REGULARITY OF COBOUNDARIES FOR NONUNIFORMLY EXPANDING MARKOV MAPS

SÉBASTIEN GOUÉZEL

(Communicated by Michael Handel)

Abstract. We prove that solutions $u$ of the equation $f = u - u \circ T$ are automatically Hölder continuous when $f$ is Hölder continuous and $T$ is nonuniformly expanding and Markov. This result applies in particular to Young towers and to intermittent maps.

1. Results

Let $(X, m)$ be a probability space and let $T : X \to X$ be an ergodic measure-preserving transformation. Also let $G$ be a locally compact abelian group, endowed with an invariant metric that we denote by $|x - y|$. It is often important to know whether a function $f : X \to G$ is a measurable coboundary, i.e., there exists a measurable function $u : X \to G$ such that

$$f = u - u \circ T$$

almost everywhere. For $G = \mathbb{R}$, this condition is indeed often the only obstruction to have a nondegenerate central limit theorem for the Birkhoff sums of $f$ (see e.g. [Leo60], [GH88], [Liv96]). For $G = S^1$, it is relevant to prove local limit theorems (see [AD01] and [ADSZ04] when $f$ is locally constant, in the Markov and non-Markov case).

When $T$ is uniformly hyperbolic and $f$ is Hölder continuous, the Livšic regularity theorem ([Liv72]) states that $u$ must have a Hölder continuous version, for which holds everywhere. In particular, if there exists a point $x$ such that $T^n(x) = x$ and $\sum_{k=0}^{n-1} f(T^k x) \neq 0$, then $f$ is not a measurable coboundary. Hence, it is possible to prove in practice that a function is not a coboundary (see also [PY99] and [NS03]).

In this note, we extend the aforementioned result of Livšic to nonuniformly expanding Markov dynamical systems, without any additional assumption on the functions $f$ or $u$. The result will first be given in the abstract setting of Gibbs-Markov maps (see [Aar97]). Applications to Young towers, intermittent maps in dimension 1 and positive recurrent Markov shifts will also be described.

The proof is quite flexible since it is completely elementary and does not use spectral theory. Hence, the same kind of arguments may be used in other settings.

Received by the editors July 16, 2004.

2000 Mathematics Subject Classification. Primary 37A20, 37D25.

©2005 American Mathematical Society
Reverts to public domain 28 years from publication
1.1. Results for Gibbs-Markov maps. In this paragraph, we will work in the setting of Gibbs-Markov maps, defined in [Aar97, Section 4.7].

Let us recall briefly the definitions. Let \((X, d, \mathcal{B}, m)\) be a bounded metric space endowed with its Borel \(\sigma\)-algebra and a probability measure. A nonsingular map \(T : X \to X\) is Gibbs-Markov if there exists a partition \(\alpha\) of \(X\) (modulo 0) by sets of positive measure, such that

1. For all \(a \in \alpha\), \(T(a)\) is a union (modulo 0) of elements of \(\alpha\) and \(T : a \to T(a)\) is invertible.
2. There exists a finite subset \(\{a_1, \ldots, a_n\}\) of \(\alpha\) with the following property: for any \(a \in \alpha\), there exist \(i, j \in \{1, \ldots, n\}\) such that \(a \subset T(a_i)\) and \(a_j \subset T(a)\) (modulo 0).
3. Expansion: there exists \(\lambda > 1\) such that \(\forall a \in \alpha\), for almost all \(x, y \in a\), \(d(Tx, Ty) \geq \lambda d(x, y)\).
4. Distortion: for \(a \in \alpha\), let \(g\) be the inverse of the jacobian of \(T\) on \(a\), i.e., \(g(x) = \frac{dm_{T^{-1}}}{dm_{T}}(x)\) for \(x \in a\). Then there exists \(C\) such that, for all \(a \in \alpha\), for almost all \(x, y \in a\), \[1 - \frac{g(x)}{g(y)}\leq Cd(Tx, Ty)\] where \(\tau\) is the separation time of \(\alpha\).

Remarks. Property (2), also known as the BIP (big images and preimages) property, is apparently stronger than the usual big image property \(\inf_{a \in \alpha} m(Ta) > 0\). However, when (4) is satisfied and \(T\) is probability preserving, these two properties are equivalent by [Sar03].

Usually, Gibbs-Markov maps are endowed with a distance given by \(d(x, y) = \tau\sigma(x, y)\), where \(\tau \in (0, 1)\) and \(\sigma(x, y)\) is the separation time of \(x\) and \(y\). Here we have chosen to use a general distance, since it will be more convenient in the applications: our main result will say that a function is Lipschitz continuous with respect to \(d\), which means that having more freedom to choose the distance will give more precise results. In particular, when the Gibbs-Markov map is obtained by coding another dynamical system, it is natural to use the distance induced by the original distance (see Sections 1.2 and 1.3 for illustrations of this phenomenon).

For \(a_0, \ldots, a_n = \alpha\), let \([a_0, a_{n-1}] = \bigcap_{i=0}^{n-1} T^{-i}(a_i)\). It is a cylinder of length \(n\). For \(f : X \to G\) and \(Z \subset X\), set
\[Df(Z) = \inf \{ C > 0 : \exists \Omega \subset Z \text{ with } m(Z \setminus \Omega) = 0 \text{ such that } \forall x, y \in \Omega, |f(x) - f(y)| \leq Cd(x, y) \}\.

The main result of this note is the following theorem:

**Theorem 1.1.** Let \((X, T, m, \alpha)\) be a probability-preserving Gibbs-Markov map. Let \(f : X \to G\) satisfy \(\sum_{a \in \alpha} m(a)Df(a) < +\infty\). Let \(u : X \to G\) be a measurable function such that \(f = u - u \circ T\) almost everywhere.

Then \(\sup_{a \in \alpha} Du(a) < \infty\), where \(\alpha_*\) is the partition generated by the images of the elements of \(\alpha\). Moreover, the function \(u\) is essentially bounded.

**Remarks.**

1. Since \(T\) is Markov, \(\alpha_*\) is coarser than \(\alpha\). In particular, \(\sup_{a \in \alpha} Du(a) < \infty\), i.e., \(u\) has a version which is uniformly Lipschitz on each element of the partition \(\alpha\).

2. The map \(T\) is also Gibbs-Markov for the distance \(d(x, y)\), when \(\gamma \in (0, 1]\).

Hence, Theorem 1.1 implies a similar statement for Hölder functions.
Proof. By \[\text{You99}\], we can assume without loss of generality that 
\[m \leq C \sum_{a \in \alpha} m(a) Df(a)\]. In particular, when \(f\) is constant on each element of \(\alpha\), we get \(Df(a) = 0\), i.e., \(u\) is essentially constant on the elements of \(\alpha_*\). When \(G = S^1\), we get a completely different proof of [AD01, Theorem 3.1].

(4) The proof would be easier under the stronger assumption 
\[\sup_{a \in \alpha} Df(a) < \infty.\]

However, this assumption is too strong, since it is not compatible with the induction process which will enable us to extend Theorem [11] to nonuniformly expanding settings.

In this paper, \(N = \{n \in \mathbb{Z}, n \geq 0\}\) and \(N^* = \mathbb{N}\setminus\{0\}\).

1.2. Application to Young towers. Let \((X, d, m)\) be a probability space endowed with a bounded metric \(d\). A map \(T: X \to X\) is a Young tower \((\text{You99})\) if there exist integers \(R_i \in N^*\) and a partition \(\{\Delta_{k,i}\}_{l \in N, k \in \{0, \ldots, R_i - 1\}}\) of \(X\) such that

(1) For all \(l\) and \(k < R_i - 1\), \(T\) is a measurable isomorphism between \(\Delta_{k,i}\) and \(\Delta_{k+1,i}\), preserving \(m\).

(2) For all \(l\), \(T\) is a measurable isomorphism between \(\Delta_{R_i-1,i}\) and \(\Delta_0 := \bigcup_m \Delta_{0,m}\).

(3) There exists \(\lambda > 1\) such that, for all \(l\), for all \(x, y \in \Delta_{0,i}\), \(d(T^{R_i}x, T^{R_i}y) \geq \lambda d(x, y)\).

(4) There exists \(C > 0\) such that, for all \(l\) and \(k < R_i\), for all \(x, y \in \Delta_{k,i}\), \(d(x, y) \leq Cd(T^{R_i-k}x, T^{R_i-k}y)\).

(5) For \(x \in \Delta_{R_i-1,i}\), let \(g(x)\) be the inverse of the distortion of \(T\) at \(x\), i.e.,
\[g(x) = \frac{d(x, y)}{d(y, T^k x)} \quad \text{such that, for all } l, \text{ for all } x, y \in \Delta_{R_i-1,i}, \quad 1 - \frac{g(x)}{g(y)} \leq Cd(x, y).\]

The third and fifth conditions mean that the returns to the basis are expanding and have a controlled distortion. Hence, Young towers are a good model for many nonuniformly expanding maps: the map has good properties, but after some waiting time, which can be arbitrarily long.

Theorem 1.2. Let \((X, T, m, d)\) be a Young tower, and let \(f: X \to G\) satisfy 
\[\sum m(\Delta_{k,i}) Df(\Delta_{k,i}) < \infty.\]

If \(u: X \to G\) is such that \(f = u - u \circ T\) almost everywhere, then the function \(u\) has a version which is Lipschitz on \(\Delta_0\), i.e., there exists \(C > 0\) such that, for almost all \(x, y \in \Delta_0\), \(|u(x) - u(y)| \leq C d(x, y)\).

This result applies in particular when the function \(f\) is Lipschitz.

Proof. By [You99], we can assume without loss of generality that \(m\) is invariant.

Let \(Y = \Delta_0\) with the partition \(\alpha = \{\Delta_{0,1}\}\), let \(\varphi: Y \to N^*\) be the first return time to \(Y\) (i.e., on \(\Delta_{0,t}, \varphi = R_t\)), and let \(T_Y = T^\varphi\) be the map induced by \(T\) on \(Y\).

Also define a distance \(d'\) on \(\Delta_{0,t} \in \alpha\) by \(d'(x, y) = d(T^{R_t}x, T^{R_t}y)\). If \(x\) and \(y\) are in two different elements of the partition \(\alpha\), also set \(d'(x, y) = \lambda \sup_{x, y} d\).

Then \((Y, T_Y, m_Y/m(Y), d')\) is a Gibbs-Markov map for the partition \(\alpha\). Moreover, \(T_Y\) preserves the measure \(m_Y/m(Y)\), and the partition \(\alpha_*\) is the trivial partition.
Let $f : X \to G$ satisfy $\sum m(\Delta_{k,l}) Df(\Delta_{k,l}) < \infty$, and assume that $f = u - u \circ T$. Define a function $f_Y$ on $Y$ by $f_Y(x) = \sum_{k=0}^{\varphi(x)-1} f(T^k x)$. On $\Delta_{0,l}$,

$$|f_Y(x) - f_Y(y)| \leq \sum_{k=0}^{R_l-1} |f(T^k x) - f(T^k y)| \leq \sum_{k=0}^{R_l-1} d(T^k x, T^k y) Df(\Delta_{k,l})$$

$$\leq Cd(T^{R_l} x, T^{R_l} y) \sum_{k=0}^{R_l-1} Df(\Delta_{k,l}) = Cd'(x, y) \sum_{k=0}^{R_l-1} Df(\Delta_{k,l}).$$

Hence, $\sum_{\alpha \in \alpha} m(a) Df_Y(a) \leq C \sum m(\Delta_{k,l}) Df(\Delta_{k,l}) < \infty$. Moreover, $f_Y = u - u \circ T_Y$.

Theorem 1.2 applies and proves that $u$ is almost everywhere Lipschitz on each element of $\Delta_0$, for the distance $d'$. In particular, on any element $\Delta_{0,l}$ of $\alpha$, we get $|u(x) - u(y)| \leq Ed'(x, y)$.

Finally take $x', y' \in \Delta_0$. They have preimages $x, y$ under $T^{R_l}$ in $\Delta_{0,l}$. As $f_Y(x) = u(x) - u(x')$ and $f_Y(y) = u(y) - u(y')$, we get

$$|u(x') - u(y')| \leq |f_Y(x) - f_Y(y)| + |u(x) - u(y)| \leq C d'(x, y) + Ed'(x, y) = C d'(x', y'). \quad \Box$$

1.3. Applications to intermittent maps. For $\alpha \in (0, 1)$, let $T$ be the map from $[0, 1]$ to itself given by

$$T(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{if } 0 \leq x \leq 1/2, \\ 2x - 1 & \text{if } 1/2 < x \leq 1. \end{cases}$$

This map has been studied by [LSV99]. It is nonuniformly expanding since the fixed point $0$ satisfies $T'(0) = 1$, and admits an absolutely continuous invariant probability measure.

**Proposition 1.3.** Let $f : [0, 1] \to G$ be Hölder with exponent $\gamma > 0$ on the intervals $[0, 1/2]$ and $(1/2, 1]$. If $u : [0, 1] \to G$ is measurable and satisfies $f = u - u \circ T$ Lebesgue almost everywhere, then there exists a function $\tilde{u}$, equal to $u$ almost everywhere, Hölder with exponent $\gamma$, and such that $f = \tilde{u} - \tilde{u} \circ T$ everywhere.

**Proof.** Let $Y = (1/2, 1]$, let $\varphi$ be the first return time from $Y$ to itself and let $T_Y : Y \to Y$ be the induced map. Then $T_Y$ is Gibbs-Markov for the partition $B_n = \{ y \in Y : \varphi(y) = n \}$, by [LSV99]. Hence, the arguments in the proof of Theorem 1.2 apply and prove that $u$ is a.e. Hölder on $Y$. As $T : (1/2, 1] \to (0, 1]$ is Lipschitz and has Lipschitz inverse, the coboundary equation implies that $u$ is a.e. Hölder on $(0, 1]$, i.e., there exists a set $V$ of full measure and a constant $C$ such that, for all $x, y \in V$, $|u(x) - u(y)| \leq C|x - y|\gamma$.

The function $u$ is uniformly continuous on $V$, whence it can be extended to a continuous—and even Hölder—function $\tilde{u}$ on $[0, 1]$. On $V \cap T^{-1}(V)$, which is dense, we have $f(x) = \tilde{u}(x) - \tilde{u}(T x)$. Since both members of this equality are continuous on the intervals $[0, 1/2]$ and $(1/2, 1]$, this equality in fact holds everywhere. \( \Box \)

In particular, if $f$ is a measurable coboundary, it satisfies $\sum_{0}^{n-1} f(T^k x) = \tilde{u}(T^n x) - \tilde{u}(x) = 0$ at any point $x$ such that $T^n(x) = x$.

**Corollary 1.4.** If $f : [0, 1] \to \mathbb{R}$ is Hölder continuous on $[0, 1/2]$ and $(1/2, 1]$ and satisfies $f(0) \neq 0$, then $f$ is not a measurable coboundary.
This solves a conjecture stated in [PHV03]; in this article, the authors need to know that \( f = \log |T'| - \int \log |T'| \) is not a coboundary to get a nonzero variance in the central limit theorem. As \( f \) is \( \alpha \)-Hölder on \([0, 1/2] \) and \((1/2, 1] \), and \( f(0) = -\int \log |T'| < 0 \), the corollary applies and proves that it is indeed never the case.

Using Theorem 1.1 with \( G = S^1 \), in the same way we can get a stronger result:

**Corollary 1.5.** The function \( f(x) = \log |T'| - \int \log |T'| \) cannot be written as \( f = u - u \circ T + \lambda q + \mu \) almost everywhere, where \( u : [0, 1] \to \mathbb{R} \) is measurable, \( q : X \to \mathbb{Z} \) and \( \lambda, \mu \in \mathbb{R} \).

The proof is the same, using the behavior at the fixed points 0 and 1 to get a contradiction. This is a strong aperiodicity result on the function \( f \). By [God03, Theorem 1.2], it implies that \( f \) satisfies a local limit theorem when \( \alpha < 1/2 \).

### 1.4. Application to positive recurrent Markov shifts

Let \( T : X \to X \) be a positive recurrent Markov shift with Hölder potential, as defined in [Sar01], preserving the probability measure \( m \). The map \( T \) satisfies the same assumptions as a Gibbs-Markov map, except the BIP property. We also assume that the distance \( d \) is given by \( d(x, y) = \tau(x, y) \), where \( \tau \in (0, 1) \) and \( s(x, y) \) is the separation time of \( x \) and \( y \). Such maps have in general more complicated combinatorics than Young towers, but they enjoy uniform expansion (since \( d(Tx, Ty) = \tau^{-1}d(x, y) \) for all \( x, y \) in the same element of \( \alpha \)), while Young towers are expanding only after many iterates.

**Theorem 1.6.** Let \( f : X \to G \) satisfy \( \sum_{a \in \alpha} m(a)Df(a) < \infty \). Let \( u : X \to G \) be a measurable function such that \( f = u - u \circ T \) almost everywhere. Then, for all \( a \in \alpha \), \( Du(a) < \infty \). Moreover, if \( T \) is transitive, \( \sum_{a \in \alpha} m(a)Du(a) < \infty \).

**Proof.** For \( a \in \alpha \), let \( T_a \) be the map induced by \( T \) on \([a] \). It is Gibbs-Markov. Using Theorem 1.1, we show as in the proof of Theorem 1.2 that \( Du(a) < \infty \). If \( T \) is transitive, the proof of Lemma 2.3 applies and gives \( \sum m(a)Du(a) < \infty \). \qed

## 2. Proof of Theorem 1.1

A Gibbs-Markov map is **transitive** if, for all \( a, b \in \alpha \), there exists \( n \) such that \( b \in T^n(a) \mod 0 \). When \( T \) preserves a probability measure, there exists a finite decomposition \( \alpha = \alpha_1 \cup \ldots \cup \alpha_n \) such that the image of an element of \( \alpha_i \) is contained in \( X_i = \bigcup_{a \in \alpha_i} a \), and such that \( T \) is a transitive Gibbs-Markov map on \( X_i \) ([Ama97]). To prove the theorem, it is sufficient to prove it on each \( X_i \). We can therefore assume that \( T \) is transitive.

The main step of the proof is the following lemma:

**Lemma 2.1.** There exists \( \alpha_1 \in \alpha \) such that \( Du(\alpha_1) < \infty \).

**Proof.** Let \( \Phi(x) = Df(a) \) when \( x \in a \). This function is integrable by assumption. In particular, there exists a set \( X_1 \) of full measure such that the Birkhoff sums \( S_n \Phi(x) = \sum_{k=0}^{n-1} \Phi(T^k x) \) satisfy \( S_n \Phi(x) = O(n) \) when \( x \in X_1 \).

There exists \( X_2 \) of full measure such that, if \( x \in X_2 \), all its iterates satisfy: for almost all \( y \) in the same element of partition \( a \) as \( T^n x \), \( |f(y) - f(T^n x)| \leq Df(a)d(y, T^n x) \).

The martingale convergence theorem implies that almost every point is a measurable continuity point of \( u \): there exists \( X_3 \) of full measure such that, if \( x \in X_3 \) and
\(a_0, a_1, \ldots\) denotes the sequence of elements of \(\alpha\) containing respectively \(x, Tx, \ldots\), then, for all \(\varepsilon > 0\),
\[
\frac{m\{y \in [a_0, \ldots, a_{n-1}] : |u(y) - u(x)| > \varepsilon\}}{m[a_0, \ldots, a_{n-1}]} \to 0.
\]

As \(T\) is Gibbs-Markov, all its iterates have a bounded distortion ([Aar97 Proposition 4.3.1]). Hence, there exists \(B > 0\) such that, for any measurable set \(Z\) and for any cylinder of length \(k\),
\[
B^{-1} \frac{m(T(a_{k-1}) \cap Z)}{m(Ta_{k-1})} \leq \frac{m([a_0, \ldots, a_{k-1}] \cap T^{-k}Z)}{m[a_0, \ldots, a_{k-1}]} \leq B \frac{m(T(a_{k-1}) \cap Z)}{m(Ta_{k-1})}.
\]

Since \(T\) has the big image property, this implies that there exists \(B' > 0\) such that
\[
\frac{m([a_0, \ldots, a_{k-1}] \cap T^{-k}Z)}{m[a_0, \ldots, a_{k-1}]} \leq B' m(Z).
\]

Let \(\lambda > 1\) be the expansion factor of \(T\) and let \(K > 0\) be large enough so that
\[
K \log \lambda > 3.
\]

Let \(\alpha_1, \ldots, \alpha_N\) be a finite number of elements of \(\alpha\) such that \(m(X \setminus \bigcup \alpha_i) \leq \varepsilon_0\) where \(\varepsilon_0\) satisfies \(K \log (1 - B' \varepsilon_0) \geq -1/2\). Write
\[
Z_n = \{x : \forall n^3 \leq k < n^3 + |K \log n|, T^n(x) \in \alpha_1 \cup \ldots \cup \alpha_N\}.
\]

Finally let \(X_4\) be the set of points belonging to infinitely many \(Z_n\).

**Lemma 2.2.** The set \(X_4\) has nonzero measure.

**Proof.** Write \(A = \alpha_1 \cup \ldots \cup \alpha_N\). Let us first bound \(m(Z_n)\) from below. For any cylinder \([a_0, \ldots, a_{k-1}]\), we apply (3) to \(X \setminus A\), of measure at most \(\varepsilon_0\), and we get
\[
m([a_0, \ldots, a_{k-1}] \cap T^{-k}A) \geq (1 - B' \varepsilon_0) m([a_0, \ldots, a_{k-1}]).
\]

Summing these inequalities for \(a_{k-1} = \alpha_1, \ldots, \alpha_N\) yields
\[
m([a_0, \ldots, a_{k-2}] \cap T^{-k+1}A \cap T^{-k}A) \geq (1 - B' \varepsilon_0) m([a_0, \ldots, a_{k-2}] \cap T^{-k+1}A).
\]

This last term is larger than \((1 - B' \varepsilon_0)^2 m[a_0, \ldots, a_{k-2}]\), again by (3). In this way we get by induction
\[
m([a_0, \ldots, a_l] \cap T^{-l+1}A \cap \ldots \cap T^{-k}A) \geq (1 - B' \varepsilon_0)^{k-l} m[a_0, \ldots, a_l].
\]

In particular, for \(l = -1\) and \(k = [K \log n] - 1\), we get using the invariance of \(m\) that
\[
m(Z_n) \geq (1 - B' \varepsilon_0)^{K \log n} = n^{K \log (1 - B' \varepsilon_0)} \geq \frac{1}{\sqrt{n}}.
\]

Hence, \(\sum m(Z_n) = \infty\). We will use a version of the Borel-Cantelli Lemma to conclude. Since the sets \(Z_n\) are not independent, we will use the following version of this lemma, due to Lamperti ([Spi64 Proposition 6.26.3]):

If \(\sum m(Z_n) = \infty\) and
\[
\lim \inf_{n \to \infty} \frac{\sum_{k=1}^{n} m(Z_j \cap Z_k)}{(\sum_{k=1}^{n} m(Z_k))^2} < \infty,
\]

then the set of points belonging to infinitely many \(Z_n\) has nonzero measure.

To estimate \(m(Z_j \cap Z_k)\), we will use the transfer operator \(\tilde{T}\), defined on \(L^2\) as the adjoint of the composition by \(T\). It acts continuously on the space \(\mathcal{L}\) of
functions which are bounded and Lipschitz on any element of α. Moreover, by [Aar97, Proposition 4.7.3], there exist \( M > 0 \) and \( \eta < 1 \) such that, for any \( h \in \mathcal{L} \),
\begin{equation}
\| \hat{T}^p h \| \mathcal{L} \leq M (\eta^p \| h \| \mathcal{L} + \| h \|_1).
\end{equation}

Let \( \chi \) be the characteristic function of \( A \), and let \( \gamma_n = \prod_{0 \leq k < |K \log n|} \chi \circ T^k \). Hence, \( m(Z_n) = \int \gamma_n \circ T^n = \int \gamma_n \), and \( m(Z_n \cap Z_p) = \int \gamma_n \circ T^n \cdot \gamma_p \circ T^p \). The function \( \chi \) belongs to \( \mathcal{L} \). For \( k > j \),
\begin{equation}
m(Z_j \cap Z_k) = \int \gamma_j \circ T_j^3 \cdot \gamma_k \circ T^k = \int \hat{T}^{k^3-j^3}(\gamma_j) \cdot \gamma_k \leq \| \hat{T}^{k^3-j^3}(\gamma_j) \| \mathcal{L} m(Z_k).
\end{equation}

As \( \hat{T} \) acts continuously on \( \mathcal{L} \), the function
\[\delta_j = \hat{T}^{[K \log j]}(\gamma_j) = \hat{T}(\chi \hat{T} \cdots \hat{T}(\chi) \cdots)\]
satisfies \( \| \delta_j \| \mathcal{L} \leq (2M)^{K \log j} \). The inequality (5) applied to \( p = k^3 - j^3 - [K \log j] \) and \( h = \delta_j \) yields
\begin{equation}
\| \hat{T}^{k^3-j^3} \gamma_j \| \mathcal{L} \leq M \left( \eta^{k^3-j^3-K \log j} \| \delta_j \| \mathcal{L} + \| \delta_j \|_1 \right) \\
\leq M \left( \eta^{k^3-j^3-K \log j} (2M)^{K \log j} + m(Z_j) \right)
\end{equation}

since \( \| \delta_j \|_1 = \int \delta_j = \int \gamma_j \), for all these functions are nonnegative. Hence, (6) and (7) give
\[|m(Z_j \cap Z_k)| \leq M \eta^{k^3-j^3} (2M/\eta)^{K \log j} + M m(Z_j) m(Z_k).
\]

Finally,
\[\sum_{j < k \leq n} m(Z_j \cap Z_k) \leq M \sum_{j < k} m(Z_j) m(Z_k) + M \sum_{j=1}^{\infty} \eta^{-j^3} (2M/\eta)^{K \log j} \sum_{k=j+1}^{\infty} \eta^{k^3} \\
\leq M \left( \sum_{k \leq n} m(Z_k) \right)^2 + M \sum_{j=1}^{\infty} \eta^{-j^3} (2M/\eta)^{K \log j} \sum_{l=(j+1)^3}^{\infty} \eta^l.
\]
The last sum is bounded by
\[M \sum_{j=1}^{\infty} \eta^{-j^3} (2M/\eta)^{K \log j} \eta^{(j+1)^3} (1-\eta)^{3(j+1)} < \infty,
\]
which shows that the aforementioned Borel-Cantelli lemma applies. \qed

We can take \( x_0 \in X_1 \cap X_2 \cap X_3 \cap X_4 \) since this set has positive measure. Let \( m_k \to \infty \) be such that \( x_0 \in Z_{m_k} \), and \( n_k = m_k^3 + [K \log m_k] - 1 \). Then \( T^{n_k}(x_0) \) belongs to one of the sets \( \alpha_1, \ldots, \alpha_N \). In particular, one of these sets is used infinitely many times, and taking a further subsequence we can for example assume that \( T^{m_k}(x_0) \in \alpha_1 \) for all \( k \). We will show that \( D \alpha_1 < \infty \). Denote by \( a_0, a_1, \ldots \) the elements of \( \alpha \) containing respectively \( x_0, T(x_0), \ldots \). Let \( [a_n] = [a_0, \ldots, a_{n-1}] \), and let \( v_n : Ta_{n-1} \to [a_n] \) be the inverse of \( T^n : [a_n] \to Ta_{n-1} \).

Let \( \varepsilon > 0 \). As \( x_0 \in X_3 \),
\[\frac{m \{ y \in [a_{n_k}] : |u(y) - u(x_0)| \geq \varepsilon \}}{m([a_{n_k}])} \to 0.
\]
Taking a further subsequence of \( n_k \), we can assume that
\[
\sum \frac{m\{y \in [a_{n_k}] : |u(y) - u(x_0)| > \varepsilon\}}{m[a_{n_k}]} < \infty.
\]
For all \( k \in \mathbb{N} \), the distortion control (2) implies that
\[
\frac{m\{y \in T^{n_{k-1}} : |u(v_{n_k}y) - u(x_0)| > \varepsilon\}}{m[T^{n_{k-1}}]} \geq \frac{m\{y \in [a_{n_k}] : |u(y) - u(x_0)| > \varepsilon\}}{m[a_{n_k}]}.\]
Hence, \( \sum k m\{y \in T^{n_{k-1}} : |u(v_{n_k}y) - u(x_0)| > \varepsilon\} < +\infty \). Therefore, \( U_\varepsilon := \{ y \in X : \exists \kappa, \forall k \geq \kappa, \text{ if } y \in T^{n_{k-1}}, \text{ then } |u(v_{n_k}y) - u(x_0)| \leq \varepsilon \} \) has full measure.

Let \( y_1, y_2 \in U_\varepsilon \cap \alpha_1 \). If \( k \) is large enough, the preimages \( y_1' \) and \( y_2' \) of \( y_1 \) and \( y_2 \) under \( T^{n_k} \) in \([a_{n_k}]\) satisfy \( |u(y_1') - u(x_0)| \leq \varepsilon \), whence \( |u(y_1') - u(y_2')| \leq 2\varepsilon \). Then
\[
|u(y_1) - u(y_2)| = |u \circ T^{n_k}(y_1') - u \circ T^{n_k}(y_2')| 
\leq \sum_{i=0}^{n_k-1} |f \circ T^i(y_1') - f \circ T^i(y_2')| + |u(y_1') - u(y_2')|.
\]
(8)
Recall that \( n_k = m_k^3 + [K \log m_k] - 1 \), and that \( \Phi \) is defined by \( \Phi(x) = Df(a) \) when \( x \in a \). Then
\[
\sum_{i=0}^{m_k^3-1} |f \circ T^i(y_1') - f \circ T^i(y_2')| 
\leq \sum_{i=0}^{m_k^3-1} \Phi(T^i(x_0))d(T^iy_1', T^iy_2') 
\leq \sum_{i=0}^{m_k^3-1} \Phi(T^i(x_0))\lambda^{i-n_k}d(T^{n_k}y_1', T^{n_k}y_2') 
\leq \lambda^{-K \log m_k + 2S_{D_k}3^3} \Phi(x_0)\lambda\lambda d(y_1, y_2).
\]
Since \( x_0 \in X_1 \), there exists \( C \) such that \( S_{D_k}(x_0) \leq Cn \) for all \( n \). As \( -K \log \lambda \leq -3 \) by (4), we get that (6) tends to 0.

Finally, set \( D = \sup Df(\alpha_j) \) for \( 1 \leq j \leq N \). By definition of \( m_k \), we have \( T^i(x_0) \in \bigcup_{1 \leq j \leq N} \alpha_j \) for all \( m_k^3 \leq i < n_k \), whence
\[
\sum_{i=m_k^3}^{n_k-1} |f \circ T^i(y_1') - f \circ T^i(y_2')| 
\leq \sum_{i=m_k^3}^{n_k-1} Dd(T^iy_1', T^iy_2') 
\leq D \sum_{i=m_k^3}^{n_k-1} \lambda^{i-n_k}d(y_1, y_2) 
\leq \frac{D}{\lambda-1} d(y_1, y_2).
\]
Equation (8) then yields
\[
|u(y_1) - u(y_2)| \leq o(1) + \frac{D}{\lambda-1} d(y_1, y_2) + 2\varepsilon.
\]
Finally, on \( \alpha_1 \cap \bigcap_{\varepsilon>0} U_\varepsilon \), we have \( |u(y_1) - u(y_2)| \leq \frac{D}{\lambda-1} d(y_1, y_2) \). \( \square \)

**Lemma 2.3.** We have \( \sum_{a \in \alpha} m(a)Du(a) < \infty \).

**Proof.** Let us show that, for any \( a \in \alpha \), \( Du(a) < \infty \). As \( T \) is transitive, there exists \( n \) such that \( a \subset T^n(\alpha_1) \). Let \([a_0, \ldots, a_{n-1}]\) be a cylinder included in \( \alpha_1 \) such
that $a \subset T(a_{n-1})$. For $y_1, y_2 \in a$, let $y'_1$ and $y'_2$ be their preimages under $T^n$ in $[a_0, \ldots, a_{n-1}]$. Then

$$|u(y_1) - u(y_2)| \leq \sum_{i=0}^{n-1} |f(T^n y'_1) - f(T^n y'_2)| + |u(y'_1) - u(y'_2)|$$

(10)

$$\leq \sum_{i=0}^{n-1} Df(a_i) \lambda^{i-n} d(y_1, y_2) + \lambda^{-n} Du(a_1)d(y_1, y_2),$$

which proves that $Du(a) < \infty$.

Let $\beta$ be a finite nonempty subset of $\alpha$. For $a \in \alpha \setminus \beta$, let us show that

$$m(a) = \sum_{n=1}^{\infty} \sum_{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta} m[a_0, a_1, \ldots, a_{n-1}, a].$$

Let $Y = \bigcup_{b \in \beta} b$. Write $A_0 = a$, and $A_{n+1} = T^{-1}(A_n) \setminus Y$ and $B_{n+1} = T^{-1}(A_n) \cap Y$. We get

$$A_n = \bigcup_{a_0, \ldots, a_{n-1} \in \alpha \setminus \beta}[a_0, \ldots, a_{n-1}, a]$$

and

$$B_n = \bigcup_{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta}[a_0, \ldots, a_{n-1}, a].$$

Thus, we want to show that $m(a) = \sum_n m(B_n)$. The equality $T^{-1}(A_n) = A_{n+1} \cup B_{n+1}$ implies $m(A_n) = m(A_{n+1}) + m(B_{n+1})$. By induction, we get $m(a) = m(B_1) + \ldots + m(B_n) + m(A_n)$. It remains to prove that $m(A_n) \to 0$. Note that $A_n \subset C_n = \{x : \forall 0 \leq k \leq n, T^k(x) \notin Y\}$. We will show that $m(C_n) \to 0$ by proving that $C = \bigcap C_n$ has 0 measure. Since the measure is invariant and $C \subset T^{-1}(C)$, $C = T^{-1}(C) \mod 0$, whence $m(C) = 0$ or 1 by ergodicity (Aar97 Theorem 4.4.7). The set $C$ does not intersect $Y$, which has nonzero measure, hence $m(C) = 0$. This proves (11).

Let $[a_0, \ldots, a_{n-1}, a]$ be a cylinder of nonzero measure. By (10),

$$Du(a) \leq \sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) + \lambda^{-n} Du(a_0).$$

Hence, (11) yields

$$m(a)Du(a) \leq \sum_{n=1}^{\infty} \sum_{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta} m[a_0, \ldots, a_{n-1}, a] \left( \sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) \right) + m(a) \sup_{b \in \beta} Du(b).$$

As $\sum \sup_{b \in \beta} Du(b) < \infty$, we will show that $\sum m(a)Du(a) < \infty$ by showing that

$$\sum_{n=1}^{\infty} \sum_{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta} m[a_0, \ldots, a_{n-1}] \left( \sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) \right)$$

(12)
is finite. In this expression, for \( a' \in \alpha \setminus \beta \), the prefactor of a term \( \lambda^{-k}Df(a') \) is

\[
\sum_{n=1}^{\infty} \sum_{\substack{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta \\ a_{n+1}, \ldots, a_{n+k-1} \in \alpha \setminus \beta}} m[a_0, \ldots, a_{n-1}, a', a_{n+1}, \ldots, a_{n+k-1}]
\]

\[
\leq \sum_{n=1}^{\infty} \sum_{\substack{a_0 \in \beta, a_1, \ldots, a_{n-1} \in \alpha \setminus \beta}} m[a_0, \ldots, a_{n-1}, a'].
\]

By (11), this last term is equal to \( m(a') \). In (12), the prefactor of a term \( \lambda^{-k}Df(a') \) with \( a' \in \beta \) is also at most \( m(a') \). Hence,

\[
(12) \leq \sum_{a' \in \alpha} \sum_{k=1}^{\infty} m(a') \lambda^{-k}Df(a'),
\]

which is finite since \( \sum m(a')Df(a') < \infty \). \( \square \)

**Proof of Theorem 1.1** For almost all \( x \), \( \sum_{y=x} T^{-1}(y) g(y) = 1 \). Let us write \( T^{-1}(x) = \{x_0, x_1, \ldots\} \), and let \( a_i \) be the element of \( \alpha \) containing \( x_i \). By bounded distortion and the big image property, there exists \( C > 0 \) such that, for all \( n \), \( g(x_n) \leq Cm(a_n) \).

As \( \sum g(x_n) = 1 \), this implies \( C \sum m(a_n) \geq 1 \).

Let \( a_\ast \) be an element of \( \alpha_\ast \). Let \( x, y \in a_\ast \). By definition of \( \alpha_\ast \), their preimages \( x_0, x_1, \ldots \) and \( y_0, y_1, \ldots \) belong to the same elements \( a_0, a_1, \ldots \) of \( \alpha \). Since \( f = u - u \circ T \), we have for any \( n \)

\[
|u(x) - u(y)| \leq |f(x_n) - f(y_n)| + |u(x_n) - u(y_n)| \leq (Df(a_n) + Du(a_n))d(x_n, y_n)
\]

\[
\leq (Df(a_n) + Du(a_n))\lambda^{-1}d(x, y).
\]

Hence,

\[
|u(x) - u(y)| \leq C \sum m(a_n)|u(x) - u(y)|
\]

\[
\leq C \sum m(a_n)(Df(a_n) + Du(a_n))\lambda^{-1}d(x, y).
\]

Finally, \( Du(a_\ast) \leq \frac{C}{\lambda} \sum_{a \in \alpha} m(a)(Df(a) + Du(a)) \), which is finite by Lemma 2.3.

To prove that \( u \) is essentially bounded, we use the big preimage property. Let \( a_1, \ldots, a_n \in \alpha \) be such that every element of \( \alpha \) is contained in the image of some \( a_i \). Let \( a \in \alpha \), and let \( i \) be such that \( a \subset T(a_i) \). For \( x \in a \), let \( x' \) be its preimage in \( a_i \). Then we get

\[
|u(x)| = |u(x') - f(x')| \leq \|u|a|\|_{\infty} + \|f|a|\|_{\infty}.
\]

This last quantity is uniformly bounded. \( \square \)

**References**


Département de mathématiques et applications, École Normale Supérieure, 45 rue d’Ulm, 75005 Paris, France

E-mail address: sebastien.gouezel@ens.fr