Open access • Journal Article • DOI:10.1017/S0143385703000579

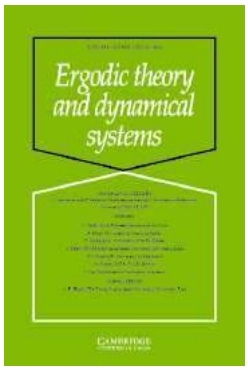
Lyapunov exponents and rates of mixing for one-dimensional maps

José F. Alves, Stefano Luzzatto, Vilton Pinheiro

Institutions: Imperial College London, Federal University of Bahia

Published on: 01 Jun 2004 - Ergodic Theory and Dynamical Systems (Cambridge University Press)

Topics: Lyapunov exponent, Almost everywhere, Mixing (mathematics), Invariant measure and Zero (complex analysis)



Lyapunov exponents and rates of mixing for one-dimensional maps

JOSÉ F. ALVES, STEFANO LUZZATTO and VILTON PINHEIRO

Ergodic Theory and Dynamical Systems / Volume 24 / Issue 03 / June 2004, pp 637 - 657
DOI: 10.1017/S0143385703000579, Published online: 04 May 2004

Ergod. Th. & Dynam. Sys. (2004), **24**, 637–657
DOI: 10.1017/S0143385703000579

Lyapunov exponents and rates of mixing for one-dimensional maps

JOSÉ F. ALVES†, STEFANO LUZZATTO‡ and VILTON PINHEIRO§

† *Departamento de Matemática Pura, Faculdade de Ciências do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal (e-mail: jfalves@fc.up.pt)*

‡ *Mathematics Department, Imperial College, 180 Queen's Gate, London SW7, UK (e-mail: stefano.luzzatto@ic.ac.uk)*

§ *Departamento de Matemática, Universidade Federal da Bahia, Av. Ademar de Barros s/n, 40170-110 Salvador, Brazil (e-mail: viltonj@ufba.br)*

(*Received* 13 *February* 2003 *and accepted in revised form* 2 *September* 2003)

Abstract. We show that one-dimensional maps f with strictly positive Lyapunov exponents almost everywhere admit an absolutely continuous invariant measure. If f is topologically transitive, some power of f is mixing and, in particular, the correlation of Hölder continuous observables decays to zero. The main objective of this paper is to show that the rate of decay of correlations is determined, in some situations, by the average rate at which typical points start to exhibit exponential growth of the derivative.

1. *Introduction and statement of results*

1.1. *Lyapunov exponents.*

The purpose of this paper is to study the statistical properties of one-dimensional maps $f : I \to I$ with positive Lyapunov exponents, where I may be the circle $\mathbb{S}^1$ or an interval. Such maps satisfy asymptotic exponential estimates for growth of the derivative, but do not necessarily exhibit exponential estimates for other features of the dynamics such as the decay of correlations. Our main objective here is to identify some criterion which distinguishes different degrees of expansivity and which is reflected in the rate of decay of correlations of the system.

Definition 1. We say that a map $f : I \to I$ has *positive Lyapunov exponents* almost everywhere if there exists some $\lambda > 0$ such that

$$\liminf_{n\to\infty} \frac{1}{n} \log |(f^n)'(x)| \geq \lambda > 0 \qquad (*)$$

for Lebesgue almost every point $x \in I$.

1.2. *Decay of correlations.* Positive Lyapunov exponents are known to be a cause of sensitive dependence on initial conditions and other dynamical features which give rise to a degree of *chaoticity* or *stochasticity* in the dynamics. We can formalize this idea through the notion of *mixing with respect to some invariant measure*.

Definition 2. A probability measure μ defined on the Borel sets of I is said to be f*-invariant* if $\mu(f^{-1}(A)) = \mu(A)$ for every Borel set $A \subset I$.

Definition 3. A map f is said to be *mixing* with respect to some f-invariant probability measure μ if

$$|\mu(f^{-n}(A) \cap B) - \mu(A)\mu(B)| \to 0, \quad \text{when } n \to \infty,$$

for any measurable sets A, B.

One interpretation of this property is that the conditional probability of B given $f^{-n}(A)$, i.e. the probability that the event A is a consequence of the event B having occurred at some time in the past, is asymptotically the same as if the two events were completely independent. This is sometimes referred to as a property of *loss of memory*, and thus in some sense of *stochasticity*, of the system. A natural question of interest both for application and for intrinsic reasons, therefore, is the speed at which such loss of memory occurs. Standard counterexamples show that, in general, there is no specific rate: it is always possible to choose sets A and B for which mixing is arbitrarily slow. However, this notion can be generalized in the following way.

Definition 4. For a map $f : I \to I$ preserving a probability measure μ and functions $\varphi, \psi \in L^1(\mu)$, we define the *correlation function*

$$\mathcal{C}_n = \mathcal{C}_n(\varphi, \psi) = \left| \int (\varphi \circ f^n)\psi \, d\mu - \int \varphi \, d\mu \int \psi \, d\mu \right|.$$

Note that choosing these observables to be characteristic functions of Borel sets gives the well-known definition of mixing when $\mathcal{C}_n \to 0$. By restricting the set of allowed *observables*, for example to the class of Hölder continuous functions, it is sometimes possible to obtain specific upper bounds for the decay of the correlation function $\mathcal{C}_n$ which depend only on the map f (up to a multiplicative constant which is allowed to depend on φ and ψ). Indeed, it is generally possible to allow $\psi \in L^\infty(\mu)$ and only restrict the choice of φ to the class of Hölder continuous functions. We make several statements below concerning the decay of correlations for Hölder continuous observables meaning that φ is Hölder continuous and $\psi \in L^\infty(\mu)$.

Not more than a decade ago, the only examples for which a specific rate of decay of correlations was known were uniformly expanding maps in one dimension and, more generally, uniformly hyperbolic systems in higher dimensions [**14–16, 36**]. In these cases we always have exponential decay of correlations for Hölder continuous observables, an estimate which can be expected due to the fact that essentially everything is exponential in these cases. Relaxing uniform expansion conditions has proved extremely hard and until now estimates on the decay for correlations for one-dimensional systems satisfying the asymptotic exponential expansion condition $(*)$ but strictly not uniformly expanding have

only been known in some fairly specific classes of examples such as maps with indifferent fixed points [**11, 12, 15, 24**] or non-flat critical points [**5, 6, 22**] where the rate of decay depends quite explicitly on the features of the neutral fixed point and the critical points, respectively. See [**18**] for a more detailed discussion and references.

1.3. *Degree of non-uniformity of the expansion.* A natural question is what general characteristics of a system satisfying $(*)$ determine the rate of decay of correlations? The purpose of the present paper is to give a partial solution to this question by relating the rate of decay of correlations to the *degree of non-uniformity* of the expansivity, i.e. the time we have to wait for typical points to start behaving as though the system were uniformly expanding.

Definition 5. For $0 < \lambda' < \lambda$, we define the *expansion time* function $\mathcal{E}$ by

$$\mathcal{E}(x) = \min\left\{N : \frac{1}{n}\log|(f^n)'(x)| \geq \lambda', \quad \forall n \geq N\right\}.$$

Condition $(*)$ implies that $\mathcal{E}$ is defined and finite almost everywhere. We think of this as the *waiting time* before the exponential derivative growth kicks in. A map is uniformly expanding if $\mathcal{E}(x)$ is uniformly bounded for every x. In general, however, $\mathcal{E}(x)$ is defined only almost everywhere and unbounded, indicating that some points may exhibit no growth or even arbitrarily large loss of growth of the derivative along their orbit for an arbitrarily long time before starting to exhibit the exponential growth implied by condition $(*)$. Our results show that at least in certain situations the properties of the function $\mathcal{E}(x)$ are closely related to the rate of mixing of the map.

1.4. *Local diffeomorphisms.* We begin with the simplest situation in order to highlight the main idea of the results.

THEOREM 1. *Let $f : I \to I$ be a C^2 local diffeomorphism with some point having dense pre-orbit. Suppose that f satisfies condition $(*)$ and that there exists $\gamma > 1$ such that for some $0 < \lambda' < \lambda$ we have*

$$|\{\mathcal{E}(x) > n\}| \leq \mathcal{O}(n^{-\gamma}).$$

Then there exists an absolutely continuous, f-invariant, probability measure μ on I. Some finite power of f is mixing with respect to μ and the correlation function $\mathcal{C}_n$ for Hölder continuous observables on I satisfies

$$\mathcal{C}_n \leq \mathcal{O}(n^{-\gamma+1}).$$

We emphasize here that the asymptotic statements here *do not depend* on the choice of λ'. Thus the rate of decay of correlations depends essentially on the average time with which *some* given uniform exponential rate of expansion is attained and *does not* depend on what turn out to be significantly more subtle characteristics of the system such as the actual rate of convergence of the Lyapunov exponents to the limit.

Remark 1. The absolute continuity and ergodicity of μ follow from [**2**]. Our argument gives an alternative proof of the absolute continuity of μ and allows us to obtain the estimates on the rate of decay of correlations which are the main purpose of this paper.

Remark 2. The statement about the rate of decay of the correlations is of interest even if an absolutely continuous mixing f-invariant, probability measure μ is given to begin with. Then condition $(*)$ is just equivalent to the integrability condition

$$\int \log|f'|\,d\mu > 0,$$

since Birkhoff's ergodic theorem then implies that the limit

$$\lambda = \lim_{n\to\infty} \frac{1}{n}\log|(f^n)'(x)| = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\log|f'(x)| = \int \log|f'|\,d\mu > 0$$

exists for μ-almost every $x \in I$. In particular, the expansion time function $\mathcal{E}(x)$ is also defined and finite almost everywhere and the conclusions of the theorem hold under the given conditions on the rate of decay of $|\{\mathcal{E}(x) > n\}|$.

1.5. *Multimodal maps.* We can generalize our result to C^2 maps with non-flat critical points if we assume that almost all orbits have *slow approximation* to the critical set $\mathcal{C}$. Let $\mathrm{dist}_\delta(x,\mathcal{C})$ denote the δ-*truncated* distance from x to $\mathcal{C}$ defined as $\mathrm{dist}_\delta(x,\mathcal{C}) = \mathrm{dist}(x,\mathcal{C})$ if $\mathrm{dist}(x,\mathcal{C}) \le \delta$ and $\mathrm{dist}_\delta(x,\mathcal{C}) = 1$ otherwise.

Definition 6. We say that a map $f : I \to I$, with a critical set $\mathcal{C}$, satisfies the *slow recurrence condition* if given any $\epsilon > 0$ there exists $\delta > 0$ such that for Lebesgue almost every $x \in I$

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} -\log \mathrm{dist}_\delta(f^j(x),\mathcal{C}) \le \epsilon. \qquad (**)$$

Condition $(**)$ is an asymptotic statement, just like condition $(*)$, and we have no *a priori* knowledge about how fast this limit is approached or with what degree of uniformity for different points x. Thus we introduce the analogue of the expansion time function as follows.

Definition 7. The *recurrence time* function is defined by

$$\mathcal{R}(x) = \min\left\{N \ge 1 : \frac{1}{n}\sum_{i=0}^{n-1} -\log \mathrm{dist}_\delta(f^j(x),\mathcal{C}) \le 2\epsilon,\ \forall n \ge N\right\}.$$

Condition $(**)$ implies that $\mathcal{R}$ is well defined and finite almost everywhere in I. For stating our results we also need the following.

Definition 8. For each $n \ge 1$ define the *tail of non-uniformity* by

$$\Gamma_n = \{x : \mathcal{E}(x) > n \text{ or } \mathcal{R}(x) > n\}$$

This is the set of points which at time n have not yet achieved either the uniform exponential growth or the uniform slow approximation to $\mathcal{C}$ given by $(*)$ and $(**)$.

THEOREM 2. *Let $f : I \to I$ be a C^2 map with a set $\mathcal{C}$ of non-flat critical points and with some point having dense pre-orbit in $I \setminus \mathcal{C}$. Suppose that f satisfies the non-uniform expansivity condition* $(*)$ *and the slow approximation condition* $(**)$ *to $\mathcal{C}$ and suppose that there exists $\gamma > 1$ such that*

$$|\Gamma_n| \leq \mathcal{O}(n^{-\gamma}).$$

Then there exists an absolutely continuous, f-invariant, probability measure μ. Some finite power of f is mixing with respect to μ and the correlation function $\mathcal{C}_n$ for Hölder continuous observables on I satisfies

$$\mathcal{C}_n \leq \mathcal{O}(n^{-\gamma+1}).$$

Remark 3. As for condition $(*)$, condition $(**)$ also admits a formulation which may appear more natural if an ergodic f-invariant, absolutely continuous probability measure μ is given to begin with. Indeed, it then just reduces to an integrability condition on the log of the distance function to the critical set:

$$\int |\log \operatorname{dist}(x, \mathcal{C})| \, d\mu < \infty.$$

This can be thought of as saying that the invariant measure does not give disproportionate weight to neighbourhoods of $\mathcal{C}$. In particular, the bounded recurrence condition $(*)$ is satisfied quite generally in smooth maps for which the density of the absolutely continuous invariant measure is known to be quite regular.

Remark 4. The existence of an f-invariant, absolutely continuous probability measure under assumption $(*)$ only, has been known for some time in the unimodal case and is due to Keller [**13**].

1.6. *Maps with singular points.* The way in which the control of the recurrence near critical points comes into play in our argument can be used to deal similarly with other kinds of *criticalities* or *singularities*, such as discontinuities and/or points with infinite derivative. In fact, it is becoming increasingly clear (see, for example, the references given above as well as [**14, 17**]) that there is a kind of duality in one-dimensional maps between critical points and discontinuities with infinite derivative. Regions in which the derivative is unbounded may appear at first to contribute positively in the direction of proving some properties of the map which often follow from the expansivity. However, it has been known for some time that discontinuities and unbounded derivatives create non-trivial obstructions to the geometric structures and distortion estimates which usually form the basis for the study of dynamical properties. Therefore, these regions need to be avoided, or at least the recurrence in these regions needs to be controlled, in a very similar way to the way in which the recurrence near critical points is controlled. In fact, it is possible to formulate a general set of conditions in which we can talk about a *singular set* $\mathcal{S}$ without specifying the nature of the singularities which may be critical points, discontinuities or points at which the norm of the derivative goes to infinity. In this context we need to generalize the notion of *non-flatness* for critical points.

Definition 9. We say that the singular set $\mathcal{S}$ is *non-degenerate* if:

(s_1) f' *behaves like a power of the distance* to $\mathcal{S}$: there are constants $B > 1$ and $\beta > 0$ such that for every $c \in \mathcal{S}$ and $x \in I \setminus \mathcal{S}$

$$\frac{1}{B}|x-c|^{\beta} \le |f'(x)| \le B|x-c|^{-\beta};$$

(s_2) $\log|f'|$ is *locally Lipschitz* at points $x \in I \setminus \mathcal{S}$ with Lipschitz constant depending on the distance to $\mathcal{S}$: for every $c \in \mathcal{S}$ and $x, y \in I \setminus \mathcal{S}$ with $|x-y| < |x-c|/2$ we have

$$|\log|f'(x)| - \log|f'(y)|| \le \frac{B}{|x-c|^{\beta}}|x-y|.$$

We can define the slow recurrence condition $(**)$ with respect to a general non-degenerate singular set in exactly the same way as for the simpler case of non-flat smooth critical points. Similarly, we extend the definition of the recurrence function $\mathcal{R}$ and of the tail of non-uniformity Γ_n. With these definitions we can now formulate our results in the most general setting. Theorems 1 and 2 are then special cases of the following.

THEOREM 3. *Let $f : I \to I$ be a C^2 local diffeomorphism outside a non-degenerate critical/singular set $\mathcal{S}$ with some point having dense pre-orbit in $I \setminus \mathcal{S}$. Suppose that f satisfies the non-uniform expansivity condition $(*)$ and the slow approximation condition $(**)$ to the critical set and suppose that there exists $\gamma > 1$ such that*

$$|\Gamma_n| \le \mathcal{O}(n^{-\gamma}).$$

Then there exists an absolutely continuous, f-invariant, probability measure μ. Some finite power of f is mixing with respect to μ and the correlation function C_n for Hölder continuous observables on I satisfies

$$C_n \le \mathcal{O}(n^{-\gamma+1}).$$

We expect this relation between the rate of decay of the tails of the expansivity and recurrence functions and the rate of decay of correlations to also hold for faster rates such as exponential and stretched exponential. For the moment, however, the kind of statistical arguments we use do not allow us to include these cases in our statements.

1.7. *The Central Limit Theorem.* As a by-product of our results we do however get some further interesting statistical properties of the maps under consideration. In particular we can apply some general results, e.g. [**16, 24**], which show that under certain conditions including sufficiently fast decay of correlations, we can prove the Central Limit Theorem which states that the probability of a given deviation of the average values of an observable along an orbit from the asymptotic average is essentially given by a normal distribution: given a Hölder continuous function ϕ which is not a coboundary ($\phi \ne \psi \circ f - \psi$ for any ψ) there exists $\sigma > 0$ such that for every interval $J \subset \mathbb{R}$,

$$\mu\left\{x \in M : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\left(\phi(f^j(x)) - \int \phi\, d\mu\right) \in J\right\} \to \frac{1}{\sigma\sqrt{2\pi}}\int_J e^{-t^2/2\sigma^2}\, dt.$$

In our context we have the following.

THEOREM 4. *Let $f : I \to I$ be a C^2 local diffeomorphism outside a non-degenerate critical/singular set $\mathcal{S}$ with some point having dense pre-orbit in $I \setminus \mathcal{S}$. Suppose that f satisfies the non-uniform expansivity condition* $(*)$ *and the slow approximation condition* $(**)$ *to $\mathcal{S}$ and suppose that*

$$|\Gamma_n| \leq \mathcal{O}(n^{-\gamma}), \quad \text{for some } \gamma > 2.$$

Then there exists an absolutely continuous, f-invariant, probability measure μ and the Central Limit Theorem holds.

Note that we have strengthened our condition on the rate of decay of the tail of $|\Gamma_n|$ to $\gamma > 2$ rather than just $\gamma > 1$.

Remark 5. The hypotheses of f having a point whose pre-orbit is dense and disjoint from the singular set is not needed in all its strength in all our theorems; cf. Remark 6. Actually, we can prove that this is a consequence of the transitivity of the map.

1.8. *Strategy and overview of the paper.* There exist at least two basic methodologies for studying the statistical properties of dynamical systems and, in particular, the decay of correlations. The first is the functional-analytic approach where the problem is formulated in terms of the spectral properties of the Ruelle–Perron–Frobenius operator. This approach has been extremely successful in a variety of examples and general situations [**10, 19, 20, 22**]. Another approach is more geometric/probabilistic and uses the idea of coupling measures in a somewhat more direct approach, taking advantage of the idea of the decay of correlations as being related to the convergence to equilibrium of arbitrary densities. Both methods necessarily rely to some extent on the geometric structure of the system under consideration; for example, the pioneering work of Sinai–Ruelle–Bowen for uniformly hyperbolic systems combined the information on the existence of finite Markov partitions with the functional-analytic approach. More recently, Young [**23, 24**], has made the profound observation that the rate of decay of correlations can be deduced via a coupling argument from the existence of a *generalized Markov structure* and some analytic information about this structure. The existence of such a structure and the necessary analytic estimates are highly non-trivial in general and of independent interest. This is the approach which we take in this paper. More specifically, we show that there exists an interval $\Delta \subset I \setminus \mathcal{S}$, a countable partition $\mathcal{P}$ (mod 0) of Δ into intervals J and a return time function $R : \Delta \to \mathbb{N}$ piecewise constant on elements of $\mathcal{P}$ satisfying the following properties.

(1) *Markov*: for each $J \in \mathcal{P}$ and $R = R(J)$, $f^R : J \to \Delta_0$ is a C^2 diffeomorphism (and, in particular, a bijection). Thus an induced map

$$F : \Delta_0 \to \Delta_0 \quad \text{given by } F(x) = f^{R(x)}(x)$$

is defined almost everywhere and satisfies the classical Markov property.

(2) *Uniform expansivity*: there exists $\lambda_0 > 1$ such that for almost all $x \in \Delta_0$ we have

$$|F'(x)| \geq \lambda_0.$$

In particular the separation time $s(x, y)$ given by the maximum integer such that $F^i(x)$ and $F^i(y)$ belong to the same element of the partition $\mathcal{P}$ for all $i \leq s(x, y)$, is defined and finite for almost every pair of points $x, y \in \Delta_0$.

(3) *Bounded distortion*: there exists a constant $K > 0$ such that for any pair of points $x, y \in \Delta_0$ with $\infty > s(x, y) \geq 1$ we have

$$\left|\frac{F'(x)}{F'(y)} - 1\right| \leq K\hat{\lambda}^{-s(F(x),F(y))}.$$

(4) *Polynomial decay of tail of return times*:

$$|\{R > n\}| \leq \mathcal{O}(n^{-\gamma}).$$

Our results then follow immediately from this construction and the abstract results of Young [**24**] which state that precisely under these conditions there exists an absolutely continuous invariant probability measure with decay of correlation of order $\mathcal{O}(n^{-\gamma+1})$ and for which the Central Limit Theorem holds when $\gamma > 2$.

The argument is divided into two natural sections. In §2 we give the basic algorithm for choosing the base interval Δ_0 and for constructing the induced map of Δ_0 to itself. This includes various estimates concerning the time it takes for small regions to grow to some fixed size and for these to return to Δ_0 and relies essentially on the notion of *hyperbolic time* introduced by Alves in [**1**]. We also prove that the partition elements satisfy the required expansivity and distortion estimates. In §3 we address the more analytical/statistical issues in the argument. We analyze the relative sizes of various combinatorially distinct regions and conclude that almost every point effectively belongs to some region of Δ_0 which eventually has a good return. We also give the statistical argument which allows us to relate more explicitly the tail of the return time function for the induced map to the tail of non-uniformity of the system.

2. *An induced Markov structure*

2.1. *Attaining large scale: hyperbolic times.* The key idea underlying our strategy is that of *hyperbolic times* first introduced in [**1**]. Let $B > 1$ and $\beta > 0$ be as in the hypotheses (s_1) and (s_2). In the following b is any fixed constant satisfying $0 < b < \min\{1/2, 1/(4\beta)\}$.

Definition 10. Given $\sigma < 1$, we state that n is a (σ)-*hyperbolic time* for a point $x \in I$ if for all $c \in \mathcal{S}$ and $1 \leq k \leq n$,

$$|(f^k)'(f^{n-k}(x))| \geq \sigma^{-k} \quad \text{and} \quad |f^{n-k}(x) - c| \geq \sigma^{bk}.$$

For each $n \geq 1$ we define

$$H_n = H_n(\sigma) = \{x \in I : n \text{ is a } (\sigma)\text{-hyperbolic time for } x\}.$$

LEMMA 1. *Given $\sigma < 1$, there exist $\delta_1, \kappa, D_1 > 0$, depending only on σ and f, such that for any $x \in I$ and $n \geq 1$ a (σ)-hyperbolic time for x, there exists an open interval $V_n(x)$ containing x with the following properties:*

(1) *f^n maps $V_n(x)$ diffeomorphically onto $(f^n(x) - \delta_1, f^n(x) + \delta_1)$;*

(2) *for $1 \leq k < n$ and $y, z \in V_n(x)$, $|f^{n-k}(y) - f^{n-k}(z)| \leq \sigma^{k/2}|f^n(y) - f^n(z)|$;*

(3) *$V_n(x)$ is contained in $B(x, \kappa^{-n})$;*
(4) *$f^n|V_n(x)$ has distortion bounded by D_1: for every $y, z \in V_n(x)$,*

$$\frac{1}{D_1} \leq \frac{|(f^n)'(y)|}{|(f^n)'(z)|} \leq D_1.$$

For a proof, see [**2**, Lemma 5.2] and [**2**, Corollary 5.3].

We often refer to the sets $V_n(x)$ as *hyperbolic pre-intervals* and to their images $f^n(V_n(x))$ as *hyperbolic intervals*. Note that the latter are intervals of length $2\delta_1$. The existence of hyperbolic times for points satisfying conditions $(*)$ and $(**)$ is assured by the following lemma.

LEMMA 2. *There exist $\theta > 0$ and $0 < \sigma < 1$, depending only on f, λ and λ', such that for every $x \in I \setminus \Gamma_n$ there exist (σ)-hyperbolic times $1 \leq n_1 < \cdots < n_l \leq n$ for x with $l \geq \theta n$.*

For a proof, see [**2**, Lemma 5.4].

2.2. *The partition.* We start by outlining the order in which the main constants used in the construction are chosen. We take the two constants $\theta > 0$ and $0 < \sigma < 1$ given by Lemma 2, which depend only on the map f and the constants λ and λ' which we consider to be given. We then fix a small constant $\delta_1 = \delta_1(\sigma) > 0$ given by Lemma 1. We finally take small constants $\delta_0 = \delta_0(\delta_1, \theta) > 0$ and $\varepsilon = \varepsilon(\delta_0) > 0$, and a large positive integer $R_0 = R_0(\theta)$. We adjust δ_0 at several places in the proof, but only a finite number of times.

Let us now fix once and for all $p \in I$ and $N_0 \in \mathbb{N}$ for which $\bigcup_{j=0}^{N_0} f^{-j}(\{p\})$ is $\delta_1/3$ dense in I and disjoint from the critical set $\mathcal{S}$.

Remark 6. Incidentally this is the only point where we use the existence of a point with dense pre-orbit. Actually, $\delta_1/3$ dense is enough for our purposes.

Then we introduce the intervals $\Delta_0^0 = \Delta_0 = (p - \delta_0, p + \delta_0)$ and

$$\Delta_0^1 = (p - 2\delta_0, p + 2\delta_0), \quad \Delta_0^2 = (p - \sqrt{\delta_0}, p + \sqrt{\delta_0}), \quad \Delta_0^3 = (p - 2\sqrt{\delta_0}, p + 2\sqrt{\delta_0}).$$

Let

$$I_k = \{x \in \Delta_0^1 : \delta_0(1 + \sigma^{k/2}) < |x - p| < \delta_0(1 + \sigma^{(k-1)/2})\}, \quad k \geq 1,$$

be a partition (mod 0) into countably many rings of $\Delta_0^1 \setminus \Delta_0$. The next lemma shows that there is some 'scale' after which we can guarantee a good return to Δ_0 within a fixed number of iterates.

LEMMA 3. *There exist constants $C_0 > 1$ and $D_0 > 0$ depending only on f, σ, δ_1 and the point p, such that for any interval $U \subset I$ of length $2\delta_1$, there exist an open interval $V \subset U$ and an integer $0 \leq m \leq N_0$ such that:*
(1) *f^m maps V diffeomorphically onto Δ_0^3;*
(2) *$f^m|V$ has distortion bounded by D_0;*
(3) *the m-preimages of Δ_0^3 are uniformly bounded away from $\mathcal{S}$ for all $1 \leq m \leq N_0$ and for x belonging to any such m-preimage we have*

$$\frac{1}{C_0} \leq |(f^m)'(x)| \leq C_0.$$

Proof. Since $\bigcup_{j=0}^{N_0} f^{-j}(\{p\})$ is $\delta_1/3$ dense in I and disjoint from $\mathcal{S}$, choosing $\delta_0 > 0$ sufficiently small we have that each connected component of the preimages of Δ_0^3 up to time N_0 are bounded away from the critical set $\mathcal{S}$ and are contained in an interval of length $2\delta_1/3$. This immediately implies that any interval $U \subset I$ of length $2\delta_1$ contains a preimage V of Δ_0^3 which is mapped diffeomorphically onto Δ_0^3 in at most N_0 iterates. Moreover, since the number of iterations and the distance to the critical region are uniformly bounded, the distortion is uniformly bounded.

Observe that δ_0 and N_0 have been chosen in such a way that all the connected components of the preimages of Δ_0^3 up to time N_0 satisfy the conclusions of the lemma. In particular, they are uniformly bounded away from the critical set $\mathcal{S}$ and so there is some constant $C_0 > 1$ depending only on f and δ_1 such that

$$\frac{1}{C_0} \le |(f^m)'(x)| \le C_0$$

for all $1 \le m \le N_0$ and x belonging to an m-preimage of Δ_0^3. □

Now we describe the inductive construction of the partition $\mathcal{P}$ of Δ_0. For the sake of clearness let us explain the meaning of the objects that we introduce below. At each step n we take $\Delta_n = \Delta_0 \setminus \{R \le n\}$ and write Δ_n as the disjoint union of sets A_n and B_n. A_n may be understood as the set of points that is ready for returning at time $n+1$ and $t_n : \Delta_n \to \mathbb{N}$ gives the *waiting time* for points in B_n. We also introduce a neighborhood A_n^ε of A_n whose interest will become clear in §3.

We ignore any dynamics occurring up to time R_0. We assume that sets Δ_i, A_i, A_i^ε, B_i, $\{R = i\}$ and functions $t_i : \Delta_i \to \mathbb{N}$ are defined for all $i \le n-1$. For $i \le R_0$, we just let $A_i = A_i^\varepsilon = \Delta_i = \Delta_0$, $B_i = \{R = i\} = \emptyset$ and $t_i \equiv 0$. Now let $(J_{n,j}^3)_j$ be the connected components of $f^{-n}(\Delta_0) \cap A_{n-1}^\varepsilon$ contained in hyperbolic pre-intervals V_k, with $n - N_0 \le k \le n$, which are mapped onto Δ_0^3 by f^n. Take $J_{n,j}^i = J_{n,j}^3 \cap f^{-n}\Delta_0^i$ for $i = 0, 1, 2$ and set $R(x) = n$ for $x \in J_{n,j}^0$. Also take $\Delta_n = \Delta_{n-1} \setminus \{R = n\}$. The function $t_n : \Delta_n \to \mathbb{N}$ is defined as follows:

$$t_n(x) = \begin{cases} s & \text{if } x \in J_{n,j}^1 \setminus J_{n,j}^0 \text{ and } f^n(x) \in I_s \text{ for some } j, \\ 0 & \text{if } x \in A_{n-1} \setminus \bigcup_j J_{n,j}^1, \\ t_{n-1}(x) - 1 & \text{if } x \in B_{n-1} \setminus \bigcup_j J_{n,j}^1. \end{cases}$$

Finally let

$$A_n = \{x \in \Delta_n : t_n(x) = 0\}, \quad B_n = \{x \in \Delta_n : t_n(x) > 0\}$$

and

$$A_n^\varepsilon = \{x \in \Delta_n : \operatorname{dist}(f^{n+1}(x), f^{n+1}(A_n)) < \varepsilon\}.$$

At this point we have described the inductive construction of the sets A_n, A_n^ε, B_n and $\{R = n\}$. The proof that this indeed gives a partition (mod 0) of Δ_0 is left to §3.

Associated to each component J_{n-k}^0 of $\{R = n-k\}$, for some $k > 0$, we have the two intervals of $J_{n-k}^1 \setminus J_{n-k}^0$ on both sides of it; knowing that the new components of $\{R = n\}$

'do not intersect too much' $J^1_{n-k} \setminus J^0_{n-k}$ is important for preventing overlaps on the new intervals constructed. This is indeed the case as long as $\varepsilon > 0$ is taken small enough.

LEMMA 4. *If $\varepsilon > 0$ is sufficiently small, then $J^1_n \cap \{t_{n-1} \geq 1\} = \emptyset$ for each J^1_n.*

Proof. Take some $k > 0$ and let J^0_{n-k} be a component of $\{R = n-k\}$. Let Q_k be the part of J^1_{n-k} that is mapped by f^{n-k} onto I_k and assume that Q_k intersects some J^3_n. Recall that, by construction, Q_k is precisely the part of J^1_{n-k} on which t_{n-1} takes the value 1. Letting q_1 and q_2 be the two endpoints of one (of the two) connected components of Q_k, we have by Lemmas 1 and 3

$$|f^{n-k}(q_1) - f^{n-k}(q_2)| \leq C_0 \sigma^{(k-N_0)/2} |f^n(q_1) - f^n(q_2)|. \tag{1}$$

We also have

$$\begin{aligned} |f^{n-k}(q_1) - f^{n-k}(q_2)| &\geq \delta_0(1 + \sigma^{(k-1)/2}) - \delta_0(1 + \sigma^{k/2}) \\ &= \delta_0 \sigma^{k/2}(\sigma^{-1/2} - 1), \end{aligned}$$

which combined with (1) gives

$$|f^n(q_1) - f^n(q_2)| \geq C_0^{-1} \sigma^{N_0/2} \delta_0 (\sigma^{-1/2} - 1).$$

On the other hand, since $J^3_n \subset A^{\varepsilon}_{n-1}$ by the construction of J^3_n, taking

$$\varepsilon < C_0^{-1} \sigma^{N_0/2} \delta_0 (\sigma^{-1/2} - 1)$$

we have $J^3_n \cap \{t_{n-1} > 1\} = \emptyset$. This implies $J^1_n \cap \{t_{n-1} \geq 1\} = \emptyset$. □

2.3. *Expansivity.* Recall that by construction, the return time $R(J)$ for J an element of the partition $\mathcal{P}$ of Δ_0, is formed by a certain number n of iterations given by the hyperbolic time of a hyperbolic pre-interval $V_n \supset J$, and a certain number $m \leq N_0$ of additional iterates which is the time it takes to go from $f^n(V_n)$ (which could be anywhere in I) to $f^{n+m}(V_n)$ (which covers Δ_0 completely). By choosing R_0 sufficiently large it then follows from Lemma 3 that there exists a constant $\hat{\lambda} > 1$ and a time n_0 such that for any hyperbolic time $n \geq n_0$ and any point $x \in V_n$ and $1 \leq m \leq N_0$, we have

$$|(f^{n+m})'(x)| \geq \hat{\lambda} > 1.$$

We immediately have the required uniform expansivity property

$$|F'(x)| = |(f^{R(x)})'(x)| \geq \hat{\lambda} > 1.$$

In particular, this implies that for any $x, y \in \Delta_0$ which have the same combinatorics, i.e. which remain in the same elements of the partition $\mathcal{P}$ for some number $s(x, y)$ of iterates of the induced map F, we have

$$|x - y| \leq \hat{\lambda}^{-s(x,y)}. \tag{2}$$

2.4. *Bounded distortion.* The distortion estimate required follows immediately from (2) above and the following more classical formulation of the bounded distortion property.

LEMMA 5. *There exists a constant $D > 0$ such that for any x, y belonging to the same element $J \in \mathcal{P}$ with return time R, we have*

$$\log \left| \frac{F'(x)}{F'(y)} \right| = \log \left| \frac{\left(f^R\right)'(x)}{\left(f^R\right)'(y)} \right| \leq D|f^R(x) - f^R(y)|.$$

Proof. Recall that by construction, the return time $R(J)$ for J an element of the partition $\mathcal{P}$ of Δ_0, is formed by a certain number n of iterations given by the hyperbolic time of a hyperbolic pre-interval $V_n \supset J$, and a certain number $m \leq N_0$ of additional iterates which is the time it takes to go from $f^n(V_n)$ (which could be anywhere in I), to $f^{n+m}(V_n)$ (which covers Δ completely). Some standard formal manipulation based on the chain rule gives

$$\log \left| \frac{(f^R)'(x)}{(f^R)'(y)} \right| = \log \left| \frac{(f^{R-n})'(f^n(x))}{(f^{R-n})'(f^n(y))} \right| + \log \left| \frac{(f^n)'(x)}{(f^n)'(y)} \right|.$$

Since $f^i(x)$ and $f^i(y)$ are uniformly bounded away from $\mathcal{S}$ for $n \leq i \leq R$ (recall Lemma 3), we may write

$$\log \left| \frac{(f^{R-n})'(f^n(x))}{(f^{R-n})'(f^n(y))} \right| \leq B_1|f^R(x) - f^R(y)|$$

where B_1 is some constant not depending on x, y or R. On the other hand, by the construction of V_n (see the proof of Lemma 5.2 in [**2**]), there must be some $z \in V_n$ for which n is a hyperbolic time and such that, for $0 \leq j < n$, the distance from $f^j(z)$ to either $f^j(x)$ or $f^j(y)$ is smaller than $|f^n(x) - f^n(y)|\sigma^{(n-j)/2}$, which is much smaller than $\sigma^{b(n-j)} \leq \operatorname{dist}(f^j(z), \mathcal{S})$. Thus, by (s_2) we have

$$\log \left| \frac{(f^n)'(x)}{(f^n)'(y)} \right| \leq \sum_{j=0}^{n-1} \log \left| \frac{f'(f^j(x))}{f'(f^j(y))} \right| \leq |f^n(x) - f^n(y)| \sum_{j=0}^{n-1} 2B \frac{\sigma^{(n-j)/2}}{\sigma^{b\beta(n-j)}}.$$

Since $b\beta < 1/2$, there must be some $B_2 > 0$ such that

$$\log \left| \frac{(f^n)'(x)}{(f^n)'(y)} \right| \leq B_2|f^n(x) - f^n(y)|.$$

Using again that $f^i(y)$ and $f^i(y)$ are uniformly bounded away from $\mathcal{S}$ for $n \leq i \leq R$ (cf. Lemma 3), it follows that

$$|f^n(x) - f^n(y)| \leq B_2|f^R(x) - f^R(y)|,$$

where B_2 is some constant not depending on x, y or R. This completes the proof of the lemma. □

3. *The statistical argument*

We now come to the main analytic estimates, i.e. the relation between the decay of $|\Delta_n|$ and that of $|\Gamma_n|$.

3.1. *Proportion of points in $A_{n-1} \cap H_n$ returning at time of order n.* In the first part of the argument we use the strategy introduced in [**23**] for uniformly expanding maps, together with the properties of hyperbolic times and give a lower bound for the flow of mass from $A_{n-1} \cap H_n$ to the new intervals of the partition. We start with an auxiliary lemma.

LEMMA 6. *For each $\varepsilon > 0$ there exists $N_\varepsilon > 0$ such that any interval $B \subset I$ with $|B| \geq 2\varepsilon$ contains a hyperbolic pre-interval V_n with $n \leq N_\varepsilon$.*

Proof. Given $\varepsilon > 0$ and an interval $(z - \varepsilon, z + \varepsilon)$, choose N'_ε large enough so that any hyperbolic pre-interval V_n associated to a hyperbolic time $n \geq N'_\varepsilon$ will be contained in an interval of length $\varepsilon/5$ ($N'_\varepsilon \sim \kappa^{-1} \log(10\varepsilon^{-1})$). Now note that almost every point has an infinite number of hyperbolic times and therefore we have that

$$\left| I \setminus \bigcup_{j=N'_\varepsilon}^{n} H_j \right| \to 0 \quad \text{as } n \to \infty.$$

Therefore, it is possible to choose

$$N_\varepsilon = \min \left\{ n \geq N'_\varepsilon : \left| I \setminus \bigcup_{j=N'_\varepsilon}^{n} H_j \right| \leq \varepsilon/10 \right\}.$$

This ensures that there is a point $\hat{x} \in (z - \varepsilon/2, z + \varepsilon/2)$ with a hyperbolic time $n \leq N_\varepsilon$ and associated hyperbolic pre-ball $V_n(x) \subset (z - \varepsilon, z + \varepsilon)$. □

PROPOSITION 1. *There exist $c_0 > 0$ and $N = N(\varepsilon)$ such that for every $n \geq 1$*

$$\left| \bigcup_{i=0}^{N} \{R = n + i\} \right| \geq c_0 |A_{n-1} \cap H_n|.$$

Proof. Take $r = 5\delta_0 C_0^{N_0}$, where N_0 and C_0 are given by Lemma 3. Let $\{z_j\}$ be a maximal set in $f^n(A_{n-1} \cap H_n)$ with the property that the intervals $(z_j - r, z_j + r)$ are pairwise disjoint. By maximality we have

$$\bigcup_j (z_j - 2r, z_j + 2r) \supset f^n(A_{n-1} \cap H_n).$$

Let x_j be a point in H_n such that $f^n(x_j) = z_j$ and consider the hyperbolic pre-interval $V_n(x_j)$ associated to x_j. Observe that f^n sends $V_n(x_j)$ diffeomorphically onto an interval of length $2\delta_1$ centered at z_j, as in Lemma 1. In the following, given $B \subset (z_j - \delta_1, z_j + \delta_1)$, we simply denote $(f^n | V_n(x_j))^{-1}(B)$ by $f^{-n}(B)$.

Our aim now is to prove that $f^{-n}((z_j - r, z_j + r))$ contains some component of $\{R = n + k_j\}$ with $0 \leq k_j \leq N_\varepsilon + N_0$. We start by showing that

$$t_{n+k_j} | f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) > 0, \quad \text{for some } 0 \leq k_j \leq N_\varepsilon + N_0. \tag{3}$$

Assume for a contradiction that $t_{n+k_j} | f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) = 0$, for all $0 \leq k_j \leq N_\varepsilon + N_0$. This implies that

$$f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) \subset A^{\varepsilon}_{n+k_j}, \quad \text{for all } 0 \leq k_j \leq N_\varepsilon + N_0. \tag{4}$$

Using Lemma 6 we may find a hyperbolic pre-interval $V_m \subset (z_j - \varepsilon, z_j + \varepsilon)$ with $0 < m \leq N_\varepsilon$. Now, since $B = f^m(V_m)$ is an interval of length $2\delta_1$, it follows from Lemma 3 that there is some $V \subset B$ and $m' \leq N_0$ with $f^{m'}(V) = \Delta_0$. Thus, taking $k_j = m + m'$ we have that $0 \leq k_j \leq N_\varepsilon + N_0$ and $f^{-n}(V_m)$ is an element of $\{R = n + k_j\}$ inside $f^{-n}((z_j - \varepsilon, z_j + \varepsilon))$. Note that, by (4), the interval $f^{-n}(V_m)$ satisfies the requirements for the construction of partition elements described earlier. This contradicts the fact that $t_{n+k_j} | f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) = 0$ for all $0 \leq k_j \leq N_\varepsilon + N_0$ and so (3) holds.

Let k_j be the smallest integer $0 \leq k_j \leq N_\varepsilon + N_0$ for which $t_{n+k_j} | f^{-n}(B(z_j, \varepsilon)) > 0$. Since

$$f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) \subset A^\varepsilon_{n-1} \subset \{t_{n-1} \leq 1\},$$

there must be some element $J^0_{n+k_j}(j)$ of $\{R = n + k_j\}$ for which

$$f^{-n}((z_j - \varepsilon, z_j + \varepsilon)) \cap J^1_{n+k_j}(j) \neq \emptyset.$$

Recall that by definition f^{n+k_j} sends $J^1_{n+k_j}(j)$ diffeomorphically onto Δ^1_0, the ball of radius $(1 + s)\delta_0$ around p. From time n to $n + k_j$ we may have some final 'bad' period of length at most N_0 where the derivative of f may contract, however, being bounded from below by $1/C_0$ in each step. Thus, the diameter of $f^n(J^1_{n+k_j}(j))$ is at most $4\delta_0 C_0^{N_0}$. Since $(z_j - \varepsilon, z_j + \varepsilon)$ intersects $f^n(J^1_{n+k_j}(j))$ and $\varepsilon < \delta_0 < \delta_0 C_0^{N_0}$, we have by the definition of r

$$f^{-n}((z_j - r, z_j + r)) \supset J^0_{n+k_j}(j).$$

Thus we have shown that $f^{-n}((z_j - r, z_j + r))$ contains some component of $\{R = n + k_j\}$ with $0 \leq k_j \leq N_\varepsilon + N_0$. Moreover, since n is a hyperbolic time for x_j, we have by the distortion control given by Lemma 1

$$\frac{|f^{-n}((z_j - 2r, z_j + 2r))|}{|f^{-n}((z_j - r, z_j + r))|} \leq D_1 \frac{4r}{2r} = 2D_1 \tag{5}$$

and

$$\frac{|f^{-n}((z_j - r, z_j + r))|}{|J^0_{n+k_j}(j)|} \leq D_1 \frac{2r}{|f^n(J^0_{n+k_j}(j))|}. \tag{6}$$

Here we are implicitly assuming that $2r < \delta_1$. This can be done just by taking δ_0 small enough. Note that the estimates on N_0 and C_0 improve when we diminish δ_0.

From time n to time $n + k_j$ we have at most $k_j = m_1 + m_2$ iterates with $m_1 \leq N_\varepsilon$, $m_2 \leq N_0$ and $f^n(J^0_{n+k_j}(j))$ contains some point $w_j \in H_{m_1}$. By the definition of (σ, δ)-hyperbolic time we have that $\mathrm{dist}_\delta(f^i(x), \mathcal{S}) \geq \sigma^{bN_\varepsilon}$ for every $0 \leq i \leq m_1$, which by the uniform distortion control implies that there is some constant $D = D(\varepsilon) > 0$ such that $|(f^i)'(x)| \leq D$ for $0 \leq i \leq m_1$ and $x \in f^n(J^0_{n+k_j}(j))$. On the other hand, since the first N_0 preimages of Δ_0 are uniformly bounded away from $\mathcal{S}$ we also have some $D' > 0$ such that $|(f^i)'(x)| \leq D'$ for every $0 \leq i \leq m_2$ and x belonging to any i preimage of Δ_0. Hence,

$$|f^n(J^0_{n+k_j}(j))| \geq \frac{1}{DD'} |\Delta_0|,$$

which combined with (6) gives

$$|f^{-n}((z_j - r, z_j + r))| \leq C |J^0_{n+k_j}(j)|,$$

with C only depending on D_1, D, D' and δ_0. We also deduce from (5) that

$$|f^{-n}((z_j - 2r, z_j + 2r))| \le C'|f^{-n}((z_j - r, z_j + r))|$$

with C' only depending on D_1.

Finally, let us compare the measure of the sets $\bigcup_{i=0}^{N}\{R = n + i\}$ and $A_{n-1} \cap H_n$. We have

$$|A_{n-1} \cap H_n| \le \sum_j |f^{-n}((z_j - 2r, z_j + 2r))| \le C' \sum_j |f^{-n}((z_j - r, z_j + r))|.$$

On the other hand, by the disjointness of the intervals $(z_j - r, z_j + r)$ we have

$$\sum_j |f^{-n}((z_j - r, z_j + r))| \le C \sum_j |J^0_{n+k_j}(j)| \le C\left|\bigcup_{i=0}^{N}\{R = n + i\}\right|.$$

We just have to take $c_0^{-1} = CC'$. □

3.2. *Relative proportion of A_n and B_n in Δ_n.* We now prove a couple of lemmas. The first gives a lower bound for the flow of mass from B_{n-1} to A_n and second gives a lower bound for the flow of mass from A_{n-1} to B_n and $\{R = n\}$.

LEMMA 7. *There exists $a_1 > 0$ such that for every $n \ge 1$*

$$|B_{n-1} \cap A_n| \ge a_1|B_{n-1}|.$$

Moreover, a_1 is bounded away from zero independently from δ_0.

Proof. It is enough to see this for each component of B_{n-1} at a time. Let C be a component of B_{n-1} and Q be its outer interval corresponding to $t_{n-1} = 1$. Observe that by Lemma 4 we have $Q = C \cap A_n$. Moreover, there must be some $k < n$ and a component J_k^0 of $\{R = k\}$ such that f^k maps C diffeomorphically onto $\bigcup_{i=k}^{\infty} I_i$ and Q onto I_k, both with uniform bounded distortion (not depending on δ_0 or n). Thus, it is sufficient to compare the lengths of $\bigcup_{i=k}^{\infty} I_i$ and I_k. We have

$$\frac{|I_k|}{\left|\bigcup_{i=k}^{\infty} I_i\right|} = \frac{\delta_0(1 + \sigma^{(k-1)/2}) - \delta_0(1 + \sigma^{k/2})}{\delta_0(1 + \sigma^{(k-1)/2}) - \delta_0} \approx 1 - \sigma^{1/2}.$$

Clearly this proportion does not depend on δ_0. □

The second item of the lemma below is apparently counterintuitive, since our main goal is to make the points in Δ_0 have small return times. However, this is needed to keep $|A_n|$ uniformly much bigger than $|B_n|$. This will help us in the statistical estimates of the last section.

LEMMA 8. *There exist $b_1 = b_1(\delta_0) > 0$ and $c_1 = c_1(\delta_0) > 0$ with $b_1 + c_1 < 1$ such that for every $n \ge 1$:*

(1) $|A_{n-1} \cap B_n| \le b_1|A_{n-1}|$;

(2) $|A_{n-1} \cap \{R = n\}| \le c_1|A_{n-1}|$.

Moreover, $b_1 \to 0$ and $c_1 \to 0$ as $\delta_0 \to 0$.

Proof. It is enough to prove this for each neighborhood of a component J_n^0 of $\{R = n\}$. Observe that by construction we have $J_n^3 \subset A_{n-1}^{\varepsilon}$, which means that $J_n^2 \subset A_{n-1}$, because $\varepsilon < \delta_0 < \sqrt{\delta_0}$. Using the uniform bounded distortion of f^n on J_n^3 given by Lemmas 6 and 3, we obtain

$$\frac{|J_n^1 \setminus J_n^0|}{|J_n^2 \setminus J_n^1|} \approx \frac{|\Delta_0^1 \setminus \Delta_0^0|}{|\Delta_0^2 \setminus \Delta_0^1|} \approx \frac{\delta_0}{\sqrt{\delta_0}} \ll 1,$$

which gives the first estimate. Moreover,

$$\frac{|J_n^0|}{|J_n^2 \setminus J_n^1|} \approx \frac{|\Delta_0^0|}{|\Delta_0^2 \setminus \Delta_0^1|} \approx \frac{\delta_0}{\sqrt{\delta_0}} \ll 1,$$

and this gives the second estimate. □

The next result is a consequence of the estimates we obtained in the last two lemmas. The proof is essentially the same as in the uniformly hyperbolic case (see [**23**]), although here we need to be more careful with the estimates.

LEMMA 9. *There exists $a_0 = a_0(\delta_0) > 0$ such that for every $n \geq 1$*

$$|B_n| \leq a_0 |A_n|.$$

Moreover, $a_0 \to 0$ as $\delta_0 \to 0$.

Proof. We have by Lemma 8

$$|A_n \cap A_{n-1}| \geq \eta |A_{n-1}| \tag{7}$$

where $\eta = 1 - b_1 - c_1$. Then we define

$$\widehat{a} = \frac{b_1 + c_1}{a_1}, \quad a_0 = \frac{(1 + a_1)b_1 + c_1}{a_1 \eta}.$$

The fact that $a_0 \to 0$ when $\delta_0 \to 0$ is a consequence of $b_1 \to 0$ and $c_1 \to 0$ when $\delta_0 \to 0$ and a_1 being bounded away from zero. Observe that $0 < \eta < 1$ and $\widehat{a} < a_0$. Now the proof of the proposition follows by induction. The result obviously holds for n up to R_0. Assuming that it holds for $n - 1 \geq R_0$, we show that it also holds for n by separately considering the cases $|B_{n-1}| > \widehat{a}|A_{n-1}|$ and $|B_{n-1}| \leq \widehat{a}|A_{n-1}|$.

Assume first that $|B_{n-1}| > \widehat{a}|A_{n-1}|$. We may write

$$|B_{n-1}| = |B_{n-1} \cap A_n| + |B_{n-1} \cap B_n|,$$

which by Lemma 7 gives

$$|B_{n-1} \cap B_n| \leq (1 - a_1)|B_{n-1}|. \tag{8}$$

Since we also have

$$|B_n| = |B_n \cap B_{n-1}| + |B_n \cap A_{n-1}|,$$

it follows from (8) and Lemma 8 that

$$|B_n| \leq (1 - a_1)|B_{n-1}| + b_1 |A_{n-1}|,$$

which according to the case we are considering leads to

$$|B_n| \leq (1-a_1)|B_{n-1}| + \frac{b_1 a_1}{b_1+c_1}|B_{n-1}| < |B_{n-1}|. \tag{9}$$

On the other hand, we have $|A_n| = |A_n \cap A_{n-1}| + |A_n \cap B_{n-1}|$, which together with Lemma 7 and (7) gives $|A_n| \geq \eta|A_{n-1}| + a_0|B_{n-1}|$. Again by the case we are considering we have

$$|A_n| \geq \eta|A_{n-1}| + a_0\widehat{a}|A_{n-1}| \geq |A_{n-1}|. \tag{10}$$

Inequalities (9) and (10) together with the inductive hypothesis yield the result in this first case.

Assume now that $|B_{n-1}| \leq \widehat{a}|A_{n-1}|$. Since we have

$$|B_n| = |B_n \cap B_{n-1}| + |B_n \cap A_{n-1}|,$$

it follows from Lemma 8 that

$$|B_n| \leq |B_{n-1}| + b_1|A_{n-1}|.$$

From (7) we also know that $|A_n| \geq \eta|A_{n-1}|$. Hence

$$\frac{|B_n|}{|A_n|} < \frac{|B_{n-1}| + b_1|A_{n-1}|}{\eta|A_{n-1}|} \leq \frac{\widehat{a}+b_1}{\eta} = a_0,$$

which also gives the result in this case.

Assume now that $|B_{n-1}| \leq \widehat{a}|A_{n-1}|$. Since we have

$$|B_n| = |B_n \cap B_{n-1}| + |B_n \cap A_{n-1}|,$$

it follows from Lemma 8 that

$$|B_n| \leq |B_{n-1}| + b_1|A_{n-1}|.$$

Hence,

$$\frac{|B_n|}{|A_n|} < \frac{|B_{n-1}| + b_1|A_{n-1}|}{\eta|A_{n-1}|} \leq \frac{\widehat{a}+b_1}{\eta} = a_0,$$

which also gives the result in this case. □

Recalling that $\Delta_n = A_n \cup B_n$, it easily follows from Proposition 1 and Lemma 9 that there is some constant $b_0 > 0$ not depending on n such that

$$\left|\bigcup_{i=0}^{N}\{R = n+i\}\right| \geq b_0 \frac{|A_{n-1} \cap H_n|}{|A_{n-1}|}|\Delta_{n-1}|.$$

This immediately implies that

$$|\Delta_{n+N}| \leq \left(1 - b_0\frac{|A_{n-1} \cap H_n|}{|A_{n-1}|}\right)|\Delta_{n-1}|.$$

It is not hard to deduce from this last formula that

$$|\Delta_{n+N}| \leq \exp\left(-\frac{b_0}{N+1}\sum_{j=R_0}^{n}\frac{|A_{j-1} \cap H_j|}{|A_{j-1}|}\right)|\Delta_0|. \tag{11}$$

On the other hand, it follows from Lemmas 8 and 9 that there is some constant $c_2 > 0$ not depending on n such that for every $n \geq 1$,

$$|\Delta_n| \leq c_2|\Delta_{n+1}|. \tag{12}$$

3.3. *The tail estimates.* Recall that $\theta > 0$ was obtained in Lemma 2 and gives a lower bound for the frequency of hyperbolic times; it only depends on the non-uniform expansion coefficient λ and the map f. Let us now derive a consequence of Lemma 2.

COROLLARY 1. *For every $n \geq 1$ and every $A \subset I \setminus \Gamma_n$ with $|A| > 0$ we have*

$$\frac{1}{n}\sum_{j=1}^{n}\frac{|A \cap H_j|}{|A|} \geq \theta.$$

Proof. Take $n \geq 1$ and $A \subset I \setminus \Gamma_n$ with positive Lebesgue measure. Let ξ_n be the measure in $\{1, \ldots, n\}$ defined by $\xi_n(J) = \#J/n$, for each subset J. Then, using Fubini's theorem,

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n}|A \cap H_j| &= \int \left(\int_A \chi(x,i)\,dx\right) d\xi_n(i)\\ &= \int_A \left(\int \chi(x,i)\,d\xi_n(i)\right) dx,\end{aligned}$$

where $\chi(x,i) = 1$ if $x \in H_i$ and $\chi(x,i) = 0$ otherwise. Now, Lemma 2 means that the integral with respect to $d\xi_n$ is larger than $\theta > 0$. So, the last expression above is bounded from below by $\theta|A|$. □

Let $\gamma > 0$ be some positive number (to be determined later) and take $0 < \alpha < (\theta/12)^{\gamma+1}$. Then we choose $\delta_0 > 0$ small so that $a_0 = a(\delta_0) < 2\alpha$. We define, for each $n \geq 1$,

$$E_n = \left\{ j \leq n \colon \frac{|A_{j-1} \cap H_j|}{|A_{j-1}|} < \alpha \right\},$$

and

$$F = \left\{ n \in \mathbb{N} \colon \frac{\#E_n}{n} > 1 - \frac{\theta}{12} \right\}.$$

In the following proposition we establish the relation between the statistics of hyperbolic times and the geometrical structure of sets arising from the construction of the partition described above. In the polynomial case this establishes an essentially optimal link between the rate of decay of the expansion/recurrence function and the rate of decay of correlations. The nature of the argument does not immediately extend to the exponential case.

PROPOSITION 2. *Take any $n \in F$ with $n \geq R_0 > 12/\theta$. If $|A_n| \geq 2|\Gamma_n|$, then there is some $0 < k = k(n) < n$ for which $|A_n|/|A_k| < (k/n)^{\gamma}$.*

Proof. We have for $j \leq n$

$$\frac{|A_n \cap H_j|}{|A_n|} \geq \frac{|A_n \setminus \Gamma_n|}{|A_n|}\frac{|(A_n \setminus \Gamma_n) \cap H_j|}{|A_n \setminus \Gamma_n|} \geq \frac{1}{2}\frac{|(A_n \setminus \Gamma_n) \cap H_j|}{|A_n \setminus \Gamma_n|},$$

which together with the conclusion of Corollary 1 for the set $A_n \setminus \Gamma_n$ gives

$$\frac{1}{n}\sum_{j=1}^{n}\frac{|A_n \cap H_j|}{|A_n|} \geq \frac{\theta}{2}. \tag{13}$$

Let

$$G_n = \left\{ j \in E_n \colon \frac{|A_{j-1}|}{|A_n|} > \frac{\theta}{12\alpha} \right\}.$$

Since $n \in F$, we have

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n} \frac{|A_n \cap H_j|}{|A_n|} &\le \frac{\theta}{12} + \frac{1}{n}\sum_{j\in E_n} \frac{|A_n \cap H_j|}{|A_n|} \\ &\le \frac{\theta}{12} + \frac{1}{n}\sum_{j\in E_n\setminus G_n} \frac{|A_n \cap H_j|}{|A_n|} + \frac{\#G_n}{n}.\end{aligned}$$

Now, for $j \in E_n \setminus G_n$,

$$\begin{aligned}\frac{|A_n \cap H_j|}{|A_n|} &= \frac{|A_n \cap H_j|}{|A_{j-1}|}\frac{|A_{j-1}|}{|A_n|} \\ &\le \left(\frac{|A_n \cap A_{j-1} \cap H_j|}{|A_{j-1}|} + \frac{|(A_n \setminus A_{j-1}) \cap H_j|}{|A_{j-1}|} \right) \frac{|A_{j-1}|}{|A_n|} \\ &\le \left(\frac{|A_{j-1} \cap H_j|}{|A_{j-1}|} + a_0 \right) \frac{\theta}{12\alpha}.\end{aligned}$$

For this last inequality we used the fact that $(A_n \setminus A_{j-1}) \subset B_{j-1}$ and $j \notin G_n$. Hence,

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n} \frac{|A_n \cap H_j|}{|A_n|} &\le \frac{\theta}{12} + \frac{1}{n}\sum_{j\in E_n\setminus G_n} \frac{|A_{j-1} \cap H_j|}{|A_{j-1}|}\frac{\theta}{12\alpha} + a_0\frac{\theta}{12\alpha} + \frac{\#G_n}{n} \\ &< \frac{\theta}{12} + \alpha\frac{\theta}{12\alpha} + a_0\frac{\theta}{12\alpha} + \frac{\#G_n}{n}.\end{aligned}$$

By the choice of a_0 we have that the third term in the last sum above is smaller than $\theta/6$. So, using (13), we obtain

$$\frac{\#G_n}{n} > \frac{\theta}{6}. \tag{14}$$

Now, defining

$$k = \max(G_n) - 1,$$

we have

$$|A_n| < \frac{12\alpha}{\theta}|A_k|.$$

It follows from (14) that $k + 1 > \theta n/6$ and so $k/n > \theta/12$, because $n \ge R_0 > 12/\theta$. Since we have chosen $\alpha < (\theta/12)^{\gamma+1}$, it follows that

$$\left(\frac{k}{n}\right)^{\gamma} > \frac{12}{\theta}\left(\frac{\theta}{12}\right)^{\gamma+1} > \frac{12\alpha}{\theta}.$$

This completes the proof of the result. □

By Lemma 9, we have $|\Delta_n| \le \text{const}|A_n|$ and so it is enough to derive the tail estimate for $|A_n|$ in the place of $|\{R > n\}| = |\Delta_n|$. Given any large integer n, we consider the following situations.

(1) If $n \in \mathbb{N} \setminus F$, then by (11) and (12) we have

$$|\Delta_n| \le c_2^N \exp\left(-\frac{b_0\theta\alpha}{12(N+1)}(n - R_0)\right)|\Delta_0|.$$

(2) If $n \in F$, then we distinguish the following two cases:
   (a) if $|A_n| < 2|\Gamma_n|$, then nothing has to be done;
   (b) if $|A_n| \ge 2|\Gamma_n|$, then we apply Proposition 2 and get some $k_1 < n$ for which

$$|A_n| < \left(\frac{k_1}{n}\right)^\gamma |A_{k_1}|.$$

The only situation we are left to consider is (2)(b). In this case, either k_1 is in situation (1) or (2)(a), or by Proposition 2 we can find $k_2 < k_1$ for which

$$|A_{k_1}| < \left(\frac{k_2}{k_1}\right)^\gamma |A_{k_2}|.$$

Arguing inductively we show that there is a sequence of integers $0 < k_s < \cdots < k_1 < n$ for which one of the following cases eventually holds:

(I)
$$|A_n| < \left(\frac{k_s}{n}\right)^\gamma c_2^N \exp\left(-\frac{b_0\theta\alpha}{12(N+1)}(k_s - R_0)\right)|\Delta_0|;$$

(II)
$$|A_n| < \left(\frac{k_s}{n}\right)^{\gamma|\Gamma_{k_s}|};$$

(III)
$$|A_n| < \left(\frac{R_0}{n}\right)^{\gamma|\Delta_0|}.$$

Case (III) corresponds to falling into situation (2)(b) repeatedly until $k_s \le R_0$. Observe that until now $\gamma > 0$ was arbitrary. So, the only case we are left to consider is (II).

Assume that $|\Gamma_n| \le \mathcal{O}(n^{-\gamma})$ for some $\gamma > 0$. In this case there must be some $C > 0$ such that $k^\gamma|\Gamma_k| \le C$ for all $k \in \mathbb{N}$, which applied to k_s in case (II) leads to $|A_n| \le \mathcal{O}(n^{-\gamma})$.

Acknowledgements. This work was carried out at the Federal University of Bahia, University of Porto and Imperial College, London. It was partially supported by CMUP, PRODYN, SAPIENS and UFBA. We would like to thank the referee for a careful reading of the paper and several useful suggestions.

References

[1] J. F. Alves. SRB measures for non-hyperbolic systems with multidimensional expansion. *Ann. Sci. École Norm. Sup.* **33**(4) (2000), 1–32.

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

[3] J. F. Alves, S. Luzzatto and V. Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré, Anal. Non Linéaire* to appear.
[4] J. F. Alves and M. Viana. Statistical stability for robust classes of maps with non-uniform expansion. *Ergod. Th. & Dynam. Sys.* **22** (2002), 1–32.
[5] V. Baladi and M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. *Ann. Sci. École Norm. Sup*. **29** (1996), 483–517.
[6] H. Bruin, S. Luzzatto and S. van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup*. **36** (2003), 621–646.
[7] C. Bonatti and M. Viana. SRB measures for partially hyperbolic systems with mostly contracting central direction. *Israel J. Math.* **115** (2000), 157–193.
[8] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math*. **29** (1975), 181–202.
[9] R. Bowen. Markov partitions for Axiom diffeomorphisms. *Amer. J. Math*. **92** (1970), 725–747.
[10] X. Bressaud. Subshifts on an infinite alphabet. *Ergod. Th. & Dynam. Sys.* **19**(5) (1999), 1175–1200.
[11] M. Holland. Slowly mixing systems and intermittency maps. *Preprint* 2002.
[12] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Preprint* 1999.
[13] G. Keller. Exponents, attractors and Hopf decompositions for interval maps. *Ergod. Th. & Dynam. Sys.* **10**, (1990), 717–744.
[14] T. Krüger and S. Troubetzkoy. Markov partitions and shadowing for non-uniformly hyperbolic systems with singularities. *Ergod. Th. & Dynam. Sys.* **12**(3) (1992), 487–508.
[15] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergod. Th. & Dynam. Sys.* **19** (1999), 671–685.
[16] C. Liverani. Central limit theorem for deterministic systems. Eds. F. Ledrappier, J. Lewowicz and S. Newhouse. *Proc. Int. Conf. on Dynamical Systems (Montevideo 1995)—A Tribute to Ricardo Mañé*, Pitman, Boston, 1996, pp. 56–75.
[17] S. Luzzatto and W. Tucker. Non-uniformly expanding dynamics in maps with singularities and criticalities. *Publ. Math. IHES* **89** (1999), 180–226.
[18] S. Luzzatto. Absolutely continuous invariant measures and decay of correlations for nonuniformly expanding maps. *Handbook of Dynamical Systems*, Vol. 1B. Eds. B. Hasselblatt and A. Katok. Elsevier to appear.
[19] V. Maume-Deschamps. Correlation decay for Markov maps on a countable state space. *Ergod. Th. & Dynam. Sys.* **21** (2001), 165–196.
[20] O. Sarig. Subexponential decay of correlations. *Invent. Math.* **150**(3) (2002), 629–653.
[21] Y. Sinai. Markov partitions and U-diffeomorphisms. *Funk. Anal. Prilozen 2*. **32** (1968), 70–80.
[22] L.-S. Young. Decay of correlations for certain quadratic maps. *Comm. Math. Phys*. **146** (1992), 123–138.
[23] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math*. **147** (1998), 585–650.
[24] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math*. **110** (1999), 153–188.