Bull. Soc. math. France
134 (1), 2006, p. 1–31

DECAY OF CORRELATIONS FOR NONUNIFORMLY EXPANDING SYSTEMS

BY SÉBASTIEN GOUËZEL

ABSTRACT. — We estimate the speed of decay of correlations for general nonuniformly expanding dynamical systems, using estimates on the time the system takes to become really expanding. Our method can deal with fast decays, such as exponential or stretched exponential. We prove in particular that the correlations of the Alves-Viana map decay in $O(\mathrm{e}^{-c\sqrt{n}})$.

RÉSUMÉ (*Décroissance des corrélations d'un système non uniformément dilatant*)
On montre comment estimer la vitesse de mélange d'un système dynamique non uniformément dilatant, à partir d'estimées sur le temps dont le système a besoin pour être vraiment dilatant. Cette méthode permet d'obtenir des vitesses rapides, par exemple exponentielles gauches ou exponentielles. Comme application, on obtient en particulier le fait que les corrélations des applications d'Alves-Viana décroissent en $O(\mathrm{e}^{-c\sqrt{n}})$.

1. Results

1.1. Decay of correlations and asymptotic expansion. — When $T : M \to M$ is a map on a compact space, the asymptotic behavior of Lebesgue-almost every point of M under the iteration of T is related to the existence

Texte reçu le 16 février 2004, révisé le 2 août 2004 et le 13 décembre 2004.

SÉBASTIEN GOUËZEL, Département de Mathématiques et Applications, École Normale Supérieure, 45 rue d'Ulm 75005 Paris (France). • *E-mail :* `Sebastien.Gouezel@ens.fr`

2000 Mathematics Subject Classification. — 37A25, 37D25.

Key words and phrases. — Decay of correlations, Young tower, non uniformly expanding maps.

of absolutely continuous (or more generally SRB) invariant probability measures μ. To understand more precisely the mixing properties of the system, an essential feature is the speed at which the correlations

$$\mathrm{Cor}(f, g\circ T^n) := \int f\cdot g\circ T^n \,\mathrm{d}\mu - \int f\,\mathrm{d}\mu \int g\,\mathrm{d}\mu$$

tend to 0. In a uniformly expanding setting, the decay is exponential, but little is known when the expansion is non uniform.

Recently, [6] introduced a quantitative way to measure the non-uniform expansion of a map, and showed that this "measure of non-uniformity" makes it possible to control the speed of decay of correlations. More precisely, when the measure of non-uniformity decays polynomially, [6] shows that the decay of correlations is also polynomial, using hyperbolic times techniques (see [2]) and Young towers (see [14]). As a consequence of this result, the correlations of the Alves-Viana map (see [12]) decay faster than any polynomial (which implies for example a central limit theorem). However, all the estimates of [12] are in $O(\mathrm{e}^{-c\sqrt{n}})$, which is stronger. A precise study of the recurrence makes it in fact possible to show that the correlations also decay in $O(\mathrm{e}^{-c\sqrt{n}})$ (see [8], [9]). However, this direct approach relies strongly on the specificities of the Alves-Viana map, contrary to the approach of [6], which uses only some general abstract properties, and can therefore be extended to many other cases. The goal of this article is to extend the results of [6] (using a substantially different method) to speeds of $\mathrm{e}^{-c\sqrt{n}}$ (among others), which implies that the results of [9] hold in a much wider setting.

Let M be a compact Riemannian manifold (possibly with boundary) and $T : M \to M$. We assume that there exists a closed subset $S \subset M$, with zero Lebesgue measure (containing possibly discontinuities or critical points of T, and with $\partial M \subset S$), such that T is a C^2 local diffeomorphism on $M \setminus S$, and is non uniformly expanding: there exists $\lambda > 0$ such that, for Lebesgue almost every $x \in M$,

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \log \|DT(T^k x)^{-1}\|^{-1} \geqslant \lambda. \tag{1}$$

We also need non-degeneracy assumptions close to S, similar to the assumptions in [5] or [6]: we assume that there exist $B > 1$ and $\beta > 0$ such that, for any $x \in M \setminus S$ and every $v \in T_x M \setminus \{0\}$,

$$\frac{1}{B}\,\mathrm{dist}(x,S)^{\beta} \leqslant \frac{\|DT(x)v\|}{\|v\|} \leqslant B\,\mathrm{dist}(x,S)^{-\beta}. \tag{2}$$

Assume also that, for all $x, y \in M$ with $\mathrm{dist}(x,y) < \frac{1}{2}\,\mathrm{dist}(x,S)$,

$$\Big|\log\|DT(x)^{-1}\| - \log\|DT(y)^{-1}\|\Big| \leqslant B\frac{\mathrm{dist}(x,y)}{\mathrm{dist}(x,S)^{\beta}} \tag{3}$$

and

$$\big|\log|\det DT(x)^{-1}| - \log|\det DT(y)^{-1}|\big| \leqslant B\frac{\operatorname{dist}(x,y)}{\operatorname{dist}(x,S)^{\beta}}, \tag{4}$$

i.e. $\log\|DT^{-1}\|$ and $\log|\det DT^{-1}|$ are locally Lipschitz, with a constant which is controlled by the distance to the critical set. This implies that the singularities are at most polynomial, and in particular that the critical points are not flat.

We assume that the critical points come subexponentially close to S in the following sense. For $\delta > 0$, set $\operatorname{dist}_\delta(x,S) = \operatorname{dist}(x,S)$ if $\operatorname{dist}(x,S) < \delta$, and $\operatorname{dist}_\delta(x,S) = 1$ otherwise. We assume that, for all $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that, for Lebesgue almost every $x \in M$,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} -\log\operatorname{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant \varepsilon. \tag{5}$$

We will need to control more precisely the speed of convergence in (1) and (5). As [6], we consider for this the following function, which measures the non-uniformity of the system

$$h^1_{(\varepsilon_1,\varepsilon_2)}(x) = \inf\Big\{N\in\mathbb{N}^* \mid \forall n\geqslant N,\ \frac{1}{n}\sum_{k=0}^{n-1}\log\|DT(T^k x)^{-1}\|^{-1} \geqslant \frac{\lambda}{2}$$
$$\text{and for } i=1,2,\ \frac{1}{n}\sum_{k=0}^{n-1} -\log\operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i\Big\}.$$

It is important to have two indexes ε_1 and ε_2 to guarantee the existence of hyperbolic times (see Lemma 2.2). To simplify the notations, we will write $\varepsilon = (\varepsilon_1, \varepsilon_2)$. The points x such that $h^1_\varepsilon(x) = n$ are "good" for times larger than n. Hence, the lack of expansion of the system at time n is evaluated by

$$\operatorname{Leb}\{x \mid h^1_\varepsilon(x) > n\}, \tag{6}$$

and it is natural to try to estimate the speed of decay of correlations using this quantity. This is done in [6] in the polynomial case: if (6) $= O(1/n^\gamma)$ for some $\gamma > 1$, then the correlations of Hölder functions decay at least like $1/n^{\gamma-1}$.

Set $\Lambda = \bigcap_{n\geqslant 0} T^n(M)$. We will say that T is topologically transitive on the attractor Λ if, for every nonempty open subsets U, V of Λ, there exists n such that $T^{-n}(U)\cap V$ contains a nonempty open set (the precise formulation is important since T may not be continuous on S).

We will say that a sequence $(u_n)_{n\in\mathbb{N}}$ has *polynomial decay* if there exists $C > 0$ such that, for all $\frac{1}{2}n \leqslant k \leqslant n$, $0 < u_k \leqslant Cu_n$. This implies in particular that u_n does not tend too fast to 0: there exists $\gamma > 0$ such that $1/n^\gamma = O(u_n)$ (for example $\gamma = \log C/\log 2$).

Finally, the *basin* of a probability measure μ is the set of points x such that $n^{-1}\sum_{k=0}^{n-1}\delta_{T^k x}$ converges weakly to μ, where δ_y is the Dirac mass at y.

THEOREM 1.1. — *Assume that all the iterates of T are topologically transitive on Λ and that, for all $\varepsilon=(\varepsilon_1,\varepsilon_2)$, there exists a sequence $u_n(\varepsilon)$ with $\sum u_n(\varepsilon)<+\infty$ and $\mathrm{Leb}\{x \mid h^1_\varepsilon(x)>n\}=O(u_n(\varepsilon))$. Assume moreover that $u_n(\varepsilon)$ satisfies one of the following properties:*

1) *$u_n(\varepsilon)$ has polynomial decay.*

2) *There exist $c(\varepsilon)>0$ and $\eta(\varepsilon)\in(0,1]$ such that $u_n(\varepsilon)=\mathrm{e}^{-c(\varepsilon)n^{\eta(\varepsilon)}}$.*

Then T preserves a unique (up to normalization) absolutely continuous (with respect to Lebesgue) measure μ. Moreover, this is a mixing probability measure, whose basin contains Lebesgue-almost every point of M.

Finally, there exists $\varepsilon^0=(\varepsilon^0_1,\varepsilon^0_2)$ such that, if $f,g:M\to\mathbb{R}$ are two functions with f Hölder and g bounded, their correlations

$$\mathrm{Cor}(f,g\circ T^n)=\int f\cdot g\circ T^n\,\mathrm{d}\mu-\int f\,\mathrm{d}\mu\int g\,\mathrm{d}\mu$$

decay at the following speed:

1) $|\mathrm{Cor}(f,g\circ T^n)|\leqslant C\sum_{p=n}^{\infty}u_p(\varepsilon^0)$ *in case* 1).

2) *There exists $c'>0$ such that $|\mathrm{Cor}(f,g\circ T^n)|\leqslant C\,\mathrm{e}^{-c'n^{\eta(\varepsilon^0)}}$ in case* 2).

In fact, ε^0 can be chosen *a priori*, depending only on λ and T. It would then be sufficient to have (5) for ε^0_1 and ε^0_2 to get the theorem. However, in practical cases, it is often not harder to prove (5) for all values of ε than to prove it for a specific value of ε. This is why, as in [5] and [6], we have preferred to state the theorem in this more convenient way.

In the first case, taking $u_n=1/n^\gamma$, we get another proof of the result of [6].

The main problem of this theorem is that (6) is often difficult to estimate, since $h^1_\varepsilon(x)$ states a condition on *all* iterates of x, and not only a finite number of them.

1.2. The Alves-Viana map.

— Theorem 1.1 applies to the Alves-Viana map, given by

$$T:\begin{cases} S^1\times I \longrightarrow S^1\times I,\\ (\omega,x)\ \longmapsto \big(16\omega, a_0+\varepsilon\sin(2\pi\omega)-x^2\big),\end{cases}\tag{7}$$

where $a_0\in(1,2)$ is a Misiurewicz point (i.e. the critical point 0 is preperiodic for $x\mapsto a_0-x^2$), ε is small enough and I is a compact interval of $(-2,2)$ such that T sends $S^1\times I$ into its interior.

This map has been introduced by Viana in [12]. He shows that T (and in fact any map close enough to T in the C^3 topology) has almost everywhere two positive Lyapunov exponents, even though there are critical points in the fibers. More precisely, Viana shows that the points that do not see the expansion in the fiber have a measure decaying like $O(\mathrm{e}^{-c\sqrt{n}})$. In [3], Alves and Araújo obtain from this information that, for every $\varepsilon = (\varepsilon_1, \varepsilon_2)$, for every $c < \frac{1}{4}$,

$$\mathrm{Leb}\big\{x \mid h^1_\varepsilon(x) > n\big\} = O(\mathrm{e}^{-c\sqrt{n}}). \tag{8}$$

Moreover, [7] shows that all the iterates of T are topologically transitive on Λ.

A consequence of the results of [6] is that the correlations of the Alves-Viana map decay faster than any polynomial. However, their method of proof can deal only with polynomial speeds (see paragraph 1.4), and hence can not reach the conjectural upper bound of $\mathrm{e}^{-c'\sqrt{n}}$. Theorem 1.1 implies this conjecture (already announced in [8]):

Theorem 1.2. — *The correlations of Hölder functions for any map close enough (in the C^3 topology) to the Alves-Viana map decay at least like* $\mathrm{e}^{-c'\sqrt{n}}$ *for some $c' > 0$.*

This result applies also if the expansion coefficient 16 is replaced by 2, according to [10]. Note that the specific method of [9], which proves Theorem 1.2, can not be directly used when 16 is replaced by 2, since it uses in particular the specific form of admissible curves. On the other hand, the abstract method of this article applies immediately, since [10] proves essentially (8).

1.3. Decorrelation and expansion in finite time. — The function $h^1_\varepsilon(x)$ takes into account the expansion at x for large enough times, and is consequently hard to estimate in general. It is more natural to consider the first time with enough expansion. For technical reasons, we will need three parameters to get results in this setting (see the proof of Lemma 2.1). Set

$$h^2_{(\varepsilon_1,\varepsilon_2,\varepsilon_3)}(x) = \inf\Big\{n \in \mathbb{N}^* \mid \frac{1}{n}\sum_{k=0}^{n-1} \log \|DT(T^k x)^{-1}\|^{-1} \geqslant \frac{\lambda}{2}$$
$$\text{and for } i = 1,2,3,\ \frac{1}{n}\sum_{k=0}^{n-1} -\log \mathrm{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i\Big\}.$$

This definition takes only the first n iterates of x into account, and can consequently be checked in finite time. We will write $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$. The time h^2_ε is related to the notion of *first hyperbolic time* studied for example in [4].

If there were only two parameters in the definition of h^2, we would have $h^2 \leqslant h^1$. However, since there are three parameters, h^1 and h^2 can rigourously not be compared.

We will estimate the speed of decay of correlations using $\mathrm{Leb}\{x \mid h^2_\varepsilon(x) > n\}$. Our main result is the following theorem:

Theorem 1.3. — *Assume that all the iterates of T are topologically transitive on Λ and that, for all $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$, there exists a sequence $u_n(\varepsilon)$ with*

$\sum(\log n)u_n(\varepsilon) < +\infty$ and $\operatorname{Leb}\{x \mid h^2_\varepsilon(x) > n\} = O(u_n(\varepsilon))$. Assume moreover that $u_n(\varepsilon)$ satisfies one of the following properties:

1) *$u_n(\varepsilon)$ has polynomial decay.*
2) *there exist $c(\varepsilon) > 0$ and $\eta(\varepsilon) \in (0,1]$ such that $u_n(\varepsilon) = \mathrm{e}^{-c(\varepsilon)n^{\eta(\varepsilon)}}$.*

Then T preserves a unique (up to normalization) absolutely continuous invariant measure μ. Moreover, this measure is a mixing probability measure, whose basin contains Lebesgue almost every point of M.

Finally, there exists $\varepsilon^0 = (\varepsilon^0_1, \varepsilon^0_2, \varepsilon^0_3)$ such that, if $f, g : M \to \mathbb{R}$ are two functions with f Hölder and g bounded, their correlations

$$\operatorname{Cor}(f, g \circ T^n) = \int f \cdot g \circ T^n \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu \int g \,\mathrm{d}\mu$$

decay at the following speed:

1) *$|\operatorname{Cor}(f, g \circ T^n)| \leqslant C \sum_{p=n}^{\infty} (\log p) u_p(\varepsilon^0)$ in case 1).*
2) *There exists $c' > 0$ such that $|\operatorname{Cor}(f, g \circ T^n)| \leqslant C \mathrm{e}^{-c' n^{\eta(\varepsilon^0)}}$ in case 2).*

For example, when $\operatorname{Leb}\{x \mid h^2_\varepsilon(x) > n\} = O(1/n^\gamma)$ with $\gamma > 1$, the correlations decay like $\log n / n^{\gamma-1}$. In the first case (polynomial decay), note that there is a loss of $\log n$ between Theorem 1.1 and Theorem 1.3. It is not clear whether this loss is real, or due to the technique of proof.

The comments on the choice of ε^0 following Theorem 1.1 are still valid here. It is even possible to take the same value for ε^0_1 and ε^0_2 in both theorems.

We will return later to the existence of invariant measures (Theorems 3.2 and 4.3). Without transitivity assumptions, we will get a spectral decomposition: T admits a finite number of absolutely continuous invariant ergodic probability measures, and each of these measures has a finite number of components which are mixing for an iterate of T, with the same bounds on the decay of correlations as in Theorems 1.1 and 1.3: these theorems correspond to the case where the spectral decomposition is trivial.

Remark. — If u_n has polynomial decay and $u_n = O(1/n^\gamma)$ for some $\gamma > 1$, then $\sum_{p=n}^{\infty} (\log p) u_p = O((\log n) \sum_{p=n}^{\infty} u_p)$, which simplifies a little the bound on the decay of correlations.

Remark. — In the stretched exponential case (i.e. $0 < \eta < 1$), the conclusions of Theorems 1.1 and 1.3 are true for any $c' < c(\varepsilon^0)$. This can easily be checked in all the following proofs (except in the proof of Lemma 4.2, where slightly more careful estimates are required).

1.4. Strategy of proof. — As it is often the case when one wants to estimate the decay of correlations, the strategy of proof consists in building a Young tower (see [14]), i.e. selecting a subset B of M and building a partition $B = \bigcup B_i$ such that T^{R_i} is an isomorphism between B_i and B, for some return time

R_i. Then [14] gives estimates on the decay of correlations, depending on the measure of points coming back to B after time n, i.e., $\text{Leb}\left(\bigcup_{R_i>n} B_i\right)$. To construct the sets B_i, we will use *hyperbolic times*. Denote by H_n the set of points for which n is a hyperbolic time.

This strategy is implemented in [6]. We will describe quickly their inductive construction, in a somewhat incorrect way but giving the essential ideas. Before time n, assume that some sets B_i have already been constructed, with a return time R_i satisfying $R_i < n$. At time n, consider $H_n \setminus \left(\bigcup_{R_i<n} B_i\right)$, and construct new sets B_j covering a definite proportion of this set, with return time $R_j = n$. Using some information about the repartition of hyperbolic times (the *Pliss Lemma*), it is then possible to prove that $\text{Leb}\left(\bigcup_{R_i>n} B_i\right)$ decays at least polynomially. The main limitation of this strategy is that, at time n, it can deal only with a fraction of H_n. Since the repartition of hyperbolic times is *a priori* unknown (except for the Pliss Lemma), we may have to wait a long time ($\sim n$) to see another hyperbolic time. This makes it impossible to prove that the decorrelations decay faster than $\mathrm{e}^{-c(\log n)^2}$ without further information.

To avoid this problem, we will deal with all points of H_n at time n, and not only a fraction. To do this, we will consider a fixed partition $U_1, \dots, U_N$ of the space (with N fixed) and use $T^{-n}(U_1), \dots, T^{-n}(U_N)$ to partition H_n. In this way, we will get a partition $\mathcal{B}_i$ of U_i (for each i), and each element of $\mathcal{B}_i$ will be sent on some (possibly different) U_j by an iterate of T. Moreover, we will keep a precise control on the measure of points having long return times.

Using this auxiliary partition, it will be quite easy to build a Young tower, using an inductive process: select some U_i, for example U_1. While a point does not fall into U_1, go on iterating, so that it falls in some U_j, then some U_k, and so on. Most points will come back to U_1 after a finite (and well controlled) number of iterates, and this will give the required partition of U_1.

Finally, to estimate the decay of correlations, it will not be possible to apply directly the results of [14], since they are slightly too weak (in the case of $\mathrm{e}^{-cn^{\eta}}$ with $0 < \eta < 1$, Young proves only a decay of correlations of $\mathrm{e}^{-c'n^{\eta'}}$ for any $\eta' < \eta$, which is weaker than the results of Theorems 1.1 and 1.3). However, the combinatorial techniques used in the construction of the partition will easily enable us to strengthen the results of [14], to obtain the required estimates.

The main difficulty of the proof will be to get the estimates on the auxiliary partition $U_1, \dots, U_N$, in Section 3 (for example, the logarithmic loss between Theorems 1.3 and 1.1 will appear there). Then we will build the Young tower in Section 4, and estimate the decay of correlations in paragraph 4.2. We will prove at the same time Theorems 1.1 and 1.3.

Acknowledgments. — I would like to thank V. Baladi for many enlightening discussions and explanations, and the referee for his valuable comments.

2. Hyperbolic times

We recall in this section the notion of hyperbolic times, of [2] and [5], and we describe different sets that can be built at hyperbolic times. These sets will be the basic stones used to build the auxiliary partition in Section 3.

Let b be a constant such that $0 < b < \min(1/2, 1/(4\beta))$. For $\sigma < 1$ and $\delta > 0$, we say that n is a (σ,δ)-hyperbolic time for x if, for all $1 \leqslant k \leqslant n$,

$$\prod_{j=n-k}^{n-1} \|DT(T^j x)^{-1}\| \leqslant \sigma^k \text{ and } \operatorname{dist}_\delta(T^{n-k}x, S) \geqslant \sigma^{bk}. \tag{9}$$

We will denote by $H_n = H_n(\sigma,\delta)$ the set of points for which n is a (σ,δ)-hyperbolic time.

In paragraph 2.1, we will choose carefully the constants σ and δ (as well as ε^0 given by Theorems 1.1 and 1.3). However, the reasons for this choice will not become clear before paragraph 3.3, and the reader may admit the existence of σ, δ and ε_0, and come back to paragraph 2.1 just before reading paragraph 3.3.

2.1. Frequency of hyperbolic times. — The following lemma is a slight generalization of [5, Lemma 5.4]:

LEMMA 2.1. — *Take $T : M \to M$ and $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ such that* (1) *and* (5) *are satisfied. Then there exist $\varepsilon_3 > 0$ and $\kappa > 0$ such that, for all $\varepsilon_1, \varepsilon_2 < \kappa$, there exists $\theta(\varepsilon_1,\varepsilon_2) > 0$ such that, if $x \in M$ and $n \in \mathbb{N}^*$ satisfy*

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \|DT(T^k x)^{-1}\|^{-1} \geqslant \frac{\lambda}{2}$$

and for $i = 1, 2, 3$,

$$\frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i,$$

then there exist times $1 \leqslant p_1 < \cdots < p_\ell \leqslant n$ with $\ell \geqslant \theta(\varepsilon_1,\varepsilon_2)n$ such that, for all $j \leqslant \ell$,

$$\forall k,\ 1 \leqslant k \leqslant p_j, \quad \sum_{s=p_j-k}^{p_j-1} \log \|DT(T^s x)^{-1}\|^{-1} \geqslant \frac{\lambda}{4}k \tag{10}$$
$$\text{and for } i = 1, 2, \quad \sum_{s=p_j-k}^{p_j-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^s x, S) \leqslant 2\sqrt{\varepsilon_i}\, k.$$

This means that the density of times p between 1 and n satisfying (10) is at least $\theta(\varepsilon_1,\varepsilon_2)$. Before giving the proof of the lemma, we will state another lemma with the same flavor:

LEMMA 2.2. — *Take* $T : M \to M$ *and* $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ *such that* (1) *and* (5) *are satisfied. Take also* $\kappa > 0$. *Then there exist* $\varepsilon_1, \varepsilon_2 < \kappa$ *and* $\theta > 0$ *such that, if* $x \in M$ *and* $n \in \mathbb{N}^*$ *satisfy*

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \|DT(T^k x)^{-1}\|^{-1} \geqslant \frac{\lambda}{4}$$

and for $i = 1, 2$,

$$\frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\sqrt{\varepsilon_i},$$

then there exist times $1 \leqslant p_1 < \cdots < p_\ell \leqslant n$ *with* $\ell \geqslant \theta n$ *such that, for all* $j \leqslant \ell$,

$$\forall k,\ 1 \leqslant k \leqslant p_j,\ \sum_{s=p_j-k}^{p_j-1} \log \|DT(T^s x)^{-1}\|^{-1} \geqslant \frac{\lambda}{8} k$$

and

$$\sum_{s=p_j-k}^{p_j-1} -\log \operatorname{dist}_{\delta(\varepsilon_1)}(T^s x, S) \leqslant b\frac{\lambda}{8} k.$$

Until the end of this article, we will denote by ε_3^0 the value of ε_3 given by Lemma 2.1, and by $\varepsilon_1^0, \varepsilon_2^0$ the values of ε_1 and ε_2 given by Lemma 2.2. We will also set $\sigma = \mathrm{e}^{-\frac{1}{8}\lambda} < 1$. Finally, write $\delta = \delta(\varepsilon_1^0)$.

Hence, the times p_j given by the conclusion of Lemma 2.2 are (σ, δ)-hyperbolic. In the same way, the times p_j satisfying (10) are also (σ, δ)-hyperbolic (if κ is small enough), but they are more than that since they guarantee a control at the same time for ε_1^0 and for ε_2^0 (whence Lemma 2.2 can be applied to them): we will say that a time which satisfies (10) for ε_1^0 and ε_2^0 is a *super hyperbolic time*. We will write SH_n for the set of points for which n is a super hyperbolic time, and $H_n = H_n(\sigma, \delta)$ for the set of points for which n is a (σ, δ)-hyperbolic time. In particular, $SH_n \subset H_n$.

In the following proof, we will see why an index ε is lost: it is used to obtain the conclusion on $\sum_{s=p_j-k}^{p_j-1} \log \|DT(T^s x)^{-1}\|^{-1}$, since Pliss Lemma can not be applied directly (since this sequence is not bounded), whence another control is needed.

Proof of Lemma 2.1. — The proof is essentially the proof of Lemma 5.4 of [5]: they first show that there exist $\varepsilon_3 > 0$ (which can be taken arbitrarily small) and $\theta_1 > 0$ such that, if

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \|DT(T^k x)^{-1}\|^{-1} \geqslant \frac{\lambda}{2} \quad \text{and} \quad \frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_3)}(T^k x, S) \leqslant 2\varepsilon_3,$$

then there is a proportion at least $\theta_1 > 0$ of times p between 1 and n such that

$$\forall k,\ 1 \leqslant k \leqslant p,\ \sum_{s=p-k}^{p-1} \log \|DT(T^s x)^{-1}\|^{-1} \geqslant \frac{\lambda}{4} k.$$

Moreover, [5, Lemma 3.1] also shows that, for $\varepsilon > 0$, if x satisfies

$$\frac{1}{n} \sum_{k=0}^{n-1} - \log \operatorname{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant 2\varepsilon,$$

then there exists a proportion at least $\theta(\varepsilon) = 1 - \sqrt{\varepsilon}$ of times p between 1 and n such that

$$\forall k,\ 1 \leqslant k \leqslant p,\ \sum_{s=p-k}^{p-1} - \log \operatorname{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant 2\sqrt{\varepsilon}\, k.$$

When $\varepsilon \to 0$, $\theta(\varepsilon) \to 1$. Hence, if κ is small enough, for all $\varepsilon_1, \varepsilon_2 < \kappa$, we will have $\theta(\varepsilon_1, \varepsilon_2) := \theta_1 + \theta(\varepsilon_1) + \theta(\varepsilon_2) - 2 > 0$, which gives the conclusion of the lemma. □

The proof of Lemma 2.2 is similar.

2.2. Constructions at hyperbolic times. — The following lemma refines [5, Lemma 5.2] and [6, Lemma 4.1]:

LEMMA 2.3. — *There exist $\delta_2, D_1, \lambda_1 < 1$ such that, if $x \in M$ and n is a (σ, δ)-hyperbolic time for x, there exists a unique neighborhood $V_n(x)$ of x with the following properties:*

1) *T^n is a diffeomorphism between $V_n(x)$ and the ball $B(T^n x, \delta_2)$.*
2) *For $1 \leqslant k \leqslant n$ and $y, z \in V_n(x)$,*
$$\operatorname{dist}(T^{n-k} y, T^{n-k} z) \leqslant \sigma^{\frac{1}{2}k} \operatorname{dist}(T^n y, T^n z).$$
3) *For all $y, z \in V_n(x)$,*
$$\left| \frac{\det DT^n(y)}{\det DT^n(z)} - 1 \right| \leqslant D_1 \operatorname{dist}(T^n y, T^n z).$$
4) *$V_n(x) \subset B(x, \lambda_1^n)$.*
5) *If $n \leqslant m$, $y \in H_m$ and $V_n(x) \cap V_m(y) \neq \varnothing$, then T^n is injective on $V_n(x) \cup V_m(y)$.*

Note that the third assertion of the lemma implies that the volume-distortion of T^n is bounded by $D_2 := 2\delta_2 D_1 + 1$, i.e., for all $U, V \subset V_n(x)$,

$$D_2^{-1} \frac{\operatorname{Leb}(T^n(U))}{\operatorname{Leb}(T^n(V))} \leqslant \frac{\operatorname{Leb}(U)}{\operatorname{Leb}(V)} \leqslant D_2 \frac{\operatorname{Leb}(T^n(U))}{\operatorname{Leb}(T^n(V))}. \tag{11}$$

Proof. — Lemma 5.2 of [5] shows that there exists $\delta_1 > 0$ such that, if x belongs to $H_n(\sigma, \delta)$, then there exists a neighborhood $V_n'(x)$ mapped diffeomorphically by T^n to $B(T^n x, \delta_1)$. We set $V_n(x) = V_n'(x) \cap T^{-n}(B(T^n x, \frac{1}{4}\delta_1))$, and $\delta_2 = \frac{1}{4}\delta_1$. As $V_n(x) \subset V_n'(x)$, the first and second assertion of the lemma come from Lemma 5.2 of [5], and the third one from Lemma 4.1 of [6]. The fourth one is a consequence of the second one (for $\lambda_1 = \sigma^{\frac{1}{2}}$).

For the uniqueness, note that two distinct neighborhoods $V_n^1(x)$ and $V_n^2(x)$ would give two different lifts by T^n of a path from $T^n(x)$ to a point in $B(T^n(x), \frac{1}{4}\delta_1)$, which is not possible.

Finally, assume that $V_n(x) \cap V_m(y)$ contains a point z. Then

$$\operatorname{diam}\big(T^n\big(V_m(y)\big)\big) \leqslant \operatorname{diam}\big(T^m\big(V_m(y)\big)\big) = \tfrac{1}{2}\delta_1,$$

whence $T^n(V_m(y)) \subset B(T^n x, \delta_1)$. We build a set

$$W_m(y) = T^{-n}\big(T^n\big(V_m(y)\big)\big) \cap V_n'(x).$$

By definition of $V_n'(x)$, T^n is an isomorphism between $W_m(y)$ and $T^n(V_m(y))$. But T^n is also an isomorphism between $V_m(y)$ and $T^n(V_m(y))$. As $V_m(y)$ and $W_m(y)$ both contain z, the previous uniqueness argument implies that $V_m(y) = W_m(y)$. In particular, $V_m(y) \subset V_n'(x)$. As T^n is injective on $V_n'(x)$, it is also injective on $V_n(x) \cup V_m(y)$. □

Take $\mathcal{U} = \{U_1, \dots, U_N\}$ a finite partition of M by sets of diameter at most $\frac{1}{10}\delta_2$, with nonempty interiors and piecewise smooth boundaries (for example a triangulation of M). Hence, there exist constants $C_2 > 0$ and $\lambda_2 < 1$ such that

$$(12) \qquad \forall i,\ 1 \leqslant i \leqslant N,\ \forall n \in \mathbb{N}, \quad \operatorname{Leb}\big\{x \in U_i \mid \operatorname{dist}(x, \partial U_i) \leqslant \lambda_1^n\big\} \leqslant C_2 \lambda_2^n.$$

We will write $U_i' = \{x \in M \mid \operatorname{dist}(x, U_i) \leqslant \delta_2/10\}$. Increasing C_2 and λ_2 if necessary, we can also assume that

$$(13) \qquad \forall i,\ 1 \leqslant i \leqslant N,\ \forall n \in \mathbb{N},$$
$$\operatorname{Leb}\big\{x \in M \mid \operatorname{dist}(x, \partial U_i') \leqslant \tfrac{1}{2}\delta_2 \sigma^{\frac{1}{2}n}\big\} \leqslant C_2 \lambda_2^n \operatorname{Leb}(U_i).$$

We will finally assume that, for any ball $B(x, \delta_2)$ of radius δ_2 and for all $1 \leqslant i \leqslant N$,

$$(14) \qquad \operatorname{Leb} B(x, \delta_2) \leqslant C_2 \operatorname{Leb}(U_i).$$

Take $x \in H_n$. Then $T^n x$ belongs to a unique $U_i =: U(x, n)$, included in $B(T^n x, \delta_2) = T^n(V_n(x))$. We will write $I_\infty^n(x) = T^{-n}(U_i) \cap V_n(x)$. In the construction of the auxiliary partition in Section 3, the partition elements will be such sets $I_\infty^n(x)$. In the construction, if we choose $I_\infty^n(x)$ and then $I_\infty^{n+1}(y)$ while $y \notin I_\infty^n(x)$ but y is very close to the boundary of $I_\infty^n(x)$, the two sets $I_\infty^n(x)$ and $I_\infty^{n+1}(y)$ may have a nonempty intersection, which we want to avoid since we are building a partition. As in [13], we will have to introduce a waiting time telling when it is not dangerous to select y, ensuring that $I^n_\infty(x) \cap I^m_\infty(y) = \varnothing$. We thus set, for $m > n$,

$$I^n_m(x) = \left\{ y \in V_n(x) \mid \tfrac{1}{10}\delta_2 \sigma^{\frac{1}{2}(m-n)} < \operatorname{dist}(T^n y, U(x,n)) \leqslant \tfrac{1}{10}\delta_2 \sigma^{\frac{1}{2}(m-n-1)} \right\}$$

and

$$I^n_{\geqslant m}(x) = \bigcup_{m \leqslant t < \infty} I^n_t(x).$$

These are the points which are not allowed to be selected at time m, because they could interfere with x at time n (this choice will be justified by Lemma 2.5, and (15)). We will say that a point of $I^n_{\geqslant m}(x)$ is forbidden by the time n, at the time m. We will also write

$$\widetilde{I}^n_{\geqslant m}(x) = \bigcup_{m \leqslant t \leqslant \infty} I^n_t(x),$$

i.e. we add the "core" $I^n_\infty(x)$. The main difference with [14] or [6] is that, in these articles, the combinatorial estimates are less precise, whence they can afford to forget the time by which a point is forbidden (the n in $I^n_{\geqslant m}$).

Lemma 2.4. — *If $0 < n \leqslant m$ and $\widetilde{I}^n_{\geqslant n+1}(x) \cap \widetilde{I}^m_{\geqslant m+1}(y) \neq \varnothing$, then*

$$\widetilde{I}^n_{\geqslant n+1}(x) \cup \widetilde{I}^m_{\geqslant m+1}(y) \subset V_n(x).$$

Note that, when we write $\widetilde{I}^n_{\geqslant n+1}(x)$ (for example), it is implicit that this set is well defined, i.e. that $x \in H_n$.

Proof. — Take $z \in \widetilde{I}^n_{\geqslant n+1}(x) \cap \widetilde{I}^m_{\geqslant m+1}(y)$. By Lemma 2.3,

$$T^n\big(\widetilde{I}^m_{\geqslant m+1}(y)\big) \subset B(T^n z, \tfrac{1}{2}\delta_2) \subset B(T^n x, \delta_2).$$

In particular, every $u \in T^n(\widetilde{I}^m_{\geqslant m+1}(y))$ has a preimage u' under T^n in $V_n(x)$. We have to see that u' belongs to $\widetilde{I}^m_{\geqslant m+1}(y)$. Otherwise, u would have another preimage u'' in $\widetilde{I}^m_{\geqslant m+1}(y)$. As $V_n(x) \cap V_m(y)$ contains z, the fifth assertion of Lemma 2.3 gives that T^n is injective on $V_n(x) \cup V_m(y)$. This is a contradiction since $u' \neq u''$ but $T^n(u') = T^n(u'')$. □

Lemma 2.5. — *There exists $P > 0$ such that, for $0 < n < m$, $x \in H_n$ and $y \in H_m \setminus \widetilde{I}^n_{\geqslant m}(x)$,*

$$\widetilde{I}^n_{\geqslant m+P}(x) \cap \widetilde{I}^m_{\geqslant m+P}(y) = \varnothing.$$

This means that, if it not forbidden by x to choose y at time m, then there is no interaction between x and y after time $m + P$. Thus, the waiting time P makes it possible to separate completely the two points (which will be used in Lemma 3.6). In particular,

$$I^n_\infty(x) \cap I^m_\infty(y) = \varnothing, \qquad (15)$$

which implies that the sets we will select in the construction of the auxiliary partition will be disjoint.

Proof. — Set $U_i = T^n(I^n_\infty(x))$. Assume that $\widetilde{I}^n_{\geqslant m+P}(x) \cap \widetilde{I}^m_{\geqslant m+P}(y) \neq \varnothing$, and take a point z in this intersection. Then

$$\operatorname{dist}(T^n z, U_i) \leqslant \tfrac{1}{10}\delta_2 \sigma^{\frac{1}{2}(m+P-n-1)} \text{ and } \operatorname{dist}(T^m z, T^m y) \leqslant \tfrac{1}{10}\delta_2\big(1+\sigma^{\frac{1}{2}(P-1)}\big).$$

Note also that, since $y, z \in V_m(y)$, Lemma 2.3 implies that

$$\operatorname{dist}(T^n y, T^n z) \leqslant \sigma^{\frac{1}{2}(m-n)} \operatorname{dist}(T^m y, T^m z).$$

Hence,

$$\begin{aligned}
\operatorname{dist}(T^n y, U_i) &\leqslant \operatorname{dist}(T^n y, T^n z) + \operatorname{dist}(T^n z, U_i)\\
&\leqslant \sigma^{\frac{1}{2}(m-n)} \operatorname{dist}(T^m y, T^m z) + \operatorname{dist}(T^n z, U_i)\\
&\leqslant \sigma^{\frac{1}{2}(m-n)} \tfrac{1}{10}\delta_2\big(1+\sigma^{\frac{1}{2}(P-1)}\big) + \tfrac{1}{10}\delta_2\sigma^{\frac{1}{2}(m+P-n-1)}\\
&= \tfrac{1}{10}\delta_2\sigma^{\frac{m-n}{2}}\big(1+2\sigma^{\frac{1}{2}(P-1)}\big).
\end{aligned}$$

If P is large enough so that $1+2\sigma^{\frac{1}{2}(P-1)} \leqslant \sigma^{-\frac{1}{2}}$, we get $\operatorname{dist}(T^n y, U_i) \leqslant \frac{1}{10}\delta_2\sigma^{\frac{1}{2}(m-n-1)}$. As $y \in V_n(x)$ by Lemma 2.4, we finally get $y \in \widetilde{I}^n_{\geqslant m}(x)$. □

LEMMA 2.6. — *There exists a positive sequence c_n such that, for all $n \in \mathbb{N}^*$, for every $x \in H_n$, $\operatorname{Leb} I^n_\infty(x) \geqslant c_n$.*

Proof. — The condition $x \in H_n$ implies that, for $k \leqslant n$, $\operatorname{dist}(T^k x, S) \geqslant \alpha_n > 0$, and T is a local diffeomorphism on $M \setminus S$ by definition of S. As T is C^1 on $\{y \mid \operatorname{dist}(y, S) \geqslant \alpha_n\}$, there exists a constant C_n which bounds $\det DT^n(x)$ for $x \in H_n$. Since the volume-distortion is bounded by D_2 on $V_n(x)$, we get that, for any $y \in V_n(x)$, $|\det DT^n(y)| \leqslant D_2 C_n$. In particular, $\operatorname{Leb} I^n_\infty(x) \geqslant \operatorname{Leb}(T^n(I^n_\infty(x)))/(D_2 C_n)$. But $T^n(I^n_\infty(x))$ is one of the U_i, whence its measure is uniformly bounded away from 0. □

LEMMA 2.7. — *There exists a positive constant $C > 0$ such that, for any measurable set A, for any $n \in \mathbb{N}^*$, $\operatorname{Leb}(H_n \cap T^{-n}(A)) \leqslant C \operatorname{Leb}(A)$.*

Proof. — The sets $I^n_\infty(x)$, for $x \in H_n$, cover H_n, and are equal or disjoint. By Lemma 2.6, there is a finite number of them, say $I^n_\infty(x_1), \dots, I^n_\infty(x_k)$ (where k depends on n).

For $1 \leqslant j \leqslant k$, the distortion is bounded by D_2 on $I^n_\infty(x_j)$, whence

$$\frac{\operatorname{Leb}(I^n_\infty(x_j) \cap T^{-n}A)}{\operatorname{Leb}(I^n_\infty(x_j))} \leqslant D_2 \frac{\operatorname{Leb}(A)}{\operatorname{Leb}(T^n(I^n_\infty(x_j)))}.$$

But $T^n(I^n_\infty(x_j))$ is one of the U_i, and its measure is consequently $\geqslant c$ for some positive c. Summing over j, we get

$$\operatorname{Leb}\big(H_n \cap T^{-n}(A)\big) \leqslant \frac{D_2}{c} \operatorname{Leb}(A) \operatorname{Leb}(M). \qquad \square$$

3. The auxiliary partition

In this section, we will show the following result (without any transitivity assumption on T):

Theorem 3.1. — *Let T be a map on a compact manifold M and $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ be such that* (1) *and* (5) *are satisfied. Let ε^0 be given by Lemmas* 2.1 *and* 2.2. *We assume that T satisfies one of the following conditions:*

1) $\mathrm{Leb}\{x \mid h^1_{\varepsilon^0}(x) > n\} = O(u_n)$ *where u_n has polynomial decay and tends to* 0.
2) $\mathrm{Leb}\{x \mid h^1_{\varepsilon^0}(x) > n\} = O(u_n)$ *where* $u_n = \mathrm{e}^{-cn^\eta}$ *with* $\eta \in (0,1]$.
3) $\mathrm{Leb}\{x \mid h^2_{\varepsilon^0}(x) > n\} = O(u_n)$ *where u_n has polynomial decay and* $(\log n)u_n \to 0$.
4) $\mathrm{Leb}\{x \mid h^2_{\varepsilon^0}(x) > n\} = O(u_n)$ *where* $u_n = \mathrm{e}^{-cn^\eta}$ *with* $\eta \in (0,1]$.

Then there exist a finite partition $U_1, \dots, U_N$ of M, another finer partition (modulo a set of zero Lebesgue measure) $W_1, W_2, \dots$ and times $R_1, R_2, \dots$ such that, for all j,

1) T^{R_j} *is a diffeomorphism between W_j and one of the U_i.*
2) $T^{R_j}_{|W_j}$ *expands the distances of at least* $\sigma^{-\frac{1}{2}} > 1$.
3) *The volume-distortion of $T^{R_j}_{|W_j}$ is Lipschitz, i.e. there exists a constant C (independent of j) such that, for every $x, y \in W_j$,*

$$\Big|1 - \frac{\det DT^{R_j}(x)}{\det DT^{R_j}(y)}\Big| \leqslant C \operatorname{dist}(T^{R_j}x, T^{R_j}y).$$

4) *For $x, y \in W_j$ and $n \leqslant R_j$,* $\operatorname{dist}(T^n x, T^n y) \leqslant \operatorname{dist}(T^{R_j}x, T^{R_j}y)$.

Moreover, there exists $c' > 0$ such that, under the different assumptions, the following estimates on the tails hold:

$$\mathrm{Leb}\Big(\bigcup_{R_j > n} W_j\Big) = \begin{cases} O(u_n) & \text{in the first case,} \\ O((\log n)u_n) & \text{in the third case,} \\ O(\mathrm{e}^{-c'n^\eta}) & \text{in the second and fourth cases.} \end{cases}$$

In the proof of the theorem, it will be sufficient to work on U_1, since the same construction can then be made on each U_j.

The fact that the W_j form a partition of M modulo a set of zero Lebesgue measure will come from the estimates on the size of the tails, and is not at all trivial from the construction.

This theorem implies the following result on invariant measures:

Theorem 3.2. — *Under the assumptions of Theorem 3.1, assume moreover that $\sum u_n < \infty$ in the first case, $\sum (\log n) u_n < \infty$ in the third case. Then there exists a finite number of invariant absolutely continuous ergodic probability measures $\mu_1, \dots, \mu_k$. Moreover, their basins cover almost all M. Finally, there* *exist disjoint open subsets* $O_1, \dots, O_k$ *such that* μ_i *is equivalent to* Leb *on* O_i *and vanishes on* $M \setminus O_i$.

In particular, if T is topologically transitive on Λ, there exists a unique absolutely continuous invariant measure.

Proof of Theorem 3.2. — We build an extension of M, similar to a Young tower except that the basis will be constituted of the finite number of sets $U_1, \dots, U_N$. More precisely, set

$$X = \{(x,i) \mid x \in W_j,\ i < R_j\},$$

and let $\pi : X \to M$ be given by $\pi(x,i) = T^i(x)$. We set, for $x \in W_j$,

$$T'(x,i) = \begin{cases} (x, i+1) & \textit{if } i+1 < R_j, \\ (T^{R_j}(x), 0) & \textit{if } i+1 = R_j. \end{cases}$$

Thus, $\pi \circ T' = T \circ \pi$. Let m be the measure on X given by $m(A \times \{i\}) = \mathrm{Leb}(A)$ when $A \subset W_j$ and $i < R_j$, so that $\pi_*(m)$ is equivalent to Lebesgue measure. The condition on the tails ensures that m is of finite mass.

On X, the map T' is Markov, and the map T'_Y induced by T on the basis $Y = \{(x,0)\}$ is Markov with a Lipschitz volume-distortion and the big image property. Classical arguments (see [1, Section 4.7]) show that T'_Y admits a finite number of invariant ergodic absolutely continuous probability measures $\rho_1, \dots, \rho_\ell$. Moreover, each of these measures is equivalent to m on a union Y_j of some sets $U_i \times \{0\}$ (the Y_j are exactly the transitive subsystems for the map T'_Y). Finally, almost every point of Y lands in one of these Y_j after a finite number of iterations of T'_Y. Inducing (see [1, Prop. 1.5.7]), we get a finite number of absolutely continuous invariant ergodic measures $\nu_1, \dots, \nu_\ell$, whose basins cover almost all X. The condition on the measure of the tails ensures that the ν_i are still of finite mass, whence we can assume that they are probability measures.

The measures $\pi_*(\nu_i)$ are not necessarily all different. Let $\mu_1, \dots, \mu_k$ be these measures without repetition. They are ergodic, and their basins cover almost all M, whence there is no other absolutely continuous invariant ergodic measure.

Let $\mu = \pi_*(\nu)$ be one of the measures μ_j. Since ν is equivalent to m on some set $U_i \times \{0\}$, μ is equivalent to Leb on U_i. We will construct the open set $O(\mu)$ of the statement of the theorem. Let Ω_0 be the interior of U_i (it is nonempty by construction). By induction, if Ω_n is defined and open, set $\Omega_{n+1} = T(\Omega_n \setminus S) \cup \Omega_n$. As S is closed and T is a local diffeomorphism outside of S, Ω_{n+1} is still an open set. Set $O = \bigcup \Omega_n$. As μ is invariant, we check by induction that μ is equivalent to Leb on Ω_n, whence on O. Let us show that, if $A \subset M \setminus O$, then $\mu(A) = 0$. Otherwise, by ergodicity, there would exist n

such that $\mu(T^{-n}(A) \cap \Omega_0) > 0$. As $\mu(S) = 0$ (since $\mathrm{Leb}(S) = 0$), we get

$$\mu(T^{-n}(A) \cap (\Omega_0 \setminus S)) > 0,$$

whence $\mu(T^{-(n-1)}(A) \cap \Omega_1) > 0$. By induction, $\mu(A \cap \Omega_n) > 0$, which is a contradiction. □

This result is a first step towards the spectral decomposition of T. It was already known, under weaker assumptions (see in [5] the remark following Corollary D). We will get later a complete spectral decomposition: each measure μ_i has a finite number of components which are mixing (and even exact) for an iterate of T (Theorem 4.3, which also gives the speed of decay of correlations).

3.1. Description of the construction. — To prove Theorem 3.1, we will build a partition of U_1 by sets $W_1, W_2, \dots$ such that, for every n, there exists a return time R_n such that T^{R_n} is an isomorphism between W_n and one of the U_i, expanding of at least $\sigma^{-\frac{1}{2}n}$ and whose volume-distortion is D_1-Lipschitz. In fact, W_n will be some set $I_\infty^{R_n}(x)$. Set

$$H_n(U_1) = H_n \cap \{y \in U_1 \mid \mathrm{dist}(y, \partial U_1) \geqslant \lambda_1^n\}.$$

Hence, if $x \in H_n(U_1)$, we have $V_n(x) \subset U_1$ by the fourth assertion of Lemma 2.3.

We build in fact points $x_1^1, \dots, x_{\ell(1)}^1$ at time 1, and $x_1^2, \dots, x_{\ell(2)}^2$ at time 2, and so on. They will satisfy the following properties:

- $x_1^n, \dots, x_{\ell(n)}^n$ belong to $H_n(U_1) \setminus \bigcup_{i<n,\, j \leqslant \ell(i)} \widetilde{I}^i_{\geqslant n}(x_j^i)$, and this set is covered by $\bigcup_j I_\infty^n(x_j^n)$;
- the sets $I_\infty^n(x_j^n)$ (for $n \in \mathbb{N}^*$ and $1 \leqslant j \leqslant \ell(n)$) are disjoint, and included in U_1.

We will take for W_j the sets $I_\infty^n(x_i^n)$, and the corresponding return time R_j will be n.

Construction of x_i^n. — The construction is by induction on n. At time n, note that, if $x, y \in H_n(U_1)$, then $I_\infty^n(x)$ and $I_\infty^n(y)$ are either disjoint or equal. Hence, there exists a system $I_\infty^n(x_1^n), \dots, I_\infty^n(x_{\ell(n)}^n)$ of representatives of the sets $I_\infty^n(x)$ for $x \in H_n(U_1) \setminus \bigcup_{i<n,\, j \leqslant \ell(i)} \widetilde{I}^i_{\geqslant n}(x_j^i)$ (and it is finite by Lemma 2.6).

By construction, two sets $I_\infty^n(x_i^n)$ constructed at the same time are disjoint. Take $m > n$, and $x_k^m \in H_m(U_1) \setminus \bigcup_{i<m,\, j \leqslant \ell(i)} \widetilde{I}^i_{\geqslant m}(x_j^i)$. Then x_k^m belongs to $H_m \setminus \widetilde{I}^n_{\geqslant m}(x_i^n)$, whence Lemma 2.5 ensures that $I_\infty^m(x_k^m)$ is disjoint from $I_\infty^n(x_i^n)$.

Finally, to see that $I_\infty^n(x_i^n) \subset U_1$, we use the fact that $x_i^n \in H_n(U_1)$, whence $\mathrm{dist}(x_i^n, \partial U_1) \geqslant \lambda_1^n$. As $V_n(x_i^n) \subset B(x_i^n, \lambda_1^n)$, this implies that $I_\infty^n(x_i^n) \subset U_1$. □

The properties of hyperbolic times given in Lemma 2.3 imply that the expansion and distortion requirements of Theorem 3.1 are satisfied. It only remains to estimate $\mathrm{Leb}\{x \mid \exists j,\ x \in W_j \text{ and } R_j > n\}$.

3.2. Measure of points which are forbidden many times. — We will denote by I_n the set of points which are forbidden at the instant n, i.e.

$$I_n = \bigcup_{\substack{i<n \\ j\leqslant \ell(i)}} \widetilde{I}^i_{\geqslant n}(x^i_j),$$

and I^n the set of points which are forbidden by the instant n, i.e.

$$I^n = \bigcup_{j\leqslant \ell(n)} \widetilde{I}^n_{\geqslant n+1}(x^n_j).$$

In particular, $I^n \subset I_{n+1}$. Finally, set

$$S_n = \bigcup_{\substack{i\leqslant n \\ j\leqslant \ell(i)}} I^i_\infty(x^i_j). \tag{16}$$

This is the set of points which are selected before the instant n. In this paragraph, the word "time" will be used only for durations, and "instant" will be used otherwise.

In this paragraph, we will prove Lemma 3.7, which says that the set of points which are forbidden at k instants without being selected has a measure which decays exponentially fast. The argument is combinatorial: if a point is forbidden by few instants, then it will be forbidden for a long time at many of these instants, and it is easily seen that this gives a small measure (Lemma 3.6). Otherwise, the point is forbidden by many instants, and we have to see that each of these instants enables us to gain a multiplicative factor $\lambda < 1$. We will treat two cases: either the forbidden sets are included one in each other, whence only a proportion < 1 is kept at each step, which concludes (Lemma 3.4), or the forbidden sets intersect each other close to their respective boundaries, and we just have to ensure that these boundaries are small enough (Lemma 3.3).

We will write B for a set $\widetilde{I}^n_{\geqslant n+1}(x^n_i)$, i.e. a "forbidden ball" (where x^n_i is one of the points defined in the construction of paragraph 3.1). Then $t(B)$ will denote the instant n by which it is forbidden, while the "core" $C(B) = I^n_\infty(x^n_i)$ is the inner part of B, corresponding to points which are really selected. If $T^{t(B)}(C(B)) = U_i$, then $T^{t(B)}(B) = \{x \mid \mathrm{dist}(x, U_i) \leqslant \frac{1}{10}\delta_2\}$, whence $\mathrm{diam}\, T^{t(B)}(B) \leqslant \frac{3}{10}\delta_2 \leqslant \frac{1}{2}\delta_2$. In all the statements and proofs of this paragraph, the sets denoted by B_i or B'_i will implicitly be such forbidden balls. We will define in the following lemmas sets $Z^1, \dots, Z^6$ of "points which are forbidden at many instants", and we will see that each of them has an exponentially small measure.

LEMMA 3.3. — *Let $Q \in \mathbb{N}^*$. Set*

$$Z^1(k,B_0) = \Big\{x \mid \exists B_1', B_1, \dots, B_r', B_r \text{ with } \forall i,\ 1 \leqslant i \leqslant r,\ t(B_{i-1}) \leqslant t(B_i') \leqslant t(B_i) - Q,\ B_i \not\subset B_i',\ \sum_{i=1}^{r} \Big\lfloor \frac{t(B_i)-t(B_i')}{Q} \Big\rfloor \geqslant k, \text{ and } x \in \bigcap_{i=0}^{r} B_i \cap \bigcap_{i=1}^{r} B_i' \Big\}.$$

Then there exists a constant C_3 (independent of Q) such that for all k and B_0,

$$\operatorname{Leb}\big(Z^1(k,B_0)\big) \leqslant C_3 (C_3 \lambda_2^Q)^k \operatorname{Leb}\big(C(B_0)\big).$$

Recall that λ_2 is a constant satisfying (12) and (13).

Proof. — Let C_3 be such that, for $1 \leqslant i \leqslant N$,

$$\operatorname{Leb}\big\{x \mid \operatorname{dist}(x,U_i) \leqslant \tfrac{1}{10}\delta_2\big\} \leqslant \frac{C_3}{D_2} \operatorname{Leb}(U_i),$$

and such that $C_3^{-1}/(1-C_3^{-1})(C_2D_2)^2 \leqslant 1$. We will prove that C_3 satisfies the assertion of the lemma, by induction on k.

Take $k=0$. Let $n = t(B_0)$, and i be such that $T^n(C(B_0)) = U_i$. Then $Z^1(0,B_0) = B_0$, whence $T^n(Z^1(0,B_0)) = \{x \mid \operatorname{dist}(x,U_i) \leqslant \frac{1}{10}\delta_2\}$. This gives $\operatorname{Leb}(T^n(Z^1(0,B_0))) \leqslant C_3/D_2 \operatorname{Leb}(T^n(C(B_0)))$. As the distortion of T^n is bounded by D_2, by (11), we get $\operatorname{Leb}(Z^1(0,B_0)) \leqslant C_3 \operatorname{Leb}(C(B_0))$.

Take now $k \geqslant 1$. Then, decomposing according to the value of B_1', we get

$$Z^1(k,B_0) \subset \bigcup_{t=1}^{k} \bigcup_{B_1' \cap B_0 \neq \varnothing} \bigcup_{\substack{B_1 \cap B_1' \neq \varnothing,\ B_1 \not\subset B_1' \\ \lfloor (t(B_1)-t(B_1'))/Q \rfloor \geqslant t}} Z^1(k-t, B_1).$$

Let us show that, if $t(B_1) - t(B_1') = n$, then B_1 is included in an annulus of size $\sigma^{\frac{1}{2}n}$ around B_1'. More precisely, set $p = t(B_1')$, $U_i' = T^p(B_1')$, and let us show that

$$T^p(B_1) \subset \big\{y \mid \operatorname{dist}(y, \partial U_i') \leqslant \tfrac{1}{2}\delta_2 \sigma^{\frac{1}{2}n}\big\}. \tag{17}$$

Note that B_1 contains a point of $\partial B_1'$, since it is connected and intersects B_1' and its complement. Thus, $T^p(B_1)$ contains a point of $\partial U_i'$. Moreover,

$$\operatorname{diam} T^p(B_1) \leqslant \sigma^{\frac{1}{2}n} \operatorname{diam} T^{n+p}(B_1) \leqslant \tfrac{1}{2}\delta_2 \sigma^{\frac{1}{2}n}.$$

This shows (17). Note that (13) gives an upper bound for the measure of (17).

Since the distortion is bounded by D_2 at hyperbolic times, and the cores $C(B_1)$ are disjoint by construction, we get by (17) and (13) that

$$\sum_{\substack{B_1 \cap B_1' \neq \varnothing,\ B_1 \not\subset B_1' \\ \lfloor (t(B_1)-t(B_1'))/Q \rfloor \geqslant t}} \operatorname{Leb}\big(C(B_1)\big) \leqslant C_2 \lambda_2^{Qt} D_2 \operatorname{Leb}\big(C(B_1')\big). \tag{18}$$

Finally, write $q = t(B_0)$. Let x be such that $C(B_0) = I^q_\infty(x)$. The sets $C(B'_1)$ are pairwise disjoint by construction, and included in $V_q(x)$ by Lemma 2.4. Moreover, T^q is a diffeomorphism on $V_q(x)$ and its distortion is bounded by D_2. Since $T^q(C(B_0))$ is a set U_i and $T^q(V_q(x)) = B(T^q x, \delta_2)$, we have $\operatorname{Leb}(T^q(V_q(x))) \leqslant C_2 \operatorname{Leb}(T^q(C(B_0)))$ by (14). By bounded distortion, we obtain

$$\sum_{B'_1\cap B_0\neq\varnothing} \operatorname{Leb}\bigl(C(B'_1)\bigr) \leqslant C_2 D_2 \operatorname{Leb}\bigl(C(B_0)\bigr). \tag{19}$$

Using the induction assumption, we finally obtain

$$\begin{aligned}
\operatorname{Leb} Z^1(k, B_0) &\leqslant \sum_{t=1}^{k} \sum_{B'_1\cap B_0\neq\varnothing} \sum_{\substack{B_1\cap B'_1\neq\varnothing,\, B_1\not\subset B'_1\\ \lfloor (t(B_1)-t(B'_1))/Q\rfloor\geqslant t}} \operatorname{Leb} Z^1(k-t, B_1)\\
&\leqslant \sum_{t=1}^{k} \sum_{B'_1\cap B_0\neq\varnothing} \sum_{\substack{B_1\cap B'_1\neq\varnothing,\, B_1\not\subset B'_1\\ \lfloor (t(B_1)-t(B'_1))/Q\rfloor\geqslant t}} C_3 (C_3\lambda_2^Q)^{k-t} \operatorname{Leb}\bigl(C(B_1)\bigr)\\
&\leqslant \sum_{t=1}^{k} \sum_{B'_1\cap B_0\neq\varnothing} C_3 (C_3\lambda_2^Q)^{k-t} C_2 \lambda_2^{Qt} D_2 \operatorname{Leb}\bigl(C(B'_1)\bigr)\\
&\leqslant C_3 \lambda_2^{Qk} C_3^k (C_2D_2)^2 \Bigl(\sum_{t=1}^{k} C_3^{-t}\Bigr) \operatorname{Leb}\bigl(C(B_0)\bigr).
\end{aligned}$$

By definition of C_3, we have $(C_2D_2)^2\bigl(\sum_{t=1}^k C_3^{-t}\bigr) \leqslant 1$. This concludes the induction. ☐

Lemma 3.4. — *Set*

$$Z^2_{k,N} = \bigl\{x \mid \exists\, B_1 \supsetneq B_2 \supsetneq \cdots \supsetneq B_k \text{ with } t(B_k)\leqslant N \text{ and } x\in (B_1\cap\cdots\cap B_k)\setminus S_N\bigr\}.$$

Then there exists a constant $\lambda_3 < 1$ *such that* $\operatorname{Leb}(Z^2_{k,N}) \leqslant \lambda_3^k \operatorname{Leb}(M)$.

Proof. — We fix N once and for all in this proof, and we will omit all indexes N. We will show that $\lambda_3 = C_2D_2/(C_2D_2+1)$ satisfies the conclusion of the lemma. Note that, for every B,

$$\operatorname{Leb}(B) \leqslant C_2 D_2 \operatorname{Leb}\bigl(C(B)\bigr) \tag{20}$$

by (14) and the bounded distortion of hyperbolic times.

We will write $\mathcal{B}_1$ for the sets of balls B with $t(B)\leqslant N$ which are not included in any other ball B'. Write also $\mathcal{B}_2$ for the set of balls $B\notin\mathcal{B}_1$ with $t(B)\leqslant N$ which are included only in balls of $\mathcal{B}_1$, and so on. We will say that a ball of $\mathcal{B}_i$

has rank i. Every ball B has finite rank, since a ball which is constructed at time n has at most rank n.

Set $S'_k = \bigcup_{i=1}^k \bigcup_{B\in\mathcal{B}_i} C(B)$: these are the points which are selected in balls of rank at most k. Set

$$Z_k^3 = \Big(\bigcup_{B\in\mathcal{B}_k} B\Big) \setminus S'_k.$$

Let us show that $Z_k^2 \subset Z_k^3$.

Take $x \in Z_k^2$, it is in a set $(B_1\cap\dots\cap B_k)\setminus S_N$ with $B_1 \supsetneq B_2 \supsetneq \dots \supsetneq B_k$ and $t(B_k) \leqslant N$. In particular, B_k is of rank $r \geqslant k$. Take $B'_1 \supsetneq B'_2 \supsetneq \dots \supsetneq B'_{r-1} \supsetneq B'_r$ a sequence with $B'_i \in \mathcal{B}_i$ and $B'_r = B_k$. In particular, $x \in B'_k$. Moreover, $S'_k \subset S_N$. As $x \notin S_N$, we get $x \notin S'_k$. This shows that $x \in Z_k^3$.

Let us estimate $\mathrm{Leb}(Z_{k+1}^3)$ using $\mathrm{Leb}(Z_k^3)$. Consider $B_{k+1} \in \mathcal{B}_{k+1}$. Let B_k be a ball of rank k containing B_{k+1}. As the cores of different balls are disjoint, $C(B_{k+1})\cap S'_k = \varnothing$. Thus, $C(B_{k+1}) \subset B_k \setminus S'_k \subset Z_k^3$. However, $C(B_{k+1}) \subset S'_{k+1}$ by definition, whence $C(B_{k+1}) \cap Z_{k+1}^3 = \varnothing$. This shows that $C(B_{k+1}) \subset Z_k^3 \setminus Z_{k+1}^3$.

Finally, by (20),

$$\begin{aligned}\mathrm{Leb}(Z_{k+1}^3) \leqslant \sum_{B_{k+1}\in\mathcal{B}_{k+1}} \mathrm{Leb}(B_{k+1}) &\leqslant C_2D_2 \sum_{B_{k+1}\in\mathcal{B}_{k+1}} \mathrm{Leb}\big(C(B_{k+1})\big)\\ &\leqslant C_2D_2 \,\mathrm{Leb}(Z_k^3 \setminus Z_{k+1}^3)\end{aligned}$$

since the $C(B_{k+1})$ are disjoint. Hence,

$$\begin{aligned}(C_2D_2+1)\,\mathrm{Leb}(Z_{k+1}^3) &\leqslant C_2D_2\,\mathrm{Leb}(Z_{k+1}^3) + C_2D_2\,\mathrm{Leb}(Z_k^3\setminus Z_{k+1}^3)\\ &= C_2D_2\,\mathrm{Leb}(Z_k^3).\end{aligned}$$

We obtain by induction that $\mathrm{Leb}(Z_k^3) \leqslant (C_2D_2/(C_2D_2+1))^k\,\mathrm{Leb}(M)$, which gives the same inequality for $\mathrm{Leb}(Z_k^2)$ since $Z_k^2 \subset Z_k^3$. □

Lemma 3.5. — *Set*

$$Z^4(k,N) = \big\{x \mid \exists\, t_1 < \dots < t_k \leqslant N,\ x \in I^{t_1}\cap\dots\cap I^{t_k}\big\}\setminus S_N.$$

There exist constants $C_4 > 0$ and $\lambda_4 < 1$ such that, for all $1 \leqslant k \leqslant N$,

$$\mathrm{Leb}\big(Z^4(k,N)\big) \leqslant C_4\lambda_4^k.$$

This lemma means that the points forbidden by at least k instants have an exponentially small measure.

Proof. — Take Q large enough so that $C_3\lambda_2^Q < 1$ in Lemma 3.3. Write $N = rQ + s$ with $s < Q$.

Let $x \in Z^4(k,N)$, forbidden by the instants $t_1 < \dots < t_k$. For $0 \leqslant u < r$, we choose in each interval $[uQ, (u+1)Q)$ the first instant t_i (if there exists one), which gives a sequence $t'_1 < \dots < t'_{k'}$, with $Qk' + s \geqslant k$. Then we keep the instants with an odd index, which gives a sequence of instants $u_1 < \dots < u_\ell$

with $2\ell \geqslant k'$, whence $\ell \geqslant k/(2Q) - s$. Moreover, $u_{i+1} - u_i \geqslant Q$ for all i. Let $B_1, \dots, B_\ell$ be balls constructed at the instants u_i and forbidding x.

Set $I = \{1 \leqslant i \leqslant \ell,\ B_i \subset B_1 \cap \dots \cap B_{i-1}\}$ and $J = [1, \ell] \setminus I$. If $\operatorname{Card} I \geqslant \frac{1}{2}\ell$, we keep only the balls whose indexes are in I. Since there are at least $\frac{1}{2}\ell$ such balls, $x \in Z^2_{\frac{1}{2}\ell, N}$ (where Z^2 is defined in Lemma 3.4). This lemma implies that the points obtained in this way have an exponentially small measure (in ℓ, whence in k).

Otherwise, $\operatorname{Card} J \geqslant \frac{1}{2}\ell$. Let $j_0 = \sup J$, and $i_0 = \inf\{i < j_0,\ B_{j_0} \not\subset B_i\}$. Let $j_1 = \sup\{j \leqslant i_0,\ j \in J\}$, and $i_1 = \inf\{i < j_1,\ B_{j_1} \not\subset B_i\}$, and so on: the construction stops at some step, say i_n. Then $J \subset \bigcup (i_s, j_s]$ by construction, whence $\sum (j_s - i_s) \geqslant \operatorname{Card} J \geqslant \frac{1}{2}\ell$, which implies that

$$\sum \Big\lfloor \frac{t(B_{j_s}) - t(B_{i_s})}{Q} \Big\rfloor = \sum \Big\lfloor \frac{u_{j_s} - u_{i_s}}{Q} \Big\rfloor \geqslant \frac{1}{2}\ell,$$

since two instants u_j and u_i are separated by at least $Q(j-i)$ by construction. Hence, the sequence $B_{i_n}, B_{i_n}, B_{j_n}, \dots, B_{i_0}, B_{j_0}$ shows that $x \in Z^1(\frac{1}{2}\ell, B_{i_n})$. Summing the estimates given by Lemma we also get an exponentially small measure (since the cores are disjoint). □

Lemma 3.6. — *For a ball $B_1 = \widetilde{I}^{t_1}_{\geqslant t_1+1}(x_1)$, set*

$$Z^5(n_1, \dots, n_k, B_1) = \{x \mid \exists\, t_2, \dots, t_k \text{ with } t_1 < \dots < t_k \text{ and } x_2, \dots, x_k \text{ such that } \forall i,\ 1 \leqslant i \leqslant k,\ x \in I^{t_i}_{\geqslant t_i + n_i}(x_i)\}.$$

There exists a constant C_5 (independent of $B_1, n_1, \dots, n_k$) such that, when $n_1, \dots, n_k > P$ (given by Lemma 2.5),

$$\operatorname{Leb}\big(Z^5(n_1, \dots, n_k, B_1)\big) \leqslant C_5 (C_5 \lambda_2^{n_1}) \cdots (C_5 \lambda_2^{n_k}) \operatorname{Leb}\big(C(B_1)\big).$$

In fact, $Z^5(n_1, \dots, n_k, B_1)$ is the set of points which are forbidden for a time at least n_1 by B_1, and then for a time at least n_2 by another ball B_2, and so on.

Proof. — The proof is by induction on k.

Let $x \in Z^5(n_1, \dots, n_k, B_1)$. There exists by definition a ball $B_2 = \widetilde{I}^{t_2}_{\geqslant t_2+1}(x_2)$, constructed at an instant $t_2 > t_1$, such that $x \in Z^5(n_2, \dots, n_k, B_2)$. The point x_2 is not forbidden at the instant t_2 (otherwise, x_2 could not be selected at the instant t_2 according to the construction of paragraph 3.1). Hence, Lemma 2.5 yields that $\widetilde{I}^{t_1}_{\geqslant t_2+P}(x_1) \cap \widetilde{I}^{t_2}_{\geqslant t_2+P}(x_2) = \varnothing$. But x is forbidden by the instant t_2 for a time at least $n_2 > P$, whence $x \in \widetilde{I}^{t_2}_{\geqslant t_2+P}(x_2)$. Thus, $x \notin \widetilde{I}^{t_1}_{\geqslant t_2+P}(x_1)$. As $x \in \widetilde{I}^{t_1}_{\geqslant t_1+n_1}(x_1)$, we get $t_1 + n_1 < t_2 + P$, i.e. $t_2 - t_1 > n_1 - P$.

Set $U_i = T^{t_1}(C(B_1))$. The expansion at hyperbolic times gives

$$\operatorname{diam}\big(T^{t_1}(B_2)\big) \leqslant \sigma^{\frac{1}{2}(t_2 - t_1)} \operatorname{diam}\big(T^{t_2}(B_2)\big) \leqslant \sigma^{\frac{1}{2}(n_1 - P)} \tfrac{1}{2}\delta_2.$$

As $\operatorname{dist}(T^{t_1}(x), \partial U_i) \leqslant \frac{1}{10}\delta_2\sigma^{\frac{1}{2}(n_1-1)}$ since x if forbidden for a time at least n_1, we have proved that there exists a constant C_6 such that

$$T^{t_1}(B_2) \subset \mathcal{C} := \big\{y \mid \operatorname{dist}(y, \partial U_i) \leqslant C_6\sigma^{\frac{1}{2}n_1}\big\}.$$

By the induction hypothesis,

$$\operatorname{Leb}\big(Z^5(n_2, \dots, n_k, B_2)\big) \leqslant C_5(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \operatorname{Leb} C(B_2).$$

As the distortion is bounded, we get

$$\operatorname{Leb}\big(T^{t_1}\big(Z^5(n_2, \dots, n_k, B_2)\big)\big) \leqslant D_2C_5(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \operatorname{Leb}\big(T^{t_1}\big(C(B_2)\big)\big).$$

The sets $C(B_2)$ are disjoint by construction and included in $V_{t_1}(x_1)$ by Lemma 2.4. Since T^{t_1} is injective on $V_{t_1}(x_1)$ by Lemma 2.3, the sets $T^{t_1}(C(B_2))$ are still pairwise disjoint. Moreover, they are all included in the annulus $\mathcal{C}$. Hence,

$$\begin{aligned}\operatorname{Leb}\big(T^{t_1}\big(Z^5(n_1, \dots, n_k, B_1)\big)\big) &\leqslant \sum_{B_2} \operatorname{Leb}\big(T^{t_1}\big(Z^5(n_2, \dots, n_k, B_2)\big)\big)\\ &\leqslant C_5D_2(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \sum_{B_2} \operatorname{Leb}\big(T^{t_1}\big(C(B_2)\big)\big)\\ &\leqslant C_5D_2(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \operatorname{Leb}(\mathcal{C}).\end{aligned}$$

By (13), there exists C_7 such that $\operatorname{Leb}(\mathcal{C}) \leqslant C_7\lambda_2^{n_1} \operatorname{Leb}(U_i)$. Hence,

$$\operatorname{Leb}\big(T^{t_1}\big(Z^5(n_1, \dots, n_k, B_1)\big)\big) \leqslant C_5C_7D_2\lambda_2^{n_1}(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \operatorname{Leb}(U_i).$$

The distortion of the map T^{t_1} is bounded by D_2 on B_1. Since $U_i = T^{t_1}(C(B_1))$, the previous equation implies

$$\operatorname{Leb}\big(Z^5(n_1, \dots, n_k, B_1)\big) \leqslant C_5C_7D_2^2\lambda_2^{n_1}(C_5\lambda_2^{n_2}) \cdots (C_5\lambda_2^{n_k}) \operatorname{Leb}\big(C(B_1)\big).$$

This concludes the proof, if $C_5 \geqslant C_7D_2^2$ is taken large enough so that the result holds for $k = 0$. □

The following lemma will subsume all the previous lemmas: it shows that the points forbidden at k instants have an exponentially small measure.

Lemma 3.7. — *Set*

$$Z^6(k, N) = \big\{x \mid \exists\, t_1 < \cdots < t_k \leqslant N,\ x \in I_{t_1} \cap \cdots \cap I_{t_k}\big\} \setminus S_N.$$

There exist constants $C_8 > 0$ and $\lambda_5 < 1$ such that, for all $k \leqslant N$,

$$\operatorname{Leb}\big(Z^6(k, N)\big) \leqslant C_8\lambda_5^k.$$

Proof. — Take $R > P$ (given by Lemma 2.5) so that $\lambda_2 + C_5\lambda_2^R < 1$. Let $x \in Z^6(k, N)$, and consider all the instants u_i by which it is forbidden for a time $n_i \geqslant R$, ordered so that $u_1 < \cdots < u_p$. Then $x \in Z^5(n_1, \dots, n_p, B_1)$ for some ball B_1. If $\sum n_i \geqslant \frac{1}{2}k$, we do not do anything else. Otherwise, let $v_1 < \cdots < v_q$ be the other instants by which x is forbidden, for times $m_1, \dots, m_q < R$. Then

$\sum n_i + \sum m_j$ is not less than the number of instants at which x is forbidden, whence $\sum m_j \geqslant \frac{1}{2}k$. This implies that $Rq \geqslant \frac{1}{2}k$. We obtain

$$Z^6(k,N) \subset \Big(\bigcup_{B_1} \bigcup_{\substack{n_1,\dots,n_p \geqslant R \\ \sum n_i \geqslant \frac{1}{2}k}} Z^5(n_1,\dots,n_p,B_1)\Big) \cup Z^4\Big(\frac{k}{2R},N\Big).$$

Consequently, Lemmas 3.5 and 3.6 yield that

$$\operatorname{Leb}\big(Z^6(k,N)\big) \leqslant \sum_{B_1} \sum_{\substack{n_1,\dots,n_p \geqslant R \\ \sum n_i \geqslant \frac{1}{2}k}} C_5(C_5\lambda_2^{n_1})\cdots(C_5\lambda_2^{n_p}) \operatorname{Leb}\big(C(B_1)\big) + C_4\lambda_4^{k/(2R)}.$$

As the cores $C(B_1)$ are disjoint, $\sum \operatorname{Leb}(C(B_1)) \leqslant \operatorname{Leb}(M) < \infty$. To conclude, it is therefore sufficient to prove that

$$\sum_{\substack{n_1,\dots,n_p \geqslant R \\ \sum n_i \geqslant \frac{1}{2}k}} (C_5\lambda_2^{n_1})\cdots(C_5\lambda_2^{n_p})$$

decays exponentially fast.

We use generating series:

$$\sum_n \sum_{\substack{n_1,\dots,n_p \geqslant R \\ \sum n_i = n}} (C_5\lambda_2^{n_1})\cdots(C_5\lambda_2^{n_p}) z^n = \sum_{p=1}^{\infty}\Big(C_5 \sum_{n=R}^{\infty} \lambda_2^n z^n\Big)^p = \frac{C_5\lambda_2^R z^R}{1-\lambda_2 z - C_5\lambda_2^R z^R}.$$

As $\lambda_2 + C_5\lambda_2^R < 1$, this function has no pole in a neighborhood of the unit disk in $\mathbb{C}$. Hence, its coefficients decay exponentially fast, i.e. there exist constants $C_9 > 0$ and $\lambda_6 < 1$ such that

$$\sum_{\substack{n_1,\dots,n_p \geqslant R \\ \sum n_i = n}} (C_5\lambda_2^{n_1})\cdots(C_5\lambda_2^{n_p}) \leqslant C_9\lambda_6^n.$$

We just have to sum over $n \geqslant \frac{1}{2}k$ to conclude. □

3.3. Proof of Theorem 3.1. — We check in the four cases of Theorem 3.1 that the conclusions on the measures of the tails hold. In this proof, the precise choice of σ, δ and ε^0 in paragraph 2.1 is important. From the previous paragraph, we will only use Lemma 3.7.

Proof of the first and second cases. — Recall that $\operatorname{Leb}\{x \mid h^1_{\varepsilon^0}(x) > n\} = O(u_n)$. Recall also that S_n is the set of points selected before time n, and that θ is defined in Lemma 2.2. Let us show that

$$U_1 \setminus S_n \subset \big\{x \in U_1 \mid h^1_{\varepsilon^0}(x) > n\big\} \cup \big\{x \in U_1 \mid \operatorname{dist}(x, \partial U_1) \leqslant \lambda_1^{\frac{1}{2}\theta n}\big\} \cup Z^6(\tfrac{1}{2}\theta n, n).$$

This will conclude the proof, since the second and third sets have an exponentially small measure, by (12) and Lemma 3.7.

Take x in $U_1 \setminus S_n$, which does not belong either to $\{h^1_{\varepsilon^0}(x) > n\}$ or to $\{\mathrm{dist}(x, \partial U_1) \leqslant \lambda_1^{\frac{1}{2}\theta n}\}$. By Lemma 2.2, x has at least θn hyperbolic times between 1 and n, whence at least $\frac{1}{2}\theta n$ between $\frac{1}{2}\theta n$ and n. We will denote them by $t_1 < \cdots < t_k \leqslant n$. As $\mathrm{dist}(x, \partial U_1) > \lambda_1^{\frac{1}{2}\theta n}$, we have in fact $x \in H_{t_i}(U_1)$ for all these instants. If x was not forbidden at the instant t_i, then it would be selected at the instant t_i by construction, which is not possible since $x \notin S_n$. Hence, $x \in I_{t_i}$. We obtain in this way at least $\theta n/2$ instants at which x is forbidden, whence $x \in Z^6(\frac{1}{2}\theta n, n)$. □

Proof of the third and fourth case. — Denote by $N(x, n)$ the number of hyperbolic times of x between 1 and n.

Lemma 3.8. — *Let $n \in \mathbb{N}^*$ and $k(n) \in [1, \theta n]$. Then*

$$\mathrm{Leb}\big\{x \mid N(x,n) < k(n)\big\} \leqslant Ck(n)/\theta\, \mathrm{Leb}\big\{x \mid h^2_{\varepsilon^0}(x) > n - k(n)/\theta\big\}.$$

Proof. — Write SH^*_ℓ for the set of points whose first positive super hyperbolic time is ℓ. If a point x has a super hyperbolic time j between $k(n)/\theta$ and n, then it will have at least $\theta j \geqslant k(n)$ hyperbolic times between 1 and j, by Lemma 2.2. Hence,

$$\big\{x \mid N(x,n) < k(n)\big\} \subset M \setminus \bigcup_{k(n)/\theta \leqslant j \leqslant n} SH_j.$$

Denote by $k \in [0, k(n)/\theta)$ the last super hyperbolic time of x before $k(n)/\theta$. We get

$$\begin{aligned}\mathrm{Leb}\big(M \setminus \bigcup_{k(n)/\theta \leqslant j \leqslant n} SH_j\big) &\leqslant \sum_{k=0}^{k(n)/\theta} \mathrm{Leb}\big(SH_k \cap T^{-k}\big(\bigcup_{\ell > n-k} SH^*_\ell\big)\big)\\ &\leqslant C \sum_{k=0}^{k(n)/\theta} \mathrm{Leb}\big(\bigcup_{\ell > n-k} SH^*_\ell\big),\end{aligned}$$

using the inclusion $SH_k \subset H_k$ and Lemma 2.7 for the last inequality.

By Lemma 2.1, a point x has at least one super hyperbolic time between 1 and $h^2_{\varepsilon^0}(x)$, whence $\bigcup_{\ell > n-k} SH^*_\ell \subset \{x \mid h^2_{\varepsilon^0}(x) > n - k\}$. This concludes the proof of the lemma. □

For any $k(n)$, the same arguments as in the proof of the first and second cases imply that

$$U_1 \setminus S_n \subset \big\{x \mid N(x,n) < k(n)\big\} \cup \big\{x \mid \mathrm{dist}(x, \partial U_1) \leqslant \lambda_1^{\frac{1}{2}k(n)}\big\} \cup Z^6\big(\tfrac{1}{2}k(n), n\big).$$

By (12), Lemma 3.7 and Lemma 3.8, we get

$$
\begin{aligned}
(21)\qquad & \operatorname{Leb}(U_1 \setminus S_n) \\
& \leqslant C\frac{k(n)}{\theta} \operatorname{Leb}\left\{ h^2_{\varepsilon^0}(x) > n - \frac{k(n)}{\theta} \right\} + C_2 \lambda_2^{\frac{1}{2}k(n)} + C_8 \lambda_5^{\frac{1}{2}k(n)}.
\end{aligned}
$$

To conclude the proof, we just have to choose correctly the sequence $k(n)$.

Assume that $\operatorname{Leb}\{x \mid h^2_{\varepsilon^0}(x) > n\} = O(u_n)$ where u_n has polynomial decay. Choose K large enough so that $k(n) := \lfloor K \log n \rfloor$ satisfies $\lambda_5^{\frac{1}{2}k(n)} = O(u_n)$ and $\lambda_2^{\frac{1}{2}k(n)} = O(u_n)$. Then (21) gives

$$
\operatorname{Leb}(U_1 \setminus S_n) = O\big((\log n) u_{n-k(n)/\theta}\big) = O\big((\log n) u_n\big).
$$

Assume finally that $\operatorname{Leb}\{x \mid h^2_{\varepsilon^0}(x) > n\} = O(\mathrm{e}^{-cn^\eta})$ with $\eta \in (0,1]$. Choose $k(n) = \lfloor n^\eta \rfloor$ if $\eta < 1$, and $k(n) = \left\lfloor \frac{1}{2}\theta n \right\rfloor$ if $\eta = 1$. Then (21) gives $\operatorname{Leb}(U_1 \setminus S_n) = O(\mathrm{e}^{-c' n^\eta})$ for some $c' > 0$. □

The logarithmic loss in the polynomial case comes from the factor $k(n)$ in Lemma 3.8.

4. The Young tower

Using Theorem 3.1, it is possible to prove directly the estimates on the decay of correlations (under a mixing assumption): the coupling arguments of [14] apply to the "tower" built from the partition W_j (the only difference with the towers of [14] is that the returns to the basis do not cover the whole basis, but only one of the sets U_i). This is for example shown in [11]. However, in view of the existing literature, it seems more economical to build a true Young tower, in order to apply directly the results of [14] (or rather a small improvement of these results, since the results of Young are not sharp enough in the stretched exponential case).

4.1. Construction of the Young tower. — The Young tower is given by the following theorem:

Theorem 4.1. — *Under the assumptions of Theorem* 3.2*, let* μ *be one of the invariant absolutely continuous ergodic probability measures given by this theorem. Then there exist a nonempty open set* B *on which* μ *is equivalent to Lebesgue measure, a partition* (*modulo* 0) $Z_1, Z_2, \dots$ *of* B*, and times* $R'_1, R'_2, \dots$ *such that, for all* j

1) $T^{R'_j}$ *is a diffeomorphism between* Z_j *and* B*;*
2) $T^{R'_j}_{|Z_j}$ *expands the distances of at least* $\sigma^{-\frac{1}{2}} > 1$*;*
3) *the volume-distortion of* $T^{R'_j}_{|Z_j}$ *is Lipschitz;*

4) *for* $x, y \in Z_j$ *and* $n \leqslant R'_j$, $\operatorname{dist}(T^n x, T^n y) \leqslant \operatorname{dist}(T^{R'_j} x, T^{R'_j} y)$.

Moreover, the estimates on the size of tails as given in Theorem 3.1 *still hold.*

Proof. — Let X be the extension of M constructed in the proof of Theorem 3.2 using the auxiliary partition, and ν one of the invariant ergodic measures on X such that $\pi_*(\nu) = \mu$. We identify each set U_i in M with $U_i \times \{0\}$ in X.

On one U_i (let us say U_1), the measure ν is equivalent to m. The basis B of the Young tower will be U_1. Write $U_2, \dots, U_s$ for the other sets U_i on which ν is equivalent to m. Let T'_Y be the map induced by T' on $Y = \{(x, 0)\} \subset X$, i.e., on an element W_j of the partition $\mathcal{B}$ given by Theorem 3.1, with return time R_j, we set $T'_Y(x, 0) = (T^{R_j}(x), 0)$. We define a partition $\mathcal{B}^n$ of Y by

$$\mathcal{B}^n = \bigcap_0^{n-1} (T'_Y)^{-i}(\mathcal{B}).$$

Thus, an element of $\mathcal{B}^n$ is sent by $T'_Y, \dots, (T'_Y)^{n-1}$ on subsets of elements of $\mathcal{B}$, and by $(T'_Y)^n$ on a set U_i. As ν is ergodic, there exists $L > 0$ such that every U_i (with $i \leqslant s$) contains an element of $\mathcal{B}^n$, for some $n < L$, whose image under $(T'_Y)^n$ is U_1.

For $x \in \bigcup_1^s U_i$, we define a sequence of times $t_0(x) = 0, t_1(x), t_2(x), \dots$ and an integer $k(x)$ (corresponding to the number of iterations before x is selected) in the following way: let $B_0 \in \mathcal{B}$ contain x, and let R_1 be its return time. Set $t_1(x) = R_1$. If $T'_Y(B_0) = U_1$, we set $k(x) = 1$ and we stop here. Otherwise, $T^{R_1}(B_0)$ is one of the sets U_i with $2 \leqslant i \leqslant s$. We consider the set B_1 of the partition $\mathcal{B}$ containing $T^{R_1}(x)$, with a return time R_2. Set $t_2(x) = t_1(x) + R_2$. If $T^{t_2(x)}(x)$ is in U_1, we set $k(x) = 2$ and we stop here. Otherwise we consider the next iterate of $T^{t_2(x)}$, that we denote by $T^{t_3(x)}$, and we go on. More formally, $k(x) = k(T^{R_1}x) + 1$ and $t_j(x) = t_{j-1}(T^{R_1}(x)) + t_1(x)$ for every $j \leqslant k(x)$. By definition, $k(x)$ is the smallest integer $n \geqslant 1$ such that the element of $\mathcal{B}^n$ containing x is sent on U_1 by $(T'_Y)^n$.

The elements of the final partition will be the sets Z_j constructed in this way, included in U_1, and the corresponding return time will be $t_{k(x)}(x)$ for $x \in Z_j$ (this is independent of x). By construction, $T^{t_{k(x)}(x)}(Z_j) = U_1$, and we have a Young tower.

In the end, almost every point will be selected (we will see later that the measure of the tails tends to 0). The distortion and expansion properties of the partition $\mathcal{B}$ ensure that these properties will remain satisfied by the Young tower. We just have to prove the estimates on the measures of the tails to conclude.

Set $\tau(x) = t_{k(x)}(x)$. In at most L steps, an element of every U_i is selected to come back to U_1, by definition of L. Since the distortion is bounded, there exists $\varepsilon > 0$ such that

$$\operatorname{Leb}\bigl(\tau = t_j \text{ or } \dots \text{ or } \tau = t_{j+L-1} \mid t_1, \dots, t_{j-1},\ \tau > t_{j-1}\bigr) \geqslant \varepsilon. \tag{22}$$

Moreover, still by bounded distortion,

$$\mathrm{Leb}\big\{t_{j+1}-t_j>n \mid t_1,\dots,t_j\big\} \leqslant C \sum_{W_k\in\mathcal{B},\, R_k>n} \mathrm{Leb}(W_k), \tag{23}$$

this last term being estimated by Theorem 3.1. We want to obtain estimates on the measure of the tails, i.e. on $\mathrm{Leb}\{x \mid \tau(x)>n\}$, and we will use (22) and (23) to get them. The following lemma is indeed sufficient to conclude the proof. □

Lemma 4.2. — *Let (X,μ) be a space endowed with a finite measure, $k : X \to \mathbb{N}$ and $t_0, t_1, t_2, \dots : X \to \mathbb{N}$ measurable functions such that $0=t_0<t_1<t_2<\cdots$ almost everywhere. Set $\tau(x)=t_{k(x)}(x)$, and assume that there exist $L>0$ and $\varepsilon>0$ such that*

$$\mu\big\{\tau=t_j \text{ or } \dots \text{ or } \tau=t_{j+L-1} \mid t_1,\dots,t_{j-1},\, \tau>t_{j-1}\big\} \geqslant \varepsilon. \tag{24}$$

Assume moreover that there exist a positive sequence u_n and a constant C such that

$$\mu\big\{t_{j+1}-t_j>n \mid t_1,\dots,t_j\big\} \leqslant C u_n. \tag{25}$$

Then

1) *If u_n has polynomial decay, $\mu\{\tau>n\}=O(u_n)$.*
2) *If $u_n=\mathrm{e}^{-cn^\eta}$ with $c>0$ and $\eta\in(0,1]$, then there exists $c'>0$ such that $\mu\{\tau>n\}=O(\mathrm{e}^{-c'n^\eta})$.*

Proof. — Young [14] considers a problem which is *a priori* completely different: she wants to estimate the speed of decay of correlations in towers. However, she introduces a sequence of times $t_n(x)$ which satisfies the assumptions of the lemma, and she uses only the properties (24) and (25) to obtain estimates on the set $\mu\{\tau>n\}$. In particular, in the fourth section of [14], she proves our lemma when $u_n=\mathrm{e}^{-cn}$, and when u_n has polynomial decay. She assumes $L=1$, but her proofs can easily be adapted to the general case. Moreover, for the polynomial case, she only deals with the case $u_n=1/n^\gamma$, but the same proof works directly in the general case, using that $u_{n/i}\leqslant u_n i^\gamma$ for some $\gamma>0$.

However, in the stretched exponential case (i.e. $0<\eta<1$), the estimates of Young give only $\mu\{\tau>n\}=O(\mathrm{e}^{-n^{\eta'}})$ for any $\eta'<\eta$, which is weaker than the result of our lemma. We will give a different proof in this case.

When w^1 and w^2 are two real sequences, we will write $w^1\star w^2$ for their convolution, given by $(w^1\star w^2)_n=\sum_{a+b=n} w^1_a w^2_b$. When w is a sequence, we will also write $w^{\star\ell}$ for the sequence obtained by convolving ℓ times w with itself.

Write $v_n=C\mathrm{e}^{-cn^\eta}$, so that $\mu\{t_j-t_{j-1}=n \mid t_{j-1},\dots,t_1\}\leqslant v_n$. Let us show that, for large enough K, the sequence $w_n=1_{n\geqslant K}\, v_n$ satisfies

$$\forall p\in\mathbb{N},\quad (w\star w)_p \leqslant w_p. \tag{26}$$

Note that, on $[0,\frac{1}{2}]$, the function $(x^\eta+(1-x)^\eta-1)/x^\eta$ is continuous (it tends to 1 at 0), and positive, whence larger than some constant $\gamma>0$. Hence, $x^\eta+(1-x)^\eta\geqslant 1+\gamma x^\eta$. For $p<2K$, $(w\star w)_p=0$. Take $p\geqslant 2K$. Then

$$(w\star w)_p\leqslant 2C^2\sum_{K\leqslant j\leqslant\frac{1}{2}p}\mathrm{e}^{-cj^\eta}\,\mathrm{e}^{-c(p-j)^\eta}=2C^2\sum_{K\leqslant j\leqslant\frac{1}{2}p}\mathrm{e}^{-cp^\eta((j/p)^\eta+(1-j/p)^\eta)}.$$

For $x=j/p$, we have $x\in[0,\frac{1}{2}]$, whence

$$(w\star w)_p\leqslant 2C^2\sum_{K\leqslant j\leqslant\frac{1}{2}p}\mathrm{e}^{-cp^\eta(1+\gamma(j/p)^\eta)}\leqslant 2C^2\,\mathrm{e}^{-cp^\eta}\sum_{j\geqslant K}\mathrm{e}^{-c\gamma j^\eta}.$$

Taking K large enough so that $2C\sum_{j\geqslant K}\mathrm{e}^{-c\gamma j^\eta}\leqslant 1$, we obtain (26).

Let $k\geqslant 0$ and $A\subset\{1,\dots,k\}$. For $j\in A$, take $n_j\geqslant 1$. Set

$$Y(A,n_j)=\bigl\{x\mid k(x)\geqslant\sup A\text{ and }\forall j\in A,\ t_j(x)-t_{j-1}(x)=n_j\bigr\}.$$

Conditioning successively with respect to the different times, we get by (25) and the definition of v_n,

$$\mu\bigl(Y(A,n_j)\bigr)\leqslant\prod_{j\in A}\mu\bigl\{t_j-t_{j-1}=n_j\mid t_{j-1},\dots,t_1\bigr\}\leqslant\prod_{j\in A}v_{n_j}.$$

Set $q(n)=\lfloor\alpha n^\eta\rfloor$, where α will be chosen later. Take x such that $\tau(x)>n$. If $k(x)>q(n)$, i.e. x is selected after more than $q(n)$ steps, we do not do anything. Otherwise, let $\ell=k(x)\leqslant q(n)$, and let $n_j=t_j(x)-t_{j-1}(x)$ for $j\leqslant\ell$. Write $A=\{j\mid n_j\geqslant K\}$. Thus, $x\in Y(A,n_j)$. Moreover, as $\sum n_j=\tau(x)>n$, we have $\sum_{j\in A}n_j\geqslant n-Kq(n)\geqslant\frac{1}{2}n$ if n is large enough. We have shown that

$$\bigl\{x\mid\tau(x)>n\bigr\}\subset\bigl\{k(x)>q(n)\bigr\}\cup\bigcup_{A\subset\{1,\dots,q(n)\}}\ \bigcup_{\substack{n_j\geqslant K\\ \sum_A n_j\geqslant\frac{1}{2}n}}Y(A,n_j).\tag{27}$$

By (24), $\mu\{k(x) > q(n)\} \leqslant (1-\varepsilon)^{q(n)/L} \leqslant \mathrm{e}^{-c'' n^\eta}$ for some c''. Moreover, writing $\ell = \operatorname{Card} A$ and using (26),

$$\begin{aligned}
\mu\Big(\bigcup_{A\subset\{1,\dots,q(n)\}} \bigcup_{\substack{n_j\geqslant K\\ \sum_A n_j\geqslant\frac12 n}} Y(A,n_j)\Big) &\leqslant \sum_{A\subset\{1,\dots,q(n)\}} \sum_{\substack{n_j\geqslant K\\ \sum_A n_j\geqslant\frac12 n}} \prod_{j\in A} v_{n_j}\\
&\leqslant \sum_{0\leqslant\ell\leqslant q(n)} \binom{q(n)}{\ell} \sum_{\substack{n_1,\dots,n_\ell\geqslant K\\ \sum n_j\geqslant\frac12 n}} v_{n_1}\cdots v_{n_\ell}\\
&= \sum_{0\leqslant\ell\leqslant q(n)} \binom{q(n)}{\ell} \sum_{\frac12 n}^{\infty} (w^{\star\ell})_p\\
&\leqslant \sum_{0\leqslant\ell\leqslant q(n)} \binom{q(n)}{\ell} \sum_{\frac12 n}^{\infty} w_p = 2^{q(n)} \sum_{\frac12 n}^{\infty} w_p.
\end{aligned}$$

As $w_n = O(\mathrm{e}^{-cn^\eta})$, one proves (comparing to an integral) that

$$\sum_{\frac12 n}^{\infty} w_p = O(n^{1-\eta}\,\mathrm{e}^{-c(\frac12 n)^\eta}).$$

Hence, if α is small enough, $2^{q(n)}\sum_{\frac12 n}^{\infty} w_p = O(\mathrm{e}^{-c'n^\eta})$ for some $c'>0$. By (27), we have proved that $\mu\{\tau(x)>n\} = O(\mathrm{e}^{-c'n^\eta})$. □

4.2. Consequences

THEOREM 4.3. — *Let T satisfy the assumptions of Theorem 3.2, μ be one of the invariant ergodic absolutely continuous probability measures given by this theorem, and O be an open set such that μ is equivalent to $\mathrm{Leb}_{|O}$.*

Then there exists a finite partition (modulo 0) $\Omega_0,\dots,\Omega_{d-1}$ of O in open sets, such that $T(\Omega_i)=\Omega_{i+1}$ (modulo 0) for $i\leqslant d-1$ (Ω_d is identified with Ω_0), and such that, on each Ω_i, the map T^d is mixing (and even exact) for the measure μ.

Finally, for every functions $f,g : M\to\mathbb{R}$ with f Hölder and g bounded, there exists a constant C such that, for $0\leqslant i\leqslant d-1$, for all $n\in\mathbb{N}$, the correlations $\mathrm{Cor}_{\Omega_i}(f, g\circ T^{dn}) := \int_{\Omega_i} f\cdot g\circ T^{dn}\,\mathrm{d}\mu - \left(\int_{\Omega_i} f\,\mathrm{d}\mu\right)\left(\int_{\Omega_i} g\,\mathrm{d}\mu\right)$ satisfy

$$\left|\mathrm{Cor}_{\Omega_i}(f,g\circ T^{dn})\right| \leqslant \begin{cases} C\sum_{p=n}^{\infty} u_p & \text{in the first case,}\\ C\sum_{p=n}^{\infty}(\log p)u_p & \text{in the third case,}\\ C\,\mathrm{e}^{-c'n^\eta} & \text{in the second.}\\ & \text{and fourth cases.}\end{cases} \tag{28}$$

When all the iterates of T are topologically transitive, there exist a unique measure μ and a unique set Ω. This proves Theorems 1.1 and 1.3.

Proof. — Theorem 4.1 makes it possible to construct an abstract Young tower $X = \{(x,i) \mid x \in Z_j,\ i < R'_j\}$, a projection $\pi : X \to M$ given by $\pi(x,i) = T^i(x)$, and a map T' on X such that $\pi \circ T' = T \circ \pi$, as in the proof of Theorem 3.2 (but using the partition given by Theorem 4.1 instead of the partition given by Theorem 3.1).

By [14], T' admits a unique absolutely continuous invariant probability measure ν. The measure $\pi_*(\nu)$ is absolutely continuous with respect to μ, whence $\pi_*(\nu) = \mu$ by ergodicity.

Set $d_1 = \gcd(R'_j)$, and write $X_k = \{(x,i) \in X \mid i \equiv k \mod d_1\}$, for $0 \leqslant k \leqslant d_1 - 1$. Thus, T' maps X_k to X_{k+1} for $k < d_1$ (taking k modulo d_1). The system $(X_k, (T')^{d_1})$ is then a Young tower whose return times are relatively prime, and whose invariant measure is $\nu_k := \nu_{|X_k}$. [14, Theorem 1] implies that ν_k is exact for $(T')^{d_1}$. Moreover, the correlations of Hölder functions (as defined in [14]) decay as indicated in (28): in the exponential case, this is proved in [14]. Young treats the case of $1/n^\gamma$, but her proof can easily be adapted to the polynomial case. It remains to treat the stretched exponential case, which is given by the following lemma:

Lemma 4.4. — *Let (X, T') be a mixing Young tower, and assume that the return time on the basis R satisfies $m(R > n) = O(\mathrm{e}^{-cn^\eta})$ for some $0 < \eta < 1$. Then, if f is Hölder and g is bounded, the correlations of f and g are bounded by $\mathrm{e}^{-c'n^\eta}$ for some $c' > 0$.*

Proof. — This is a consequence of [14, Section 3.5] and Lemma 4.2. □

These results are true on X, we still have to come back to M.

The measures $\lambda_k = \pi_*(\nu_k)$ satisfy $T_*\lambda_k = \lambda_{k+1}$, and are invariant and ergodic for T^{d_1}. In particular, two such measures are either equal or mutually singular. Hence, there exists d (dividing d_1, let us say $d_1 = sd$) such that $\lambda_k = \lambda_\ell$ if and only if $k \equiv \ell \mod d$. Using the same argument as in the proof of Theorem 3.2, we check that the measures λ_k (for $0 \leqslant k < d$) are supported on disjoint open sets Ω_k. Moreover, $T_*(\lambda_k) = \lambda_{k+1}$, whence $T(\Omega_k) = \Omega_{k+1}$ modulo 0.

Let us show that λ_k is exact for T^d. Let $A \subset \Omega_k$ have nonzero measure, such that A can we written as $T^{-dn}(A_n)$ for any n. Hence, $A' = \pi^{-1}(A)$ is equal to $(T')^{-dn}(A'_n)$, where $A'_n = \pi^{-1}(A_n)$. In particular, since X_k is invariant under $(T')^{d_1}$, we get $A' \cap X_k = (T')^{-nd_1}(A'_{sn} \cap X_k)$. As (X_k, ν_k) is exact, this proves that $A' \cap X_k$ has full ν_k-measure, which concludes the proof.

Let finally f, g be two functions on M such that f is Hölder and g is bounded. Write $f' = f \circ \pi$ and $g' = g \circ \pi$: the function f' is Hölder on X, and g' is bounded. For $n \in \mathbb{N}$, write $n = ps + r$ with $0 \leqslant r < s$. Then

$$\int_{\Omega_k} f \cdot g \circ T^{dn} = \int_{X_k} f' \cdot (g' \circ (T')^{dr}) \circ (T')^{pds} = \int_{X_k} f' \cdot (g' \circ (T')^{dr}) \circ (T')^{pd_1}.$$

The function $g' \circ (T')^{dr}$ is bounded on X_k, whence the estimate on the speed of decay of correlations for ν_k on X_k gives the same estimate for the decay of correlations of f and g on M. □

BIBLIOGRAPHY

[1] J. Aaronson – *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, 1997.

[2] J. F. Alves – SRB measures for non-hyperbolic systems with multidimensional expansion, *Ann. Sci. École Norm. Sup.* **33** (2000), p. 1–32.

[3] J. F. Alves & V. Araújo – Random perturbations of nonuniformly expanding maps, *Astérisque* **286** (2003), p. 25–62.

[4] ______, Hyperbolic times: frequency versus integrability, *Ergodic Theory Dynam. Systems* **24** (2004), p. 329–346.

[5] J. F. Alves, C. Bonatti & M. Viana – SRB measures for partially hyperbolic systems whose central direction is mostly expanding, *Invent. Math.* **140** (2000), p. 351–398.

[6] J. F. Alves, S. Luzzatto & V. Pinheiro – Markov structures and decay of correlations for non-uniformly expanding dynamical systems, Preprint, 2002.

[7] J. F. Alves & M. Viana – Statistical stability for robust classes of maps with non-uniform expansion, *Ergodic Theory Dynam. Systems* **22** (2002), p. 1–32.

[8] V. Baladi & S. Gouëzel – Stretched exponential bounds for the correlations of the Viana-Alves skew product, `www.math.jussieu.fr/~baladi`, to appear, Proceedings Workshop on Dynamics and Randomness, Universidad de Chile, Santiago de Chile, 2002.

[9] ______, A note on stretched exponential decay of correlations for the Viana-Alves map, `arXiv.org/math.DS/0311189`, 2003.

[10] J. Buzzi, O. Sester & M. Tsujii – Weakly expanding skew-products of quadratic maps, *Ergodic Theory Dynam. Systems* **23** (2003), p. 1401–1414.

[11] S. Gouëzel – Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes, Thèse, Université Paris Sud, 2004.

[12] M. Viana – Multidimensional nonhyperbolic attractors, *Publ. Math. IHÉS* **85** (1997), p. 63–96.

[13] L.-S. Young – Statistical properties of dynamical systems with some hyperbolicity, *Ann. of Math.* **147** (1998), p. 585–650.

[14] ______, Recurrence times and rates of mixing, *Israel J. Math.* **110** (1999), p. 153–188.