# La condition de Walters

Thierry Bousch[1], Janvier 2000

**Résumé.** On étudie une condition de régularité introduite par Walters en 1978 pour le formalisme thermodynamique. Nous montrons que cette condition est la “bonne” condition de régularité à imposer pour une classe assez large de problèmes, incluant aussi bien le formalisme thermodynamique que l'étude des mesures maximisantes, ainsi que la recherche de formes normales modulo les cobords de fonctions continues.

**Abstract.** We discuss a regularity condition introduced by Walters in 1978 for the needs of the thermodynamic formalism. We show that it is the “right” regularity condition to require in a large class of problems, including the thermodynamic formalism as well as the study of maximising measures, and the search of normal forms modulo coboundaries of continuous functions.

**Codes matière AMS (1991) :** 46N10, 49J27, 58F15

## 1. Hypothèses, notations et définitions

Dans tout cet article, $(X, d)$ est un espace métrique compact muni d'une application continue surjective $T : X \to X$. Soit $C(X)$ l'ensemble des fonctions continues de $X$ dans $\mathbb{R}$. Pour $n > 0$, on notera $d_n(x, y) = \max_{0 \leqslant k < n} d(T^k x, T^k y)$ et $S_n f = \sum_{0 \leqslant k < n} f \circ T^k$. On posera également $S_0 f = 0$.

Rappelons que le système dynamique $(X, T)$ est dit *transitif* si pour toute paire $U, V$ d'ouverts non vides de $X$, il existe $n \in \mathbb{N}$ tel que $U \cap T^{-n} V \neq \varnothing$. Une propriété équivalente est l'existence d'une “orbite dense”, c.à.d. d'un point $x$ tel que $\omega(x) = X$.

On dira qu'une fonction $f \in C(X)$ est un *cobord topologique* s'il existe $u \in C(X)$ telle que $f = u \circ T - u$. On dira simplement “cobord” pour abréger, dans tout cet article. On notera $\mathrm{Cob}(X, T)$ l'ensemble des cobords, qui est évidemment un espace vectoriel réel, et on dira que $f, g \in C(X)$ sont *cohomologues* si leur différence est un cobord.

**Définition.** *On dira que $f \in C(X)$ vérifie la condition de Walters (ou simplement que $f$ est Walters) si pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que*

$$\forall n > 0 \quad \forall x, y \in X \quad d_n(x, y) \leqslant \eta \implies |S_n f(x) - S_n f(y)| \leqslant \epsilon$$

On notera $\mathrm{Wal}(X, T)$ l'ensemble des fonctions Walters. C'est évidemment un espace vectoriel réel, et on vérifie que la définition donnée ci-dessus ne dépend que de la topologie de $X$ ; ainsi, si deux systèmes dynamiques $(X, T)$ et $(X', T')$ sont topologiquement conjugués, alors $\mathrm{Wal}(X, T)$ et $\mathrm{Wal}(X', T')$ sont isomorphes.

[1] Laboratoire de Mathématique (UMR 8628 du CNRS), bât. 425, Université de Paris-Sud, 91405 Orsay Cedex, France. E-mail: `Thierry.Bousch@math.u-psud.fr`

**Proposition.** *Tout cobord vérifie la condition de Walters.*

La démonstration est évidente, mais ce résultat est très important.

Cette proposition conduit naturellement à s'intéresser à la structure de l'espace quotient $\mathrm{Wal}(X,T)/\mathrm{Cob}(X,T)$, c.à.d. l'espace des classes de cohomologie, à chercher des éléments remarquables (en un sens restant à définir) dans chaque classe de cohomologie.

Pour les systèmes dynamiques expansifs ou hyperboliques (comme par exemple les décalages de Bernoulli), il existe une condition suffisante simple pour vérifier que $f$ est Walters :

**Définition-Proposition.** *On dira que $f \in C(X)$ est à variation sommable si son module de continuité $h$ vérifie*

$$\int_0^1 \frac{h(s)}{s} ds \ < \ \infty$$

*En particulier, toute fonction hölderienne est à variation sommable. Si $(X,T)$ est hyperbolique, alors toute fonction à variation sommable vérifie la condition de Walters.*

En effet, deux orbites qui restent $\eta$-proches pendant un certain temps (avec $\eta$ petit) sont obligées de se rapprocher exponentiellement dans la partie centrale de l'orbite, ce qui permet de majorer $|S_n f(x) - S_n f(y)|$ par

$$2[h(\eta) + h(\eta/\lambda) + h(\eta/\lambda^2) + \cdots]$$

pour un certain $\lambda > 1$, et cette quantité peut être bornée indépendamment de $n$ si $f$ est à variation sommable. Ainsi, toute fonction à variation sommable est dans $\mathrm{Wal}(X,T)$.

La condition de variation sommable est d'une simplicité séduisante, et d'ailleurs on la rencontre fréquemment dans la littérature. Bien que d'une grande utilité pratique, la condition de variation sommable présente le gros inconvénient de ne pas être préservée par l'ajout d'un cobord. Elle ne définit donc pas un espace naturel du point de vue des problèmes que nous allons considérer dans cet article.

La condition de Walters ne présente pas ce problème car elle inclut naturellement les cobords, et nous verrons au chapitre 5 que l'espace des fonctions Walters dispose d'une structure naturelle d'espace de Banach, du moins dans le cas d'une dynamique hyperbolique. Ce n'est, en revanche, ni une algèbre, ni un espace de Riesz : le produit et le maximum de deux fonctions Walters ne sont pas Walters en général.

**Définition.** *On dira que la fonction $h : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$ est un module de Walters pour $f$ si $h$ est croissante, nulle et continue à droite en zéro, et telle que*

$$\forall s \in \mathbb{R}^+ \quad \forall n > 0 \quad \forall x, y \in X \qquad d_n(x,y) \leqslant s \Longrightarrow |S_n f(x) - S_n f(y)| \leqslant h(s)$$

Il est clair que $f$ est Walters si et seulement si elle admet un module de Walters.

## 2. Le lemme de Mañé, version expansive

Soit $\mathcal{M}$ l'ensemble des probabilités boréliennes sur $X$ qui sont invariantes par $T$. Cet ensemble est convexe et compact pour la topologie de la convergence vague. Ainsi, étant donné $f \in C(X)$, l'application $\mu \mapsto \int f\mu$ atteint son maximum sur $\mathcal{M}$. Les probas invariantes $\mu$ pour lesquelles $\int f\mu$ est maximum seront dites *maximisantes* (pour $f$).

Il est clair que les mesures maximisantes, ainsi que la valeur du maximum, ne changent pas quand on ajoute à $f$ un cobord. D'autre part, le lemme sous-additif entraîne facilement le résultat classique suivant :

$$\tfrac{1}{n}\max_{x\in X} S_n f(x) \;\longrightarrow\; \max_{\mu\in\mathcal{M}} \int f\mu \quad \text{quand} \quad n\to\infty\,.$$

Il est impossible de préciser davantage la vitesse de convergence sans faire d'hypothèse supplémentaire sur la fonction $f$ ou sur le système dynamique. On notera $\beta = \max_{\mu\in\mathcal{M}} \int f\mu$.

Dans le cadre des flots lagrangiens, R. Mañé (voir [Man]) avait démontré que que sous certaines hypothèses, $f-\beta$ pouvait s'écrire comme un cobord sur le support des mesures maximisantes ergodiques. Ensuite, A. Fathi a donné dans [Fat] une forme forte de cet énoncé, permettant de majorer $f-\beta$ par un cobord sur $X$ entier, et avec une démonstration bien plus simple que celle originale de Mañé.

Le cadre de cet article est différent, mais l'idée de majorer $f-\beta$ par un cobord est généralement attribuée à Mañé, d'où le nom générique de “lemme de Mañé” pour tout énoncé de ce type. Le théorème suivant entre dans cette catégorie. Avant de l'énoncer, il nous faut donner une définition.

Notons $\mathcal{K}(X)$ l'ensemble des compacts non vides de $X$, muni de la distance de Hausdorff $d_H$. Nous allons définir une condition technique, la propriété d'*expansivité faible* (Weak Expansivity) comme suit :

**Définition.** *On dira que $(X,d,T)$ vérifie la propriété d'expansivité faible, ou propriété (WE), si l'application $T^{-1}:\mathcal{K}(X)\to\mathcal{K}(X)$ est 1-lipschitz pour la distance de Hausdorff. Autrement dit :*

$$\forall x,b\in X \quad \exists y\in T^{-1}(b) \quad d(x,y)\leqslant d(Tx,Ty)$$

Cette condition est extrêmement peu contraignante ; elle inclut en particulier tous les sous-décalages de type fini (quitte à modifier légèrement la distance $d$), mais aussi des applications aussi peu “expansives” que... les isométries de $X$ (par exemple une rotation irrationnelle sur un tore).

**Théorème.** *On suppose que le système dynamique $(X,T)$ est transitif et vérifie la propriété d'expansivité faible (WE). Soit $f\in \mathrm{Wal}(X,T)$ et $\beta=\max_{\mu\in\mathcal{M}}\int f\mu$. Alors il existe $u\in C(X)$ telle que*

$$\forall y\in X \qquad u(y) = -\beta + \max_{Tx=y}(f+u)(x)$$

*Remarques.* Une formulation équivalente du théorème ci-dessus (obtenue en posant $h = f + u - u \circ T$) consiste à dire que $f$ est cohomologue à une fonction $h$ telle que

$$\forall y \in X \qquad \max_{Tx=y} h(x) = \beta.$$

En particulier $h$ est Walters, et $h \leqslant \beta$. Intégrant cette inégalité par une proba invariante quelconque $\mu$, il vient $\int f\mu = \int h\mu \leqslant \beta$, l'égalité étant atteinte si et seulement si $h = \beta$ presque partout, autrement dit :

$$\mu \text{ est } f\text{-maximisante} \quad \Longleftrightarrow \quad \mathrm{supp}(\mu) \subset h^{-1}(\beta)$$

On note en particulier que la propriété que $\mu$ soit maximisante ou non ne dépend que de son support. Plus précisément, on a le "principe de subordination" suivant : *si $\mu$ et $\nu$ sont deux probas invariantes telles que $\nu$ soit maximisante et* $\mathrm{supp}(\mu) \subset \mathrm{supp}(\nu)$, *alors $\mu$ est également maximisante.* En particulier, Coelho a remarqué (voir [Coe]) que le support d'une mesure maximisante ne peut jamais être $X$ tout entier, sauf si $f$ est cohomologue à une constante. Ces résultats reposent de façon essentielle sur le "lemme de Mañé" ci-dessus, et donc sur l'hypothèse que $f$ est Walters ; sans cette hypothèse, on peut mettre en défaut le principe de subordination (voir la remarque au chapitre 8).

Le lemme de Mañé est également d'un grand intérêt pratique pour la détermination explicite des probas maximisantes dans des exemples particuliers, notamment les polynômes trigonométriques de degré un sur le cercle (voir [Bou]).

*Preuve.* Pour $\lambda \in [0, 1]$, définissons $L_\lambda$ comme l'opérateur (non linéaire) de $C(X)$ dans lui-même défini par

$$\forall y \in X \quad [L_\lambda u](y) = \max_{Tx=y}(f + \lambda u)(x)$$

Cet opérateur est $\lambda$-lipschitz pour la norme uniforme sur $C(X)$. Pour $\lambda < 1$, il est contractant ; soit $u_\lambda \in C(X)$ son unique point fixe. On veut déterminer le module de continuité $h_\lambda$ de la fonction $u_\lambda$, à savoir

$$h_\lambda(s) = \max_{d(x,y)\leqslant s} u_\lambda(x) - u_\lambda(y)$$

Soit $s > 0$, et $x_0, y_0 \in X$ deux points tels que $d(x_0, y_0) \leqslant s$. On peut construire par récurrence une suite $(x_i)_{i\geqslant 0}$ prolongeant $x_0$ telle que

$$\forall i \in \mathbb{N} \quad x_i = Tx_{i+1} \text{ et } u_\lambda(x_i) = (f + \lambda u_\lambda)(x_{i+1})$$

D'autre part, d'après la propriété (WE) on peut construire une suite $(y_i)_{i\geqslant 0}$ prolongeant $y_0$ telle que

$$\forall i \in \mathbb{N} \quad y_i = Ty_{i+1} \text{ et } d(x_i, y_i) \geqslant d(x_{i+1}, y_{i+1})$$

Des inégalités $u_\lambda(y_i) \geqslant (f + \lambda u_\lambda)(y_{i+1})$ on déduit

$$\forall i \in \mathbb{N} \quad u_\lambda(x_i) - u_\lambda(y_i) \leqslant f(x_i) - f(y_i) + \lambda\big[u_\lambda(x_{i+1}) - u_\lambda(y_{i+1})\big]$$

et par conséquent

$$u_\lambda(x_0) - u_\lambda(y_0) \leqslant \sum_{i=0}^{\infty} \lambda^i \big[f(x_i) - f(y_i)\big]$$

Appliquant la transformation d'Abel au membre de droite, on obtient la majoration

$$\begin{aligned}
\sum_{i=0}^{\infty} \lambda^i \big[f(x_i) - f(y_i)\big] &= (1-\lambda) \sum_{n\geqslant 0} \lambda^n \sum_{i=0}^{n} f(x_i) - f(y_i) \\
&\leqslant \sup_{n\geqslant 0} \sum_{i=0}^{n} f(x_i) - f(y_i) \\
&= \sup_{n\geqslant 0} \; S_{n+1}f(x_n) - S_{n+1}f(y_n) \\
&\leqslant h_W(s)
\end{aligned}$$

où $h_W$ est un module de Walters pour $f$.

Ainsi, toutes les fonctions $u_\lambda$ (pour $0 \leqslant \lambda < 1$) admettent $h_W$ comme module de continuité commun. Toutefois cette estimation d'équicontinuité est seulement locale : $h_W(s)$ peut être infini pour $s$ assez grand.

Si on veut démontrer que la famille $u_\lambda$ est précompacte dans $C(X)/\mathbb{R}$, il nous reste à majorer l'oscillation de $u_\lambda$ par une quantité indépendante de $\lambda$. Pour cela, notons que si $M_\lambda = \max u_\lambda$ , alors l'équation fonctionnelle sur $u_\lambda$ implique

$$\begin{aligned}
\forall x \in X \qquad (M_\lambda - u_\lambda)(Tx) &\leqslant \mathrm{Osc}(f) + \lambda \cdot (M_\lambda - u_\lambda)(x) \\
&\leqslant \mathrm{Osc}(f) + (M_\lambda - u_\lambda)(x)
\end{aligned}$$

Soit $\alpha > 0$ tel que $h_W(\alpha) \leqslant 1$, et soit $\mathcal{R}$ un ensemble fini de boules fermées de rayon $\alpha$ recouvrant $X$. Soient $\mathbf{b}_1, \mathbf{b}_2$ deux éléments de $\mathcal{R}$. Comme $T$ est transitive, il existe $(x,n) \in X \times \mathbb{N}$ tel que $x \in \mathbf{b}_1$ et $T^n x \in \mathbf{b}_2$. De l'inégalité

$$(M_\lambda - u_\lambda)(T^n x) \leqslant n\,\mathrm{Osc}(f) + (M_\lambda - u_\lambda)(x)$$

on déduit

$$\max_{(s_1,s_2)\in \mathbf{b}_1\times\mathbf{b}_2} u_\lambda(s_1) - u_\lambda(s_2) \;\leqslant\; 2h_W(\alpha) + n\,\mathrm{Osc}(f)$$

En écrivant cette inégalité pour tous les couples $(\mathbf{b}_1, \mathbf{b}_2) \in \mathcal{R}^2$, et notant $N$ le maximum de tous les entiers $n$ apparaissant ci-dessus, on obtient

$$\forall \lambda \in [0,1[ \qquad \mathrm{Osc}\; u_\lambda \;\leqslant\; 2 + N\,\mathrm{Osc}\, f$$

et les fonctions $u_\lambda$ sont bien d'oscillation bornée. Par conséquent, la famille $u^*_\lambda = u_\lambda - \min u_\lambda$ est précompacte dans $C(X)$ et vérifie

$$\forall \lambda \in [0,1[ \quad \exists b_\lambda \in \mathbb{R} \qquad u^*_\lambda = -b_\lambda + L_\lambda u^*_\lambda$$

Si $u \in C(X)$ désigne une valeur d'adhérence quelconque des $(u_\lambda)$ quand $\lambda \to 1$, alors on aura $u = -b + L_1 u$, avec une certaine constante $b \in \mathbb{R}$.

Il reste à voir que cette constante $b$ n'est autre que $\beta = \max_{\mu \in \mathcal{M}} \int f\mu$. Pour cela, posons $h = f + u - u \circ T$ ; cette fonction est cohomologue à $f$ et vérifie

$$\forall y \in X \qquad b = \max_{Tx=y} h(x)$$

Comme $h \leqslant b$, on aura $\int f\mu = \int h\mu \leqslant b$ pour toute proba invariante $\mu$, donc $\beta \leqslant b$. Mais d'autre part, le compact $K = h^{-1}(b)$ vérifie $TK = X$, et ceci implique que le compact invariant $K' = \bigcap_{n \geqslant 0} T^{-n}K$ est non vide. Il porte donc une proba invariante $\mu$, pour laquelle $h = b$ presque partout, d'où $\int f\mu = \int h\mu = b$, et donc $b \leqslant \beta$. On a donc bien $b = \beta$, et le théorème est démontré.

## 3. Propriétés de la limite projective

J'aurai besoin du résultat topologique suivant, dont la démonstration (un peu délicate) est donnée en annexe.

**Proposition A.** *Soit $\pi : X \to Y$ une surjection continue entre deux espaces topologiques compacts, et $\Delta = \big\{(x,y) \in X^2 : \pi(x) = \pi(y)\big\}$. On suppose qu'on a une fonction $u \in C(\Delta)$ vérifiant la "relation de Chasles"*

$$\forall x, y, z \in X \qquad \pi(x) = \pi(y) = \pi(z) \quad \Longrightarrow \quad u(x,y) + u(y,z) + u(z,x) = 0$$

*Alors il existe $f \in C(X)$ telle que $\|f\|_\infty = \frac{1}{2}\|u\|_\infty$ et*

$$\forall (x,y) \in \Delta \qquad u(x,y) = f(x) - f(y)\,.$$

Rappelons que la *limite projective* (aussi appelée *extension naturelle*) du système dynamique $(X,T)$ est un autre système dynamique $(\hat{X}, \hat{T})$ défini de la façon suivante : $\hat{X}$ est l'ensemble des suites $(x_i)_{i \in \mathbb{N}}$ telles que $\forall i \; x_i = Tx_{i+1}$, muni de la topologie induite par la topologie produit de $X^{\mathbb{N}}$. L'ensemble $\hat{X}$ est donc un compact métrisable, que nous munirons de la distance

$$d\big[(x_i),(y_i)\big] = \max_{n \in \mathbb{N}} \; 2^{-n} d(x_n, y_n)$$

L'application $\hat{T} : \hat{X} \to \hat{X}$ est définie par

$$\hat{T}\big[(x_0, x_1, \ldots)\big] = (Tx_0, x_0, x_1, \ldots)$$

Il est clair que $\hat{T} : \hat{X} \to \hat{X}$ est une bijection, dont l'inverse est l'application de décalage, et que le diagramme suivant est commutatif :

$$\begin{array}{ccc} \hat{X} & \xrightarrow{\hat{T}} & \hat{X} \\ \big\downarrow{\scriptstyle\pi} & & \big\downarrow{\scriptstyle\pi} \\ X & \xrightarrow{T} & X \end{array}$$

où $\pi : \hat{X} \to X$ est la projection définie par $\pi[(x_i)] = x_0$.

En particulier, il existe une injection naturelle de $C(X)$ dans $C(\hat{X})$, définie par $f \mapsto f \circ \pi$, et cette injection permet d'identifier $C(X)$ à un sous-espace de $C(\hat{X})$. Cependant, il n'est pas clair que cette identification soit compatible avec la condition de Walters : peut-on dire qu'une fonction $f \in C(X)$ est Walters sans préciser si le système dynamique sous-jacent est $(X, T)$ ou $(\hat{X}, \hat{T})$ ? La proposition suivante répond par l'affirmative, ce qui légitime l'identification $\mathrm{Wal}(X, T) = C(X) \cap \mathrm{Wal}(\hat{X}, \hat{T})$.

**Proposition.** *Soit $f \in C(X)$. Alors $f \in \mathrm{Wal}(X, T) \iff f \circ \pi \in \mathrm{Wal}(\hat{X}, \hat{T})$.*

*Preuve.* L'implication $\Longrightarrow$ est une conséquence directe de la semi-conjugaison entre $(\hat{X}, \hat{T})$ et $(X, T)$. Soit $n > 0$ et $u, v \in \hat{X}$ quelconques, et $\hat{f} = f \circ \pi$. Comme $\pi$ est 1-lipschitz, on a

$$\forall k \in [0, n[ \qquad d(T^k \pi u, T^k \pi v) = d(\pi \hat{T}^k u, \pi \hat{T}^k v) \leqslant d(\hat{T}^k u, \hat{T}^k v) \leqslant d_n(u, v)$$

et donc $d_n(\pi u, \pi v) \leqslant d_n(u, v)$. Et d'autre part,

$$S_n \hat{f}(u) = \sum_{0 \leqslant k < n} (f \circ \pi \circ \hat{T}^n)(u) = \sum_{0 \leqslant k < n} (f \circ T^n)(\pi u) = S_n(f)(\pi u)$$

Ceci entraîne immédiatement que $f \in \mathrm{Wal}(X, T) \Longrightarrow \hat{f} \in \mathrm{Wal}(\hat{X}, \hat{T})$.

L'implication $\Longleftarrow$ est plus délicate. Supposons que $\hat{f} \in \mathrm{Wal}(\hat{X}, \hat{T})$. Soit $\epsilon > 0$ quelconque. On peut trouver un $\eta > 0$ tel que

$$\forall n > 0 \quad \forall u, v \in \hat{X} \qquad d_n(u, v) \leqslant \eta \Longrightarrow \left| S_n \hat{f}(u) - S_n \hat{f}(v) \right| \leqslant \epsilon$$

Choisissons un entier $p \geqslant 1$ tel que $2^{-p} \mathrm{diam}(X) \leqslant \eta$. Et soit $\eta_1 \in ]0, \eta]$ tel que

$$\forall n \in [0, p[ \quad \forall x_0, y_0 \in X \qquad d(x_0, y_0) \leqslant \eta_1 \Longrightarrow |S_n f(x_0) - S_n f(y_0)| \leqslant \epsilon$$

Alors j'affirme que

$$\forall n > 0 \quad \forall x_0, y_0 \in X \qquad d_n(x_0, y_0) \leqslant \eta_1 \Longrightarrow |S_n f(x_0) - S_n f(y_0)| \leqslant 2\epsilon$$

Vu la définition de $\eta_1$, il suffit évidemment de démontrer cette inégalité pour $n \geqslant p$. Soient $x, y \in \hat{X}$ tels que $\pi x = x_0$ et $\pi y = y_0$, et puis $u = \hat{T}^{p-1}(x)$ et $v = \hat{T}^{p-1}(y)$. Alors pour tout $s \in \mathbb{N}$, on a

$$\begin{aligned}
d(\hat{T}^s u, \hat{T}^s v) &= d(\hat{T}^{s+p-1} x, \hat{T}^{s+p-1} y) \\
&= \max_{k \geqslant 0} 2^{-k} d\big[\pi(\hat{T}^{s+p-1-k} x), \pi(\hat{T}^{s+p-1-k} y)\big] \\
&\leqslant \max\Big[\max_{k \geqslant p} \big[2^{-k} \mathrm{diam}(X)\big], \max_{0 \leqslant k < p} 2^{-k} d\big[\pi(\hat{T}^{s+p-1-k} x), \pi(\hat{T}^{s+p-1-k} y)\big]\Big] \\
&\leqslant \max\big[2^{-p} \mathrm{diam}(X), \max_{0 \leqslant k < p} d(T^{s+p-1-k} x_0, T^{s+p-1-k} y_0)\big] \\
&\leqslant \max\big[\eta, d_{s+p}(x_0, y_0)\big]
\end{aligned}$$

Par conséquent, pour tout $n \geqslant p$ on a

$$d_{n-p+1}(u,v) \leqslant \max\bigl(\eta, d_n(x_0,y_0)\bigr)$$

Ainsi $d_n(x_0,y_0) \leqslant \eta_1$ entraîne $d_{n-p+1}(u,v) \leqslant \eta$ et

$$\Bigl|\sum_{p-1\leqslant k<n} fT^k x_0 - fT^k y_0\Bigr| = \Bigl|S_{n-p+1}\hat{f}(u) - S_{n-p+1}\hat{f}(v)\Bigr| \leqslant \epsilon$$

ce qui, combiné avec

$$\Bigl|\sum_{0\leqslant k<p-1} fT^k x_0 - fT^k y_0\Bigr| \leqslant \epsilon$$

donne bien $|S_n f(x_0) - S_n f(y_0)| \leqslant 2\epsilon$. La fonction $f$ est donc Walters, ce qui prouve la réciproque, et termine la preuve de la proposition.

Signalons un résultat analogue mais plus facile, dont la démonstration est laissée au lecteur :

**Proposition.** *Soit* $f \in C(X)$. *Alors* $f \in \mathrm{Cob}(X,T) \iff f\circ\pi \in \mathrm{Cob}(\hat{X},\hat{T})$.

Les deux propositions précédentes entraînent l'existence d'une application injective naturelle de $\mathrm{Wal}(X,T)/\mathrm{Cob}(X,T)$ dans $\mathrm{Wal}(\hat{X},\hat{T})/\mathrm{Cob}(\hat{X},\hat{T})$ ; le théorème suivant affirme que cette application est un isomorphisme d'espaces vectoriels.

**Théorème.** *Toute fonction* $f \in \mathrm{Wal}(\hat{X},\hat{T})$ *est cohomologue à (au moins) une fonction de la forme* $h\circ\pi$, *avec* $h \in \mathrm{Wal}(X,T)$.

*Preuve.* Soit $f \in \mathrm{Wal}(\hat{X},\hat{T})$. Comme dans la Proposition A, on note

$$\Delta = \bigl\{(x,y)\in \hat{X}^2 : \pi(x)=\pi(y)\bigr\}$$

Considérons la fonction $u \in C(\Delta)$ définie par

$$\forall (x,y)\in\Delta \qquad u(x,y) = \sum_{n=0}^{\infty} f\hat{T}^n x - f\hat{T}^n y$$

Notons d'abord que cette définition a un sens. En effet, si $\pi(x)=\pi(y)$ alors $d(\hat{T}^n x, \hat{T}^n y) \leqslant 2^{-n-1}\mathrm{diam}(X)$. Comme $f$ est Walters, la série ci-dessus est de Cauchy et donc convergente dans $C(\Delta)$ pour la topologie de la convergence uniforme. D'autre part, on a évidemment

$$\forall x,y,z\in\hat{X} \qquad \pi(x)=\pi(y)=\pi(z) \implies u(x,y)+u(y,z)+u(z,x)=0$$

En vertu de la Proposition A, on peut trouver une fonction $v\in C(\hat{X})$ telle que $u(x,y) = v(x)-v(y)$ pour tout $(x,y)\in\Delta$. Par ailleurs, $u$ vérifie également l'équation fonctionnelle

$$\forall (x,y)\in\Delta \qquad u(x,y) = f(x)-f(y)+u(\hat{T}x,\hat{T}y)$$

En termes de la fonction $v$, cela s'écrit

$$\forall x, y \in \hat{X} \qquad \pi(x) = \pi(y) \implies f(x) + v(\hat{T}x) - v(x) = f(y) + v(\hat{T}y) - v(y)$$

Autrement dit, la fonction $\hat{h} = f + v \circ \hat{T} - v$ est constante sur les fibres de $\pi$. Elle peut donc s'écrire $\hat{h} = h \circ \pi$ avec $h \in C(X)$. La fonction $\hat{h}$ est cohomologue à $f$ donc Walters, et donc $h$ est Walters également. Le théorème est démontré.

Coelho et Quas ont donné un analogue de ce théorème pour les fonctions à variation sommable quand $(X, T)$ est un décalage de Bernoulli, voir [CQ] théorème 2. Pour les fonctions Hölder, c'est un résultat très classique dans la théorie des mesures de Gibbs, voir en particulier [Sin].

## 4. Le lemme de Mañé, version hyperbolique

Pour des systèmes dynamiques inversibles, la propriété (WE) ne sera en général pas vérifiée (parce que l'application $T$ sera expansive dans certaines directions seulement) et on utilisera plutôt la propriété suivante, analogue mais plus faible :

**Définition.** *On dira que $(X, T)$ vérifie la propriété de produit local faible, ou propriété (WLP), si pour tout $\epsilon > 0$, on peut trouver $\eta > 0$ vérifiant la propriété suivante : quelles que soient les orbites $(x_i)_{i \leqslant 0}$ et $(y_i)_{i \geqslant 0}$ telles que $d(x_0, y_0) \leqslant \eta$, il existe une orbite $(z_i)_{i \in \mathbb{Z}}$ telle que $\forall i \leqslant 0$ $d(x_i, z_i) \leqslant \epsilon$ et $\forall i \geqslant 0$ $d(y_i, z_i) \leqslant \epsilon$.*

*Remarque.* Dans la définition ci-dessus, on appelle “orbite” une famille indicée $(x_i)_{i \in A}$ d'éléments de $X$, où $A$ est une partie de $\mathbb{Z}$, telle que $Tx_i = x_{i+1}$ pour tout $i$ tel que $i \in A$ et $i + 1 \in A$.

**Théorème.** *On suppose que $(X, T)$ est transitif et vérifie la propriété de produit local faible (WLP). Soit $f \in \mathrm{Wal}(X, T)$, et $\beta = \max_{\mu \in \mathcal{M}} \int f \mu$. Alors il existe $u \in C(X)$ telle que $f \leqslant \beta + u \circ T - u$.*

Une formulation équivalente est que $f$ est cohomologue à une fonction (en l'occurence, $h = f + u - u \circ T$) majorée par $\beta$. On a, en particulier, le “principe de subordination” discuté plus haut.

*Preuve du théorème.* Quitte à translater $f$, on peut supposer $\beta = 0$. Dans ce cas, nous avons vu que les sommes de Birkhoff vérifient $\frac{1}{n} \max S_n f \to 0$ quand $n \to \infty$. Nous devons tout d'abord affiner cette estimation.

**Lemme.** *Sous les hypothèses du théorème, avec $\beta = 0$, les sommes de Birkhoff sont uniformément majorées : il existe une constante $M \in \mathbb{R}$ telle que $S_m f(x) \leqslant M$ pour tous $x \in X$ et $m \in \mathbb{N}$.*

*Preuve du lemme.* D'après la condition de Walters, on peut trouver $\epsilon > 0$ tel que

$$\forall x, y \in X \quad \forall n > 0 \qquad d_n(x, y) \leqslant \epsilon \implies |S_n f(x) - S_n f(y)| \leqslant 1$$

Ensuite, appliquant deux fois la propriété (WLP), on voit qu'il existe $\eta > 0$ vérifiant la propriété suivante : quelles que soient les orbites $(x_i)_{a \leqslant i \leqslant b}$, $(y_i)_{b \leqslant i \leqslant c}$ et $(z_i)_{c \leqslant i \leqslant d}$ telles que $d(x_b, y_b) \leqslant \eta$ et $d(y_c, z_c) \leqslant \eta$, il existe une orbite $(t_i)_{a \leqslant i \leqslant d}$ qui est $\epsilon$-proche de $(x_i)$, $(y_i)$ et $(z_i)$.

La transitivité de $T$ et la compacité de $X^2$ impliquent l'existence d'une partie finie $\mathcal{L}$ de $X \times \mathbb{N}$ vérifiant

$$\forall x, y \in X \quad \exists (z, \ell) \in \mathcal{L} \qquad d(x, z) \leqslant \eta \quad \text{et} \quad d(y, T^\ell z) \leqslant \eta$$

Posons alors $L = \max\{\ell : (z, \ell) \in \mathcal{L}\}$ et $K = \min\{S_\ell f(z) : (z, \ell) \in \mathcal{L}\}$. J'affirme que

$$\forall x \in X \quad \forall n > 0 \qquad S_n f(x) \leqslant 3 - K.$$

Pour cela, on va démontrer que pour tout $(x, n) \in X \times \mathbb{N}$, on peut trouver $(x', n') \in X \times \mathbb{N}$ tel que $n' \leqslant 2n + L$ et $S_{n'} f(x') \geqslant 2 S_n f(x) + K - 3$. Il n'est pas difficile de voir que ceci entraîne l'affirmation ci-dessus (exercice).

Soit $(x, n) \in X \times \mathbb{N}$. On peut trouver $(y, \ell) \in \mathcal{L}$ tel que $d(T^n x, y) \leqslant \eta$ et $d(x, T^\ell y) \leqslant \eta$. Par définition de $\eta$, on peut recoller les trois orbites

$$\begin{gathered}(x, Tx, \dots T^n x)\\ (y, Ty, \dots T^\ell y)\\ (x, Tx, \dots T^n x)\end{gathered}$$

en une orbite $(t_i)_{0 \leqslant i \leqslant 2n+\ell}$ telle que $d_n(x, t_0) \leqslant \epsilon$, $d_\ell(y, t_n) \leqslant \epsilon$ et $d_n(x, t_{n+\ell}) \leqslant \epsilon$. Par conséquent, on aura les inégalités

$$\begin{gathered}\sum_{i=0}^{n-1} f(t_i) \geqslant S_n f(x) - 1\\ \sum_{i=n}^{n+\ell-1} f(t_i) \geqslant K - 1\\ \sum_{i=n+\ell}^{2n+\ell-1} f(t_i) \geqslant S_n f(x) - 1\end{gathered}$$

et donc finalement $S_{2n+\ell} f(t_0) \geqslant 2 S_n f(x) + K - 3$, ainsi le couple $(x', n') = (t_0, 2n + \ell)$ vérifie bien les propriétés requises. Ceci termine la preuve du lemme.

Revenons à la preuve du théorème. Pour toute paire $U, V$ d'ouverts de $X$, définissons

$$\tilde{u}(U, V) = \sup \left\{ S_n f(x) : n \geqslant 0, x \in U, T^n x \in V \right\}$$

Cette fonction est manifestement croissante en $U$ et $V$, elle ne prend pas la valeur $-\infty$ à cause de la transitivité de $T$, et elle est bornée supérieurement d'après le lemme ci-dessus. Définissons maintenant, pour tout $(x, y) \in X^2$,

$$u(x, y) = \inf \left\{ \tilde{u}(U, V) : x \in U, y \in V,\ U \text{ et } V \text{ ouverts} \right\}$$

Par construction, $u$ est automatiquement une fonction semi-continue supérieurement de $X^2$ dans $\mathbb{R} \cup \{-\infty\}$ vérifiant $S_n f(x) \leqslant u(x, T^n x)$ pour tous $x \in X$ et $n \geqslant 0$. En fait, il est facile de se convaincre que $u$ est la plus petite fonction s.c.s. vérifiant cette propriété.

**Lemme.** *On a $u(x,y) + u(y,z) \leqslant u(x,z)$ pour tous $x, y, z \in X$.*

*Preuve du lemme.* Soient $x, y, z$ quelconques dans $X$. L'inégalité ci-dessus est évidente si $u(x,y)$ ou $u(y,z)$ vaut $-\infty$, et nous supposerons donc que ce n'est pas le cas.

Soit $\epsilon > 0$ et $\Theta > u(x,z)$ quelconques. Par définition de $u$, il existe $\theta_0 > 0$ tel que

$$u(x,z) \leqslant \tilde{u}\big(B(x, 2\theta_0), B(z, 2\theta_0)\big) \leqslant \Theta$$

D'autre part, d'après la condition de Walters, on peut trouver $\theta \in \,]0, \theta_0]$ tel que

$$\forall p, q \in X \quad \forall n \in \mathbb{N} \qquad d_n(p,q) \leqslant \theta \implies |S_n f(p) - S_n f(q)| \leqslant \epsilon$$

Et d'après la propriété (WLP), il existe $\eta \in \,]0, 2\theta]$ tel que pour toute paire d'orbites $(x_i)_{a \leqslant i \leqslant b}$, $(y_i)_{b \leqslant i \leqslant c}$ telles que $d(x_b, y_b) \leqslant \eta$, il existe une orbite $(z_i)_{a \leqslant i \leqslant c}$ qui est $\theta$-proche de $(x_i)$ et $(y_i)$.

Par définition de $u$, on peut trouver $(s,n) \in X \times \mathbb{N}$ tel que $s \in B(x, \eta/2)$, $T^n s \in B(y, \eta/2)$, $S_n f(s) \geqslant u(x,y) - \epsilon$, et $(t,p) \in X \times \mathbb{N}$ tel que $t \in B(y, \eta/2)$, $T^p t \in B(z, \eta/2)$ et $S_p f(t) \geqslant u(y,z) - \epsilon$. Les deux orbites $(s, Ts, \ldots T^n s)$ et $(t, Tt, \ldots T^p t)$ peuvent être approchées à $\theta$ près par une orbite $(\alpha, T\alpha, \ldots T^{n+p}\alpha)$. On a en particulier $d_n(s, \alpha) \leqslant \theta$ et $d_p(t, T^n\alpha) \leqslant \theta$, d'où

$$S_n f(\alpha) \geqslant S_n f(s) - \epsilon \geqslant u(x,y) - 2\epsilon$$
$$S_p f(T^n \alpha) \geqslant S_p f(t) - \epsilon \geqslant u(y,z) - 2\epsilon$$

et donc

$$S_{n+p} f(\alpha) \geqslant u(x,y) + u(y,z) - 4\epsilon$$

Mais d'autre part, on a $\alpha \in B(x, \eta/2 + \theta) \subset B(x, 2\theta_0)$ et $T^{n+p}\alpha \in B(z, \eta/2 + \theta) \subset B(z, 2\theta_0)$, donc

$$S_{n+p} f(\alpha) \leqslant \tilde{u}\big(B(x, 2\theta_0), B(z, 2\theta_0)\big)$$

et par suite

$$\Theta \geqslant u(x,y) + u(y,z) - 4\epsilon$$

Faisant tendre $\epsilon \to 0$ et $\Theta \to u(x,z)$ il vient finalement $u(x,z) \geqslant u(x,y) + u(y,z)$ et le lemme est démontré.

Pour terminer la preuve du théorème, nous aurons besoin du résultat suivant, dont la preuve est donnée en annexe :

**Proposition B.** *Soit $X$ un espace topologique compact, et $u : X^2 \to \mathbb{R} \cup \{-\infty\}$ une fonction semi-continue supérieurement vérifiant*

$$\forall x, y, z \in X \qquad u(x,y) + u(y,z) \leqslant u(x,z)$$

*Alors il existe $f \in C(X)$ telle que $\|f\|_\infty = \frac{1}{2}(\max u)^+$ et*

$$\forall x, y \in X \qquad u(x,y) \leqslant f(x) - f(y).$$

D'après cette proposition, il existe une fonction $g \in C(X)$ telle que $u(x,y) \leqslant g(y) - g(x)$ pour tous $x, y \in X$. Alors $f(x) = S_1 f(x) \leqslant u(x, Tx) \leqslant g(Tx) - g(x)$ pour tout $x \in X$ et le théorème est démontré (pour $\beta = 0$).

Indiquons une conséquence importante de ce théorème, à savoir, une caractérisation des cobords :

**Théorème.** *On suppose $(X, T)$ transitif et vérifiant la propriété de produit local faible. Soit $f \in C(X)$. Alors $f$ est un cobord si et seulement si $f$ est Walters et de moyenne nulle pour toute proba invariante.*

*Preuve.* Les deux conditions sont évidemment nécessaires ; montrons qu'elles sont suffisantes. Soit $f \in \mathrm{Wal}(X, T)$ telle que $\int f\mu = 0$ pour tout $\mu \in \mathcal{M}$. On a donc $\beta = 0$ et d'après le théorème précédent, il existe $u \in C(X)$ telle que $f \leqslant u \circ T - u$. Appliquant le même argument à la fonction $-f$, on trouve $v \in C(X)$ telle que $-f \leqslant v \circ T - v$. Ajoutant ces deux inégalités, il vient $(u + v) \circ T \geqslant u + v$, ce qui n'est possible que si $u + v$ est constante, compte tenu de la transitivité de $T$. On a donc $f = u \circ T - u$ et le théorème est démontré.

Ce théorème étend un résultat classique de Livšic (voir [Liv]), qui étudie le cas où $f$ est hölderienne, et $(X, T)$ est Axiome A et transitif.

*Remarque.* Les deux conditions nécessaires et suffisantes pour que $f \in \mathrm{Cob}(X, T)$, à savoir que (i) $f$ soit Walters et (ii) de moyenne nulle pour toute proba invariante, sont indépendantes en ce sens qu'aucune des deux n'implique l'autre. En particulier on peut trouver des fonctions continues d'intégrale nulle pour toute proba invariante et qui ne sont pas Walters, et donc ne sont pas des cobords (M. Zinsmeister, communication privée).

## 5. Normes sur l'espace des fonctions Walters

De façon générale, $\mathrm{Wal}(X, T)$ n'admet pas de topologie naturelle. Dans ce chapitre, nous allons voir comment le munir d'une structure d'espace de Banach, moyennant une hypothèse sur la dynamique.

**Définition-Proposition.** *Si $X$ n'est pas réduit à un point, alors il existe un plus grand réel $S \geqslant 0$ vérifiant la propriété suivante : si $(x_k)_{k\in\mathbb{Z}}$ et $(y_k)_{k\in\mathbb{Z}}$ sont deux orbites quelconques pour $T$, alors*

$$\sup_{k\in\mathbb{Z}} d(x_k, y_k) < S \quad \Longrightarrow \quad \forall k \in \mathbb{Z} \;\; x_k = y_k$$

*Cette constante sera appelée constante de séparation du système dynamique $(X, T)$ et notée* $\mathrm{sep}(X, d, T)$. *Si $d'$ est une autre distance définissant la même topologie, alors*

$$\mathrm{sep}(X, d, T) > 0 \quad \Longleftrightarrow \quad \mathrm{sep}(X, d', T) > 0.$$

*Preuve.* Soit $\mathcal{I}$ l'ensemble des $S \in \mathbb{R}^+$ vérifiant la propriété de l'énoncé. On voit facilement que $\mathcal{I}$ est un sous-intervalle de $\mathbb{R}^+$ contenant 0, majoré par $\operatorname{diam} X$, et contenant son supremum ; on a donc bien $\mathcal{I} = [0, S]$ pour un certain $S \in \mathbb{R}^+$. Quant à l'équivalence

$$\mathrm{sep}(X, d, T) > 0 \quad \Longleftrightarrow \quad \mathrm{sep}(X, d', T) > 0,$$

elle découle de l'équivalence des structures uniformes définies par $d$ et $d'$ ; nous écrirons simplement $\operatorname{sep}(X,T) > 0$ avec un abus de notation évident, et nous dirons que la dynamique (ou l'application) est *séparante*[2].

Notons que cette propriété n'est pas affectée par le passage à la limite projective : $(\hat{X},\hat{T})$ est séparant si et seulement si $(X,T)$ est séparant. Il en est de même du passage à l'application inverse, quand la dynamique est inversible.

**Lemme.** *Soit $(X,T)$ de constante de séparation $S > 0$. Alors pour tout $s \in \,]0,S[$, il existe une fonction $\theta_s : \mathbb{N} \to \mathbb{R}^+$ décroissante, tendant vers $0$ et telle que, quelles que soient les orbites $(x_i)_{a\leqslant i\leqslant b}$ et $(y_i)_{a\leqslant i\leqslant b}$, on ait*

$$\max_{a\leqslant k\leqslant b} d(x_k,y_k) \leqslant s \quad \Longrightarrow \quad \forall k \in [a,b] \;\; d(x_k,y_k) \leqslant \theta_s\big[\min(k-a,b-k)\big]$$

*Preuve.* Définissons $\theta_s(n)$ comme le maximum de $d(x_0,y_0)$ sur toutes les paires d'orbites $(x_i)_{-n\leqslant i\leqslant n}$ et $(y_i)_{-n\leqslant i\leqslant n}$ vérifiant la condition $\max_{-n\leqslant k\leqslant n} d(x_k,y_k) \leqslant s$. Il est facile de voir que $\theta_s$ vérifie bien les propriétés annoncées.

**Corollaire.** *Soit $(X,T)$ de constante de séparation $S > 0$, et $f \in \operatorname{Wal}(X,T)$. Alors la quantité*

$$\sup_{n>0} \max_{d_n(x,y)\leqslant s} |S_n f(x) - S_n f(y)|$$

*est finie pour tout $s < S$. En d'autres termes, tout élément de $\operatorname{Wal}(X,T)$ admet un module de Walters $h_W$ tel que $h_W(s) < \infty$ pour tout $s < S$.*

**Proposition.** *On suppose que la constante de séparation $S$ du système dynamique $(X,T)$ est non nulle. Alors $\operatorname{Wal}(X,T)$ muni de l'une quelconque des normes $W_s$ définies par*

$$W_s(f) = 2\,\|f\|_\infty + \sup_{n>0} \max_{d_n(x,y)\leqslant s} |S_n f(x) - S_n f(y)|$$

*avec $s \in \,]0,S[$, est un espace de Banach. De plus, toutes ces normes sont équivalentes.*

*Preuve.* Le lecteur vérifiera aisément que les $W_s$ sont des normes sur $\operatorname{Wal}(X,T)$ et qu'elles en font un espace complet. D'autre part, la majoration $W_s \leqslant W_t$ quand $s \leqslant t$ implique $W_s \sim W_t$ en vertu du théorème de Banach. Ainsi, toutes ces normes définissent la même topologie, que nous appellerons la *topologie de Walters*.

Dans toute la suite, $\operatorname{Wal}(X,T)$ sera supposé muni de cette topologie (sous réserve que la dynamique soit séparante). Voici un premier résultat très simple, et vrai en toute généralité, justifiant l'introduction de cette topologie :

**Proposition.** *On suppose que la dynamique est séparante. Alors l'opérateur de cobord $u \mapsto u\circ T - u$ est continu de $C(X)$ dans $\operatorname{Wal}(X,T)$.*

[2] L'auteur est bien conscient que cette terminologie est non-standard ; une telle dynamique est habituellement appelée *expansive*, avec tous les risques de confusion imaginables (aussi bien en français qu'en anglais) sur le sens de ce mot.

## 6. Un cas particulier, le décalage de Bernoulli

Ici nous considérons $X = A^{\mathbb{N}}$, où $A$ est un alphabet fini, muni de la distance $d(a,b) = 2^{-n}$, où $n$ est le plus petit indice tel que $a_n \neq b_n$, et $\sigma : X \to X$ est l'application de décalage. Cette dynamique est transitive, séparante, et (WE) pour la distance $d$. La constante de séparation vaut $\operatorname{sep}(X,d,\sigma) = \operatorname{diam} X = 1$.

Soit $\mathcal{E}$ l'opérateur de transfert particulier défini par

$$\forall y \in X \qquad [\mathcal{E}f](y) \;=\; \frac{1}{\sharp A}\sum_{\sigma x = y} f(x) \;=\; \frac{1}{\sharp A}\sum_{a\in A} f(ay)$$

et notons $\operatorname{Wal}^{\dagger}_{\mathcal{E}}(X,T)$ l'ensemble des $f \in \operatorname{Wal}(X,T)$ tels que $\mathcal{E}f = 0$. Pour abréger, nous omettrons l'indice $\mathcal{E}$ dans la suite.

**Théorème.** *L'application*

$$\begin{array}{ccc} \big[C(A^{\mathbb{N}})/\mathbb{R}\big] \times \operatorname{Wal}^{\dagger}(A^{\mathbb{N}},\sigma) \times \mathbb{R} & \longrightarrow & \operatorname{Wal}(A^{\mathbb{N}},\sigma) \\ (u,f,b) & \longmapsto & (u\circ\sigma - u) + f + b \end{array}$$

*est un isomorphisme d'espaces de Banach. En particulier, on a la décomposition en somme directe topologique*

$$\operatorname{Wal}(A^{\mathbb{N}},\sigma) = \operatorname{Cob}(A^{\mathbb{N}},\sigma) \oplus \operatorname{Wal}^{\dagger}(A^{\mathbb{N}},\sigma) \oplus \mathbb{R}$$

*et l'opérateur de cobord envoie isomorphiquement $C(A^{\mathbb{N}})/\mathbb{R}$ sur $\operatorname{Cob}(A^{\mathbb{N}},\sigma)$.*

**Corollaire.** *L'espace quotient $\operatorname{Wal}(A^{\mathbb{N}},\sigma)/\operatorname{Cob}(A^{\mathbb{N}},\sigma)$ est isomorphe à $\operatorname{Wal}^{\dagger}(A^{\mathbb{N}},\sigma) \oplus \mathbb{R}$.*

*Preuve du théorème.* L'application est manifestement continue. Démontrons d'abord l'injectivité ; supposons $u\circ\sigma - u + f + b = 0$, avec $u \in C(X)$, $f \in \operatorname{Wal}^{\dagger}(X,T)$ et $b \in \mathbb{R}$. Soit $\mu \in \mathcal{M}$ la proba invariante d'entropie maximale (qui est aussi la mesure de Bernoulli où toutes les lettres de $A$ sont équiprobables). Comme $u\circ\sigma - u + f$ est de moyenne nulle pour $\mu$, on doit avoir $b = 0$. D'autre part, $f$ et $u\circ\sigma$ sont orthogonales dans $L^2(\mu)$, et l'identité $u\circ\sigma + f = u$ implique donc, pour les normes $L^2$,

$$\|u\|^2 = \|u\circ\sigma\|^2 + \|f\|^2 = \|u\|^2 + \|f\|^2$$

ce qui entraîne que $f$ est nulle et $u$ constante. L'application est donc bien injective.

Démontrons maintenant la surjectivité. Soit $f \in \operatorname{Wal}(X,T)$ ; on veut trouver $u \in C(X)$ et $b \in \mathbb{R}$ tels que $\mathcal{E}(f - b + u - u\circ\sigma) = 0$. Cette équation s'écrit, de façon équivalente,

$$\forall y \in X \qquad u(y) = -b + \frac{1}{\sharp A}\sum_{a\in A}(f+u)(ay)$$

Pour $\lambda \in [0,1]$, notons $L_\lambda$ l'opérateur de $C(X)$ dans lui-même défini par

$$\forall y \in X \qquad [L_\lambda u](y) = \frac{1}{\sharp A} \sum_{a\in A} (f+\lambda u)(ay)$$

Cet opérateur est $\lambda$-lipschitz pour la norme uniforme ; notons $u_\lambda$ son unique point fixe pour $\lambda \in [0,1[$. Posons $H_\lambda(x,y) = u_\lambda(x) - u_\lambda(y)$. Alors pour tous $x, y \in X$ on a

$$\begin{aligned} H_\lambda(x,y) &= \frac{1}{\sharp A} \sum_{a\in A} f(ax) - f(ay) + \lambda H_\lambda(ax, ay) \\ &\leqslant \max_{a\in A}\ f(ax) - f(ay) + H_\lambda(ax, ay) \end{aligned}$$

On en déduit

$$\begin{aligned} H_\lambda(x,y) &\leqslant \max_{\theta\in A^{\mathbb{N}}} \sum_{n=0}^{\infty} \lambda^n \big[f(\theta_n \ldots \theta_0 x) - f(\theta_n \ldots \theta_0 y)\big] \\ &= \max_{\theta\in A^{\mathbb{N}}} (1-\lambda) \sum_{m=0}^{\infty} \lambda^m \sum_{n=0}^{m} f(\theta_n \ldots \theta_0 x) - f(\theta_n \ldots \theta_0 y) \\ &\leqslant \max_{\theta\in A^{\mathbb{N}}} \sup_{m\geqslant 0} \sum_{n=0}^{m} f(\theta_n \ldots \theta_0 x) - f(\theta_n \ldots \theta_0 y) \\ &\leqslant h_W\big[d(x,y)/2\big] \end{aligned}$$

où $h_W$ est un module de Walters pour $f$. On a vu au chapitre précédent qu'on pouvait toujours supposer $h_W(1/2) < \infty$. La famille $u_\lambda$ est donc équicontinue et d'oscillation bornée pour $\lambda \to 1$ ; on peut donc en prendre une valeur d'adhérence dans $C(X)/\mathbb{R}$, et on obtient ainsi une fonction $u \in C(X)$ vérifiant $u = -b + L_1 u$, où $b \in \mathbb{R}$ est une certaine constante. Ceci prouve la surjectivité de l'opérateur donné dans l'énoncé du théorème, et termine la preuve.

*Remarque.* Le théorème ci-dessus n'est pas vraiment spécifique au décalage de Bernoulli ; les seuls ingrédients nécessaires sont une dynamique expansive, et un opérateur de transfert $\mathcal{E}$ préservant les constantes (afin que l'opérateur dual $\mathcal{E}'$ préserve une proba invariante). Par exemple, dans le cas du cercle $\mathbb{T}^1$ muni de l'application de doublement de l'angle, on prendrait l'opérateur $\mathcal{E}$ défini par $\mathcal{E}f : y \mapsto \frac{1}{2}\sum_{2x=y} f(x)$ et on aurait un théorème de décomposition parfaitement analogue.

Les énoncés qui vont suivre, au contraire, sont très spécifiques au décalage de Bernoulli, puisqu'ils concernent l'ensemble $\mathcal{LC}(X)$ des fonctions localement constantes sur $X$. Nous allons nous attacher à démontrer que ce sous-espace est dense, et qu'une propriété analogue est vraie dans la limite projective.

Signalons en passant un résultat classique du folklore, que nous n'utiliserons pas, et dont nous laissons la preuve au lecteur.

**Proposition.** *Soit $u \in C(A^{\mathbb{N}})$. Alors $u \in \mathcal{LC}(A^{\mathbb{N}}) \Longleftrightarrow u \circ \sigma - u \in \mathcal{LC}(A^{\mathbb{N}})$.*

Voici une manière simple de construire des approximations localement constantes d'une fonction continue $f \in C(X)$ donnée : soit $\alpha$ une lettre distinguée de l'alphabet $A$, et définissons pour $n \in \mathbb{N}$ l'application $\tau_n : X \to X$ comme suit :

$$\tau_n(x_0\, x_1 \ldots x_{n-1}\, x_n\, x_{n+1} \ldots) = (x_0\, x_1 \ldots x_{n-1}\, \alpha\, \alpha \ldots)$$

Alors les applications $f_n = f \circ \tau_n$ sont localement constantes et convergent vers $f$ pour la topologie de la convergence uniforme. En topologie de Walters, on a un résultat plus faible :

**Théorème.** *On suppose que $f$ vérifie la condition de Walters. Alors les fonctions $f_n = f \circ \tau_n$ convergent au sens de Cesàro vers $f$ en topologie de Walters.*

*Preuve.* Soit $f \in \mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ de module de Walters $h_W$ avec $h_W(1/2) < \infty$. Posons $\theta_n = h_W(2^{-n})$ pour $n \geqslant 1$ ; cette suite tend vers zéro. Nous devons démontrer que

$$\frac{1}{n} \sum_{k=0}^{n-1} f_k \;\longrightarrow\; f \quad \text{(Walters)}$$

Posons $g_n = \sum_{k=0}^{n-1}(f - f_k)$ pour $n \geqslant 1$. La norme de Walters de $g_n$ vaut

$$W_{1/2}(g_n) = 2\,\|g_n\|_\infty + \sup_{m \geqslant 1} \max_{d_m(x,y) \leqslant 1/2} |S_m g_n(x) - S_m g_n(y)|$$

Le premier terme ne pose pas de difficulté : $\|f - f_n\|_\infty = o(1)$ donc $\|g_n\|_\infty = o(n)$. Pour estimer le second terme, donnons-nous $x, y \in A^{\mathbb{N}}$ tels que $d_m(x,y) \leqslant 1/2$ pour un certain $m \geqslant 1$ (en d'autres termes, les $m$ premières lettres de $x$ et $y$ sont identiques). Posons $p = \min(m,n)$. Alors

$$\begin{aligned}
S_m g_n(x) - S_m g_n(y) &= \sum_{\ell=0}^{m-1} \sum_{k=0}^{n-1} (f - f_k)(\sigma^\ell x) - (f - f_k)(\sigma^\ell y) \\
&= \underbrace{\sum_{\ell=0}^{m-p-1} \sum_{k=0}^{n-1} f(\sigma^\ell x) - f(\sigma^\ell y)}_{A} \\
&\quad + \underbrace{\sum_{\ell=m-p}^{m-1} \sum_{k=0}^{n-1} f(\sigma^\ell x) - f(\tau_k \sigma^\ell x) - f(\sigma^\ell y) + f(\tau_k \sigma^\ell y)}_{B}
\end{aligned}$$

La quantité $A = n \sum_{\ell=0}^{m-p-1} f(\sigma^\ell x) - f(\sigma^\ell y)$ est nulle si $m \leqslant p$ (c.à.d. $m \leqslant n$), et sinon on a $|A| \leqslant n\, h_W(2^{-p}) = n\theta_p$. Dans les deux cas, on a $|A| \leqslant n\theta_n$.

Pour estimer $B$, on posera $\tilde{x} = \sigma^{m-p}(x)$, $\tilde{y} = \sigma^{m-p}(y)$, $\tilde{x}_s = \tau_s \tilde{x}$ et $\tilde{y}_s = \tau_s \tilde{y}$. On note que $d_p(\tilde{x}, \tilde{y}) \leqslant 1/2$. Alors

$$\begin{aligned} B &= \sum_{\ell=0}^{p-1} \sum_{k=0}^{n-1} f(\sigma^\ell \tilde{x}) - f(\sigma^\ell \tau_{k+\ell}\, \tilde{x}) - f(\sigma^\ell \tilde{y}) + f(\sigma^\ell \tau_{k+\ell}\, \tilde{y}) \\ &= \sum_{s=0}^{n+p-2} \underbrace{\sum_{\ell=(s-n+1)^+}^{\min(p-1,s)} f(\sigma^\ell \tilde{x}) - f(\sigma^\ell \tilde{x}_s) - f(\sigma^\ell \tilde{y}) + f(\sigma^\ell \tilde{y}_s)}_{B(s)} \end{aligned}$$

Distinguons les cas $s \leqslant p-1$ et $s \geqslant p$. Si $s \leqslant p-1$, alors

$$\begin{aligned} |B(s)| &\leqslant \left| \sum_{\ell=0}^{s} f(\sigma^\ell \tilde{x}) - f(\sigma^\ell \tilde{y}) \right| + \left| \sum_{\ell=0}^{s} f(\sigma^\ell \tilde{x}_s) - f(\sigma^\ell \tilde{y}_s) \right| \\ &\leqslant 2\, h_W(2^{s-p}) = 2\theta_{p-s} \end{aligned}$$

Si au contraire $s \geqslant p$, alors

$$\begin{aligned} |B(s)| &\leqslant \left| \sum_{\ell=(s-n+1)^+}^{p-1} f(\sigma^\ell \tilde{x}) - f(\sigma^\ell \tilde{x}_s) \right| + \left| \sum_{\ell=(s-n+1)^+}^{p-1} f(\sigma^\ell \tilde{y}) - f(\sigma^\ell \tilde{y}_s) \right| \\ &\leqslant 2\, h_W(2^{p-1-s}) = 2\theta_{s-p+1} \end{aligned}$$

On en déduit que

$$\begin{aligned} |B| &\leqslant \sum_{s=0}^{p-1} |B(s)| + \sum_{s=p}^{n+p-2} |B(s)| \leqslant \sum_{s=0}^{p-1} 2\theta_{p-s} + \sum_{s=p}^{n+p-2} 2\theta_{s-p+1} \\ &\leqslant \sum_{s=1}^{p} 2\theta_s + \sum_{s=1}^{n-1} 2\theta_s \leqslant 4 \sum_{s=1}^{n} \theta_s \end{aligned}$$

ce qui nous donne la majoration

$$\sup_{m \geqslant 1} \max_{d_m(x,y) \leqslant 1/2} |S_m g_n(x) - S_m g_n(y)| \;\leqslant\; n\theta_n + 4 \sum_{s=1}^{n} \theta_s$$

et finalement

$$W_{1/2}(g_n) \;\leqslant\; 2\, \|g_n\|_\infty + n\theta_n + 4 \sum_{s=1}^{n} \theta_s$$

Le membre de droite est un $o(n)$ quand $n \to \infty$, donc les $f_n$ convergent effectivement vers $f$ au sens de Cesàro dans la topologie de Walters, et le théorème est démontré.

*Remarque.* On notera l'analogie entre le théorème ci-dessus et le théorème de Fejér sur les séries de Fourier. Si on se donne une fonction $f \in C(\mathbb{T}^1)$ continue sur le cercle, et qu'on note $f_n = f * D_n$ les sommes partielles de la série de Fourier, alors les $f_n$ tendent vers $f$ en norme $L^2$, mais pour la norme uniforme on a seulement convergence au sens de Cesàro. En fait, le théorème de Banach-Steinhaus (combiné au caractère non borné des noyaux de Dirichlet $D_n$) permet de voir que pour $f$ générique dans $C(X)$, la suite $f_n$ sera non bornée et donc *a fortiori* ne tendra pas vers $f$ en norme uniforme (voir [Rud] problème 5.11).

Par des arguments tout à fait similaires, on peut démontrer que quand $f$ est Walters, les $f \circ \tau_n$ ne convergent pas vers $f$ en général (en topologie de Walters). En effet, soit $\Xi_n$ l'opérateur de $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ dans lui-même défini par $\Xi_n f = f \circ \tau_n$. En regardant l'action de $\Xi_n$ sur les cobords, il est facile de voir que cette suite d'opérateurs n'est pas uniformément bornée. Ceci entraîne, en vertu du théorème de Banach-Steinhaus, que pour $f$ générique dans $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$, la suite $f \circ \tau_n$ sera non bornée en norme de Walters, et donc ne convergera pas vers $f$. En ce sens, le théorème que nous venons de démontrer (convergence au sens de Cesàro) n'est pas vraiment améliorable.

**Corollaire.** *$\mathcal{LC}(A^{\mathbb{N}})$ et $\mathcal{LC}(A^{\mathbb{Z}})$ sont denses dans $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ et $\mathrm{Wal}(A^{\mathbb{Z}}, \sigma)$ respectivement. En particulier, $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ et $\mathrm{Wal}(A^{\mathbb{Z}}, \sigma)$ sont séparables.*

*Preuve du corollaire.* Le théorème précédent implique évidemment la densité de $\mathcal{LC}(A^{\mathbb{N}})$ dans $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$, et d'autre part, la décomposition

$$\mathrm{Wal}(A^{\mathbb{Z}}, \sigma) = \mathrm{Wal}(A^{\mathbb{N}}, \sigma) + \mathrm{Cob}(A^{\mathbb{Z}}, \sigma)$$

(démontrée au chapitre 3) entraîne la densité de $\mathcal{LC}(A^{\mathbb{Z}})$ dans $\mathrm{Wal}(A^{\mathbb{Z}}, \sigma)$. Comme $\mathcal{LC}(A^{\mathbb{N}})$ et $\mathcal{LC}(A^{\mathbb{Z}})$ sont de dimension dénombrable, il en découle la séparabilité de $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ et $\mathrm{Wal}(A^{\mathbb{Z}}, \sigma)$.

## 7. Un théorème de Walters

Dans le fameux article [Wal] où est définie la condition qui porte maintenant son nom, Walters démontre notamment le résultat suivant, que nous énoncerons dans le cas du décalage de Bernoulli :

**Théorème (Walters).** *Soit $X = A^{\mathbb{N}}$, où $A$ est un alphabet fini, muni de l'application de décalage $\sigma$, et $f \in \mathrm{Wal}(X,\sigma)$. Alors il existe une fonction $h$ cohomologue à $f$ et une constante $P$ telles que*

$$\forall y \in X \qquad \sum_{\sigma x = y} e^{h(x)} \;=\; e^P$$

La constante $P$ qui apparaît ici n'est autre que la pression topologique de $f$ (ou de $h$). La démonstration classique, due à Walters, consiste à démontrer que l'opérateur de transfert de Ruelle $\mathcal{L} : C(X) \to C(X)$, défini par

$$\forall y \in X \qquad [\mathcal{L}U](y) = \sum_{\sigma x = y} e^{f(x)} U(x)$$

admet une fonction propre continue, à valeurs strictement positives. En effet, si $U = e^u$ est une telle fonction propre, de valeur propre $e^P$, alors on a

$$\forall y \in X \qquad e^P \;=\; \sum_{\sigma x = y} \exp\,(f + u - u \circ \sigma)(x)$$

et donc $h = f + u - u \circ \sigma$ répond bien au problème posé.

Walters prouve l'existence d'une direction propre en construisant dans $C(X)$ un cône convexe invariant par $\mathcal{L}$, dont l'ensemble des directions est compact. Dans [KMS], Kondah et al. utilisent une autre approche, basée sur une suite de cônes emboîtés, et le caractère contractant de $\mathcal{L}$ pour leurs métriques de Birkhoff.

Ici je propose une troisième démonstration de ce théorème, dans l'esprit des autres démonstrations du présent article, et dont la principale originalité est de *ne pas* reposer sur l'itération (même implicite) de l'opérateur de transfert : la construction de la direction propre de $\mathcal{L}$ se fait grâce à des approximants non linéaires de $\mathcal{L}$.

*Preuve.* Pour $\lambda \in [0,1]$, définissons l'opérateur non linéaire $T_\lambda : C(X) \to C(X)$ comme suit :

$$\forall y \in X \qquad \exp[T_\lambda u](y) = \sum_{\sigma x = y} \exp(f + \lambda u)(y) = \sum_{a \in A} \exp(f + \lambda u)(ay)$$

Cet opérateur est clairement $\lambda$-lipschitz pour la norme uniforme ; soit $u_\lambda$ son point fixe, pour $\lambda \in [0,1[$. Posons $H_\lambda(x,y) = u_\lambda(x) - u_\lambda(y)$. Alors, pour tous $x, y \in X$ on a

$$\begin{aligned}
\exp u_\lambda(x) &= \sum_{a \in A} \exp(f + \lambda u_\lambda)(ax) \\
&= \sum_{a \in A} \exp(f + \lambda u_\lambda)(ay)\ \exp\Big[f(ax) - f(ay) + \lambda\big[u_\lambda(ax) - u_\lambda(ay)\big]\Big] \\
&\leqslant \exp u_\lambda(y) \cdot \max_{a \in A}\ \exp\Big[f(ax) - f(ay) + \lambda\big[u_\lambda(ax) - u_\lambda(ay)\big]\Big]
\end{aligned}$$

et donc

$$H_\lambda(x,y) \leqslant \max_{a\in A} f(ax) - f(ay) + \lambda H_\lambda(ax,ay)$$

On en déduit, comme précédemment

$$\begin{aligned} H_\lambda(x,y) &\leqslant \max_{\theta\in A^{\mathbb{N}}} \sum_{n=0}^{\infty} \lambda^n \left[ f(\theta_n \dots \theta_0 x) - f(\theta_n \dots \theta_0 y) \right] \\ &\leqslant h_W \left[ d(x,y)/2 \right] \end{aligned}$$

Ici encore, les fonctions $u_\lambda$ sont équicontinues et d'oscillation bornée quand $\lambda \to 1$. En prenant une valeur d'adhérence dans $C(X)/\mathbb{R}$, on obtient une fonction $u \in C(X)$ telle que $u = -P + T_1 u$ pour une certaine constante $P \in \mathbb{R}$. Alors la fonction $h = f + u - u \circ \sigma$ est une solution au problème posé, et le théorème de Walters est démontré.

*Remarque.* La très grande similarité entre cette démonstration du théorème de Walters, le lemme de Mañé donné au chapitre 2, et le théorème de décomposition de $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ donné au chapitre 6 n'est pas fortuite ; ces deux derniers énoncés peuvent être vus comme des formes limites du théorème de Walters, aux basses températures ($f \gg 1$) et aux hautes températures ($f \ll 1$) respectivement.

## 8. Verrouillage des mesures maximisantes sur les orbites périodiques

De nombreux auteurs ont observé indépendamment sur des exemples (voir [Je2], [HO] et [Bou]) que dans le cas d'une dynamique expansive ou hyperbolique, les mesures maximisantes avaient une forte tendance à se "caler" sur des orbites périodiques de petites périodes.

Ces observations soulèvent un certain nombre de questions distinctes. Premièrement, la question de l'unicité de la mesure maximisante ; il semblerait que pour une fonction générique, il n'existe qu'une seule mesure maximisante, pourquoi ?

La seconde question est celle du verrouillage proprement dit : les observations suggèrent que quand $\mu$ est une orbite périodique, l'ensemble des $f$ pour laquelle $\mu$ est maximisante est un domaine d'intérieur non vide (intentionellement, je ne précise pas la topologie pour le moment). Au contraire, quand $\mu$ n'est pas une orbite périodique, ce domaine est d'intérieur vide, pourquoi ?

La troisième question est de savoir si les intérieurs des "domaines de verrouillage" ci-dessus sont denses dans l'espace fonctionnel considéré ; nous appellerons cette propriété la propriété de "verrouillage générique".

Ces questions ont été abordées dans [CLT] et [YH] dans le cas d'une application expansive sur le cercle, et d'une application Axiome A sur une variété compacte, respectivement. Rappelons brièvement leurs principaux résultats. Contreras et al. considèrent l'espace $C^{!\alpha}(X)$ des fonctions continues $f$ vérifiant

$$\forall \epsilon > 0 \quad \exists \eta > 0 \qquad d(x,y) \leqslant \eta \implies |f(y) - f(x)| \leqslant \epsilon\, d(x,y)^\alpha$$

muni de la topologie $C^\alpha$ (ainsi $C^{!\alpha}(X)$ est un sous-espace fermé de codimension infinie de $C^\alpha(X)$) alors que Yuan et Hunt se placent dans l'espace fonctionnel $C^\alpha(X)$, plus commun[3]. Dans les deux articles, le verrouillage sur les orbites périodiques est démontré “à la main” : pour chaque orbite périodique $\mu$, on construit une fonction $f$ atteignant son maximum précisément sur l'orbite, et on démontre que toute fonction assez voisine de $f$ admet encore $\mu$ comme mesure maximisante. C'est ensuite que les deux articles divergent. Au prix d'une démonstration particulièrement difficile, et qui ne semble pas simplifiable, Yuan et Hunt arrivent à démontrer qu'il ne peut pas y avoir verrouillage sur une orbite non périodique dans $C^\alpha(X)$, et ils n'obtiennent pas la propriété de verrouillage générique. De leur côté, Contreras et al. obtiennent sans difficulté aucune la propriété de verrouillage générique dans $C^{!\alpha}(X)$, et comme sous-produit l'impossibilité du verrouillage sur une orbite non périodique.

C'est ainsi qu'on se rend compte que $C^{!\alpha}(X)$ et $C^\alpha(X)$ sont en fait des espaces extrêmement différents et que, à tous points de vue, l'espace $C^{!\alpha}(X)$ est celui qui a les meilleures propriétés, en particulier du point de vue de l'approximation. Il est séparable, ses éléments sont approximables par convolution, et les opérations du treillis $f \vee g = \max(f,g)$ et $f \wedge g = \min(f,g)$ sont continues dans $C^{!\alpha}(X)$. Toutes ces propriétés sont fausses dans $C^\alpha(X)$. Le mauvais comportement de $C^\alpha(X)$ du point de vue de l'approximation est certainement ce qui explique l'extrême difficulté du problème étudié par Yuan et Hunt.

Ici nous étudierons ces questions dans l'espace $\mathrm{Wal}(X,T)$, dans le cas particulier où $(X,T)$ est un décalage de Bernoulli. Nous montrerons comment le verrouillage sur les orbites périodiques, et plus généralement sur les sous-décalages de type fini, est une conséquence de la condition de Walters. Puis nous ferons apparaître le verrouillage générique comme conséquence d'un théorème d'approximation dans $\mathrm{Wal}(X,T)$.

Pour $f$ continue (ou davantage) on notera $\mathcal{M}_{\max}(f)$ l'ensemble des probas invariantes qui maximisent l'intégrale de $f$.

La première observation est que le phénomène de verrouillage n'existe pas dans $C(X)$ : une mesure maximisante, périodique ou non, peut être “cassée” par des perturbations arbitrairement petites en norme uniforme :

**Lemme.** *Pour $\mu \in \mathcal{M}$ quelconque, l'ensemble*

$$\{f \in C(A^{\mathbb{N}}) : \mu \text{ est } f\text{-maximisante}\}$$

*est un fermé d'intérieur vide dans $C(A^{\mathbb{N}})$.*

*Preuve.* Soit $\mu \in \mathcal{M}$ et $f \in C(X)$ telle que $\mu$ soit $f$-maximisante. Il existe une suite $(\mu_n)_{n\geqslant 1}$ de probas invariantes, tendant vaguement vers $\mu$ et étrangères à $\mu$, si bien que $|\mu - \mu_n|(X) = 2$. On peut donc trouver une suite $g_n$ de fonctions continues telles que

[3] Je suggère de baptiser *petit espace de Hölder* et *grand espace de Hölder* d'exposant $\alpha$ les espaces $C^{!\alpha}(X)$ et $C^\alpha(X)$ respectivement.

$\|g_n\|_\infty \leqslant 1$ et $\int g_n(\mu_n - \mu) \geqslant 1$. Posons maintenant $\alpha_n = 1/n + \left(\int f(\mu - \mu_n)\right)^+$, et $f_n = f + \alpha_n g_n$. Cette suite tend vers $f$, mais

$$\begin{aligned}\int f_n \mu_n - \int f_n \mu &= \int f(\mu_n - \mu) + \alpha_n \int g_n(\mu_n - \mu) \\ &\geqslant \int f(\mu_n - \mu) + \alpha_n \\ &\geqslant 1/n\end{aligned}$$

donc $\mu$ n'est pas maximisante pour $f_n$ , ce qui prouve le lemme.

**Proposition.** *Pour $f$ générique dans $C(A^{\mathbb{N}})$, il existe une unique proba invariante maximisante, et ce n'est pas une orbite périodique.*

*Preuve.* Nous allons d'abord démontrer que pour $f$ générique dans $C(X)$, la mesure maximisante est unique. La démonstration donnée ci-dessous n'est guère spécifique à $C(X)$ et se généralise facilement à d'autres espaces fonctionnels.

Soit $(e_i)_{i\in\mathbb{N}}$ une famille dénombrable d'éléments de $C(X)$, telle que l'espace vectoriel engendré soit dense dans $C(X)$. Alors deux probas $\mu, \nu$ seront égales si et seulement si $\int e_i \mu = \int e_i \nu$ pour tout $i \in \mathbb{N}$, et donc

$$\begin{aligned}\Omega &= \left\{f \in C(X) : \mathcal{M}_{\max}(f) \text{ n'est pas un singleton}\right\} \\ &= \left\{f \in C(X) : \exists \mu, \nu \in \mathcal{M}_{\max}(f)\ \exists n \in \mathbb{N}\ \textstyle\int e_n(\mu - \nu) > 0\right\} \\ &= \bigcup_{m=1}^{\infty} \bigcup_{n=0}^{\infty} \left\{f \in C(X) : \exists \mu, \nu \in \mathcal{M}_{\max}(f)\ \textstyle\int e_n(\mu - \nu) \geqslant 1/m\right\}\end{aligned}$$

Ceci montre que $\Omega$ est un $F_\sigma$ et il reste à prouver que les

$$\Omega_{m,n} = \{f \in C(X) : \exists \mu, \nu \in \mathcal{M}_{\max}(f)\ \textstyle\int e_n(\mu - \nu) \geqslant 1/m\}$$

sont d'intérieur vide. Pour cela, il suffit de remarquer que si $f \in \Omega_{m,n}$ , alors on aura $f + re_n \notin \Omega_{m,n}$ pour $r > 0$ suffisamment petit (c'est évident si on fait un dessin). On a donc bien, génériquement, unicité de $\mathcal{M}_{\max}(f)$.

D'autre part, pour toute orbite périodique $\mu$, l'ensemble des $f \in C(X)$ tels que $\mu \in \mathcal{M}_{\max}(f)$ est un fermé d'intérieur vide, et l'ensemble des orbites périodiques est de cardinal dénombrable, donc l'ensemble des $f \in C(X)$ admettant une mesure maximisante périodique constitue un $F_\sigma$ d'intérieur vide. Ceci termine la preuve de la proposition.

*Remarque.* La proposition ci-dessus ne permet pas d'apprécier pleinement toute la pathologie de la situation générique dans $C(A^{\mathbb{N}})$. En fait, par des arguments analogues à ceux donnés dans le lemme, on peut démontrer que l'ensemble des $f \in C(A^{\mathbb{N}})$ admettant une proba maximisante dont le support évite un ouvert donné est d'intérieur vide ; en d'autres termes, la proba maximisante est génériquement de support plein dans $A^{\mathbb{N}}$ ! Inutile de dire que dans ces conditions, le “principe de subordination” ne s'applique plus.

Avant d'entrer dans l'étude du comportement générique dans $\mathrm{Wal}(X, \sigma)$, donnons un principe général qui permet de simplifier légèrement les énoncés. Nous le donnons ici dans le cadre de $\mathrm{Wal}(X, \sigma)$ mais le lecteur vérifiera qu'il s'applique aussi bien dans $C(X)$.

**Lemme.** *Soit $\mathcal{M}_1$ une partie convexe et fermée de $\mathcal{M}$ (pour la topologie vague). Alors les ensembles*

$$\big\{f \in \mathrm{Wal}(A^{\mathbb{N}},\sigma) : \mathcal{M}_{\max}(f) \cap \mathcal{M}_1 \neq \varnothing\big\} \quad \textit{et}$$
$$\big\{f \in \mathrm{Wal}(A^{\mathbb{N}},\sigma) : \mathcal{M}_{\max}(f) \subset \mathcal{M}_1\big\}$$

*ont même intérieur dans* $\mathrm{Wal}(A^{\mathbb{N}},\sigma)$.

*Preuve.* Notons $U$ et $V$ les deux ensembles ci-dessus. On a $U \supset V$ donc $U^\circ \supset V^\circ$. Inversement, soit $f \in U^\circ$, et supposons qu'il existe $\mu_0 \in \mathcal{M}_{\max}(f)$ tel que $\mu_0 \notin \mathcal{M}_1$. D'après Hahn-Banach (voir par exemple [AB] corollaire 5.59), il existe $h_0 \in C(X)$ telle que $\int h_0\mu_0 \geqslant 2$ et $\int h_0\mu \leqslant 0$ pour tout $\mu \in \mathcal{M}_1$. Soit $h \in \mathrm{Wal}(X,\sigma)$ telle que $\|h - h_0\|_\infty \leqslant 1/2$, alors

$$\forall \mu \in \mathcal{M}_1 \qquad \int h\mu_0 \geqslant 1 + \int h\mu$$

Par conséquent, pour tout $r > 0$ on a

$$\forall \mu \in \mathcal{M}_1 \qquad \int (f + rh)\,\mu_0 \;\geqslant\; r + \int (f + rh)\,\mu$$

donc $\mathcal{M}_{\max}(f+rh)\cap\mathcal{M}_1 = \varnothing$ pour des $r$ arbitrairement petits, ce qui contredit l'hypothèse que $f$ est dans l'intérieur de $U$. Nous avons ainsi prouvé que $U^\circ \subset V$ et donc $U^\circ \subset V^\circ$, et le lemme est démontré.

**Proposition.** *Soit $f \in \mathrm{Wal}(A^{\mathbb{N}},\sigma)$ localement constante. Alors $\mathcal{M}_{\max}(f)$ est l'ensemble des probas invariantes portées par un certain sous-décalage de type fini. En outre, pour tout $g$ suffisamment proche de $f$ en topologie de Walters, on aura $\mathcal{M}_{\max}(g) \subset \mathcal{M}_{\max}(f)$.*

**Corollaire.** *Soit $K$ un compact invariant non vide ; $K$ est supposé localement maximal dans $A^{\mathbb{N}}$, c'est-à-dire admettant un voisinage $U$ tel que $K = \bigcap_{n\geqslant 0}\sigma^{-n}U$. On note $\mathcal{M}_K$ l'ensemble des $\mu \in \mathcal{M}$ tels que* $\mathrm{supp}\,\mu \subset K$. *Alors l'ensemble*

$$\big\{f \in \mathrm{Wal}(A^{\mathbb{N}},\sigma) : \mathcal{M}_{\max}(f) \subset \mathcal{M}_K\big\}$$

*est d'intérieur non vide dans* $\mathrm{Wal}(A^{\mathbb{N}},\sigma)$.

*Preuve de la proposition.* Commençons par la deuxième partie de l'énoncé ("en outre…"). Considérons une suite $f_n \in \mathrm{Wal}(X,\sigma)$ tendant vers $f$ en topologie de Walters. Il est facile de voir que les $f_n$ admettent un module de Walters commun $h_W$, avec $h_W(1/2) < \infty$ (ceci traduit la précompacité de l'ensemble des $f_n$ en topologie de Walters). D'après le lemme de Mañé, il existe des fonctions $u_n \in C(X)$ telles que

$$\forall y \in X \qquad u_n(y) = -\beta_n + \max_{\sigma x = y}(f_n + u_n)(x)$$

avec $\beta_n \to \beta$. De cette équation fonctionnelle on déduit (exercice)

$$\forall x, y \in X \qquad |u_n(x) - u_n(y)| \leqslant h_W\big[d(x,y)/2\big]$$

Les $u_n$ sont donc équicontinues et d'oscillation bornée. Par conséquent, il existe une suite d'indices $\kappa(n)$ telle que la suite $u_{\kappa(n)}$ converge en norme uniforme (modulo les constantes) vers une certaine fonction continue $u$. Elle vérifie

$$\forall y \in X \qquad u(y) = -\beta + \max_{\sigma x = y}(f+u)(x)$$

et comme $f$ est localement constante, ceci implique (par les mêmes arguments que plus haut) que $u$ est localement constante. Posons alors $h = f + u - u \circ \sigma$, et $V = h^{-1}(\beta)$. L'ensemble $V$ est ouvert et fermé, et

$$\mathcal{M}_{\max}(f) = \{\mu \in \mathcal{M} : \operatorname{supp} \mu \subset V\}$$

De même, posant $h_n = f_n + u_n - u_n \circ \sigma$ et $V_n = h_n^{-1}(\beta_n)$, on obtient

$$\mathcal{M}_{\max}(f_n) = \{\mu \in \mathcal{M} : \operatorname{supp} \mu \subset V_n\}$$

La convergence uniforme $h_{\kappa(n)} \to h$ implique que tout voisinage de $V$ contient les $V_{\kappa(n)}$ pour $n$ assez grand. Mais $V$ est un voisinage de lui-même, donc $V_{\kappa(n)} \subset V$ pour $n$ assez grand, et

$$\mathcal{M}_{\max}(f_n) \subset \mathcal{M}_{\max}(f)$$

pour une infinité de $n$. Ceci est vrai quelle que soit la suite $(f_n)$ convergeant vers $f$, donc l'ensemble

$$\big\{g \in \mathrm{Wal}(X,\sigma) : \mathcal{M}_{\max}(g) \subset \mathcal{M}_{\max}(f)\big\}$$

doit contenir $f$ comme point intérieur. En passant, nous avons également prouvé que $\mathcal{M}_{\max}(f)$ était l'ensemble des probas invariantes contenues dans $K = \bigcap_{n \geqslant 0} \sigma^{-n} V$ où $V$ est un certain ouvert fermé, c.à.d. un sous-décalage de type fini (ce résultat est classique, voir par exemple [Zie]). La proposition est démontrée.

*Preuve du corollaire.* Par hypothèse, $K$ admet un voisinage $U$ tel que $K = \bigcap_{n \geqslant 0} \sigma^{-n} U$. Quitte à restreindre $U$, on peut le supposer ouvert et fermé. Alors sa fonction caractéristique $\chi(U)$ est localement constante et vérifie $\mathcal{M}_{\max}\big[\chi(U)\big] = \mathcal{M}_K$, et on est ramené au problème précédent.

*Remarque.* Parmi les ensembles invariants localement maximaux, il y a en particulier les orbites périodiques. Si $K$ est une orbite périodique (au sens ensembliste), alors $\mathcal{M}_K$ est un singleton $\{\nu\}$, et on peut affirmer que

$$\big\{f \in \mathrm{Wal}(A^{\mathbb{N}},\sigma) : \mathcal{M}_{\max}(f) = \{\nu\}\big\}$$

est d'intérieur non vide dans $\mathrm{Wal}(A^{\mathbb{N}},\sigma)$. Il y a donc bien “verrouillage” sur l'orbite périodique $\nu$ en topologie de Walters.

**Théorème.** *Soit $P$ l'ensemble des orbites périodiques sur $A^{\mathbb{N}}$, considéré comme partie de $\mathcal{M}$. Alors il existe une famille $(U_\nu)_{\nu \in P}$ d'ouverts non vides de $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$, convexes, deux à deux disjoints, dont l'union est dense, et tels que*

$$\forall \nu \in P \quad \forall f \in U_\nu \qquad \mathcal{M}_{\max}(f) = \{\nu\}.$$

*En conséquence, pour $f$ générique dans $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$, il existe une unique proba invariante maximisante, et c'est une orbite périodique.*

*Preuve.* Pour tout $\nu \in P$, définissons

$$V_\nu = \left\{ f \in \mathrm{Wal}(A^{\mathbb{N}}, \sigma) : \nu \in \mathcal{M}_{\max}(f) \right\}$$
$$W_\nu = \left\{ f \in \mathrm{Wal}(A^{\mathbb{N}}, \sigma) : \mathcal{M}_{\max}(f) = \{\nu\} \right\}$$

On a vu plus haut que $V_\nu$ et $W_\nu$ ont même intérieur, que nous noterons $U_\nu$. La proposition précédente affirme que $U_\nu$ est non vide. L'ensemble $V_\nu$ est un fermé convexe, donc son intérieur $U_\nu$ est convexe, et comme $V_\nu$ est d'intérieur non vide on a $V_\nu = \overline{U_\nu}$. Les $U_\nu$ sont disjoints, puisqu'ils sont contenus dans les $W_\nu$ qui sont disjoints.

Pour montrer que les $U_\nu$ sont d'union dense, notons que

$$\overline{\bigcup_{\nu \in P} U_\nu} \supset \bigcup_{\nu \in P} \overline{U_\nu} = \bigcup_{\nu \in P} V_\nu$$

et il suffit donc de montrer que les $V_\nu$ sont d'union dense.

Mais nous avons démontré plus haut que pour $f$ localement constante, $\mathcal{M}_{\max}(f)$ est l'ensemble des probas invariantes portées par un certain sous-décalage de type fini, et donc contient au moins une orbite périodique. Ainsi

$$\bigcup_{\nu \in P} V_\nu \supset \mathcal{LC}(A^{\mathbb{N}})$$

Enfin, on a démontré au chapitre 6 que $\mathcal{LC}(A^{\mathbb{N}})$ était dense dans $\mathrm{Wal}(A^{\mathbb{N}}, \sigma)$ ; les $V_\nu$ sont donc bien d'union dense, et ceci termine la preuve du théorème.

**Annexe A. Démonstration des propositions A et B**

On pourrait prouver ces deux propositions séparément, mais il est plus intéressant de faire apparaître la Proposition A comme un corollaire de la Proposition B.

En effet, soient $\pi : X \to Y$, $\Delta$ et $u$ comme dans la Proposition A, et soit $\tilde{u} : X^2 \to \mathbb{R}\cup\{-\infty\}$ la fonction définie par

$$\forall x, y \in X \qquad \tilde{u}(x,y) = \begin{cases} u(x,y) & \text{si } \pi(x) = \pi(y), \\ -\infty & \text{sinon.} \end{cases}$$

Alors $\tilde{u}$ vérifie les hypothèses de la Proposition B. Si celle-ci est vraie, alors on peut trouver $f \in C(X)$ telle que $u(x,y) \leqslant f(x) - f(y)$ pour tous $x, y$ dans la même fibre, et $\|f\|_\infty = \frac{1}{2} \max u$. En permutant les rôles de $x$ et $y$, et remarquant que $u(y,x) = -u(x,y)$, on voit que ceci entraîne $u(x,y) = f(x) - f(y)$ pour tous $x, y$ dans la même fibre, ce qui prouve la Proposition A.

La démonstration suivante de la Proposition B est librement inspirée d'une idée d'Adrien Douady, elle-même librement inspirée des méthodes de Ernest Michael pour construire des “selections continues”. Elle repose sur le lemme suivant.

**Lemme.** *Soient $X$ et $u$ comme dans la Proposition B, et soit $M \geqslant 0$ une constante majorant $u$. Alors on peut trouver $f \in C(X)$ tel que $\|f\|_\infty \leqslant M/12$ et $\max u^* \leqslant 5M/6$, où la fonction $u^*$ est définie par $u^*(x,y) = u(x,y) - f(x) + f(y)$.*

Regardons d'abord pourquoi ce lemme implique la Proposition B. Soit $u_0 : X^2 \to \mathbb{R}\cup\{-\infty\}$ vérifiant les hypothèses de la Proposition B, et $M = (\max u_0)^+$. D'après le lemme, on peut construire par récurrence $u_1, u_2 \ldots$ fonctions s.c.s. de $X^2$ dans $\mathbb{R} \cup \{-\infty\}$, et $f_0, f_1, f_2 \ldots$ fonctions continues de $X$ dans $\mathbb{R}$ telles que

$$\forall n \in \mathbb{N} \qquad \begin{aligned} & u_{n+1}(x,y) = u_n(x,y) - f_n(x) + f_n(y) \quad \text{pour tous } x, y \\ & \max u_n \leqslant (5/6)^n \cdot M \\ & \|f_n\| \leqslant (5/6)^n \cdot M/12 \end{aligned}$$

La suite $F_n = \sum_{k=0}^{n-1} f_k$ converge donc uniformément vers un certain $F \in C(X)$. En faisant tendre $n \to \infty$ dans l'inégalité

$$\forall n \in \mathbb{N} \quad \forall x, y \in X \qquad u_0(x,y) - F_n(x) + F_n(y) = u_n(x,y) \leqslant (5/6)^n \, M$$

il vient

$$\forall x, y \in X \qquad u_0(x,y) - F(x) + F(y) \leqslant 0.$$

Ceci entraîne en particulier $\max u_0 \leqslant 2 \, \|F\|_\infty$. D'autre part, on a

$$\|F\|_\infty \leqslant \sum_{n=0}^{\infty} \|f_n\|_\infty \leqslant \sum_{n=0}^{\infty} (5/6)^n \cdot M/12 = M/2$$

et donc $\|F\|_\infty = M/2$, et la Proposition B est prouvée.

*Preuve du lemme.* Posons $J = \max u \leqslant M$ (le maximum existe puisque $u$ est s.c.s.). Le résultat est évident pour $J \leqslant 0$ (il suffit de prendre $f = 0$) et nous supposerons donc que $J > 0$. Définissons alors $\Lambda = \{(x, y) \in X^2 : u(x, y) \geqslant 2J/3\}$, et soient $\Lambda_1 = \pi_1\Lambda$ et $\Lambda_2 = \pi_2\Lambda$ ses projections sur $X$. La semi-continuité supérieure de $u$ garantit que $\Lambda$ est fermé dans $X^2$, et donc $\Lambda_1$ et $\Lambda_2$ sont fermés dans $X$.

J'affirme que $\Lambda_1$ et $\Lambda_2$ sont disjoints. En effet, supposons que $y \in \Lambda_2 \cap \Lambda_1$ ; alors il existerait $x, z \in X$ tels que $u(x, y) \geqslant 2J/3$ et $u(y, z) \geqslant 2J/3$, mais ceci entraînerait $u(x, z) \geqslant u(x, y) + u(y, z) \geqslant 4J/3 > J$, en contradiction avec la définition de $J$.

Puisqu'ils sont disjoints, on peut trouver $f \in C(X)$ telle que $f_{|\Lambda_1} = J/12$ et $f_{|\Lambda_2} = -J/12$, et de norme $\|f\| = J/12$.

Posons $u^*(x, y) = u(x, y) - f(x) + f(y)$. Si $(x, y) \in \Lambda$, alors $u^*(x, y) = u(x, y) - J/6 \leqslant 5J/6$. Si au contraire $(x, y) \notin \Lambda$, alors $u(x, y) < 2J/3$ et par conséquent $u^*(x, y) < 2J/3 + 2\|f\| = 5J/6$. On a donc bien $\max u^* \leqslant 5J/6$ et le lemme est prouvé.

**Remerciements.** L'auteur souhaite remercier O. M. Jenkinson pour ses commentaires et suggestions sur la première version de cet article, ainsi que pour les nombreuses discussions que nous avons eues sur ces problèmes de mesures maximisantes.

**Bibliographie**


[AB] C. D. Aliprantis et K. C. Border, *Infinite dimensional analysis: a hitchhiker's guide*, deuxième édition, Springer (1999)

[Bou] T. Bousch, *Le poisson n'a pas d'arêtes*, preprint, Université d'Orsay (1999), à paraître dans Ann. IHP Proba. Stat.

[Coe] Z. Coelho, *Entropy and ergodicity of skew-products over subshifts of finite type and central limit asymptotics*, thèse, Université de Warwick (1990)

[CQ] Z. Coelho et A. N. Quas, *Criteria for $\bar{d}$-continuity*, Trans. Amer. Math. Soc. **350** (1998)

[CLT] G. Contreras, A. Lopes et P. Thieullen, *Lyapunov minimizing measures for expanding maps of the circle*, manuscrit (1999)

[Fat] A. Fathi, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Math. **324** (1997)

[HO] B. R. Hunt et E. Ott, *Optimal periodic orbits of chaotic systems occur at low period*, Phys. Rev. E **54** (1996) p. 328

[Je1] O. M. Jenkinson, *Conjugacy rigidity, cohomological triviality, and barycentres of invariant measures*, thèse, Université de Warwick (1996)

[Je2] O. M. Jenkinson, *Frequency-locking on the boundary of the barycentre set*, preprint 97-29, Institut de Mathématiques de Luminy (1997), à paraître dans Experimental Mathematics

[KMS] A. Kondah, V. Maume et B. Schmitt, *Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non höldériennes*, Ann. IHP Proba. Stat. **33** (1997)

[Liv] A. N. Livšic, *Homology properties of Y-systems*, Math. Zametki **10** (1971), traduction anglaise dans Math. Notes **10** (1971)

[Man] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996)

[Rud] W. Rudin, *Real and complex analysis*, troisième édition, McGraw-Hill (1987)

[Sin] Y. G. Sinaĭ, *Gibbs measures in ergodic theory*, Uspekhi Math. Nauk **27** (1972), traduction anglaise dans Russian Math. Surveys **27** (1972)

[Wal] P. Walters, *Invariant measures and equilibrium states for some mappings which expand distances*, Trans. Amer. Math. Soc. **236** (1978)

[YH] G. Yuan et B. R. Hunt, *Optimal orbits of hyperbolic systems*, Nonlinearity **12** (1999)

[Zie] K. Ziemian, *Rotation sets for subshifts of finite type*, Fundamenta Mathematicae **146** (1995)