $X \subset \mathbf{T}$ vérifiant les propriétés suivantes :*

i) $X = \Omega(F)$ *;*

ii) *X est la seule partie minimale de F ;*

iii) *pour tout* $x \in \mathbf{T}$, *on a* $\alpha(x) = \omega(x) = X$ *;*

iv) *toute composante connexe de* $\mathbf{T} \setminus X$ *est errante ;*

v) *si* $X \neq \mathbf{T}$, *alors* X *est un ensemble de Cantor ;*

vi) *on a* $X = \mathbf{T}$ *si et seulement si* F *est conjugué à une rotation irrationnelle.*

Preuve. Soit $X' \subset \mathbf{T}$ une partie fermée invariante non vide. Si $X' \neq \mathbf{T}$, son complémentaire est une union d'intervalles ouverts et F induit une bijection $I \mapsto F(I)$ sur l'ensemble de ces intervalles. Aucun d'eux n'est périodique car F n'a pas de point périodique et que les extrémités d'un intervalle périodique devraient être périodiques. On en déduit que toute composante connexe I de $\mathbf{T} \setminus X'$ est errante. On a donc $\Omega(F) \subset X'$. Comme conséquence, on sait que l'ensemble $X = \Omega(F)$ est contenu dans toute partie fermée invariante non vide. C'est donc un ensemble minimal, et même le seul. Pour tout point $x \in \mathbf{T}$, on sait que les ensembles $\alpha(x)$ et $\omega(x)$ sont inclus dans $\Omega(F)$, mais puisqu'ils contiennent également X, il coïncident avec X. Montrons maintenant que X est un ensemble de Cantor si $X \neq \mathbf{T}$. Dans ce cas $\mathrm{Fr}(X)$ est une partie fermée invariante non vide. Elle contient donc X. Ceci signifie que X est totalement discontinu. Puisque que F n'a pas de point périodique, X est infini et admet donc des points d'accumulation. L'ensemble des points d'accumulation est fermé et invariant, c'est donc l'ensemble X tout entier : ce dernier ensemble n'a donc pas de point isolé. □

Remarques :

1. Dans le cas où l'application $H : \mathbf{T} \to \mathbf{T}$ donnée par le théorème 2.4.1 est un homéomorphisme, on sait que $X = \mathbf{T}$. Dans le cas contraire, X est l'ensemble des points $\widehat{x}$ tels que H n'est pas constante au voisinage de $\widehat{x}$. En effet, l'ensemble X' des points vérifiant cette propriété, étant fermé et invariant par F, doit contenir X. Pour prouver l'inclusion inverse, remarquons que $H(X)$ est une partie fermée invariante par $R_{\rho(F)}$, donc égale à $\mathbf{T}$. Par conséquent H est constante sur chaque composante connexe de $\mathbf{T} \setminus X$ (rappelons que h est croissante) ce qui implique que X contient X'. On peut être plus précis. Si on pose $I_{\widehat{y}} = H^{-1}(\{\widehat{y}\})$ pour tout $\widehat{y} \in \mathbf{T}$, on sait que $I_{T_{\widehat{\rho}}(\widehat{y})} = F(I_{\widehat{y}})$ puisque H est une semi-conjugaison. Ceci implique que les intervalles $\mathrm{Int}(I_{\widehat{y}})$ sont errants, s'ils ne sont pas vides. L'orbite d'un tel intervalle est ordonnée cycliquement comme l'orbite d'un point par la rotation $R_{\rho(F)}$. Ces intervalles sont exactement les composantes connexes de $\mathbf{T} \setminus X$.

2. Un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T})$ de nombre de rotation rationnel est uniquement ergodique si et seulement s'il n'a qu'une orbite périodique.

3. Un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T})$ de nombre de rotation irrationnel $\widehat{\rho}$ est uniquement ergodique. C'est évident dans le cas où F est conjugué à $R_{\rho(F)}$. Étudions l'autre cas, en notant H la semi-conjugaison de F sur $R_{\rho(F)}$. Soit $\mu \in M_F(\mathbf{T})$. On sait que $H_*(\mu)$ est la mesure de Haar μ_0 car $T_{\widehat{\rho}}$ est uniquement ergodique. Le théorème de récurrence de Poincaré nous dit que l'adhérence de toute composante connexe I du complémentaire de $X = \Omega(F)$ est de mesure nulle car elle est errante. On en déduit en particulier que le support de μ est dans X (et même égal à X car il est invariant). Ces intervalles sont les préimages de H qui ne sont pas réduites à un point et il y en a un nombre dénombrable. La réunion de ces intervalles s'écrit $H^{-1}(Y)$ où Y est l'ensemble dénombrable des points de $\mathbf{T}$ ayant plus d'un antécédent. On a $\mu_0(Y) = \mu(H^{-1}(Y)) = 0$. L'application H induit une bijection entre $\mathbf{T} \setminus H^{-1}(Y)$ et $\mathbf{T} \setminus Y$ qui transporte μ sur μ_0.

2.5 Difféomorphismes du cercle de nombre de rotation irrationnel

On peut se demander s'il est possible de construire des difféomorphismes du cercle sans point périodique qui ne sont pas conjugués à une rotation. De tels exemples, en classe C^1 ont été donnés par Denjoy. Par contre, comme nous allons le voir, de tels exemples sont impossibles à construire en classe de différentiabilité plus grande. Pour tout $r \geq 1$, définissons le groupe $\mathrm{Diff}^r_+(\mathbf{T})$ des difféomorphismes de classe C^r de $\mathbf{T}$ préservant l'orientation. La dérivée F' de $F \in \mathrm{Diff}^r_+(\mathbf{T})$ est une application de $\mathbf{T}$ dans $\mathbf{R}$. On dira qu'une application $\psi\ : \mathbf{T} \to \mathbf{R}$ est à *variation bornée* s'il existe $C > 0$ tel que pour toute famille $(\widehat{x}_i)_{i \in \mathbf{Z}/q\mathbf{Z}}$ cycliquement ordonnée sur le cercle, on a

$$\sum_{i \in \mathbf{Z}/q\mathbf{Z}} |\psi(\widehat{x}_{i+1}) - \psi(\widehat{x}_i)| \leq C.$$

C'est le cas, par exemple, si ψ est de classe C^1 puisqu'on a

$$\sum_{i \in \mathbf{Z}/q\mathbf{Z}} |\psi(\widehat{x}_{i+1}) - \psi(\widehat{x}_i)| \leq \sum_{i \in \mathbf{Z}/q\mathbf{Z}} C\widehat{d}(\widehat{x}_i, \widehat{x}_{i+1}) \leq C,$$

où $C = \max_{\widehat{x} \in \mathbf{T}} |\psi'(\widehat{x})|$.

Le résultat suivant est dû à Denjoy :

THÉORÈME 2.5.1 : *Si la dérivée de $F \in \mathrm{Diffeo}^1_+(\mathbf{T})$ est à variation bornée et si F n'a pas de point périodique, alors F est conjugué à une rotation d'angle irrationnel.*

Preuve. Remarquons d'abord que $\ln F'$ est à variation bornée. En effet, si on pose $m = \min_{\widehat{x} \in \mathbf{T}} F'(\widehat{x}) > 0$ et si on se donne une famille cycliquement ordonnée $(\widehat{x}_i)_{i \in \mathbf{Z}/q\mathbf{Z}}$, on a

$$\sum_{i \in \mathbf{Z}/q\mathbf{Z}} |\ln F'(\widehat{x}_{i+1}) - \ln F'(\widehat{x}_i)| \leq \frac{1}{m} \sum_{i \in \mathbf{Z}/q\mathbf{Z}} |F'(\widehat{x}_{i+1}) - F'(\widehat{x}_i)| \leq \frac{1}{m} \mathrm{Var}(F').$$

LEMME 2.5.2 : *Il existe une suite croissante $(q_n)_{n \geq 0}$ d'entiers positifs, telle que pour tout $n \geq 1$ et tout $\widehat{x} \in \mathbf{T}$ il existe un intervalle fermé I_n de $\mathbf{T}$ joignant $\widehat{x}$ à $F^{q_n}(\widehat{x})$ dont les itérés $F^k(I_n)$, $0 \leq k \leq q_n$, sont disjoints deux à deux.*

Preuve. Puisque F est semi-conjugué à une rotation d'angle irrationnel R par une application qui préserve l'ordre cyclique, il suffit de prouver le lemme dans le cas d'une rotation R et du point $\widehat{0}$. Écrivons alors $\widehat{x}_k = R^k(\widehat{0})$, si $k \in \mathbf{Z}$.

Posons $q_1 = 1$ et définissons par récurrence une suite $(q_n)_{n \geq 0}$ par la condition suivante

$$q_{n+1} = \inf\{q > q_n \ | \ d(\widehat{0}, \widehat{x}_q) < d(\widehat{0}, \widehat{x}_{q_n})\}.$$

Remarquons alors que

$$1 \leq q < q_n \Rightarrow d(\widehat{0}, \widehat{x}_{q_n}) < d(\widehat{0}, \widehat{x}_q).$$

Notons I_n l'intervalle joignant $\widehat{0}$ à $\widehat{x}_{q_n}$ dont le diamètre est $d(\widehat{0}, \widehat{x}_{q_n})$. Supposons qu'il existe des entiers q et q' vérifiant $0 \leq q < q' \leq q_n$, tels que $R^q(I_n) \cap R^{q'}(I_n) \neq \emptyset$. On doit alors avoir $R^{q''}(I_n) \cap I_n \neq \emptyset$, où $q'' = q' - q$, ce qui est impossible. En effet, le point $x_{q''}$ n'appartient pas à I_n car $d(\widehat{0}, \widehat{x}_{q''}) > d(\widehat{0}, \widehat{x}_{q_n})$. Pour les mêmes raisons, il n'appartient pas à $-I_{q_n}$, ce qui implique que $x_{q_n + q''}$ n'appartient pas à I_{q_n}. $\square$

LEMME 2.5.3 : *Il existe une constante $C > 0$ telle que pour tout $\widehat{x} \in \mathbf{T}$, on a*

$$C^{-1} \leq (F^{q_n})'(\widehat{x})(F^{-q_n})'(\widehat{x}) \leq C.$$

Preuve. Remplaçant $\widehat{x}$ par $F^{q_n}(\widehat{x})$, on doit trouver une borne supérieure à

$$|\ln\left((F^{q_n})'(F^{q_n}(\widehat{x}))(F^{-q_n})'(F^{q_n}(\widehat{x}))\right)|.$$

On peut écrire cette quantité sous la forme

$$\left|\sum_{i=0}^{q_n-1}\ln F'(F^{q_n+i}(\widehat{x}) - \sum_{i=0}^{q_n-1}\ln F'(F^i(\widehat{x})\right|.$$

Elle est inférieure à

$$\sum_{i=0}^{q_n-1}|\ln F'(F^{q_n+i}(\widehat{x})) - \ln F'(F^i(\widehat{x}))|,$$

quantité qui admet $\operatorname{Var}\ln(F')$ comme majorant, d'après le lemme 2.5.2. On peut donc prendre

$$C = e^{\operatorname{Var}\ln(F')}.$$

$\square$

Preuve du théorème 2.5.1. On doit prouver que F n'a pas d'intervalle errant. Supposons que I soit un tel intervalle et notons μ la mesure de Lebesgue. On a

$$\lim_{n\to+\infty}\mu(F^{q_n}(I)) + \mu(F^{-q_n}(I)) = 0.$$

Or on sait que

$$\begin{aligned}
\mu(F^{q_n}(I)) + \mu(F^{-q_n}(I)) &= \int_I (F^{q_n})'\,d\mu + \int_I (F^{-q_n})'\,d\mu \\
&= \int_I (F^{q_n})' + (F^{-q_n})'\,d\mu \\
&\geq 2\int_I \left((F^{q_n})'(F^{-q_n})'\right)^{1/2}\,d\mu \\
&\geq 2C^{-1/2}\mu(I).
\end{aligned}$$

$\square$

3.1 Introduction

Ce chapitre peut être vu comme une introduction simple à ce qu'on appelle la théorie ergodique différentiable. Soit X un espace topologique "simple" (que l'on supposera compact et métrisable par souci de simplification) muni d'une mesure naturelle λ ou tout au moins d'une classe de mesures équivalentes deux à deux. L'exemple typique est celui d'une variété sur laquelle on peut définir naturellement la mesure de Lebesgue (à équivalence près). Soit $T : X \to X$ une application continue (ou même tout simplement mesurable) et μ une mesure invariante. On appelera *bassin de* μ l'ensemble $\mathcal{B}(\mu)$ des points $x \in X$ tels que la suite $\frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i(x)}$ converge vers μ, pour la topologie faible*. On dira que μ est une *mesure physique* si $\lambda(\mathcal{B}(\mu)) > 0$. Cela signifie que si $\varphi : M \to \mathbf{R}$ est une observable continue, l'ensemble des points tels que $\frac{1}{n}\sum_{i=0}^{n-1}\varphi(T^i(x)) = \int_M \varphi\, d\mu$ est de mesure non nulle (pour λ). Si μ est une mesure de probabilité ergodique, son bassin est de mesure 1 (pour μ). Ainsi, dans le cas particulier où μ est absolument continue par rapport à λ, le bassin de μ est de mesure positive (pour λ). C'est donc une mesure physique. La recherche de mesures invariantes absolument continues par rapport à la mesure de Lebesgue est donc un problème naturel dans l'étude de la dynamique de systèmes définis sur une variété. Nous allons voir des exemples ou de telles mesures existent.

Concluons cette introduction par une remarque importante. Soit $T : X \to X$ une application mesurable et λ une mesure ergodique (main pas nécessairement invariante). Il existe au plus une mesure de probabilité invariante μ qui soit absolument continue par rapport à λ. En effet, une tellle mesure doit être ergodique car c'est le cas de λ. On a donc $\lambda(\mathcal{B}(\mu) = 1$ car μ est ergodique, et par conséquent $\lambda(\mathcal{B}(\mu) > 0$. Ceci implique que $\lambda(X \setminus \mathcal{B}(\mu) = 0$ car λ est ergodique. Ainsi, si la mesure μ existe, elle est unique.

3.2 Transformations branchées, définitions, exemples

Si I_0 est un intervalle non trivial de $\mathbf{R}$, on appelera *partition de* I_0 *en intervalles* tout ensemble $\mathcal{P}$ d'intervalles non triviaux contenus dans I, d'intérieurs disjoints deux à deux, et dont la réunion coincide avec I_0, à un ensemble dénombrable près. Bien sûr l'ensemble $\mathcal{P}$ est au plus dénombrable. On dira qu'une partition $\mathcal{P}'$ de I_0 en intervalles est une *sous partition de* $\mathcal{P}$, si pour tout $I \in \mathcal{P}$, l'ensemble des intervalles $I' \in \mathcal{P}'$ tels que $\text{int}(I') \subset \text{int}(I)$ est une sous partition en intervalles de $\overline{I}$. Bien évidemment, si $\mathcal{P}''$ est une sous partition en intervalles de $\mathcal{P}'$ et $\mathcal{P}'$ une sous partition en intervalles de $\mathcal{P}'$, alors $\mathcal{P}''$ est une sous partition en intervalles de $\mathcal{P}'$. On écrira $\mathcal{P} \preceq \mathcal{P}'$ pour signifier que $\mathcal{P}'$ est une sous partition de $\mathcal{P}$.

Soit I_0 un intervalle non trivial de $\mathbf{R}$. Une *transformation branchée* de I_0 est une application $f : I_0 \to I_0$ vérifiant la propriété suivante: il existe une partition de I_0 en intervalles $\mathcal{P}$ telle que pour tout $I \in \mathcal{P}$, l'application $f_{|\text{int}(I)}$ est un homéomorphisme entre $\text{int}(I)$ et $\text{int}(I_0)$. On dira alors que $\mathcal{P}$ est *associée* à f. S'il existe $k \geq 1$, tel que, pour tout $I \in \mathcal{P}$, l'application $f_{|\text{int}(I)}$ est un difféomorphisme de classe C^k entre $\text{int}(I)$ et $\text{int}(I_0)$, on dira que f est une *transformation* *branchée de classe C^k*. Enfin, si chaque $f_{|\mathrm{int}(I)}$ est affine, on dira que f est une *transformation branchée affine.*

Commençons par énoncer la proposition suivante, simple, mais fondamentale.

PROPOSITION 3.2.1 : *Soit I_0 un intervalle non trivial de $\mathbf{R}$ et f une transformation branchée de I_0. Alors, pour tout $n \geq 2$, l'application f^n est une transformation branchée de I_0. Plus précisément, il existe une famille $(\mathcal{P}_n)_{n\geq 1}$ de partitions en intervalles de I_0, vérifiant $\mathcal{P}_n \preceq \mathcal{P}_{n+1}$, $n \geq 1$, où $(\mathcal{P}_n)_{n\geq 1}$ est associée à f^n. De plus, si $0 < n < n'$, alors pour tout $I' \in \mathcal{P}_{n'}$, il existe $I \in \mathcal{P}_n$ tel que l'application ${f^{n'-n}}_{|\mathrm{int}(I')}$ est un homéomorphisme entre* $\mathrm{int}(I')$ *et* $\mathrm{int}(I)$.

Preuve. On va construire la suite $(\mathcal{P}_n)_{n\geq 1}$ par récurrence. On commence par choisir pour $\mathcal{P}_1$ la partition associée à f dont les éléments sont des intervalles ouverts. Pour tout $I \in \mathcal{P}_1$, l'application $f_{|I}$ est un homéomorphisme entre I et $\mathrm{int}(I_0)$. Les ensembles $I \cap f^{-1}(I')$, $I \in \mathcal{P}_1$, $I' \in \mathcal{P}_1$, sont donc des intervalles ouverts. Ces ensembles définissent une partition en intervalles de I_0. De $f|_{I\cap f^{-1}(I')}$ est un homéomorphisme entre $I \cap f^{-1}(I')$ et I' et $f^2|_{I\cap f^{-1}(I')}$ un homéomorphisme entre $I \cap f^{-1}(I')$ et $\mathrm{int}(I_0)$. Ainsi, f^2 est une transformation branchée de I_0. On définit alors la suite $(\mathcal{P}_n)_{n\geq 1}$ par récurrence: les intervalles définissant $\mathcal{P}_{n+1}$ sont les intervalles de la forme $I \cap f^{-n}(I')$, $I \in \mathcal{P}_n$, $I' \in \mathcal{P}_1$. Autrement dit, les éléments de $I \in \mathcal{P}_n$ sont les intervalles de la forme $I_1 \cap f^{-1}(I_2) \cap \ldots \cap f^{n-1}(I_n)$, où I_1, I_2, ..., I_n appartiennent à $\mathcal{P}_1$. On vérifie alors aisément que les conclusions de la proposition 3.1.1. sont vérifiées. □

Remarquons que la preuve précédente, nous dit que si f est une transformation branchée de I_0 de classe C^k, il en est de même de f^n, pour tout $n \geq 2$. De même si f est une transformation affine.

Nous allons nous intéresser principalement aux mesures de probabilité invariantes des transformations branchées, particulièrement au problème d'existence de mesures invariante absolument continues par rapport à la mesure de Lebesgue. Rappelons qu'une mesure borélienne de probabilité μ est ergodique si pour tout borélien A vérifiant $f^{-1}(A) = A$, on a $\mu(A) = 0$ ou $\mu(I_0 \setminus A) = 1$. On ne suppose pas ici nécessairement que μ est invariante. Notons que si μ est ergodique et si $\nu << \mu$, alors ν est ergodique.

3.2.1 Les transformations branchées affines

PROPOSITION 3.2.2 : *Soit I_0 un intervalle non trivial de $\mathbf{R}$ et f une transformation branchée affine de I_0. Alors, la mesure de Lebesgue λ est invariante et ergodique.*

Preuve. Commençons par montrer l'invariance de la mesure de Lebesgue. Pour tout $I \in \mathcal{P}$ notons α_I la valeur commune de $f'(x)$, $x \in I$. On a $\alpha_I \neq 0$ et $\lambda(I_0) = |\alpha_I|\lambda(I)$. Pour tout $J \subset I_0$, on a

$$\lambda(f^{-1}(J)) = \sum_{I\in\mathcal{P}} \lambda(f^{-1}(J)\cap I) = \sum_{I\in\mathcal{P}} \lambda(J)\alpha_I^{-1} = \lambda(J)\sum_{I\in\mathcal{P}} |\alpha_I^{-1}| = \lambda(J)\lambda(I_0)^{-1}\sum_{I\in\mathcal{P}} \lambda(I) = \lambda(J).$$

On en déduit que $\lambda(f^{-1}(A) = \lambda(A)$ pour tout borélien A et donc que λ est invariante.

Montrons maintenant que λ est ergodique. Remarquons que pour tout borélien A et tout $I \in \mathcal{P}$ on a

$$\lambda(f^{-1}(A)\cap I) = |\alpha_I|^{-1}\lambda(A) = \lambda(I_0)^{-1}\lambda(I)\lambda(A)$$

et donc que pour tout ensemble B, réunion d'intervalles $I \in \mathcal{P}$, on a

$$\lambda(f^{-1}(A) \cap B) = \lambda(I_0)^{-1}\lambda(A)\lambda(B).$$

Plus généralement, pour tout $n \geq 1$ et pour tout ensemble B, réunion d'intervalles $I \in \mathcal{P}_n$, on a

$$\lambda(f^{-n}(A) \cap B) = \lambda(I_0)^{-1}\lambda(A)\lambda(B).$$

Puisque $\mathcal{P}$ est non triviale, on a $|\alpha_I| > 1$ pour tout $I \in \mathcal{P}$. On en déduit que, pour tout $n \geq 1$ et pour tout $I \in \mathcal{P}_n$, on a $\lambda(I) \leq \lambda(I_0)\alpha^{-n}$, où $\alpha = \min_{I \in \mathcal{P}} |\alpha_I|$.

Montrons maintenant que le complèté de la σ-algèbre engendrée par les intervalles $I \in \mathcal{P}_n$, $n \geq 1$, contient les boréliens.

LEMME 3.2.3 : *Pour tout borélien B et tout $\varepsilon > 0$, il existe un entier n et un ensemble B', union d'intervalles $I \in \mathcal{P}_n$, tels que $\lambda(B \Delta B') < \varepsilon$.*

Preuve. La mesure de Lebesgue étant régulière, il existe une partie compacte $C \subset B$ telle que $\lambda(B \setminus C) < \varepsilon/2$. Par compacité de C, il existe $\eta > 0$ tel que $\lambda(C^\eta \setminus C) < \varepsilon/2$, où $C^\eta = \{x \in I_0 \,|\, d(x, C) < \eta\}$. Soit n tel que $\lambda(I_0)\alpha^{-n} < \eta$. Notons B' la réunion des intervalles $I \in \mathcal{P}_n$ tels que $I \cap C \neq \emptyset$. L'ensemble B' est contenu dans C^η et contient C à ensemble dénombrable prêt. On en déduit que

$$\lambda(B \Delta B') \leq \lambda(B \setminus C) + \lambda(C^\eta \setminus C) \leq \varepsilon.$$

□

Supposons que $f^{-1}(A) = A$. Pour tout borélien B and tout $\varepsilon > 0$, il existe un entier n et un ensemble B', union d'intervalles $I \in \mathcal{P}_n$, tels que $\lambda(B \Delta B') < \varepsilon$. De l'égalité

$$\lambda(A \cap B') = \lambda(I_0)^{-1}\lambda(A)\lambda(B')$$

et des inégalités

$$|\lambda(A \cap B) - \lambda(A \cap B')| \leq \varepsilon, \;\; |\lambda(B) - \lambda(B')| \leq \varepsilon$$

on déduit que

$$|\lambda(A \cap B) - \lambda(I_0)^{-1}\lambda(A)\lambda(B)| \leq (1 + \lambda(I_0)^{-1}\lambda(A))\,\varepsilon.$$

En faisant tendre ε vers 0 on obtient

$$\lambda(A \cap B) = \lambda(I_0)^{-1}\lambda(A)\lambda(B).$$

En particulier, on a

$$\lambda(A \cap (I_0 \setminus A)) = \lambda(I_0)^{-1}\lambda(A)\lambda(I_0 \setminus A),$$

ce qui implique que $\lambda(A) = 0$ ou $\lambda(I_0 \setminus A) = 0$. Ceci signifie que λ est ergodique. □

Remarque: Notons que dans le cas où $\lambda(I_0) = 1$, alors λ est une mesure de probabilité mélangeante. En effet, soit A et B deux boréliens. Pour tout $\varepsilon > 0$, il existe un entier n et un ensemble B', union d'intervalles $I \in \mathcal{P}_n$, tels que $\lambda(B \Delta B') < \varepsilon$. Soit $n' \geq n$. De l'égalité

$$\lambda(f^{-n'}(A) \cap B') = \lambda(A)\lambda(B')$$

et des inégalités

$$|\lambda(f^{-n'}(A) \cap B) - \lambda(f^{-n'}(A) \cap B')| \leq \varepsilon, \;\; |\lambda(B) - \lambda(B')| \leq \varepsilon$$

on déduit que

$$|\lambda(f^{-n'}(A)\cap B)-\lambda(A)\lambda(B)|\leq(1+\lambda(A))\,\varepsilon.$$

L'application tente

Un exemple intéressant de transformation branchée affine (et de plus continue) est *l'application tente* tent : $[0,1]\to[0,1]$ définie ainsi:

$$\mathrm{tent}(x)=\begin{cases}2x & \text{si } x\in[0,1/2],\\ 2-2x & \text{si } x\in[1/2,1].\end{cases}$$

Elle laisse invariante la mesure de Lebesgue, qui on l'a vu est mélangeante. Elle laisse invariante beaucoup d'autre mesures comme l'exprime le résultat suivant:

PROPOSITION 3.2.4 : *L'application* tent *admet un nombre non dénombrable de mesures de probabilité invariantes mélangeantes.*

Preuve Pour tout $\alpha\in]0,1[$, définissons une application $\mathrm{tent}_\alpha:[0,1]\to[0,1]$ en posant:

$$\mathrm{tent}_\alpha(x)=\begin{cases}\alpha^{-1}x & \text{si } x\in[0,\alpha],\\ (1-\alpha)^{-1}-(1-\alpha)^{-1}x & \text{si } x\in[\alpha,1].\end{cases}$$

Nous obtenons ainsi une famille de transformations affines (également continues).

LEMME 3.2.5 : *Les applications* tent_α, $\alpha\in]0,1[$, *sont toutes conjuguées.*

Preuve Fixons $\alpha\in]0,1[$ et montrons que tent_α est conjuguée à $\mathrm{tent}_{1/2}=\mathrm{tent}$. Posons

$$I_0=[0,\alpha],\ I_1=[\alpha,1],\ I_0^*=[0,1/2],\ I_1^*=[1/2,1].$$

Fixons $n\geq 1$ et pour tout mot $(a_0,\ldots,a_{n-1})\in\{0,1\}^n$, posons

$$I_{a_0,\ldots,a_{n-1}}=I_{a_0}\cap\mathrm{tent}_\alpha^{-1}(I_{a_1})\cap\ldots\cap\mathrm{tent}_\alpha^{-n-1}(I_{a_{n-1}}),$$

et

$$I^*_{a_0,\ldots,a_{n-1}}=I^*_{a_0}\cap\mathrm{tent}^{-1}(I^*_{a_1})\cap\ldots\cap\mathrm{tent}^{-n-1}(I^*_{a_{n-1}}).$$

Notons que

$$\lambda(I_{a_0,\ldots,a_{n-1}})\leq\frac{1}{\min(\alpha,1-\alpha)^{n-1}},\ \ \lambda(I^*_{a_0,\ldots,a_{n-1}})=\frac{1}{2^n}.$$

Observons également que les intervalles $I_{a_0,\ldots,a_{n-1}}$ sont ordonnés comme les $I^*_{a_0,\ldots,a_{n-1}}$ relativement aux a_i. Il existe donc un unique homéomorphisme h_α de $[0,1]$ tel que pour tout mot $(a_0,\ldots,a_{n-1})\in\{0,1\}^n$, $n\geq 1$, on ait $h_\alpha(I_{a_0,\ldots,a_{n-1}})=(I^*_{a_0,\ldots,a_{n-1}})$. Puisque

$$\mathrm{tent}_\alpha(I_{a_0,\ldots,a_{n-1}})=I_{a_1,\ldots,a_{n-1}},\ \ \mathrm{tent}(I^*_{a_0,\ldots,a_{n-1}})=I^*_{a_1,\ldots,a_{n-1}},$$

on en déduit que $h_\alpha\circ\mathrm{tent}_\alpha=\mathrm{tent}\circ h_\alpha$. En fait pour toute suite $\mathbf{a}=(a_i)_{i\geq 0}$, les ensembles

$$\bigcap_{0\leq i\leq+\infty}\mathrm{tent}_\alpha^{-i}(I_{a_i})\ \text{ et }\ \bigcap_{0\leq i\leq+\infty}\mathrm{tent}^{-i}(I^*_{a_i})$$

sont tous deux réduits à un point, noté respectivement $\theta_\alpha(\mathbf{a})$ et $\theta(\mathbf{a})$. L'application $\theta_\alpha : \{0,1\}^{\mathbf{N}} \to [0,1]$ est une semi-conjugaison entre le décalage de Bernouilli σ sur $\{0,1\}^{\mathbf{N}}$ et tent_α et de même l'application $\theta : \{0,1\}^{\mathbf{N}} \to [0,1]$ est une semi-conjugaison entre σ et tent. De plus on a $h_\alpha \circ \theta_\alpha = \theta$. $\square$

D'après la proposition 3.1.2, la mesure de Lebesgue est invariante par les tent_α et mélangeante. L'application h_α décrite dans le lemme 3.1.4 conjugue tent_α à tent et envoie I_0 sur $I{*}_0$. Ainsi $\mu_\alpha = (h_\alpha)_*(\lambda)$ est une mesure de probabilité invariante mélangeante de tent qui vérifie $\mu_\alpha([0,1/2]) = \alpha$. Les mesures μ_α, $\alpha \in]0,1[$, sont donc toutes distinctes. Puisqu'elles sont ergodiques, elles sont étrangères deux à deux. Notons qu'il n'y a aucune mesure invariante, autre que la mesure de Lebesgue, qui soit absolument continue par rapport à la mesure de Lebesgue. $\square$

Transformations du cercle de degré ≥ 2.

Si $F : \mathbf{T} \to \mathbf{T}$ est une transformation continue et $f : \mathbf{R} \to \mathbf{R}$ un relèvement de F, il existe un entier p tel que $f(x+1) = f(x) + p$ pour tout $x \in \mathbf{R}$. Cet entier ne dépend pas du relèvement choisi, car les relèvements sont uniquement définis à un constante additive entière près. On l'appelle le *degré* de F et on le note $\deg(F)$. Par exemple, le degré de l'endomorphisme $\widehat{x} \to p\widehat{x}$ est p. On montre facilement que $\deg(F \circ G) = \deg(F)\deg(G)$, et donc que $\deg(F) = \pm 1$ si F est un homéomorphisme. On dira que F est *croissant, décroissant, strictement croissant, strictement décroissant* s'il est relevé par une application croissante, décroissante, strictement croissante, strictement décroissante respectivement. L'entier p est positif dans le premier cas, négatif dans le deuxième, strictement positif dans le troisième et strictement négatif dans le dernier. On dira que F est *de classe* C^k, $1 \leq k \leq +\infty$, s'il est relevé par des applications de classe C^k. Dans ce cas, on peut définir $F^{(k)} : \mathbf{T} \to \mathbf{R}$, en posant $F^{(k)}(x + \mathbf{Z}) = f^{(k)}(x)$, où f est un relèvement de F (l'application $F^{(k)}$ est bien définie et indépendante du relèvement choisi). On dira que F est *dilatante* si F est de classe C^1 et si $\min_{\widehat{x} \in \mathbf{T}} |F'(\widehat{x})| > 1$. C'est bien sûr une transformation strictement croissante ou strictement décroissante. De plus, par le théorème des accroissements finis, son degré p vérifie $|p| \geq 2$. Notons aussi qu'une transformation continue F de degré p a au moins $|p-1|$ points fixes. En effet, l'application $G = F - \mathrm{Id}_{\mathbf{T}}$ est relevée par $g = f - \mathrm{Id}_{\mathbf{R}}$, si F est relevée par f, et le degré de G est $p-1$. Le fait que pour tout $x \in \mathbf{R}$, on ait $g(x+1) = g(x) + p - 1$ implique que sur l'intervalle $[x, x+1)$, l'application g prend une valeur entière en au moins $|p-1|$ points. Ces points se projettent sur des points fixes de F tous distincts. Si F est une transformation de $\mathbf{T}$ de degré $p \geq 2$ strictement croissante et si $a + \mathbf{Z}$ est un point fixe de F, l'application $T = h^{-1} \circ F \circ h$ est une application branchée de $[a, a+1[$, où h est la restriction de $\pi : x \mapsto x + \mathbf{Z}$ à $[a, a+1[$, avec une partition en intervalles associée ayant p éléments de la forme $[c, d[$. De plus, si F est de classe C^k et si F' ne s'annule pas, alors T est une application branchée de classe C^k. On a des résultats similaires pour les transformations décroissantes de degré $p \leq -2$.

PROPOSITION 3.2.6 : *Toute transformation continue strictement croissante F de $\mathbf{T}$ de degré $p \geq 2$ est une extension de $F^* : \widetilde{x} \to p\widetilde{x}$ avec une semi-conjugaison H qui est une transformation croissante de degré* 1. *De plus, si F est dilatante, alors H est un homéomorphisme et les applications sont conjuguées.*

Proof. Soit $\widehat{x}$ un point fixe de F. L'application $T_{\widehat{x}}^{-1} \circ F \circ T_{\widehat{x}}$ est une transformation continue de $\mathbf{T}$ de degré $p \geq 2$ qui fixe $\widehat{0}$. Ainsi, il suffit de prouver la proposition quand F fixe $\widehat{0}$. Soit h

la restriction de $\pi : x \mapsto x + \mathbf{Z}$ à $[0,1)$. Posons $T = h^{-1} \circ F \circ h$ et $T^* = h^{-1} \circ F^* \circ h$. Notons $x_0 = 0 < x_1 < \ldots < x_p = 1$ les points de $[0,1]$ dont l'image par f est entière, où f est le relèvement de F qui fixe 0, puis posons $I_k = [x_k, x_{k+1})$, pour $0 \leq k < p$. De façon similaire, définissons $x_k^* = k/p$ pour $k \in \mathbf{Z}$ et $I_k = [x_k^*, x_{k+1}^*)$, pour $0 \leq k < p$.

Fixons $n \geq 1$ et pour tout mot $(a_0, \ldots, a_{n-1}) \in \{0, \ldots, p-1\}^n$, posons

$$I_{a_0,\ldots,a_{n-1}} = I_{a_0} \cap T^{-1}(I_{a_1}) \cap \ldots \cap T^{-n+1}(I_{a_{n-1}}),$$

et

$$I^*_{a_0,\ldots,a_{n-1}} = I^*_{a_0} \cap (T^*)^{-1}(I^*_{a_1}) \cap \ldots \cap (T^*)^{-n+1}(I^*_{a_{n-1}}).$$

Notons que

$$\lambda(I^*_{a_0,\ldots,a_{n-1}}) = \frac{1}{p^n}.$$

Note également que

$$\lambda(I_{a_0,\ldots,a_{n-1}}) \leq \frac{1}{\alpha^n},$$

si F est dilatante et si $\alpha = \min_{\widehat{x} \in \mathbf{T}} F'(\widehat{x}) > 1$. Là-encore, observons que les intervalles $I_{a_0,\ldots,a_{n-1}}$ sont ordonnés come les $I^*_{a_0,\ldots,a_{n-1}}$ relativement aux a_i. Ainsi, il existe une unique application croissante $h_\alpha : [0,1] \to [0,1]$ telle que $h_\alpha(I_{a_0,\ldots,a_{n-1}}) = I^*_{a_0,\ldots,a_{n-1}}$ pour tout mot $(a_0, \ldots, a_{n-1}) \in \{0,1\}^n$, $n \geq 1$. De plus, h est un homéomorphisme si F est dilatante. L'application h est la restriction à $[0,1]$ du relèvement d'une application croissante de $\mathbf{T}$ de degré 1. Le fait que

$$T(I_{a_0,\ldots,a_{n-1}}) = I_{a_1,\ldots,a_{n-1}}, \;\; T^*(I^*_{a_0,\ldots,a_{n-1}}) = I^*_{a_1,\ldots,a_{n-1}},$$

implique que $h \circ T = T^* \circ h$. En fait, pour toute suite $\mathbf{a} = (a_i)_{i \geq 0}$, l'ensemble

$$\bigcap_{0 \leq i \leq +\infty} (T^*)^{-i}(I^*_{a_i})$$

est réduit à un point et l'ensemble

$$\bigcap_{0 \leq i \leq +\infty} T^{-i}(I_{a_i})$$

est un segment, qui est réduit un point si F est dilatante. $\square$

Remark. On a un résultat similaire pour une transformation continue strictement décroissante F de $\mathbf{T}$ de degré $p \leq -2$. L'application est une extension de $F^*/\widetilde{x} \to p\widetilde{x}$ et il existe une semi-conjugaison H qui est croissante de degré 1. De plus si F est dilatante, alors H est un homéomorphisme.

La transformation de Gauss

Définissons sur $[0,1[$ l'application

$$T_G \;: \begin{cases} x \mapsto \{1/x\} = 1/x - [1/x]\,, & \text{si } x \in]0,1] \\ 0 \mapsto 0, \end{cases}$$

où $\{y\}$ est la partie fractionnaire et $[y]$ la partie entière d'un réel y. Si $x \neq 0$, on a donc:

$$x = \frac{1}{a(x) + T_G(x)},$$

où $a(x) = [1/x]$. Pour tout $k \geq 1$, l'ensemble $a^{-1}(\{k\})$ coïncide avec l'intervalle $]1/(k+1), 1/k]$. L'application T_G induit un difféomorphisme de classe C^∞ strictement décroissant de $]1/(k+1), 1/k]$ sur $[0, 1[$. C'est une transformation branchée de classe C^∞. Fixons maintenant $a_1 \geq 1$, $\ldots$, $a_n \geq 1$, et notons $I_{a_1,\ldots,a_n}$ l'ensemble des points $x \in [0, 1[$ tels que:

- les points x, $T_G(x)$, $\ldots$, $T_G^{n-1}(x)$, sont non nuls;
- $a(x) = a_1, \ldots, a(T_G^{n-1}(x)) = a_n$.

L'ensemble $I_{a_1,\ldots,a_n}$ est alors un intervalle semi-ouvert, et T_G^n induit un difféomorphisme de $I_{a_1,\ldots,a_n}$ sur $[0, 1[$, croissant si n est pair et décroissant si n est impair. La réciproque s'exprime explicitement. Remarquons que pour tout $x \in I_{a_1,\ldots,a_n}$, on a

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cfrac{\cdots}{a_{n-1} + \cfrac{1}{a_n + T_G^n(x)}}}}},$$

que l'on peut écrire

$$x = \frac{p_n + p_{n-1}T_G^n(x)}{q_n + q_{n-1}T_G^n(x)},$$

où les deux suites $(p_n)_{n\geq 0}$ et $(q_n)_{n\geq 0}$ sont définies par les égalités

$$\begin{pmatrix} p_0 & q_0 \\ p_1 & q_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix}$$

et

$$\begin{cases} p_n = a_n p_{n-1} + p_{n-2} \\ q_n = a_n q_{n-1} + q_{n-2} \end{cases}.$$

On peut vérifier que la suite $(q_n)_{n\geq 0}$ est strictement croissante et que

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}.$$

Les extrémités de $I_{a_1,\ldots,a_n}$ étant p_n/q_n et $(p_n + p_{n-1})/(q_n + q_{n-1})$, on en déduit que la longueur de $I_{a_1,\ldots,a_n}$ vérifie

$$\lambda(I_{a_1,\ldots,a_n}) \leq \frac{1}{q_n(q_n + q_{n-1})} \leq \frac{1}{q_n^2} \leq \frac{1}{(n+1)^2}.$$

Pour tout point x, la suite $(a(T_G^n(x))_{n\geq 0}$ détermine le développement en fraction continue de x. Il n'est pas difficile de voir que la suite des itérés $T_G^n(x)$ aboutit en 0 si et seulement x est rationnel. En d'autres termes la suite $(T_G^n(x))_{n\geq 0}$ est bien définie si et seulement si x est irrationnel et on a un système dynamique bien déterminé défini sur $[0, 1[\setminus\mathbf{Q}$ en prenant la restriction de T à cet ensemble. Puisque pour toute suite $(a_i)_{i\geq 1}$ d'entiers strictement positifs, et pour tout $n \geq 0$, on a

$$\mathrm{long}(I_{a_1,\ldots,a_n}) \leq \frac{1}{(n+1)^2},$$

on en déduit que l'application

$$\begin{aligned} H \ : [0,1[\setminus\mathbf{Q} &\to (\mathbf{N}\setminus\{0\})^{\mathbf{N}\setminus\{0\}} \\ x &\mapsto (a(T^{i-1}(x)) \end{aligned}$$

conjugue $T|_{[0,1[\setminus\mathbf{Q}}$ au décalage $(a_n)_{n\geq 1} \mapsto (a_{n+1})_{n\geq 1}$ sur $(\mathbf{N}\setminus\{0\})^{\mathbf{N}\setminus\{0\}}$.

Proposition 3.1.3 : *La transformation T_G préserve la* mesure de Gauss μ_G *de densité*

$$d\mu = \frac{1}{\ln 2}\,\frac{d\lambda.}{1+x}.$$

Preuve. Il suffit de vérifier que pour tout intervalle $[a,b] \subset [0,1[$, on a $\mu(T^{-1}([a,b]) = \mu([a,b])$. Or on a

$$\begin{aligned}
\ln 2\,\mu(T^{-1}([a,b])) &= \sum_{k=1}^{\infty} \int_{\frac{1}{b+k}}^{\frac{1}{a+k}} \frac{1}{1+t}dt \\
&= \sum_{k=1}^{\infty} \ln(1+\frac{1}{a+k}) - \ln(1+\frac{1}{b+k}) \\
&= \sum_{k=1}^{\infty} \ln(a+k+1) - \ln(a+k) - \ln(b+k+1) + \ln(b+k) \\
&= \ln(b+1) - \ln(a+1) = \ln 2\,\mu([a,b]).
\end{aligned}$$

□

Nous verrons que la mesure de Gauss est également ergodique. L'explication générale de ce phénomène sera décrite dans la prochaine section.

3.3 Transformations branchées à distorsion bornée.

Soit $f : I_0 \to I_0$ une transformation branchée de classe C^1 et $(\mathcal{P}_n)_{n\geq 1}$ la famille de partitions en intervalles de I_0 définie par la proposition 3.1.1. On dira que f est à *distorsion bornée* si :

$$\operatorname{dist}(f) = \sup_{n\geq 1} \sup_{I\in\mathcal{P}_n} \sup_{x\in I,\, y\in I} \frac{|(f^n)'(x)|}{|(f^n)'(y)|} < +\infty.$$

Notons que $\operatorname{dist}(f) = 1$ si et seulement si f est une transformation branchée affine. Le théorème qui va suivre illustre le fait fondamental qu'une transformation branchée à variation bornée n'est pas trop “éloignée” d'une transformation branchée affine.

Théorème 3.3.1 : *Soit I_0 un intervalle non trivial borné de $\mathbf{R}$ et $f : I_0 \to I_0$ une transformation branchée de classe C^1 à distorsion bornée. Alors, la mesure de Lebesgue λ est ergodique.*

Remarque. Attention, le théorème ne dit pas que λ est invariante, contrairement au cas des transformations branchées affines.

Preuve. Commençons par un lemme, qui exprime ce que mesure la distorsion.

Lemme 3.3.2 : *Soit I un intervalle de $\mathbf{R}$ et $g : I \to \mathbf{R}$ un difféomorphisme de classe C^1 de I sur $g(I)$. Pour tous intervalles $J \subset I$ et $J' \subset I$, on a*

$$D^{-1}\frac{\lambda(J')}{\lambda(J)} \leq \frac{\lambda(g(J'))}{\lambda(g(J))} \leq D\,\frac{\lambda(J')}{\lambda(J)},$$

où

$$D = \sup_{x\in I,\, y\in I} \frac{|g'(x)|}{|g'(y)|}.$$

Preuve. Grâce au théorème des accroissements finis, on sait qu'il existe $x \in J$ et $y \in J'$ tels que

$$\lambda(g(J)) = |g'(x)|\lambda(J) \text{ et } \lambda(g(J')) = |g'(y)|\lambda(J').$$

On sait d'autre part que

$$D^{-1} \leq \frac{|g'(y)|}{|g'(x)|} \leq D.$$

□

Considérons maintenant une transformation branchée f de classe C^1 à variation bornée et $(\mathcal{P}_n)_{n\geq 1}$ la famille de partitions en intervalles de I_0 définies par la proposition 3.2.1. Le lemme suivant exprime entre autre que la famille des intervalles $I \in \mathcal{P}_n$, $n \geq 1$, engendre la tribu borélienne.

LEMME 3.3.3 : *On a* $\lim_{n\to+\infty} \max_{I\in\mathcal{P}_n} \lambda(I) = 0$.

Preuve. Posons $\delta_1 = \max_{I\in\mathcal{P}_1} \lambda(I) < \lambda(I_0)$ puis définissons $\rho \in]0,1[$ par l'égalité

$$1 - \rho = \operatorname{dist}(f)^{-1} \frac{\lambda(I_0) - \delta_1}{\lambda(I_0)}.$$

Nous allons montrer par récurrence sur $n \geq 1$ que pour tout $I \in \mathcal{P}_n$, on a $\lambda(I) \leq \rho^{n-1}\delta_1$, ce qui prouvera le lemme. La propriété est clairement vraie au rang 1. Supposons-là vraie au rang $n-1$. Fixons $I \in \mathcal{P}_n$ et notons I' l'élément de $\mathcal{P}_{n-1}$ contenant I. D'après le lemme 3.3.2, appliqué à la restriction de f^{n-1} à I', on sait que

$$1 - \frac{\lambda(I)}{\lambda(I')} = \frac{\lambda(I') - \lambda(I)}{\lambda(I')} \geq \operatorname{dist}(f)^{-1} \frac{\lambda(f^{n-1}(I' \setminus I))}{\lambda(f^{n-1}(I'))} \geq \operatorname{dist}(f)^{-1} \frac{\lambda(I_0) - \delta_1}{\lambda(I_0)} = 1 - \rho$$

et donc

$$\lambda(I) \leq \rho\lambda(I') \leq \rho^{n-1}\delta_1.$$

□

On a également le résultat suivant:

LEMME 3.3.4 : *Pour tout intervalle $A \subset I_0$, pour tout $n \geq 1$ et pour tout $I \in \mathcal{P}_n$, on a*

$$\operatorname{dist}(f)^{-1}\lambda(I_0)^{-1}\lambda(I)\lambda(A) \leq \lambda(I \cap f^{-n}(A)) \leq \operatorname{dist}(f)\lambda(I_0)^{-1}\lambda(I)\lambda(A).$$

Preuve. Il suffit d'appliquer le lemme 3.3.2 avec $g = f^n{}_{|I}$, $J = I$ et $J' = I \cap f^{-n}(A) = \left(f^n{}_{|I}\right)^{-1}(A)$. □

Le théorème 3.3.1 est maintenant facile à prouver. Fixons $n \geq 1$ et $I \in \mathcal{P}_n$. La mesure borélienne

$$X \mapsto \lambda(I \cap f^{-n}(X)) - \operatorname{dist}(f)^{-1}\lambda(I_0)^{-1}\lambda(I)\lambda(X)$$

prend des valeurs positives sur tout intervalle de I_0. Elle est donc positive. Soit maintenant X un borélien tel que $f^{-1}(X) = X$. La mesure borélienne

$$Y \mapsto \lambda(Y \cap X) - \operatorname{dist}(f)^{-1}\lambda(I_0)^{-1}\lambda(Y)\lambda(X)$$

prend des valeurs positives sur tout intervalle $I \in \mathcal{P}_n$, $n \geq 1$. Ici aussi, la tribu borélienne est contenue das le complètè de la σ-algèbre engendrée par les $I \in \mathcal{P}_n$, $n \geq 1$. La mesure précédente est donc positive. En particulier, on a

$$-\mathrm{dist}(f)^{-1}\lambda(I_0)^{-1}\lambda(I_0 \setminus X)\lambda(X) = \lambda((I_0 \setminus X) \cap X) - \mathrm{dist}(f)^{-1}\lambda(I_0)^{-1}\lambda(I_0 \setminus X)\lambda(X) \geq 0.$$

On en déduit que $\lambda(X) = 0$ ou $\lambda(I_0 \setminus X) = 0$. La mesure de Lebesgue est donc ergodique. $\square$

Illustrons ce résultat avec la transformation de Gauss.

PROPOSITION 3.3.5 : *La transformation de Gauss est à distorsion bornée et la mesure de Gauss est donc ergodique*

Preuve. En gardant les notations de la section 3.2, on sait que tout intervalle I de $\mathcal{P}_n$ s'écrit $I = I_{a_1,\ldots,a_n}$, où $a_1 \geq 1, \ldots, a_n \geq 1$. C'est un intervalle dont les extrémités sont p_n/q_n et $(p_n + p_{n-1})/(q_n + q_{n-1})$, et on sait que la réciproque de $(T_G)^n{}_{|I}$ s'écrit

$$\varphi \ : t \mapsto \frac{p_n + p_{n-1}t}{q_n + q_{n-1}t}.$$

Or on a

$$|\varphi'(t)| = \left|\frac{p_{n-1}q_n - p_n q_{n-1}}{(q_n + q_{n-1}t)^2}\right| = \frac{1}{(q_n + q_{n-1}t)^2}.$$

On en déduit

$$\frac{1}{4q_n^2} \leq \frac{1}{(q_n + q_{n-1})^2} \leq |\varphi'(t)| \leq \frac{1}{q_n^2},$$

ce qui implique que

$$\sup_{x\in I,\, y\in I} \frac{|(T_G^n)'(x)|}{|(T_G^n)'(y)}| \leq 4.$$

On conclût que $\mathrm{dist}(T_G) \leq 4$.

Puisque T_G est à distorsion bornée et puisque la mesure de Gauss est absolument continue par rapport à la mesure de Lebesgue, elle est ergodique, d'après le théorème 3.3.1. $\square$.

Il est rare que l'on puisse majorer la distorsion d'une transformation branchée à l'aide de calculs explicites. Nous allons donner maintenant établir des conditions qui impliquent qu'une transformation branchée donnée est à distorsion bornée. Soit I_0 un intervalle borné, f une transformation branchée de classe C^1 de I_0 et $\mathcal{P}$ la partition en intervalles ouverts associée. On dira que f est *uniformément dilatante* s'il existe deux constantes $C > 0$ et $\alpha > 0$ telle que pour tout $n \geq 1$ et pour tout $x \in I_0$, tel que x, $f(x)$, $\ldots f^{n-1}(x)$ appartiennent à $\bigcup_{I\in\mathcal{P}} I$, on a $|(f^n)'(x)| \geq Ce^{\alpha n}$. On a alors

PROPOSITION 3.3.6 : *Soit I_0 un intervalle, f une transformation branchée de classe C^2 de I_0 et $\mathcal{P}$ une partition en intervalles associée. Supposons que f soit uniformément dilatante et que*

$$\sup_{I\in\mathcal{P}} \sup_{x\in I} \frac{|f''(x)|}{f'(x)^2} < +\infty.$$

Alors f est à distorsion bornée.

Preuve. Posons $K = \sup_{I\in\mathcal{P}} \sup_{x\in I} |f''(x)|/f'(x)^2$. Fixons $I \in \mathcal{P}$ et notons φ_I la réciproque de $f_{|I}$. Nous avons

$$(\log |f' \circ \varphi_I|)' = \frac{f'' \circ \varphi_I\, \varphi_I'}{f' \circ \varphi_I} = \frac{f'' \circ \varphi_I}{(f' \circ \varphi_I)^2}.$$

On en déduit que pour tous s et t dans I, on a

$$|\log |f'(\varphi_I(s))| - \log |f'(\varphi_I(t))|| \leq K|s-t|,$$

et donc que pour tous x et y dans I, on a

$$|\log |f'(x)| - \log |f'(y)|| \leq K|f(x) - f(y)|.$$

Considérons la suite $(\mathcal{P}_n)_{n\geq 1}$ naturellement associée à f. Fixons $I \in \mathcal{P}_n$ et x, y dans I. Puisque, pour tout $i \in \{0, \ldots, n-1\}$, les points $f^i(x)$ et $f^i(y)$ sont dans le même intervalle de $\mathcal{P}_{n-i}$ et donc de $\mathcal{P}_1$, on a donc :

$$\begin{aligned}
\log \frac{|(f^n)'(x)|}{|(f^n)'(y)|} &= \sum_{i=0}^{n-1} \log |f'(f^i(x))| - \log |f'(f^i(y))| \\
&\leq \sum_{i=0}^{n-1} |\log |f'(f^i(x))| - \log |f'(f^i(y))||, \\
&\leq K \sum_{i=1}^{n} |f^i(x) - f^i(y)|.
\end{aligned}$$

Utilisant le théorème des accroissements finis et le fait que $0 \leq i < n$, les points $f^i(x)$ et $f^i(y)$ appartiennent tous deux au même intervalle de $\mathcal{P}_{n-i}$, on peut écrire

$$f^n(x) - f^n(y) = (f^{n-i})'(\xi_i)(f^i(x) - f^i(y)),$$

où ξ_i est dans l'intérieur de l'intervalle délimité par $f^i(x)$ et $f^i(y)$. On en déduit que

$$|f^i(x) - f^i(y)| \leq C^{-1}e^{-\alpha(n-i)}|f^n(x) - f^n(y)| \leq C^{-1}e^{-\alpha(n-i)}\lambda(I_0),$$

puis que

$$\log \frac{|(f^n)'(x)|}{|(f^n)'(y)|} \leq K \sum_{i=1}^{n} C^{-1}e^{-\alpha(n-i)}|f^n(x) - f^n(y)| \leq \frac{KC^{-1}|f^n(x) - f^n(y)|}{1-e^{-\alpha}} \leq \frac{KC^{-1}\lambda(I_0)}{1-e^{-\alpha}}.$$

Ceci implique que f est à distorsion bornée. Remarquons que la première inégalité s'écrit également

$$\log \frac{|(f^n)'(\varphi_I(t))|}{|(f^n)'(\varphi_I(t'))|} \leq \frac{KC^{-1}|t-t'|}{1-e^{-\alpha}}.$$

□

On peut appliquer la proposition à la transformation de Gauss. En effet, sur chaque intervalle $]1/n+1, 1/n]$, on a

$$T_G'(x) = \frac{-1}{x^2}, \ T_G''(x) = \frac{2}{x^3}.$$

On déduit d'une part que

$$0 \leq \frac{T_G''(x)}{T_G'(x)^2} = 2x \leq 2,$$

d'autre part que

$$T'_G(x) \le -1,\ \ (T_G^2)'(x) = T'_G(x)\, T'_G(T_G(x)) \ge 9/4.$$

La dernière inégalité est évidement vraie si $x \le 2/3$, mais elle est également vraie si $x \ge 2/3$ car, on sait alors que $T_G(x) \le 2/3$. On vérifie aisément que pour tout $n \ge 1$ et pour tout $x \in I_0$, tel que x, $f(x)$, $\dots f^{n-1}(x)$ appartiennent à $\bigcup_{I\in\mathcal{P}} I$, on a $|(T_G^n)'(x)| \ge (3/2)^{n-1}$. On peut donc appliquer la proposition 3.2.6 pour montrer que T_G est à distorsion bornée.

3.4 Mesures physiques

Nous venons de voir que la mesure de Lebesgue est ergodique pour toute transformation branchée à distorsion finie, ainsi donc que toute mesure absolument continue par rapport à la mesure de Lebesgue. Dans le cas de la transformation de Gauss nous avons explicité une mesure invariante qui vérifie cette propriété. Il est exceptionnel qu'on puisse expliciter une telle mesure. Néanmoins, le résultat qui suit nous dira qu'une telle mesure existe toujours et est unique.

Théorème 3.4.1 : *Soit I_0 un intervalle borné, $f : I_0 \to I_0$ une transformation branchée de classe C^2 et $\mathcal{P}$ une partition en intervalles associée. Supposons que f soit uniformément dilatante et que*

$$\sup_{I\in\mathcal{P}} \sup_{x\in I} \frac{|f''(x)|}{f'(x)^2} < +\infty.$$

Alors f admet une unique mesure de probabilité invariante ergodique absolument continue par rapport à la mesure de Lebesgue. De plus, la densité $d\mu/d\lambda$ est une fonction lipschitzienne, et vérifie

$$0 < \inf d\mu/d\lambda \le \sup d\mu/d\lambda < +\infty.$$

Preuve. Définissons pour tout $n \ge 1$, la mesure

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

Remarquons que μ_n est absolument continue par rapport à la mesure de Lebesgue et démontrons pour cela que chaque $f_*^n(\lambda)$, $n \ge 1$, est absolument continue par rapport à la mesure de Lebesgue. En effet si $(\mathcal{P}_n)_{n\ge 1}$ est la suite de partitions en intervalles ouverts associée à f, alors pour tout borélien A, on a

$$\lambda(f^{-n}(A)) = \sum_{I\in\mathcal{P}_n} \left(f^n{}_{|I}\right)^{-1}(A).$$

Puisque pour tout $I \in \mathcal{P}_n$, l'application $f^n_{|I}$ est un difféomorphisme de I sur $]0,1[$, on en déduit que

$$\lambda(A) = 0 \Rightarrow \lambda(f^{-n}(A)) = 0.$$

Notons que la densité de $f_*^n(\lambda)$ est

$$df_*^n(\lambda)/d\lambda(x) = \sum_{I\in\mathcal{P}_n} \left|\left(\left(f^n{}_{|I}\right)^{-1}\right)'(x)\right| = \sum_{y\in f^{-n}(\{x\})} \frac{1}{|(f^n)'(y)|}.$$

Nous écrivons dorénavant $S_n(x) = \displaystyle\sum_{y\in f^{-n}(\{x\})} \frac{1}{|(f^n)'(y)|}$, pour $n \ge 0$, là où la quantité est définie.

Lemme 3.4.2 : *Il existe une constante $K > 0$ telle que, pour tout $n \geq 0$, on a:*

$$1/K \leq \inf_{n \geq 1,\, x \in I_0} S_n(x) \leq \sup_{n \geq 1,\, x \in I_0} S_n(x) \leq K$$

et telle que pour tout $n \geq 1$, et pour tous x et y dans I_0, on a

$$|S_n(x) - S_n(y)| \leq K|x - y|.$$

Preuve. On a vu que les hypothèses du théorème impliquaient que f est à distorsion bornée. On sait grâce au théorème des accroissements finis, que pour tout $n \geq 1$ et pour tout $I \in \mathcal{P}_n$, il existe $z \in I$ tel que $|(f^n)'(z)| = \lambda(I_0)/\lambda(I)$. On sait également que pour tout $y \in I$ on a

$$\operatorname{dist}(f)^{-1} \leq \frac{|(f^n)'(z)|}{|(f^n)'(y)|} \leq \operatorname{dist}(f).$$

et donc que

$$\operatorname{dist}(f)^{-1}\lambda(I)/\lambda(I_0) \leq \frac{1}{|(f^n)'(y)}| \leq \lambda(I)/\lambda(I_0)\operatorname{dist}(f).$$

On en déduit que

$$\operatorname{dist}(f)^{-1} = \operatorname{dist}(f)^{-1} \sum_{I \in \mathcal{P}_n} \lambda(I)/\lambda(I_0) \leq S_n(x) \leq \sum_{I \in \mathcal{P}_n} \lambda(I)/\lambda(I_0)\operatorname{dist}(f) = \operatorname{dist}(f).$$

Le premier point du lemme étant donc vérifié, étudions le second. On peut écrire

$$S_n(x) = \sum_{I \in \mathcal{P}_n} \frac{1}{|(f^n)'(\varphi_I(x))|},$$

où φ est la fonction réciproque de ${f^n}_{|I}$. On a vu dans la preuve de la proposition 3.2.6 qu'il existe une constante M ne dépendant que de f telle que pour tous x et x' dans I_0, on ait

$$|\log|(f^n)'(\varphi_I(x))| - |\log|(f^n)'(\varphi_I(x'))|| \leq M|x - x'|.$$

L'inégalité des accroissements finis à la fonction log, puis la relation précédente nous donne :

$$\begin{aligned}
\left|\frac{1}{|(f^n)'(\varphi_I(x))|} - \frac{1}{|(f^n)'(\varphi_I(x'))|}\right| &\leq \max(|(f^n)'(\varphi_I(x))|^{-1}, |(f^n)'(\varphi_I(x'))|^{-1})|\log|(f^n)'(\varphi_I(x))|^{-1} - \log|(f^n)'(\varphi_I(x\\
&= \max(|(f^n)'(\varphi_I(x))|^{-1}, |(f^n)'(\varphi_I(x'))|^{-1})|\log|(f^n)'(\varphi_I(x))| - \log|(f^n)'(\varphi_I(x'))\\
&\leq \operatorname{dist}(f)\lambda(I)\lambda(I_0)^{-1}|\log|(f^n)'(\varphi_I(x))| - \log|(f^n)'(\varphi_I(x'))||\\
&\leq M\operatorname{dist}(f)\lambda(I)\lambda(I_0)^{-1}|x - x'|.
\end{aligned}$$

On peut finalement écrire

$$\begin{aligned}
|S_n(x) - S_n(x')| &= \left|\sum_{I \in \mathcal{P}_n} \frac{1}{|(f^n)'(\varphi_I(x))|} - \frac{1}{|(f^n)'(\varphi_I(x'))|}\right|\\
&\leq \sum_{I \in \mathcal{P}_n} \left|\frac{1}{|(f^n)'(\varphi_I(x))|} - \frac{1}{|(f^n)'(\varphi_I(x'))|}\right|\\
&\leq \sum_{I \in \mathcal{P}_n} M\operatorname{dist}(f)\lambda(I)\lambda(I_0)^{-1}|x - x'|\\
&= M\operatorname{dist}(f)|x - x'|.
\end{aligned}$$

Il reste à poser $K = \max(\mathrm{dist}(f), M\mathrm{dist}(f))$. $\square$

On déduit du lemme 3.4.2 que la densité $R_n = \frac{1}{n}\sum_{i=0}^{n-1} S_i$ vérifie

$$1/K \leq \inf_{x\in I_0} R_n(x) \leq \sup_{x\in I_0} R_n(x) \leq K$$

et que pour tous x et y dans I_0, on a

$$|R_n(x) - R_n(y)| \leq K|x-y|.$$

En appliquant le théorème d'Ascoli Arzela sur cet ensemble, on en déduit qu'il existe une suite extraite $(R_{n_k})_{k\geq 0}$ qui converge uniformément sur X vers une fonction R telle que

$$1/K \leq \inf_{x\in X} R(x) \leq \sup_{x\in X} R(x) \leq K$$

et telle que pour tout $n \geq 1$, et pour tous x et y dans X, on a

$$|R(x) - R(y)| \leq K|x-y|.$$

La mesure μ dont la densité vaut R est absolument continue par rapport à la mesure de Lebesgue. Elle est ergodique, il reste à montrer que c'est une mesure de invariante. Le théorème de convergence dominée de Lebesgue, nous dit que pour tout borélien A, on a :

$$\mu(A) = \int_A R\, d\lambda = \int_A \lim_{k\to+\infty} R_{n_k}\, d\lambda = \lim_{k\to+\infty} \int_A R_{n_k}\, d\lambda.$$

Ainsi, on a

$$\mu(A) = \lim_{k\to+\infty} \mu_{n_k}(A) = \lim_{k\to+\infty} \frac{1}{n_k}\sum_{i=0}^{n_k-1} f_*^i(\lambda)(A) = \lim_{k\to+\infty} \frac{1}{n_k}\sum_{i=0}^{n_k-1} \lambda(f^{-i}(A)).$$

En particulier

$$\mu(f^{-1}(A)) = \lim_{k\to+\infty} \mu(A) + \frac{1}{n_k}\lambda(f^{-n_k}(A)) - \frac{1}{n_k}\lambda(A) = \mu(A).$$

$\square$

Remark Comme on l'a vu dans l'introduction la mesure $\mu = \frac{1}{\lambda(I_0)}\nu$ est la seule mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue. On peut appliquer le théorème 3.4.1 à une application branchée associée à une transformation dilatante de classe C^2 de $\mathbf{T}$. Une telle application a une unique mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue.

Concluons cette section par un résultat concernant l'entropie, qui s'applique en particulier quand f vérifie les hypothèses du théorème 3.4.1.

Théorème 3.4.3 : *Soit I_0 un intervalle borné et $f : I_0 \to I_0$ une transformation branchée de classe C^1 à distorsion bornée. Si f admet une mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue et si*

$$0 < \inf d\mu/d\lambda \leq \sup d\mu/d\lambda < +\infty,$$

alors l'entropie de μ vaut

$$h_\mu(f) = \int |f'(x)|\, d\mu(x).$$

Preuve. Nous notons $(\mathcal{P}_n)_{n\geq 1}$ la suite associée à f et formée d'intervalles ouverts. Nous allons prouver le résultat quand $\mathcal{P}_1$ est finie, ce qui est alors le cas de tous les $\mathcal{P}_n$. La raison étant que nous avons défini l'entropie métrique relativement aux partitions finies. Gardant le formalisme introduit dans la section concernant l'entropie, on remarque que pour tout $n \geq 1$, on a $\mathcal{P}_n = \vee_{i=0}^{n-1} f^{-i}(\mathcal{P}_1)$. D'après le lemme 3.3.3, on sait que la partition $\mathcal{P}_1$ est génératrice et donc que

$$h_\mu(f) = h_\mu(f, \mathcal{P}_1) = \lim_{n\to+\infty} \frac{1}{n} H_\mu(\mathcal{P}_n).$$

Rappelons que

$$H_\mu(\mathcal{P}_n) = \int -\ln(\mu(I_n(x))\, d\mu(x),$$

où $I_n(x)$ est l'intervalle de $\mathcal{P}_n$ qui contient x. Si on pose $m = \inf d\mu/d\lambda$ et $M = \sup d\mu/d\lambda$, on en déduit que

$$\int -\ln(\lambda(I_n(x))\, d\mu(x) - \ln M \leq H_\mu(\mathcal{P}_n) \leq \int -\ln(\lambda(I_n(x))\, d\mu(x) - \ln m$$

L'égalité des accroissements finis nous dit que pour tout $I \in \mathcal{P}_n$, il existe $y \in I$ tel que $\lambda(I_0) = \lambda(I)|(f^n)'(y)|$ et donc que pour tout $x \in I$, on a

$$\mathrm{dist}(f)^{-1}|(f^n)'(x)|\lambda(I_0)^{-1} \leq \lambda(I)^{-1} \leq \mathrm{dist}(f)|(f^n)'(x)|\lambda(I_0)^{-1}.$$

On en déduit que

$$\int \ln|(f^n)'(x)|\, d\mu(x) - \ln M - \ln\lambda(I_0) - \ln\mathrm{dist}(f) \leq H_\mu(\mathcal{P}_n) \leq \int \ln|(f^n)'(x)|\, d\mu(x) - \ln m - \ln\lambda(I_0) + \ln\mathrm{dist}(f)$$

Puisque

$$(f^n)'(x) = f(x)f'(f(x)))\dots f'(f^{n-1}(x))$$

et puisque μ est invariante par f, on sait que

$$\int \ln|(f^n)'(x)|\, d\mu(x) = n\int \ln|f'(x)|\, d\mu(x).$$

On peut donc conclure, en divisant par n et en faisant tendre n vers $+\infty$. $\square$

Calculons, à l'aide du théorème 2.4.3, l'entropie de la transformation de Gauss (qui vérifie les hypothèses). On obtient

$$
\begin{aligned}
h_{\mu_G}(T_G) &= \frac{1}{\ln 2}\int_0^1 \frac{-2\ln x}{1+x}\ dx. \\
&= \frac{-2}{\ln 2}\left([\ln x \ln(1+x)]_0^1 - \int_0^1 \frac{\ln(1+x)}{x}\, dx\right) \\
&= \frac{2}{\ln 2}\int_0^1 \frac{\ln(1+x)}{x}\, dx \\
&= \frac{2}{\ln 2}\int_0^1 \sum_{n=0}^{+\infty} \frac{(-1)^n x^n}{n+1}\, dx \\
&= \frac{2}{\ln 2}\sum_{n=0}^{+\infty} \frac{(-1)^n}{(n+1)^2} \\
&= \frac{2}{\ln 2}\left(\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} - \sum_{n=1}^{+\infty} \frac{1}{(2n)^2}\right) \\
&= \frac{2}{\ln 2}\left(\sum_{n=0}^{+\infty} \frac{1}{n^2} - 2\sum_{n=1}^{+\infty} \frac{1}{(2n)^2}\right) \\
&= \frac{1}{\ln 2}\sum_{n=0}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6\ln 2}
\end{aligned}
$$

3.5 Le cas de la dimension supérieure

Ce qui a été fait dans le cadre des applications branchées de l'intervalle reste encore valable dans un cadre un peu différent mais non restreint à la dimension 1. Comme nous le verrons, les preuves sont un peu plus compliquées, mais quand même très proches de ce qui a été fait dans le cadre des applications branchées.

Soit M une variété riemannienne de classe C^∞ et de dimension $r \geq 1$, et $f : M \to M$ une application de classe C^1. Pour tout point $x \in M$, on peut définir

$$\|Df(x)\| = \max_{v\in T_xM,\, \|v\|_x\leq 1} \|Df(x).v\|_{f(x)}.$$

On peut également définir $\det Df(x)$: on choisit une base orthonormée $(v_i)_{1\leq i\leq r}$ de T_xM et une base orthonormée $(w_j)_{1\leq j\leq r}$ de $T_{f(x)}M$. On définit alors $\det Df(x)$ comme étant le déterminant de la famille $(Df(x).v_i)_{1\leq i\leq r}$ dans la base $(w_j)_{1\leq j\leq r}$, en remarquant que ce nombre est indépendant des deux bases choisies. On peut remarquer également que la structure riemannienne permet de définir un volume et donc une mesure sur M qu'on appellera *mesure de Lebesgue*. Dans le cas d'une variété compacte, on obtient une mesure finie. De plus un autre choix de structure riemannienne donne une mesure équivalente. Rappelons que l'on peur définir une distance d sur M en posant

$$d(x,y) = \inf_{\gamma\in\mathcal{E}_{x,y}} \mathrm{long}(\gamma),$$

où $\mathcal{E}_{x,y}$ désigne l'ensemble des arcs C^1 par morceaux joignant x à y et $\mathrm{long}(\gamma)$ la longueur de γ.

On dira, que f est *dilatante* (où σ*-dilatante* si on veut préciser) s'il existe $\sigma > 1$ tel que pour tout $x \in M$ et pour tout $v \in T_xM$, on a $\|Df(x).v\|_{f(x)} \geq \sigma\|v\|_x$. Notons que cette condition dépend du choix de la structure choisie. On peut cependant donner une définition du caractère dilatant qui ne dépende pas de la structure riemannienne.

Un cas simple est celui où la variété M est le tore $\mathbf{T}^r = \mathbf{R}^r/\mathbf{Z}^r$ et où la structure riemannienne est définie par une norme sur $\mathbf{R}^r$. Dans ce cas, pour tout point $x \in \mathbf{T}^r$, la quantité $\|Df(x)\|$ est égale à $\|D\widetilde{f}(\widetilde{x})\|$, où $\widetilde{f} : \mathbf{R}^r \to \mathbf{R}^r$ est un relèvement de f, où $\widetilde{x} \in \mathbf{R}^r$ est un représentant de x , et où la norme considérée est la norme d'opérateur associée à la norme choisie sur $\mathbf{R}^r$. Supposons donc que f soit σ-dilatante et que $g : \mathbf{T}^r \to \mathbf{T}^r$ soit une application de classe C^1 telle que $\|Df(x) - Dg(x)\| < \sigma - 1$, pour tout $x \in \mathbf{T}^r$. Montrons que g est également dilatante. En effet, par continuité de l'application $x \mapsto Df(x) - Dg(x)$ et compacité de $\mathbf{T}^r$, on sait qu'il existe $\varepsilon < \sigma - 1$ tel que pour tout $x \in \mathbf{T}^r$, on a $\|Df(x) - Dg(x)\| \leq \varepsilon$. On en déduit que pour tout $x \in \mathbf{T}^r$ et pour tout $v \in \mathbf{R}^r$, on a

$$\|Dg(x).v\| \geq \|Df(x).v\| - \|Df(x).v - Dg(x).v\| \geq (\sigma - \varepsilon)\|v\|,$$

avec $\sigma - \varepsilon > 1$.

Le cas emblématique d'une application dilatante est celui d'un endomorphisme linéaire f dont le relèvement linéaire $\widetilde{f}$ a ses valeurs propres (réelles et complexes) toutes de module plus grand que 1. Dans ce cas $Df(x) = \widetilde{f}$ est indépendant de x et il est bien connu qu'il existe une norme $\| \ \|$ sur $\mathbf{R}^r$ et un réel $\sigma > 1$, tel que pour tout $v \in \mathbf{R}^r$, on a $\|\widetilde{f}(v)\| \geq \sigma\|v\|$. Plus précisément, on montre que si $\tau < 1$ est un majorant strict du module des valeurs propres de $\widetilde{f}^{-1}$, on peut construire une norme $\| \ \|$ sur $\mathbf{R}^r$, tel $\|\widetilde{f}^{-1}\| \leq \tau$ (avec la norme d'opérateur associée) et donc telle que pour tout $v \in \mathbf{R}^r$, on a $\|\widetilde{f}^{-1}(v)\| \leq \tau\|v\|$. Notons donc que toute transformation obtenue par une petite perturbation de f (pour la C^1-topologie) est encore dilatante. Dans le cas où $r = 1$, un endomorphisme linéaire est de la forme $x \mapsto px$, où $|p| \geq 2$, et définit naturellement une transformation branchée affine de $[0, 1]$. Ce qui sera fait dans cette section permettra de généraliser certains résultats obtenus dans le cas des transformations branchées au cas des transformations dilatantes en dimension supérieure.

Rappelons que si (X, d) est un espace métrique, une fonction $\varphi : X \to \mathbf{R}$ est dite *höldérienne* s'il existe $C > 0$ et $\alpha \in]0, 1]$ telle que pour tous x et y dans X, on a

$$|\varphi(x) - \varphi(y)| \leq Cd(x, y)^\alpha.$$

Si on veut préciser, on dira que φ est *α-höldérienne.*

Énonçons maintenant le théorème principal (en notant que toutes les hypothèses sont satisfaites si f est une application dilatante de classe C^2):

Théorème 3.5.1 : *Soit M une variété compacte connexe de classe C^∞, munie d'une structure riemannienne, et $f : M \to M$ une transformation de classe C^1 dilatante. On suppose de plus que l'application $x \mapsto \ln(|\det Df(x)|)$ est höldérienne. Il existe alors une unique mesure borélienne de probabilité invariante absolument continue par rapport à la mesure de Lebesgue. De plus, elle est ergodique, à support total, et son bassin est de mesure de Lebesgue totale.*

Nous fixons f dans cette section, vérifiant les hypothèses du théorème. Nous supposerons donc qu'il existe $\sigma > 1$, $\alpha \in]0, 1]$ et $C > 0$ tels que:

- pour tout $x \in M$ et tout $v \in T_xM$, on a $\|Df(x).v\| \geq \sigma\|v\|$;
- pour tous x et y dans M, on a $|\ln(|\det Df(x)|) - \ln(|\det Df(y)|)| \leq Cd(x, y)^\alpha$.

Nous supposerons également par souci de simplification que $\lambda(M) = 1$.

Le premier résultat préliminaire n'utilise que le caractère dilatant de f.

Proposition 3.5.2 : *Tous les points de M ont le même nombre d'antécédents p par f. De plus, il existe η_0 tel que pour tout $x \in M$, il existe une application $h : B(f(x), \eta_0) \to M$ telle que $f \circ h = \mathrm{id}_{|B(f(x),\eta_0)}$ et telle que pour tout y et y' dans $B(f(x), \eta_0)$, on ait*

$$d(h(y), h(y')) \leq \sigma^{-1} d(y, y').$$

Preuve Le fait que pour tout $x \in M$, l'application $Df(x)$ soit injective, implique que f est un difféomorphisme local en chacun de ses points. Puisque M est compact, on en déduit que $f^{-1}(y)$ est fini, pour tout $y \in M$. De plus, il existe $p(y) \in \mathbf{N}$ et $\eta(y) > 0$ tel que $f^{-1}(B(y, \eta(y))$ est la réunion disjointe de $p(y)$ ouverts tels, qu'en restriction à chacun de ces ouverts, f est un difféomorphisme sur $B(y, \eta(y))$. L'application $p(y)$ est localement constante car M est compact. On note p la valeur commune des $p(y)$, elle est bien sûr au moins égale à 1. Toujours par compacité de M, on peut recouvrir M par une famille finie $B(y_i, \eta(y_i))$, $1 \leq i \leq m$. On sait alors qu'il existe η_1 (le nombre de Lebesgue du recouvrement) tel que pour tout $y \in M$, la boule $B(y, \eta)$ est incluse dans un $B(y_i, \eta(y_i))$, $1 \leq i \leq m$ et donc que l'ensemble $f^{-1}(B(y, \eta)$ est la réunion disjointe de p ouverts tels, qu'en restriction à chacun de ces ouverts, f est un difféomorphisme sur $B(y, \eta)$. Encore par compacité, on peut trouver $\eta_2 > 0$ tel que pour tout $y \in M$, l'application $\exp(y) : T_y M \to M$ est une isométrie entre $\{v \in T_y M \mid \|v\| < \eta_2\}$ et $B(y, \eta_2)$ (on peut également poser $\eta_2 = \min_{k \in \mathbf{Z}^p \setminus \{0\}} \|k\|/4$ si M est le tore $\mathbf{T}^r$ muni d'une métrique constante). Posons $\eta_0 = \min(\eta_1, \eta_2)$. Pour tout $x \in M$, il existe une application $h : B(f(x), \eta_0) \to M$ telle que $f \circ h = \mathrm{id}_{|B(f(x),\eta_0)}$. Supposons maintenant que y et y' appartiennent à $B(f(x), \eta_0)$. On peut trouver un arc $\gamma : [0, 1] \to B(f(x), \eta_0)$ de classe C^1, joignant y à y' et de longueur $d(y, y')$. On a

$$\begin{aligned} d(h(y), h(y')) \leq \mathrm{long}(h \circ \gamma) &= \int_0^1 \|Dh(\gamma(t)).\gamma'(t)\| \, dt \\ &\leq \sigma^{-1} \int_0^1 \|Df(h(\gamma(t)).(Dh(\gamma(t)).\gamma'(t)))\| \, dt \\ &= \sigma^{-1} \int_0^1 \|\gamma'(t))\| \, dt = \sigma^{-1} \mathrm{long}(\gamma) = \sigma^{-1} d(y, y'). \end{aligned}$$

Dans la suite de la preuve, on appelera *branche de f en y* toute application $h : B(y, \eta_0) \to M$ de classe C^1 vérifiant $f \circ h = \mathrm{id}_{|(B(y),\eta_0)}$(il y en a p). Soit $(h_i)_{1 \leq i \leq n}$ une famille de branches, où h_i est une branche en y_i, et où $h_i(y_i) = y_{i+1}$, pour tout $i \in \{1 \ldots, n-1\}$. Puisque h_i contracte les longueurs, on a $h_i(B(y_i, \eta_0)) \subset B(y_{i+1}, \eta_0)$. Ainsi, l'application $H = h_n \circ \ldots h_1$ est donc bien définie, c'est une *branche de f^n en y*. Notons que pour tout y et y'' dans $B(y, \eta_0)$, on a $d(H(y'), H(y'')) \leq \sigma^{-n} d(y, y'')$.

Le lemme qui suit (qui n'utilise également que le caractère dilatant de f) sera utilisé en fin de preuve, mais il nous dit déjà que $p \geq 2$, car f ne peut pas être un difféomorphisme.

Lemme 3.5.3 : *Pour toute partie ouverte $U \subset M$ non vide, il existe $N > 0$ tel que pour tout $n \geq N$, on a $f^n(U) = M$.*

Preuve Si y et y' sont deux points de M vérifiant $d(y, y') < \eta_0$, et si $n \geq 1$, alors pour $x \in f^{-n}(\{y\})$, il existe $x' \in f^{-n}(\{y'\})$ tel que $d(x, x') \leq \sigma^{-n} d(y, y')$. Il suffit en effet de choisir $x' = H(y')$, où H est la branche de f^n en y qui vaut x en y. Puisque M est compact, il est de diamètre borné. Ainsi existe $L > 0$ tel que tout paire de points (y, y') peut être reliée par un chemin de longueur $< L$. Comme conséquence, il existe un entier l tel que pour tout y, y' dans M, il existe une famille $(y_i)_{0\leq i\leq l}$ vérifiant $y_0 = y$, $y_l = y'$ et $d(y_i, y_{i+1}) < \eta_0$ pour $0 \leq i < l$. On en déduit que si $n \geq 1$, alors pour tout $x \in f^{-n}(\{y\})$, il existe $x' \in f^{-n}(\{y'\})$ tel que $d(x, x') \leq \delta^{-n} l\eta_0$. Ainsi, pour tout $\varepsilon > 0$, il existe $N \geq 0$ tel que pour tout $x \in M$ et pour tout $n \geq N$, on a $f^n(B(x, \varepsilon)) = M$. □

La proposition qui suite utilise outre le caractère dilatant de f, l'hypothèse sur la fonction $x \mapsto \det f(x)$.

PROPOSITION 3.5.4 : *Il existe $K > 0$ tel que pour toute branche $H : B(y, \eta_0) \to M$ de f^n et pour tous points y_1, y_2 de $B(y, \eta_0)$, on a*

$$\frac{|\det DH(y_1)|}{|\det DH(y_2)|} \leq e^{Kd(y_1,y_2)^\alpha} \leq e^{K(2\eta_0)^\alpha}.$$

Preuve On peut écrire $H = h_n \circ \ldots \circ h_1$, où h_1 est une branche de f en y, et où h_i est une branche de f en $h_{i-1}(y)$, si $i \in \{2, \ldots, n\}$. Posons $H_i = h_i \circ \ldots h_1$, pour tout $i \in \{1, \ldots, n\}$ et $H_0 = \mathrm{id}$. On a alors

$$\begin{aligned}
\ln\left(\frac{|\det DH(y_1)|}{|\det DH(y_2)|}\right) &= \sum_{i=1}^{n} \ln|\det Dh_i(H_{i-1}(y_1))| - \ln|\det Dh_i(H_{i-1}(y_2))| \\
&= \sum_{i=1}^{n} -\ln|\det Df(H_i(y_1))| + \ln|\det Df(H_i(y_2))| \\
&\leq \sum_{i=1}^{n} C|H_i(y_1) - H_i(y_2)|^\alpha \\
&\leq \sum_{i=1}^{n} C\sigma^{-i\alpha}|y_1 - y_2|^\alpha \\
&\leq \sum_{i=1}^{+\infty} C\sigma^{-i\alpha}|y_1 - y_2|^\alpha = \frac{C\sigma^{-\alpha}}{1-\sigma^{-\alpha}}|y_1 - y_2|^\alpha.
\end{aligned}$$

□

On en déduit les corollaire suivants en notant λ la mesure de Lebesgue, c'est-à-dire la mesure définie par la structure riemannienne et K le réel donné par la proposition 3.5.4.:

COROLLAIRE 3.5.5 : *Pour toute branche $H : B(y, \eta_0) \to M$ de f^n et pour tous boréliens A et B dans $B(y, \eta_0)$, on a*

$$e^{-K(2\delta_0)^\alpha}\frac{\lambda(A)}{\lambda(B)} \leq \frac{\lambda(H(A))}{\lambda(H(B))} \leq e^{K(2\eta_0)^\alpha}\frac{\lambda(A)}{\lambda(B)}.$$

Preuve Par connexité de la boule $B(y, \eta_0)$, il existe y_1 et y_2 dans $B(y, \eta_0)$ tels que

$$\frac{\lambda(H(A))}{\lambda(H(B))} = \frac{\int_A |\det DH| d\lambda}{\int_B |\det DH| d\lambda} = \frac{|\det DH(y_1)|\,\lambda(A)}{|\det DH(y_2)|\,\lambda(B)}.$$

On applique la proposition 3.3.4 pour conclure. □

COROLLAIRE 3.5.6 : *Il existe $K' > 0$ tel que pour tout $n \geq 1$ et tout borélien A, on a $\lambda(f^{-n}(A)) \leq K'\lambda(A)$*

Preuve Supposons qu'il existe $y \in M$ tel que $A \subset B(y, \eta_0)$. Pour toute branche H de f^n en y, on a

$$\lambda(H(A)) \leq e^{2K(2\eta_0)^\alpha}\lambda(H(B(y,\eta_0))\lambda(A)\lambda(B(y,\eta_0))^{-1.}$$

En sommant sur toutes les branches inverses de f^n en y, on trouve

$$\begin{aligned}\lambda(f^{-n}(A)) &\leq e^{2K(2\eta_0)^\alpha}\lambda(f^{-n}(B(y,\eta_0))\lambda(A)\lambda(B(y,\eta_0))^{-1}\\ &\leq e^{2K(2\eta_0)^\alpha}\lambda(M)\lambda(B(y,\eta_0))^{-1}\lambda(A)\end{aligned}$$

Considérons maintenant un recouvrement fini $B(y_i, \eta_0)_{1\leq i\leq r}$ de M. Tout borélien A peut s'écrire $A = \bigsqcup_{1\leq i\leq r} A_i$, où $A_i \subset B(y_i, \eta_0)$. On en déduit que

$$K' = e^{2K(2\eta_0)^\alpha}\lambda(M)\left(\inf_{1\leq i\leq r}\lambda(B(y_i,\eta_0))\right)^{-1}$$

convient. □

Nous pouvons maintenant démontrer:

PROPOSITION 3.5.7 : *La mesure de Lebesgue λ est ergodique.*

Preuve Par compacité de M, il existe une famille finie $(y_i)_{1\leq i\leq r}$ telle que la famille $(B(y_i,\eta_0))_{1\leq i\leq r}$ recouvre M. Commençons par construire une partition mesurable $\mathcal{P} = (P_i)_{1\leq i\leq r}$, où chaque P_i est d'intérieur non vide et inclus dans $B(y_i, \eta_0)$. On considère $\eta < \eta_0$ tel que les boules $B(y_i, \eta)$ sont disjointes deux à deux. On définit $B'_i = B(y_i, \eta_0) \setminus B(y_i, \eta)$ puis on construit la famile $\mathcal{P}$ par récurrence sur i en posant $P_i = (B'_i \setminus \cup_{i'<i} B'_{i'}) \cup B(y_i, \eta_0)$. On peut alors définir une suite de partitions $(\mathcal{P}_n)_{n\geq 1}$ (formée de rp^n éléments) où tout élément de $\mathcal{P}_n$ est l'image d'un P_i par une branche de f^n en y_i. Notons que le diamètre de chaque élément de $\mathcal{P}_n$ a un diamètre au plus égal à $2\eta_0\delta^{-n}$.

Commençons par prouver deux lemmes.

LEMME 3.5.8 : *Pour toute mesure borélienne de probabilité μ et pour toute partie mesurable A telle que $\mu(A) > 0$, il existe une suite de partie mesurables $(Q_n)_{n\geq 1}$ telle que*

$$Q_n \in \mathcal{P}_n,\ \mu(Q_n) > 0,\ \lim_{n\to+\infty}\mu(A \cap Q_n)/\mu(Q_n) = 1.$$

Preuve Fixons $\varepsilon \in]0, \mu(A)[$. Par régularité de μ, il existe une partie fermée $K \subset A$ telle que $\mu(A \setminus K) < \varepsilon$. Il existe alors δ tel que

$$\mu\{x \in X \mid d(x,K) < \delta\} < \mu(K) + \varepsilon.$$

Puisque le diamètre de chaque élément de $\mathcal{P}_n$ a un diamètre au plus égal à $2\eta_0\delta^{-n}$, il existe N tel que pour tout $n \geq N$, la réunion A_n des éléments de $\mathcal{P}_n$ qui rencontrent K est inclus dans

$\{x \in X \mid d(x, K) < \delta\}$ et vérifie donc $\mu(A_n) < \mu(K) + \varepsilon$. Fixons $n \geq N$, et supposons que pour tout $Q \in \mathcal{P}_n$ on ait

$$\mu(K \cap Q) < \frac{\mu(A) - \varepsilon}{\mu(A) + \varepsilon}\, \mu(Q).$$

En sommant sur les ensembles Q qui rencontrent K on aurait

$$\mu(K) = \mu(K \cap A_n) < \frac{\mu(A) - \varepsilon}{\mu(A) + \varepsilon}\, \mu(A_n) < \frac{\mu(A) - \varepsilon}{\mu(A) + \varepsilon}\, (\mu(K) + \varepsilon) \leq \mu(A) - \varepsilon,$$

obtenant donc une contradiction. Par conséquent, pour tout $n \geq N$, on peut choisir au moins un élément $Q \in \mathcal{P}_n$ tel que:

$$\mu(A \cap Q) \geq \mu(K \cap Q) > \frac{\mu(A) - \varepsilon}{\mu(A) + \varepsilon}\, \mu(Q).$$

On construit alors facilement la suite $(Q_n)_{n \geq 1}$ demandée. □

Lemme 3.5.9 : *Si A est une partie borélienne telle que $f^{-1}(A) = A$ et $\lambda(A) \neq 0$, alors il existe $i \in \{1, \ldots, r\}$ tel que $\lambda(P_i \setminus A) = 0$.*

Preuve Par le lemme 3.3.8, on sait qu'il existe une suite $(Q_n)_{n \geq 1}$ telle que

$$Q_n \in \mathcal{P}_n,\ \mu(Q_n) > 0,\ \lim_{n \to +\infty} \lambda(Q_n \setminus A)/\lambda(Q_n) = 0.$$

Il existe alors une suite $(i_n)_{n \geq 1}$ telle que $f^n(Q_n) = P_{i_n}$. Appliquant le corollaire 3.3.5, on obtient

$$\frac{\lambda(P_{i_n} \setminus A)}{\lambda(U_{i_n})} = \frac{\lambda(f^n(Q_n) \setminus A)}{\lambda(f^n(Q_n))} = \frac{\lambda(f^n(Q_n \setminus A))}{\lambda(f^n(Q_n))} \leq e^{K(2\eta_0)^\alpha} \frac{\lambda(Q_n \setminus A)}{\lambda(Q_n)}.$$

Il existe $i \in \{1, \ldots, r]$ tel que $i_n = i$ pour une infinité de n. On a donc $\lambda(P_i \setminus A) = 0$. □.

Nous pouvons maintenant conclure que λ est ergodique. Soit A une partie boréliennne vérifiant $f^{-1}(A) = A$. D'après le lemme 3.3.9, il existe $i \in \{1, \ldots, r\}$ tel que $\lambda(P_i \setminus A) = 0$. D'après le lemme 3.3.3, puisque P_i est d'intérieur non vide, il existe q tel que $f^q(P_i) = M$. On a

$$M \setminus A = f^q(P_i) \setminus A = f^q(P_i \setminus A)$$

et donc $\lambda(M \setminus A) = 0$. □

Proposition 3.5.10 : *Il existe au moins une mesure borélienne de probabilité invariante qui est absolument continue par rapport à la mesure de Lebesgue*

Preuve Commençons par le lemme suivant:

Lemme 3.5.11 : *Soit X un espace métrique compact, μ une mesure borélienne de probabiliité et φ une fonction positive μ-intégrale. Si $(\nu_n)_{n \geq 0}$ est une suite de mesures boréliennes de probabilité qui converge en topologie faible* vers ν et si pour tout $n \geq 0$, on a $\nu_n \leq \varphi\mu$, alors on a $\nu \leq \varphi\mu$.*

Preuve Soit U une partie ouverte. Définissons une suite de fonctions continues $(f_k)_{k \geq 1}$ sur X, en posant

$$f_k\ : x \mapsto \max(1 - kd(x, X \setminus U), 0).$$

La suite $(f_k)_{k\geq 1}$ est décroissante et converge vers la fonction caractéristique $\mathbf{1}_{X\setminus U}$. Pour tout $k \geq 1$, on a

$$\limsup_{n\to+\infty} \nu_n(X\setminus U) \leq \limsup_{n\to+\infty} \int f_k\, d\nu_n = \lim_{n\to+\infty} \int f_k\, d\nu_n = \int f_k\, d\nu.$$

Ceci implique que

$$\limsup_{n\to+\infty} \nu_n(X\setminus U) \leq \inf_{k\geq 1} \int f_k\, d\nu = \lim_{k\to+\infty} \int f_k\, d\nu = \nu(X\setminus U).$$

et donc

$$\liminf_{n\to+\infty} \nu_n(U) \geq \nu(U).$$

On en déduit que $\nu(U) \leq \int_U \varphi\, d\mu$.

Toute mesure borélienne de probabilité étant régulière, pour tout borélien A et pour tout $\varepsilon > 0$, on peut trouver une partie ouverte U contenant X telle que $\int_U \varphi d\mu \leq \int_A \varphi d\mu + \varepsilon$. On a donc

$$\nu(A) \leq \nu(U) \leq \int_U \varphi d\mu \leq \int_A \varphi d\mu + \varepsilon.$$

En faisant tendre ε vers 0, on obtient que $\nu(A) \leq \int_A \varphi d\mu$. □

On peut maintenant prouver la proposition 3.5.10. D'après le lemme 3.5.6, il existe $K' > 0$ tel que pour tout $n \geq 1$, on a $f_*^n(\lambda) \leq K'\lambda$. On en déduit que pour tout $n \geq 1$, on a $\frac{1}{n}\sum_{i=0}^{n-1} f_*^i(\lambda) \leq K'\lambda$. D'après la proposition 3.5.6, toute valeur d'adhérence μ de la suite $\left(\frac{1}{n}\sum_{i=0}^{n-1} f_*^i(\lambda)\right)_{n\geq 1}$ vérifie $\mu \leq K'\lambda$. La mesure μ est donc absolument continue par rapport à λ avec une densité bornée par K'. La preuve du théorème de Krylov-Bogolyubov nous dit que μ est invariante. On peut donc conclure. □

Nous pouvons maintenant conclure.

Preuve du théorème 3.5.1 La proposition 3.5.10 nous dit qu'il existe au moins une mesure de probabilité invariante μ qui est absolument continue par rapport à λ. Grâce la proposition 3.5.7, nous savons qu'elle est ergodique. En effet λ est ergodique et μ est absolument continue par rapport à λ. D'après le théorème ergodique de Birkhoff, son bassin $\mathcal{B}(\mu)$ est de mesure pleine (pour μ) puisque μ est ergodique. Il est donc de mesure de Lebesgue non nulle puisque μ est absolument continue par rapport à λ. Puisque $f^{-1}(\mathcal{B}(\mu)) = \mathcal{B}(\mu)$ et puisque λ est ergodique, on a $\lambda(\mathcal{B}(\mu)) = 1$. En particulier μ est unique. □

3.6 Le théorème ergodique de Kingman

Commençons par énoncer le *Théorème ergodique de Kingman*, appelé également *Théorème ergodique sous-additif*, qui généralise le théorème ergodique de Birkhoff.

Pour toute fonction $f : X \to \mathbf{R}$ definie sur un ensemble X, on écrira $f^+ = \max(f, 0)$.

Théorème 3.6.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Soit $(f_n)_{n\geq 1}$ une suite de fonctions mesurables $f_n : X \to \mathbf{R} \cup \{-\infty\}$ telle que*

- $f_1^+ \in L^1(\mu)$;

- *pour tous entiers strictement positifs n, m, on a $f_{n+m} \leq f_n + f_m \circ T^n$.*

Il existe alors une fonction mesurable $f : X \to \mathbf{R} \cup \{-\infty\}$ telle que $f^+ \in L^1(\mu)$, vérifiant

$$f(x) = f \circ T(x), \quad f(x) = \lim_{n\to+\infty} \frac{1}{n} f_n(x),$$

pour presque tout $x \in X$. De plus, on a

$$\int f \, d\mu = \lim_{n\to+\infty} \frac{1}{n} \int f_n \, d\mu = \inf \frac{1}{n} \int f_n \, d\mu.$$

Preuve. Avant de prouver le théorème, notons que si $(f_n)_{n\geq 1}$ et $(f'_n)_{n\geq 1}$ sont deux suites qui vérifient la relation de sous-additivité, il en est alors de même des suites $(f_n + f'_n)_{n\geq 1}$ et $(\max(f_n, f'_n))_{n\geq 1}$. Notons aussi que si $f_1 : X \to \mathbf{R} \cup \{-\infty\}$ est mesurable, alors la suite $(\sum_{i=0}^{n-1} f \circ T^i)_{n\geq 1}$ vérifie cette relation (on a alors une inégalité). Dans le cas où la fonction est constante (égale à M), on obtient la suite $(nM)_{n\geq 1}$.

Revenons aux hypothèses. Pour tout $x \in X$ et tout $n \geq 1$ on a :

$$f_n(x) - \sum_{i=0}^{n-1} f_1^+(T^i(x)) \leq f_n(x) - \sum_{i=0}^{n-1} f_1(T^i(x)) \leq 0.$$

On obtient donc une suite $(f'_n)_{n\geq 1}$ de fonctions, à valeurs négatives, qui vérifie la relation de sous-additivité, en posant :

$$f'_n(x) = f_n(x) - \sum_{i=0}^{n-1} f_1^+(T^i(x)).$$

On en déduit que la suite $\left(\int f_n \, d\mu\right)_{n\geq 1}$ est bien définie, prenant ses valeurs dans $\mathbf{R} \cup \{-\infty\}$, et que pour tout $n \geq 1$, on a

$$\int f_n \, d\mu = \int f'_n \, d\mu + \int f_1^+ \, d\mu,$$

où $\int f'_n \, d\mu \in [-\infty, 0]$ et $\int f_1^+ \, d\mu \in \mathbf{R}$. Notons que pour tous entiers strictement positifs n, m, on a

$$\int f_{n+m} \, d\mu \leq \int f_n \, d\mu + \int f_m \, d\mu,$$

ce qui implique que la suite $\left(\frac{1}{n} \int f_n \, d\mu\right)_{n\geq 1}$ converge dans $\mathbf{R} \cup \{-\infty\}$ et vérifie

$$\lim_{n\to+\infty} \frac{1}{n} \int f_n \, d\mu = \inf_{n\geq 1} \frac{1}{n} \int f_n \, d\mu.$$

Grâce au théorème ergodique de Birkhoff, on sait que la convergence presque partout de la suite $\left(\frac{1}{n} f'_n(x)\right)_{n\geq 1}$ implique la convergence presque partout de $\left(\frac{1}{n} f_n(x)\right)_{n\geq 1}$. De, plus si la relation entre intégrales est vraie pour la suite $(f'_n)_{n\geq 1}$, elle est également vraie pour $(f_n)_{n\geq 1}$. On peut donc toujours supposer que les f_n sont à valeurs négatives.

Expliquons d'abord pourquoi il suffit de prouver le résultat dans le cas où il existe un entier $M \geq 1$ tel que les fonctions $f_n + nM$ sont à valeurs positives. Supposons donc que le théorème soit prouvé dans ce cas particulier et plaçons nous dans le cadre général. Pour tout entier $M \geq 1$, définissons la suite $(f_n^M)_{n\geq 1}$, où $f_n^M = \max(f_n, -nM)$. Remarquons que les fonctions f_n^M sont à valeurs négatives, que les fonctions $f_n^M + nM$ sont à valeurs positives et que la suite $(f_n^M)_{n\geq 1}$ vérifie la relation de sous-additivité. On déduit qu'il existe une fonction $f^M : X \to [-\infty, 0]$ vérifiant

$$f^M(x) = f^M \circ T(x) \text{ et } \lim_{n\to+\infty} \frac{1}{n} f_n^M(x) = f^M(x),$$

pour presque tout x, et que

$$\int f^M \, d\mu = \lim_{n\to+\infty} \frac{1}{n} \int f_n^M \, d\mu = \inf \frac{1}{n} \int f_n^M \, d\mu.$$

Notons que f^M prend ses valeurs dans $[-M, 0]$ et que

$$f^M = \max\left(\limsup_{n\to+\infty} \frac{f_n}{n}, -M\right).$$

Si on écrit $f = \limsup_{n\to+\infty} \frac{f_n}{n}$, on en déduit que $f(x) = \lim_{n\to+\infty} \frac{f_n(x)}{n}$ pour presque tout point x tel que $f(x) > -M$. En particulier, on a

$$f(x) = f \circ T(x) \text{ et } \lim_{n\to+\infty} \frac{1}{n} f^n(x) = f(x)$$

pour presque tout x. Par le théorème de convergence monotone, on sait que

$$\int f \, d\mu = \lim_{M\to+\infty} \int f^M \, d\mu = \inf_{M\geq 1} \int f^M \, d\mu,$$

ce qui implique que

$$\int f \, d\mu = \inf_{M\geq 1} \inf_{n\geq 1} \frac{1}{n} \int f_n^M \, d\mu = \inf_{n\geq 1} \inf_{M\geq 1} \frac{1}{n} \int f_n^M \, d\mu = \inf_{n\geq 1} \frac{1}{n} \int f_n \, d\mu.$$

On suppose maintenant qu'il existe un entier $M \geq 1$ tel que chaque fonction $f_n + nM$ est positive. En particulier, les fonctions f_n sont toutes bornées

De l'inégalité

$$\frac{n+1}{n} \; \frac{f_{n+1}(x)}{n+1} \leq \frac{f_1(x)}{n} + \frac{f_n(T(x))}{n},$$

on déduit que la fonction

$$f : x \mapsto \liminf_{n\to+\infty} \frac{1}{n} f_n(x)$$

est mesurable, à valeurs négatives, et vérifie $f(x) \leq f(Tx)$. En particulier, pour tout t, l'ensemble $A_t = \{x \in X \mid f(x) > t\}$ vérifie $A_t \subset T^{-1}(A_t)$. On déduit de l'égalité $\mu(A_t) = \mu(T^{-1}(A_t))$ que les ensembles A_t et $T^{-1}(A_t)$ diffèrent d'un ensemble de mesure nulle. On conclut aisément que l'égalité $f(T(x)) = f(x)$ a lieu presque partout.

Maintenant fixons $\varepsilon > 0$ et pour tout entier $N > 0$ posons

$$B(N) = \{x \mid f_l(x) > l(f(x) + \varepsilon) \text{ pour tous } l \in \{1, \dots, N\}\}$$

et notons $A(N)$ son complémentaire. Fixons $x \in X$ et $n > N$ et considérons le segment d'orbite $(T^k(x))_{0\leq k<n}$. On va décomposer ce segment. Plus précisément, on va construire de façon algorithmique, une décomposition de $\{0, \dots, n-1\}$ en "intervalles". Supposons que les p premiers

intervalles ont été construits. Soit k le premier entier de $\{0, \dots, n-1\}$ qui n'appartient à aucun intervalle (en particulier $k = 0$ au début de la récurrence). Si $T^k(x) \in B(N)$, définissons $\{k\}$ comme $p+1$-ème intervalle. Si $T^k(x) \in A(N)$, soit l le plus petit entier dans $\{1, \dots, N\}$ tel que $f_l(T^k(x)) \leq l(f(T^k(x)) + \varepsilon)$ (cet entier existe par définition de $A(N)$). Si $k + l > n$, définissons $\{k\}$ comme $p+1$-ème intervalle. Si $k + l \leq n$, définissons $\{k, \dots, k+l-1\}$ comme $p+1$-ème intervalle. Notons $I_j = \{k_j, \dots, k_j + l_j - 1\}$, $1 \leq j \leq J$, les intervalles du troisième type. Utilisant le fait que f est invariant par T presque partout, que les f_n prennent des valeurs négatives et que la suite $(f_n)_{n \geq 0}$ est sous-additive, on déduit que

$$f_n(x) \leq \sum_{j=1}^{J} f_{l_j}(T^{k_j}(x) \leq \sum_{j=1}^{J} l_j(f(T^{k_j}x) + \varepsilon) = \sum_{j=1}^{J} l_j(f(x) + \varepsilon) \leq f(x)(\sum_{j=1}^{J} l_j) + n\varepsilon.$$

Pour tout $Y \in \mathcal{B}$, notons 1_Y^* la limite des moyennes de Birkhoff de la fonction caractéristique 1_Y, définie grâce au théorème ergodique de Birkhoff. Utilisant le fait qu'il existe au plus $\sum_{k=0}^{n-1} 1_{B(N)}(T^k(x))$ intervalles du premier type et N intervalles du second type, on déduit que

$$\sum_{j=1}^{J} l_j \geq n - \sum_{k=0}^{n-1} 1_{B(N)}(T^k(x)) - N,$$

ce qui implique que

$$\liminf_{n \to +\infty} \frac{1}{n} \sum_{j=1}^{J} l_j \geq 1 - 1_{B(N)}^*(x)$$

pour presque tout x. Ainsi, a-t-on

$$\limsup_{n \to +\infty} \frac{f_n(x)}{n} \leq f(x)(1 - 1_{B(N)}^*(x)) + \varepsilon,$$

pour presque tout x. Par définition de $B(N)$, on sait que

$$\lim_{N \to +\infty} 1_{B(N)}(x) = 0$$

presque partout, ce qui implique, par le théorème ergodique de Birkhoff et le théorème de convergence dominée, que

$$\lim_{N \to +\infty} \int 1_{B(N)}^* \, d\mu = \lim_{N \to +\infty} \int 1_{B(N)} \, d\mu = 0.$$

Utilisant le lemme de Fatou, on en déduit que

$$\int \lim_{N \to +\infty} 1_{B(N)}^* \, d\mu \leq \lim_{N \to +\infty} \int 1_{B(N)}^* \, d\mu = 0,$$

et donc que

$$\lim_{N \to +\infty} 1_{B(N)}^*(x) = 0$$

presque partout. Ceci nous dit que

$$\limsup_{n \to +\infty} \frac{f_n(x)}{n} \leq f(x) + \varepsilon$$

presque partout, puis, en prenant l'infimum sur ε que

$$\limsup_{n\to+\infty} \frac{f_n(x)}{n} \leq f(x)$$

presque partout. Nous venons de montrer que la suite $\left(\frac{f_n(x)}{n}\right)_{n\geq 1}$ converge presque partout.

Établissons maintenant les inégalités sur les intégrales. En intégrant l'inégalité,

$$\frac{f_n(x)}{n} \leq f(x)(\frac{1}{n}\sum_{j=1}^{J} l_j) + \varepsilon,$$

on obtient

$$\int \frac{f_n(x)}{n}\, d\mu \leq \int f(x)\left(1 - \frac{1}{n}(\sum_{k=0}^{n-1} 1_{B(N)}(T^k(x)) - N)\right) d\mu + \varepsilon = \int f(x)(1 - 1_{B(N)}(x) - N/n)\, d\mu + \varepsilon,$$

ce qui implique que

$$\lim_{n\to+\infty} \int \frac{f_n(x)}{n}\, d\mu \leq \int f(x)(1 - 1_{B(N)}(x))\, d\mu + \varepsilon.$$

Grâce au théorème de convergence dominée de Lebesgue, si on fait tendre N vers $+\infty$, on obtient

$$\lim_{n\to+\infty} \int \frac{f_n(x)}{n}\, d\mu \leq \int f(x)\, d\mu + \varepsilon.$$

En prenant l'infimum sur ε, on obtient

$$\lim_{n\to+\infty} \int \frac{f_n(x)}{n}\, d\mu \leq \int f(x)\, d\mu.$$

L'inégalité inverse est une conséquence immédiate du lemme de Fatou puisque les fonctions $\frac{f_n}{n}$ sont uniformément minorées. □

□

Nous allons tout de suite donner un corollaire. Ici $\|\ \|$ est une norme d'opérateurs sur $L(\mathbf{R}^k, \mathbf{R}^k)$.

Corollaire 3.6.2: *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$ et $A : X \to L(\mathbf{R}^k, \mathbf{R}^k)$ une application mesurable telle que la fonction $x \mapsto (\log\|A(x)\|)^+$ appartient à $L^1(\mu)$. Alors, il existe une fonction mesurable $\alpha : X \to \mathbf{R} \cup \{-\infty\}$ définie presque partout, invariante par T, satisfaisant $\alpha^+ \in L^1(\mu)$, et telle que*

$$\lim_{n\to+\infty} \frac{1}{n} \log\|A(T^{n-1}(x) \circ \ldots \circ A(T(x)) \circ A(x)\| = \alpha(x)$$

et

$$\begin{aligned}
&\lim_{n\to+\infty} \frac{1}{n} \int \log\|A(T^{n-1}(x) \circ \ldots \circ A(T(x)) \circ A(x)\|\, d\mu \\
&= \inf \frac{1}{n} \int \log\|A(T^{n-1}(x) \circ \ldots \circ A(T(x)) \circ A(x)\|\, d\mu \\
&= \int \alpha(x)\, d\mu
\end{aligned}$$

Preuve La suite de fonctions $(f_n)_{n\geq 1}$, où

$$\begin{aligned} f_n : X &\to \mathbf{R} \\ x &\mapsto \log \|A(T^{n-1}(x) \circ \ldots \circ A(T(x)) \circ A(x)\|, \end{aligned}$$

vérifie les hypothèses du théorème de ergodique de Kingman. $\square$

Donnons une autre application.

COROLLAIRE 3.6.3 : *Soit M une variété riemannienne de classe C^∞ et f un difféomorphisme de classe C^1. On suppose qu'il existe une mesure de probabilité μ invariante par f et que la fonction $x \mapsto (\log |Df(x)|)^+$ appartient à $L^1(\mu)$. Alors, il existe une fonction mesurable $\alpha : M \to \mathbf{R} \cup \{-\infty\}$ définie presque partout, invariante par f, satisfaisant $\alpha^+ \in L^1(\mu)$, et telle que*

$$\lim_{n\to+\infty} \frac{1}{n} \log \|Df^n(x)\| = \alpha(x)$$

et

$$\begin{aligned} &\lim_{n\to+\infty} \frac{1}{n} \int \log \|Df^n(x)\| \, d\mu \\ &= \inf \frac{1}{n} \log \|Df^n(x)\| \, d\mu \\ &= \int \alpha(x) \, d\mu \end{aligned}$$

Preuve La suite de fonctions $(f_n)_{n\geq 1}$, où

$$\begin{aligned} f_n : X &\to \mathbf{R} \\ x &\mapsto \|Df^n(x)\| \end{aligned}$$

vérifie les hypothèses du théorème de ergodique de Kingman. $\square$

Notons que les conditions sont toujours vérifiées si M est une variété compacte et f un difféomorphisme de classe C^1.

Chapitre 1

Exercice 1

Soient $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$, le second étant un facteur du premier. Montrer que si $(X, \mathcal{B}, \mu, T)$ est faiblement mélangeant, il en est de même de $(Y, \mathcal{C}, \nu, S)$.

Exercice 2

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. On suppose que $(\mathcal{B}, \mu)$ admet une base dénombrable, c'est-à-dire qu'il existe une suite $(A_n)_{n\geq 0}$ dans $\mathcal{B}$, telle que pour tout $A \in \mathcal{B}$ et tout $\varepsilon > 0$, il existe $n \geq 0$ tel que $\mu(A \Delta A_n) \leq \varepsilon$. Monter que si le système est faiblement mélangeant, il existe une partie $I \subset \mathbf{N}$ de densité 1 telle que, pour tous A et B dans $\mathcal{B}$, on a

$$\lim_{n\to+\infty,\, n\in I} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B).$$

Exercice 3

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$.

1) Montrer que $(X, \mathcal{B}, \mu, T) \times (X, \mathcal{B}, \mu, T)$ est mélangeant si et seulement si $(X, \mathcal{B}, \mu, T)$ est mélangeant.

2) Montrer que les propriétés suivantes sont équivalentes :

i) le système $(X, \mathcal{B}, \mu, T)$ est faiblement mélangeant;

ii) le système $(X, \mathcal{B}, \mu, T) \times (X, \mathcal{B}, \mu, T)$ est faiblement mélangeant;

iii) le système $(X, \mathcal{B}, \mu, T) \times (X, \mathcal{B}, \mu, T)$ est ergodique.

(Pour montrer que ii) implique i), on utilisera le critère donné par le corollaire 1.2.7)

Exercice 4

Soient $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$, le second étant un facteur du premier. Montrer que si λ est une valeur propre de $U_S : L^2(\nu) \to L^2(\nu)$, alors λ est une valeur propre de $U_T : L^2(\mu) \to L^2(\mu)$.

Exercice 5

Soit X un espace topologique compact et $\mathcal{C}(X)$ l'espace de Banach formé des fonction continues $f : X \to \mathbf{C}$, muni de la norme $\|f\| = \max_{x\in X} |f(x)|$. A toute transformation continue $T : X \to X$ est associé un opérateur $U_T : \mathcal{C}(X) \to \mathcal{C}(X)$, où $U_T(f) = f \circ T$.

i) Montrer que U_T est continu et calculer sa norme. Montrer que 1 est valeur propre.

ii) Montrer que si T est surjective, alors toute valeur propre est de module 1.

iii) On suppose dans cette question que T est transitif, c'est-à-dire que pour tous ouverts non vides U et V de X, il existe $n \geq 0$ tel que $U \cap T^{-n}(V) \neq \emptyset$.

- Montrer que toute fonction propre a un module constant.

- Quel est l'espace propre associé à 1 ?

- Montrer que l'ensemble des valeurs propres est un sous groupe du cercle unité de $\mathbf{C}$.

- Montrer que si $\lambda = e^{2\imath\pi\alpha}$ est valeur propre, où $\alpha = p/q$ est rationnel, alors T admet comme facteur la rotation $\hat{x} \mapsto \hat{x} + \hat{p}$ de $\mathbf{Z}/q\mathbf{Z}$, où $\hat{p} = p + q\mathbf{Z}$.

- Montrer que si $\lambda = e^{2\imath\pi\alpha}$ est valeur propre, où α est irrationnel, alors T admet comme facteur la rotation $\hat{x} \to \hat{x} + \hat{\alpha}$ de $\mathbf{T}$, où $\hat{\alpha} = \alpha + \mathbf{Z}$.

iv) On suppose dans cette question que T est topologiquement mélangeant, c'est-à-dire que pour tous ouverts non vides U et V de X, il existe $n \geq 0$ tel que $U \cap T^{-m}(V) \neq \emptyset$, pour tout $m \geq n$. Que peut-on dire des valeurs propres de U_T et des fonctions propres associeés ?

Exercice 6

Soient $(X, \mathcal{B}, \mu)$ un espace de probabilité. On dira que deux partitions $\mathcal{P}$ et $\mathcal{Q}$ sont *indépendantes* si, pour tout $P \in \mathcal{P}$ et tout $Q \in \mathcal{Q}$, on a $\mu(P \cap Q) = \mu(P)\mu(Q)$. Montrer que les propriétés suivantes sont équivalentes :

- $\mathcal{P}$ et $\mathcal{Q}$ sont indépendantes ;
- $H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q})$;
- $H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P})$.

Exercice 7

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, où $\mu(X) = 1$. Prouver l'*inégalité de Rokhlin* : pour toutes partitions mesurables $\mathcal{P}$ et $\mathcal{Q}$, on a

$$|h(T, \mathcal{P}) - h(T, \mathcal{Q})| \leq D(\mathcal{P}, \mathcal{Q}).$$

Exercice 8

Soit $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$. On considère le produit $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \otimes \nu, T \times S)$. Montrer que

$$h_{\mu\otimes\nu}(T \times S) = h_\mu(T) + h_\nu(S).$$

Exercice 9

Soit $(X, \mathcal{B}, \mu, T)$ et $(Y, \mathcal{C}, \nu, S)$ deux systèmes dynamiques mesurés, avec $\mu(X) = \nu(Y) = 1$. On suppose que $h_\mu(T) = h_\nu(S)$. Cela implique t-il les systèmes sont conjugués ?

Exercice 10

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, avec $\mu(X) = 1$, et $A \subset \mathcal{B}$ une partie invariante telle que $(T^{-1}(A) = A)$ et $0 < \mu(A) < 1$. Montrer que

$$h_\mu(T) = \mu(A)h_{\mu_A}(T) + (1 - \mu(A))h_{\mu_{A^c}}(T).$$

Exercice 11

Soit $(X,\mathcal{B})$ un ensemble muni d'une σ-algèbre. On suppose que $T\ :X\to X$ est une transformation mesurable et que μ et ν sont deux mesures de probabilité invariantes mutuellement singulières. Montrer que pour tout $t\in[0,1]$ on a

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

Exercice 12

Soit $(X,\mathcal{B},\mu,T)$ un système dynamique mesuré, avec $\mu(X)=1$. Soit $\mathcal{P}$ une partition mesurable telle que $\mathcal{P}\in\overline{\bigvee_{m=1}^{+\infty}T^{-m}\mathcal{P}}$. Montrer que $h(T,\mathcal{P})=0$ (en fait la réciproque est vraie).

Exercice 13

Soit $T\ :X\to X$ une application continue définie sur un espace topologique métrisable compact. On rappelle que x est non errant si pour tout voisinage U de x, il existe $n\geq 1$ tel que $T^{-n}(U)\cap U\neq\emptyset$.

1) Montrer que l'ensemble $\Omega(T)$ des points non errants est fermé et vérifie $T(\Omega(T))\subset\Omega(T)$.

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

Exercice 14

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

Chapitre 2

Exercice 1

Montrer qu'un homéomorphisme $\mathbf{T}$ qui renverse l'orientation a exactement deux points fixes. Peut-il avoir des points périodiques de période 2 ? de période >2 ?

Exercice 2

On note $\mathcal{E}$ l'ensemble des applications croissantes $f\ :\mathbf{R}\to\mathbf{R}$ telles que $f-\mathrm{Id}_{\mathbf{R}}$ est périodique, de période 1. On munit $\mathcal{E}$ de la distance suivante

$$d'(f,g)=\max_{x\in\mathbf{R}}|f(x)-g(x)|.$$

1) Montrer qu'il existe un nombre réel ρ tel que pour tout $x\in\mathbf{R}$ et tout $k\in\mathbf{Z}$, on a

$$-1<f^k(x)-x-k\rho<1.$$

3) Quelles propriétés du nombre de rotation des homéomorphismes peuvent être étendues à cette situation ?

Exercice 3

On suppose que $F = H_0^{-1} \circ T_{\widehat{a}} \circ H_0 \in \mathrm{Homeo}_+(\mathbf{T})$ est conjugué à une rotation irrationnelle $T_{\widehat{a}}$, où $H_0 \in \mathrm{Homeo}_+(\mathbf{T})$. Trouver tous les homéomorphismes $H \in \mathrm{Homeo}_+(\mathbf{T})$ tels que $F = H^{-1} \circ T_{\widehat{a}} \circ H$.

Exercice 4

1) Donner un exemple de $f \in D^0(\mathbf{T})$ et $g \in D^0(\mathbf{T})$ tels que $f(x) < g(x)$, pour tout $x \in \mathbf{R}$, et pourtant tel que $\rho(f) = \rho(g)$.

2) Supposons que $f \in D^0(\mathbf{T})$ et $\rho(f) \notin \mathbf{Q}$. Montrer que pour tout ε, il existe $x \in \mathbf{R}$, $q \geq 1$ et $p \in \mathbf{Z}$ tels que $x - \varepsilon < f^q(x) - p < x$.

3) En déduire que si $f \in D^0(\mathbf{T})$ et $g \in D^0(\mathbf{T})$ sont tels que $f(x) < g(x)$ pour tout $x \in \mathbf{R}$ et $\rho(f) = \rho(g)$, alors $\rho(f) \in \mathbf{Q}$.

Exercice 5

On fixe $\alpha \in]-\dfrac{1}{2\pi}, \dfrac{1}{2\pi}[\setminus\{0\}$ et on définit pour tout $t \in \mathbf{R}$ l'application $f_t \ : x \mapsto x + \alpha \sin(2\pi x) + t$ et l'homéomorphisme $F_t \in \mathrm{Homeo}_+(\mathbf{T})$ relevé par f_t.

1) On note également $f_t \ : z \mapsto z + \alpha \sin(2\pi z) + t$ la fonction prolongée à $\mathbf{C}$. Si $q \geq 1$, prouver que la fonction $z \mapsto f_t^q(z) - z$ n'est pas constante sur $\mathbf{C}$.

2) Montrer que chaque F_t a un nombre fini de points périodiques.

3) Montrer que l'application $r \ : t \mapsto \rho(f_t)$ est continue, croissante et vérifie $r(t+1) = r(t)+1$, pour tout $t \in \mathbf{R}$. Montrer que chaque ensemble $r^{-1}(\{a\})$ est un intervalle non trivial si $a \in \mathbf{Q}$ et réduit à un point si $a \notin \mathbf{Q}$.

Exercice 6

Donner un exemple de couple (f, g) dans $D^0(\mathbf{T})$ tel que $\rho(f \circ g) \neq \rho(f) + \rho(g)$.

Exercice 7

Montrer que si $F \in \mathrm{Homeo}_+(\mathbf{T})$ et $G \in \mathrm{Homeo}_+(\mathbf{T})$ commutent et ont chacun un point fixe, alors $F \circ G$ a également un point fixe.

Exercice 8

On dit qu'un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T})$ est *rigide* s'il existe une suite strictement croissante $(n_k)_{k \geq 0}$ d'entiers telle que $\lim_{k \to +\infty} F^{n_k} = \mathrm{Id}_{\mathbf{T}}$ dans $\mathrm{Homeo}_+(\mathbf{T})$. Montrer que F est rigide si et seulement s'il est conjugué à une rotation.

Chapitre 3

Exercice 1

Pour tout $x \in [0,1] \setminus \mathbf{Q}$, on note $[0; a_1(x), a_2(x), \ldots]$ le développement en fraction continue de x.

1) Montrer que pour presque tout $x \in [0,1] \setminus \mathbf{Q}$ et pour tout entier $l \geq 1$, on a

$$\lim_{n \to +\infty} \frac{1}{n} \sharp \left(\{ k \in \{1, \ldots, n\} \mid a_k(x) = l \} \right) = \frac{1}{\ln 2} \ln \left(1 + \frac{1}{l(l+2)} \right).$$

2) Montrer qu'il existe $a \in [1,+\infty]$ tel que pour presque tout $x \in [0,1] \setminus \mathbf{Q}$, on a

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} a_k(x) = a.$$

Que vaut a ?

3) Soit $(k_n)_{n\geq 1}$ une suite de réels > 0 tels que $\sum_{n=1}^{\infty} k_n^{-1} < +\infty$. Montrer que l'ensemble

$$X = \{x \in [0,1] \setminus \mathbf{Q} \,|\, \forall n_0 \geq 1, \exists n \geq n_0, a_n(x) \geq k_n\}$$

est de mesure de Lebesgue nulle.

Exercice 2

Pour tout $x \in [0,1] \setminus \mathbf{Q}$, on note $[0; a_1(x), a_2(x), \ldots]$ le développement en fraction continue de x. Parmi les assertions suivantes, laquelle est la bonne ?

i) pour presque tout point $x \in [0,1] \setminus \mathbf{Q}$ (par rapport à la mesure de Lebesgue), il y a asymptotiquement autant d'entiers pairs que d'entiers impairs dans la suite $(a_n)_{n\geq 0}$;

ii) pour presque tout point $x \in [0,1] \setminus \mathbf{Q}$, il y a asymptotiquement plus d'entiers pairs que d'entiers impairs dans la suite $(a_n)_{n\geq 0}$;

iii) pour presque tout point $x \in [0,1] \setminus \mathbf{Q}$, il y a asymptotiquement moins d'entiers pairs que d'entiers impairs dans la suite $(a_n)_{n\geq 0}$;

iv) aucune des assertions **i)**, **ii)**, **iii)** n'est vraie.

Exercice 3

On note λ la mesure de Lebesgue sur $[0,1]$. Donner un exemple de transformation continue T de $[0,1]$ telle que la suite $(f_*^n(\lambda))_{n\geq 0}$ converge, pour la topologie faible*, et telle que la limite n'est pas absolument continue par rapport à la mesure de Lebesgue.

Exercice 4

Soit $F : \mathbf{T} \to \mathbf{T}$ une transformation continue de degré un et $f : \mathbf{R} \to \mathbf{R}$ un relèvement de F.

1) On note $\mathcal{E}$ l'ensemble des applications continues $h : \mathbf{R} \to \mathbf{R}$ telles que $h(x+1) = h(x) + 1$ pour tout $x \in \mathbf{R}$. Pour tout $h \in \mathcal{E}$ on définit une application $\Phi(h) \ : \mathbf{R} \to \mathbf{R}$ par :

$$\Phi(h)(x) = \frac{1}{p} h(f(x)).$$

Montrer que $\Phi(h)$ appartient à $\mathcal{E}$, puis que $\Phi \ : \mathcal{E} \to \mathcal{E}$ a un unique point fixe h_0. Expliquer pourquoi h_0 relève une transformation continue $H_0 \ : \mathbf{T} \to \mathbf{T}$ de degré 1, puis prouver que H_0 est une semi-conjugaison entre F et $F_* \ : \widehat{x} \to p\widehat{x}$.

7) On garde les hypothèses de la question précédente mais on suppose de plus que F est de classe C^1 et dilatante On note $\mathcal{E}'$ l'ensemble des applications $h \in \mathcal{E}$ qui sont croissantes. Pour tout $h \in \mathcal{E}'$ on définit une application $\Psi(h) \ : \mathbf{R} \to \mathbf{R}$ par :

$$\Psi(h)(x) = f^{-1}(h(px)).$$

Prouvez que $\Psi(h)$ appartient à $\mathcal{E}'$ et que l'application $\Psi \ : \mathcal{E}' \to \mathcal{E}'$ a un unique point fixe h_1. Expliquer pourquoi h_1 est injective et relève un homéomorphisme $H_1 \ : \mathbf{T} \to \mathbf{T}$ de degré 1. Montrez que F et $\widehat{F}_*$ sont conjugués.

Exercice 5 On considère les deux transformations $f\ : [0,1] \to [0,1]$ (la fonction tente) 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

$$g(x) = 4x(1 - x).$$

1) Tracer le graphe des deux fonctions. Expliquer pourquoi ce sont deux fonctions branchées de classe C^∞. Sont elles à distorsion bornée ?

2) Montrer que

$$h : x \mapsto \sin^2\left(\frac{\pi x}{2}\right)$$

définit une conjugaison entre f et g . Quelle est la réciproque de h ?

3) Montrer que la mesure de densité $\dfrac{1}{\pi\sqrt{x(1-x)}}$ par rapport à la mesure de Lebesgue est invariante par g. Expliquer pourquoi c'est la seule mesure de probabilité invariante qui soit absolument continue par rapport à la mesure de Lebesgue.

4) Montrer que pour tout point periodique de g différent de 0, on a $|(g^n)'(x)| = 2^n$, où n est la période de x.

5) Montrer que pour presque tout point, au sens de la mesure de Lebesgue, on a

$$\lim_{n\to+\infty} \frac{1}{n} \ln |(g^n)'(x)| = \ln 2.$$