

# STATISTICAL PROPERTIES OF ONE-DIMENSIONAL MAPS WITH CRITICAL POINTS AND SINGULARITIES

K. DÍAZ-ORDAZ, M. P. HOLLAND, AND S. LUZZATTO

ABSTRACT. We prove that a class of one-dimensional maps with an arbitrary number of non-degenerate critical and singular points admits an induced Markov tower with exponential return time asymptotics. In particular the map has an absolutely continuous invariant probability measure with exponential decay of correlations for Hölder observations.

## 1. INTRODUCTION AND STATEMENT OF RESULTS

It has been recognized that interval maps can exhibit a great degree of dynamical complexity. Indeed, within this class are the first rigorous examples of *deterministic dynamical systems* which exhibit *stochastic* behaviour. These maps can be characterized in terms of the existence of a *mixing* absolutely continuous (with respect to Lebesgue) invariant probability measure (*acip*). In this paper we explore some geometric conditions which give rise to *acip*'s and study their statistical properties.

1.1. **General background.** Early examples of maps where an explicit formula for the *acip* can be found include the Gauss map [15] and the Ulam-von Neumann transformation [48]. However, finding an explicit form for the density of the corresponding *acip* is not possible in most cases. Instead attention has focussed on the existence of an *acip* by giving sufficient conditions satisfied by certain classes of transformations.

Early work focussed on smooth uniformly expanding systems [18, 27, 40, 42] with the last couple of decades really seeing several developments in the direction of relaxing either the uniform expansivity assumption, by allowing critical points, or the smoothness assumption, by allowing discontinuities with possibly unbounded derivative.

In this paper we make a further step along these lines by considering maps which have both critical points and discontinuities. Before stating our results we give a brief review of the state of the field.

1.1.1. *Maps with critical points.* In the context of unimodal maps (maps with a single critical point  $c$ ) the existence of an *acip* can be deduced from assumptions on the derivative growth and recurrence of the post-critical orbit [8, 11, 21, 34–36, 41]. If  $|(f^n)'(f(c))| \rightarrow \infty$  exponentially fast (known as the Collet-Eckmann condition), it was also shown in [26], and in [50] with an additional condition on the rate of recurrence of the critical point, that the map is stochastic in a very strong sense: the *acip* is mixing and exhibits exponential decay of correlations for Hölder continuous

---

*Date:* 23 June 2006.

2000 *Mathematics Subject Classification.* 37D50, 37A25.

In writing this paper K. Díaz-Ordaz acknowledges the support of CONACYT Mexico. M. Holland acknowledges the support of the EPSRC, grant no. GR/S11862/01. S. Luzzatto acknowledges support of EPSRC grant no. GR/T0969901.

observables. The first generalization of these results to maps with *multiple critical points* was obtained in [7] where the existence of an *acip*  $\mu$  was obtained under the summability condition  $\sum |(f^n)'(f(c))|^{-1/(2\ell-1)} < \infty$  (for every critical point) and the assumption that all critical points have the same order  $\ell$ . Moreover, the techniques used in that paper made it possible to show a direct link between the mixing properties of  $\mu$ , in particular the rate of decay of correlations, and the rate of growth of  $|(f^n)'(f(c))|$ .

1.1.2. *Maps with discontinuities.* There is also an equally significant, and indeed much longer, list of papers concerned with *uniformly expanding* maps with discontinuities (but without critical points) in which similar results to those proved here are obtained, see [29] for references. We remark however that most results allow discontinuities but not unbounded derivatives. This is true for example in one of the applications in [51] in which a strategy similar to ours is applied. As far as we know the exponential decay of correlations for expanding Lorenz-like maps, i.e. maps with unbounded derivatives such as the ones we consider here but with no critical points, follows from arguments in [18, 25], see also [49]. More recently these results have been extended in [14] to obtain estimates for the decay of correlations for observables which are not Hölder continuous.

1.1.3. *Maps with critical points and discontinuities.* The main aim of this paper is to allow for the co-existence of *multiple critical points* (with possibly different critical orders) *and multiple singularities* (points at which the map may be discontinuous and the derivative may be discontinuous or unbounded), see Figure 1. For such maps we establish sufficient conditions for the existence of an ergodic *acip* and for exponential mixing. We remark that although singularities contribute to the “expansivity” of the system, they also give rise to significant technical issues, in particular related to distortion control. A noteworthy aspect of our argument is that we introduce a unified formalism and technique for dealing with both critical and singular points.

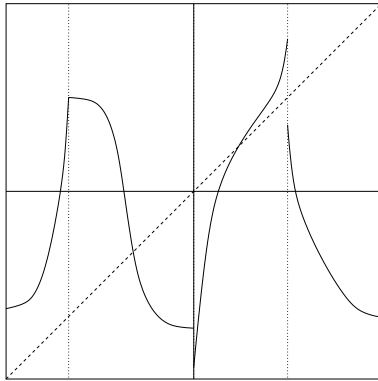


FIGURE 1. A map with a finite number of critical and singular points

Another class of results which should be mentioned also applies to very general maps with non-degenerate critical and singular points but under quite different dynamical assumptions. More specifically the expansivity and recurrence conditions are assumed to hold *asymptotically for Lebesgue almost every point* and not necessarily for the critical points. It is then possible to show the existence of an ergodic *acip* [3] and to obtain bounds for the rate of decay of correlations [4, 17].

1.1.4. *Induced Markov maps.* To obtain the required statistical properties we construct a countable partition on a reference set and analyse the recurrence time statistics through the construction of an induced Markov map. This strategy, which has already been applied successfully in several contexts [4, 5, 7, 14, 17, 19], is motivated by [51, 52] where a direct relationship between the recurrence time statistics and mixing rates is established. A similar construction is also carried out for general unimodal maps in [12].

Inducing techniques yielding countable Markov systems have been applied in other settings too. Jakobson has applied this idea extensively in the context of parameter exclusion arguments for one-dimensional maps [22, 23]. For systems admitting infinite ergodic measures, see [1] for a discussion on *Gibbs Markov maps* and their statistical properties. For maps admitting neutral fixed points see the work of [47], where inducing schemes are used to analyse the properties of the ergodic density. In connection with metric number theory and continued fraction maps, inducing schemes are used in the work of [46]. A theory of *Markov fibred systems* is developed in [2], where inducing schemes are used to provide results on Central Limit Theorems for a class of rational maps admitting neutral fixed points. An approach for analysing mixing rates using the Gibbs-Markov formalism has been developed in [16, 44]. Here, renewal theory techniques are applied to establish sharp polynomial decay of correlations for Markov systems.

A particularly interesting motivation for the construction of induced Markov maps is related to the development of a *thermodynamic formalism* for nonuniformly expanding maps based precisely on the geometrical structure of (induced) Markov maps. This is based on some recent progress on the thermodynamic formalism for abstract countable shift spaces [9, 24, 32, 43, 45], recent results on the relation between invariant measures for the system and invariant measures for the Markov tower extension [54], and the application of these results to specific classes of systems which admit induced Markov maps such as those constructed in this paper [38, 39, 53].

**1.2. Non-degenerate critical and singular points.** We now give the precise definition of the class of maps we consider. Let  $J$  be a compact interval and  $f : J \rightarrow J$  a  $C^2$  local diffeomorphism outside a finite set  $\mathcal{C} \subset \text{int}(J)$ , of *non-degenerate critical and singular* points. These are points at which  $f$  may be discontinuous or the derivative of  $f$  may vanish or be infinite. In order to treat all possibilities in a formally unified way we consider  $\lim_{x \rightarrow c^-} f(x)$  and  $\lim_{x \rightarrow c^+} f(x)$  as distinct *critical values*, thus implicitly thinking of  $c^+$  and  $c^-$  as distinct critical points. When referring to a neighbourhood of a critical point, we shall always be referring to the appropriate one-sided neighbourhood of that point. We say that the critical points are *non-degenerate* if there exists  $C > 0$  and for each  $c$  there exists a constant  $\ell_c \in (0, \infty)$  such that for each  $x$  in a neighbourhood of  $c$  we have

$$C^{-1}|x - c|^{\ell_c} \leq |f(x) - f(c)| \leq C|x - c|^{\ell_c}, \quad (1)$$

and in addition we assume that

$$C^{-1}|x - c|^{\ell_c - 1} \leq |f'(x)| \leq C|x - c|^{\ell_c - 1}, \quad (2)$$

$$C^{-1}|x - c|^{\ell_c - 2} \leq |f''(x)| \leq C|x - c|^{\ell_c - 2}. \quad (3)$$

We write

$$\mathcal{C}_c = \{c : \ell_c \geq 1\} \quad \text{and} \quad \mathcal{C}_s = \{c : 0 < \ell_c < 1\},$$

to denote the set of critical and singular points respectively. Notice that two “distinct” points  $c \in \mathcal{C}_c$  and  $s \in \mathcal{C}_s$  may actually correspond to the same point in  $J$  for which the derivative tends to zero from one side, and infinity from the other. We let

$$\ell = \max_{c \in \mathcal{C}_c} \{\ell_c\} \quad \text{and} \quad \ell^* = \max_{c \in \mathcal{C}_s} \{\ell_c\}.$$

When there is no possibility of confusion we will often use the term “critical point” to refer to a point of  $\mathcal{C}$  without necessarily specifying if  $c$  is really a critical point in the traditional sense with  $\ell_c > 1$  or whether it is a singular point with  $\ell_c \in (0, 1)$  or a “neutral” point with  $\ell_c = 1$ . For ease of exposition we assume the derivative of  $f$  at the points of discontinuity is either unbounded or zero. To accommodate bounded derivatives, we would have to slightly modify our argument to include the case of a return to a region where there is a bounded discontinuity. The derivative growth and distortion estimates would not be affected by such bounded discontinuities.

For any  $x$  let

$$\mathfrak{D}(x) = \min_{c \in \mathcal{C}} |x - c|$$

denote the distance of  $x$  from the nearest critical point, and for small  $\delta > 0$ , let

$$\Delta = \{x : \mathfrak{D}(x) \leq \delta\}$$

denote a  $\delta$ -neighbourhood of  $\mathcal{C}$ . For an arbitrary interval  $\omega$  we shall also use the notation

$$\mathfrak{D}(\omega) = \sup_{x \in \omega} \{\mathfrak{D}(x)\}$$

to denote the distance of  $\omega$  from the critical set.

**1.3. Dynamical assumptions.** For all initial values  $x \in J$  we let  $x_k = f^k(x)$  with  $k \in \mathbb{N}$  denote the iterates of  $x$ . We formulate the following three conditions concerning respectively the expansivity of  $f$  outside  $\Delta$ , the derivative growth and recurrence to  $\mathcal{C}$  of the *bona fide* critical points, and a transitivity condition on the critical orbits.

**(H1) Expansion outside  $\Delta$ :** *There exist  $\lambda > 0$  and  $\kappa > 0$  such that for every  $x$  and  $n \geq 1$  such that  $x_0 = x, \dots, x_{n-1} = f^{n-1}(x) \notin \Delta$  we have*

$$|(f^n)'(x)| \geq \kappa \delta e^{\lambda n}.$$

*Moreover, if  $x_0 \in f(\Delta)$  or  $x_n \in \Delta$  we have*

$$|(f^n)'(x)| \geq \kappa e^{\lambda n}.$$

**(H2) Bounded recurrence and derivative growth along critical orbits:**

*There exists  $\alpha > 0$  and  $\Lambda > 0$ , such that for all  $c \in \mathcal{C}_c$  and  $\forall k \geq 1$  we have*

$$\mathfrak{D}(c_k) \geq \delta e^{-\alpha k} \quad \text{and} \quad |(f^k)'(c_1)| \geq e^{\Lambda k}.$$

**(H3) Density of preimages:** *There exists  $c^*$  in  $\mathcal{C}$  whose preimages are dense in a maximal  $\hat{J} \subset J$ , where  $\hat{J}$  is a union of intervals (note that the maximal property of  $\hat{J}$  implies  $f^{-1}\hat{J} = \hat{J}$ ). In addition, this set of preimages does not contain any other point in  $\mathcal{C}$ .*

We suppose throughout that

*$f$  has a finite number of non-degenerate critical and singular points and satisfies conditions (H1)-(H3) for sufficiently small constants  $\alpha > 0$  and  $\delta > 0$  in relation to  $\Lambda, \lambda$  and  $\kappa$ .*

It was shown in [30] that these conditions are satisfied for a large (positive measure) set of parameters in an open class of one-parameter families of *Lorenz-like maps with singularities and criticalities*, in particular, our results apply to such maps. Of course they are also satisfied by many smooth maps but in these cases the results are already known [7].

**1.4. Markov structures.** Consider the system  $(f, \hat{J}, m)$ , where  $m$ , the reference measure is taken to be Lebesgue measure. The main result of this paper is that conditions (H1) to (H3) imply the existence of an induced full branched Markov map with an exponential tail of the return time function.

**Theorem 1.** *There exists a (one-sided) neighbourhood  $\Delta^* \subset \hat{J}$  of the critical point  $c^*$ , a countable partition<sup>1</sup>  $\mathcal{Q}$  of  $\Delta^*$  into subintervals, a function  $T : \Delta^* \rightarrow \mathbb{N}$  defined almost everywhere and constant on elements of the partition  $\mathcal{Q}$ , and constants  $C, \tilde{D}, \gamma, > 0$  such that for all  $\omega \in \mathcal{Q}$  and  $T = T(\omega)$  the map  $f^T : \omega \rightarrow \Delta^*$  is a  $C^2$  diffeomorphism and satisfies the following bounded distortion property: for all  $x, y \in \omega$*

$$\left| \frac{(f^T)'(x)}{(f^T)'(y)} - 1 \right| \leq \tilde{D} |f^T(x) - f^T(y)|.$$

Moreover, the “return time function”  $T$  has an exponentially decreasing tail:

$$|\{T > n\}| < Ce^{-\gamma n}.$$

Our construction also gives a couple of other interesting properties which will be used in the applications of the theorem. Namely, the induced map is uniformly expanding in the sense that there exists some  $\lambda' > 1$  such that for all  $\omega \in \mathcal{Q}$  and all  $x, y \in \omega$

$$|f^T(x) - f^T(y)| \geq \lambda' |x - y|$$

and satisfies a bounded contraction property in the sense that there exists a constant  $K > 0$  such that for all  $\omega \in \mathcal{Q}$ ,  $x, y \in \omega$  and  $1 \leq k < T = T(\omega)$

$$|f^k(x) - f^k(y)| \leq K |f^T(x) - f^T(y)|.$$

**1.5. Statistical properties.** Recent results of Young [52] link the rate of decay of the tail of the return times for an induced full branched Markov map to several statistical properties of the original system. Combining these general results with our main theorem and estimates we therefore obtain the following results.

**Theorem 2.** *There exists an absolutely continuous  $f$ -invariant probability measure  $\mu$  which is ergodic and supported on  $\hat{J}$ . Moreover  $f$  is non-uniformly expanding in the sense that the Lyapunov exponent of  $\mu$  is positive, i.e.  $\int \log |f'(x)| d\mu(x) > 0$ .*

The fact that the measure  $\mu$  has positive Lyapunov exponent implies a degree of *sensitive dependence on initial conditions* and *stochastic-like* behaviour. Indeed, in this case we can show that the dynamics is *stochastic* in a more precise and quantifiable way. We recall that a measure  $\mu$  is mixing if

$$|\mu(A \cap f^{-n}(B)) - \mu(A)\mu(B)| \rightarrow 0$$

---

<sup>1</sup>Here and for the rest of the paper we will always talk about partitions with the implicit understanding that we are referring to partitions mod 0, i.e. up to a set of zero Lebesgue measure.

as  $n \rightarrow \infty$ , for all measurable sets  $A, B$ . This corresponds to a property of asymptotic “loss of memory”. To quantify the *speed of mixing* we define, for two arbitrary  $L^2(\mu)$  functions  $\phi, \psi : J \rightarrow \mathbb{R}$ , the *correlation function*

$$C_n(\phi, \psi, \mu) = \left| \int \psi(\phi \circ f^n) d\mu - \int \psi d\mu \int \phi d\mu \right|.$$

Notice that if  $\phi, \psi$  are characteristic functions of measurable sets  $A, B$  this is exactly the quantity given above in the definition of mixing. Indeed, using standard approximation arguments, it is well known that  $C_n \rightarrow 0$  if  $\mu$  is mixing. However, in general it is not possible to obtain a uniform bound for the rate of decay of  $C_n(\phi, \psi, \mu)$  if both observables belong to a class of functions as big as  $L^2(\mu)$  or even  $L^\infty(\mu)$  which includes characteristic functions, and to obtain concrete estimates it is necessary to restrict oneself to a smaller class of functions.

**Theorem 3.** *There exists a  $k \geq 1$  such that  $(f^k, \hat{J}, \mu)$  has exponential decay of correlations for functions  $\phi \in L^\infty$  and  $\psi$  Hölder continuous.*

The exponential rate of decay represents a particularly strong form of mixing and indicates that, notwithstanding the presence of critical points and the lack of smoothness due to the presence of singularities, the system is, statistically, very similar to a uniformly expanding system. Technically, this is a consequence of the bounded recurrence and exponential growth conditions (H2) along the critical orbits. Assuming weaker growth and recurrence conditions may give rise to slower rates of mixing. This has been shown to be the case in the smooth case [7] and it would clearly be interesting to generalize that result to allow for the presence of singularities. Another extension of our work could be to study statistical properties of intermittent systems with coexisting critical points and singularities, i.e. to include the presence of a neutral periodic point.

A further natural question concerns the convergence of averages of functions along orbits to their expected values, in particular we can ask if a distributional law like the *Central Limit Theorem* (CLT) holds: for any measurable set  $A \subset \mathbb{R}$  and  $\phi : J \rightarrow \mathbb{R}$  with  $\int \phi d\mu = 0$ , there exists some  $\sigma > 0$  such that

$$\mu \left\{ x \in J : \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \phi \circ f^i(x) \in A \right\} \rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_A e^{-\frac{t^2}{2\sigma^2}} dt, \quad \text{as } n \rightarrow \infty.$$

We have the following result:

**Theorem 4.** *The Central Limit Theorem holds for  $(f, \hat{J}, \mu)$  and any Hölder observable  $\phi$  such that  $\phi \circ f \neq \varphi \circ f - \varphi$  for any  $\varphi$ .*

Notice that Theorem 4 does not depend on the mixing properties of  $f$ , see [33] and [10].

**1.6. Overview of the paper.** In Section 2 we introduce the interval  $\Delta^*$  and begin the combinatorial construction of the induced map. We define a neighbourhood  $\Delta^* \subset \Delta$  (of the critical point  $c^*$ , as in (H3)), with  $|\Delta^*| \ll \delta$ , on which we induce a Markov map. Our basic approach is to iterate  $\Delta^*$  under  $f$  and wait for  $\Delta^*$  to return to  $\Delta$ . Using the expansion and recurrence assumptions (H1) and (H2) we can show that this happens after some finite number of iterations,  $k$  say. If  $f^k(\Delta^*)$  comes close to  $\mathcal{C}_c$ , then the subinterval  $f^k(\Delta^*) \cap \Delta$  will be contracted under iteration. For

intersections with  $\mathcal{C}_s$ , we do not lose expansion, but we fail to obtain bounded distortion on future iterations of  $f^k(\Delta^*)$ ; furthermore  $f^k(\Delta^*)$  may become “cut” by the singularity. To overcome these issues, we introduce a systematic chopping procedure on  $f^k(\Delta^*)$ , using a fixed partition on  $\Delta$ . The resulting chopped up subintervals are iterated independently, and at some later time these subintervals will return back to  $\Delta$ . Using the fixed partition, we introduce a combinatorial method which keeps track of their location in relation to the critical set  $\mathcal{C}$ .

Following this combinatorial construction, a chopping method is devised in such a way that  $f$  (and its iterates) act diffeomorphically, with bounded distortion on each chopped-up subinterval. Moreover, it could be envisaged that components may get cut too fast in the chopping procedure, or fail to grow in size due to frequent returns to  $\mathcal{C}$ . This can happen, but we will show that *on average* there is a tendency to grow in size (at an exponential rate).

Sections 3 and 4 contain two fundamental technical results, one on the recovery of the loss of expansion (due to the small derivative) for returns to  $\mathcal{C}_c$  by shadowing the critical orbit, and another on some global distortion bounds which follow from the combinatorial construction. In the core Section 5, we estimate how long it takes on average for a subinterval to reach *large scale* and show that intervals grow to large scale exponentially fast. In Section 6 we collect all our results to prove an exponential tail estimate on the return time function. Finally in Section 7 we explain how the properties of the induced Markov map imply the statistical estimates given in the other theorems.

## 2. THE INDUCED MAP

In this section we give the complete algorithm for the construction of the induced Markov map as required by the statement in Theorem 1. The fact that this algorithm successfully produces an induced Markov map with the required properties is not immediate and follows from the estimates in the following sections.

**2.1. The critical partition.** For each  $c \in \mathcal{C}$  and for any integer  $r \geq 1$  we let

$$I_r(c) = [c + e^{-r}, c + e^{-r+1}] \quad \text{and} \quad I_{-r}(c) = (c - e^{-r+1}, c - e^{-r}].$$

We suppose without loss of generality that  $r_\delta = \log \delta^{-1} \in \mathbb{N}$  and, each  $c \in \mathcal{C}$  let

$$\Delta_c = \begin{cases} \{c\} \cup \bigcup_{r \geq r_\delta+1} I_r(c), & \text{if } c = c^+, \\ \{c\} \cup \bigcup_{r \leq -r_\delta-1} I_r(c), & \text{if } c = c^- \end{cases}$$

and

$$\hat{\Delta}_c = \begin{cases} \{c\} \cup \bigcup_{r \geq r_\delta} I_r(c), & \text{if } c = c^+, \\ \{c\} \cup \bigcup_{r \leq -r_\delta} I_r(c), & \text{if } c = c^-. \end{cases}$$

Notice that  $\hat{\Delta}_c$  is just  $\Delta_c$  union an extra interval of the form  $I_{\pm r_\delta}$ . We further subdivide each  $I_r \subset \Delta$  (and not the additional  $I_{\pm r_\delta} \subset \hat{\Delta} \setminus \Delta$ ) into  $r^2$  intervals  $I_{r,j}$ ,  $j \in [1, r^2]$  of equal length, this defines the *critical partition*  $\mathcal{I}$  of  $\Delta$ . Finally, for each  $r \geq r_\delta + 1$ , and  $j \in [1, r^2]$  we let  $\hat{I}_r$  denote the union of  $I_r$  and its two neighbouring intervals. In particular, if  $I_{r,j} = I_{r_\delta+1, (r_\delta+1)^2}$  is one of the two extreme intervals of  $\Delta$ , then  $\hat{I}_{r,j}$  denotes the union of this interval with the adjacent intervals  $I_{r,j-1}$  and  $I_{r_\delta}$  (which has not been subdivided into smaller subintervals).

**2.2. The binding period.** Using the critical partition defined above, we formalize, following [6], the notion of a *binding period* during which points in the critical region  $\Delta$  shadow the orbit of the critical point. For each  $r \geq r_\delta + 1$ ,  $I_r \in \mathcal{I}$  belonging to the component of  $\Delta$  containing a critical point  $c \in \mathcal{C}$ , we define

$$p(r) = \begin{cases} 0 & \text{if } c \in \mathcal{C}_s \\ \max\{k : |f^{j+1}(x) - f^{j+1}(c)| \leq \delta e^{-2\alpha_j} \forall x \in \hat{I}_r, \forall j \leq k\} & \text{if } c \in \mathcal{C}_c. \end{cases}$$

We shall show in Section 3 that the binding period is long enough for the orbit of some point  $x$  close to a critical point  $c \in \mathcal{C}_c$  to build up enough derivative growth to more than compensate the small derivative coming from its proximity to  $c$  in its starting position.

**2.3. Escape times.** Let  $I \subset \hat{J}$  be an arbitrary interval with  $|I| < \delta$ . We construct a countable partition  $\mathcal{P} = \mathcal{P}(I)$  of  $I$  into subintervals, which we call the *escape partition* of  $I$ , and a stopping time function  $E : I \rightarrow \mathbb{N}$  constant on elements of  $\mathcal{P}$ . Each element  $\omega \in \mathcal{P}$  has some combinatorial information attached to its orbit up to time  $E(\omega)$  and satisfies

$$|f^{E(\omega)}(\omega)| \geq \delta.$$

We define the construction inductively as follows. Fix  $n \geq 1$  and suppose that a certain set of subintervals of  $I$  have been defined for which  $E < n$ . Let  $\omega$  be a component of the complement of the set  $\{x \in I : E(x) < n\}$ .

*Inductive assumptions.* We suppose inductively that the following combinatorial information is also available, the meaning of which will become clear when the general inductive step of the construction is explained below:

- every iterate  $i = 1, \dots, n$  is classified as either a *free* iterate or a *bound* iterate for  $\omega$ .
- the last free iterate before a bound iterate is called either an *essential return* or an *inessential return*.
- associated to each essential and inessential return there is a positive integer called the *return depth*.

We now consider various cases depending on the length and position of the interval  $\omega_n = f^n(\omega)$  and on whether  $n$  is a free or bound iterate for  $\omega$ .

*Escape times.* If  $n$  is a free time for  $\omega$  and  $|\omega_n| \geq \delta$  we say that  $\omega$  has *escaped*. We let  $\omega \in \mathcal{P}$  and define  $E(\omega) = n$ . We call  $\omega_n$  an escape interval.

*Free times.* If  $n$  is a free time for  $\omega$  and  $|\omega_n| < \delta$  we distinguish three cases:

- (1) If  $\omega_n \cap \Delta = \emptyset$ , we basically do nothing: we do not subdivide  $\omega$  further, do not add any combinatorial information, and define  $n + 1$  to be again a free iterate for  $\omega$ .
- (2) If  $\omega_n \cap \Delta \neq \emptyset$  but  $\omega_n$  does not intersect more than two adjacent  $I_{r,j}$ 's, we do not subdivide  $\omega$  further at this moment, but add some combinatorial information in the sense that we say that  $n$  is an *inessential* return time with *return depth*  $r$  equal to the minimum  $r$  of the intervals  $I_{r,j}$  which  $\omega_n$  intersects. Moreover we define all iterates  $j = n + 1, \dots, n + p$  as *bound iterates* for  $\omega$  ( $\omega$  does not get subdivided during these iterates, see below), where  $p = p(r)$  is the binding period associated to the return depth  $r$  as defined in Section 2.2.



- (3) If  $\omega_n \cap \Delta \neq \emptyset$  and  $\omega_n$  intersects more than three adjacent  $I_{r,j}$ 's we subdivide  $\omega$  into subintervals  $\omega_{r,j}$  in such a way that each  $\omega_{r,j}$  satisfies

$$I_{r,j} \subset f^n \omega_{r,j} \subset \hat{I}_{r,j}.$$

We say that  $\omega_{r,j}$  has an *essential* return at time  $n$ , with return depth  $r$  and define the corresponding binding period as in the previous case.

*Bound times.* If  $n$  is a bound time for  $\omega$  we also basically do nothing. According to the construction above,  $n$  belongs to some binding period  $[\nu + 1, \nu + p]$  associated to a previous essential or inessential return at time  $\nu$ . So, if  $n < \nu + p$  we say that  $n + 1$  is (still) a bound iterate, if  $n = \nu + p$  then  $n + 1$  is a free iterate.

*Returns following escape times.* The notion of an escape time is meant to formalize the idea that the interval in question has reached large scale, and one intuitive consequence of this is that it should therefore “soon” make a full return to  $\Delta$ .

**Lemma 1.** *There exists  $\delta^* > 0$ ,  $t^* \in \mathbb{N}$  and  $\xi > 0$ , all depending on  $\delta$ , such that for  $\Delta^* = (c^* - \delta^*, c^* + \delta^*)$  a  $\delta^*$  neighbourhood of the point  $c^*$  (recall condition H3) and for any interval  $\omega \subset \hat{J}$  with  $|\omega| \geq \delta$ , there exists a subinterval  $\tilde{\omega} \subset \omega$  such that:*

- $f^{t_0}$  maps  $\tilde{\omega}$  diffeomorphically onto  $\Delta^*$  for some  $t_0 \leq t^*$ ,
- $|\tilde{\omega}^*| \geq \xi |\tilde{\omega}|$ ,
- both components of  $\omega \setminus \tilde{\omega}$  are of size  $\geq \delta/3$ .

*Proof.* By assumption the preimages of  $c^*$  are dense in  $\hat{J}$  and do not contain any other critical point. Therefore for any  $\varepsilon > 0$  there exists a  $t^*$  such that the set of preimages  $\{f^{-t}(c^*) : t \leq t^*\}$  of the critical point  $c^*$  is *i)*  $\varepsilon$  dense in  $\hat{J}$ , and *ii)* uniformly bounded away from  $\mathcal{C}$ . Using the  $\varepsilon$ -density and taking  $\varepsilon$  small enough (depending on  $\delta$  but not on  $\omega$ ) we can guarantee that one of these preimages belongs to  $\omega$  and in fact we can ensure that it lies arbitrarily close to the center of  $\omega$ . Then, using that fact that these preimages are uniformly bounded away from  $\mathcal{C}$  and taking  $\delta^*$  sufficiently small we can actually guarantee that a component of  $f^{-t_0}(\Delta^*)$  for some  $0 \leq t_0 \leq t^*$  is contained the central third of  $\omega$ . Since everything depends only on a fixed and finite number of intervals and iterations it follows that the proportion of this preimage in  $\omega$  is uniformly bounded below.  $\square$

*The escape partition.* We have given the complete algorithm for the construction of the escape partition  $\mathcal{P}$  of the interval  $I$ . The algorithm in itself does not show that such a partition does always exist, indeed it may be that intervals get chopped very frequently and in principle it may be that intervals never reach the “large scale”  $\delta$  required to escape. However we shall prove that this algorithm not only gives rise to a partition  $\mathcal{P}$  of  $I$  (mod 0) but in fact escapes occur exponentially fast in the following sense. Let

$$\mathcal{E}_n(\omega) = \{\omega' \subseteq \omega \text{ which have not escaped by time } n\}$$

Then we have the following

**Proposition 1.** *There exist constants  $C_1 > 0$  depending on  $\delta$  and  $\delta^*$  and  $\gamma_1 > 0$  independent of  $\delta$  and  $\delta^*$  such that for any  $\omega \subset \hat{J}$  with  $\omega = \Delta^*$  or  $\delta \geq |\omega| \geq \delta/3$  we have*

$$|\mathcal{E}_n(\omega)| \leq C_1 e^{-\gamma_1 n} |\omega|.$$

Proposition 1 will be proved in Section 5.

**2.4. The induced Markov map.** We are now ready to describe the algorithm for the construction of the final Markov induced map. We fix  $\Delta^*$  as in Lemma 1 and aim to obtain a map  $F : \Delta^* \rightarrow \Delta^*$  with a partition  $\mathcal{Q}$  and a return time function  $T : \mathcal{Q} \rightarrow \mathbb{N}$  constant on elements of  $\mathcal{Q}$  such that  $F(\omega) = f^{T(\omega)}(\omega) \in \Delta^*$  for every  $\omega \in \mathcal{Q}$ . The fact that this construction actually yields such a partition with the required properties (distortion bounds, tail estimates) will be verified in the following sections.

*First escape partition.* First of all, starting with  $\Delta^*$ , we construct the escape time partition  $\mathcal{P}(\Delta^*)$  as described in Section 2.3.

*Dealing with escaping components.* Let  $\omega \in \mathcal{P}(\Delta^*)$  with some escape time  $E(\omega) = n$ . By Lemma 1, we can subdivide its image  $\omega_n = f^n(\omega)$  into three pieces

$$\omega_n = \omega_n^L \cup \omega_n^* \cup \omega_n^R$$

with

$$\omega_{n+t_0}^* = f^{n+t_0}(\omega) = f^{t_0}(\omega_n^*) \in \Delta^*$$

for some  $t_0 \leq t^*$ , and

$$|\omega_n^L|, |\omega_n^R| > \delta/3.$$

The interval  $\omega_n^*$  becomes, by definition, an element of  $\mathcal{Q}$  and we define

$$T(\omega_n^*) = E(\omega) + t_0(\omega) = n + t_0(\omega).$$

*Iterating the argument.* The components  $\omega_n^L, \omega_n^R$  are treated as new starting intervals and we repeat the algorithm: we construct an escape partition of each of  $\omega_n^L, \omega_n^R$  and then some proportion of each escaping component returns to  $\Delta^*$  within some uniformly bounded number of iterates. Notice that if either  $|\omega_n^L| \geq \delta$  or  $|\omega_n^R| \geq \delta$  we can skip the construction of the escape partition (or, in some sense, this step is trivial) and immediately apply Lemma 1 to find a subinterval which returns to  $\Delta^*$  after some finite number of iterates bounded by  $t^*$ . As far as the construction is concerned we will only apply the escape partition algorithm to intervals  $I$  of length between  $\delta/3$  and  $\delta$ . This explains the assumptions of Proposition 1.

*The tail of the return times.* For  $n \geq 1$  we let

$$\mathcal{Q}^{(n)} = \{\omega : T(\omega) > n\}$$

denote the set of intervals which arise from the construction just described, and which have not yet had a full return at time  $n$ . In Section 6 we will prove the following

**Proposition 2.** *There exist constants  $C_2 > 0$  depending on  $\delta$  and  $\delta^*$  and  $\gamma_2 > 0$  independent of  $\delta$  and  $\delta^*$  such that for all  $n \geq 1$ :*

$$|\mathcal{Q}^{(n)}| \leq C_2 e^{-\gamma_2 n} |\Delta^*|.$$

This gives the required tail estimate and implies, in particular, that Lebesgue almost every point of  $\Delta^*$  belongs to an interval of the partition  $\mathcal{Q}$ .

## 3. THE BINDING PERIOD

In this short section we obtain some relatively simple but crucial estimates related to the binding period defined in Section 2.2. In particular this shows that the binding period defines an induced map which is uniformly expanding on each of the countable intervals of the critical partition  $\mathcal{I}$  of  $\Delta$ .

**Lemma 2.** *There exists constants  $\theta, \hat{\theta} > 0$  independent of  $\delta$  such that for all points  $x \in \hat{I}_r$ , and  $p = p(r) \geq 0$  we have*

$$|(f^{p+1})'(x)| \geq \frac{1}{\kappa} e^{\theta r} \geq \frac{1}{\kappa} e^{\hat{\theta}(p+1)}$$

where  $\kappa > 0$  is the constant in the expansivity condition (H1).

*Proof.* We consider the singular region and the critical region separately and then just take the minimum between the  $\theta$ 's which we obtain in the two cases.

*Estimates near  $\mathcal{C}_s$ .* This case is essentially trivial and follows immediately from the structure of the map near the singular points. If  $x \in \hat{I}_r \subset \Delta_c$  with  $c \in \mathcal{C}_s$ , then  $|f'(x)| \geq e^{(1-\ell_s)(r-1)}$ , and essentially any positive  $\theta < (1 - \ell_s)$  will work. Since in this case  $p = 0$  the second inequality follows as well, by taking, for example  $\hat{\theta} = \theta$ .

*Estimates near  $\mathcal{C}_c$ .* For  $x \in \hat{I}_r \subset \Delta_c$  with  $c \in \mathcal{C}_c$  we claim first of all that

$$|(f^{p+1})'(x)| \geq \frac{1}{\kappa} e^{\theta_c r} \quad \text{with} \quad \theta_c = 1 - \frac{5\alpha\ell_c}{\Lambda} > 0. \quad (4)$$

Notice that  $\alpha$  is ‘‘sufficiently small’’ by assumption, we require here  $\alpha < \Lambda/5\ell_c$  for all  $c \in \mathcal{C}_c$ . Assuming (4) we can then choose  $\theta = \min\{\theta_c : c \in \mathcal{C}_c\}$  to get the first inequality in the statement of the lemma. We proceed to show (4) and return to the second inequality in the statement at the end of the proof.

*Bounded distortion.* First of all, by a standard argument such as that in [30] there is a constant  $\mathcal{D}_1$ , independent of  $r$  and  $\delta$ , such that for all  $x_1, y_1 \in f(\hat{I}_r)$  and  $1 \leq k \leq p$ ,

$$\left| \frac{(f^k)'(x_1)}{(f^k)'(y_1)} \right| \leq \mathcal{D}_1. \quad (5)$$

*Upper bound on  $p$ .* By the definition of  $p$  we have  $\delta e^{-2\alpha(p-1)} \geq |x_p - c_p|$  and, using the Mean Value Theorem and (2) we have

$$\delta e^{-2\alpha(p-1)} \geq |x_p - c_p| \geq \mathcal{D}_1^{-1} |(f^{p-1})'(c_1)| |x_1 - c_1| \geq C^{-1} \mathcal{D}_1^{-1} e^{\Lambda(p-1)} e^{-\ell_c r}.$$

Thus  $\delta e^{-2\alpha p} e^{2\alpha} \geq \mathcal{D}_1^{-1} C^{-1} e^{\Lambda p} e^{-\Lambda} e^{-\ell_c r}$  and, rearranging,

$$p \leq \frac{\log \mathcal{D}_1 + \log \delta + 2\alpha + \Lambda + \ell_c r + \log C^{-1}}{\Lambda + 2\alpha} \leq \frac{2\ell_c r}{\Lambda} \leq \frac{2\ell r}{\Lambda}, \quad (6)$$

as long as we choose  $\delta$  so that  $r_\delta$  is sufficiently large in comparison to the other constants, none of which depend on  $\delta$ .

*Derivative estimates in terms of the return depth.* Finally, to prove (4), we use once again the definition of binding, the Mean Value Theorem, and (1), to get

$$C\mathcal{D}_1 e^{-\ell_c r} |(f^p)'(x_1)| \geq \mathcal{D}_1 |x_1 - c_1| |(f^p)'(x_1)| \geq |x_{p+1} - c_{p+1}| \geq \delta e^{-2\alpha p}$$

Rearranging to get a lower bound for  $|(f^p)'(x_1)|$ , and using (6), we have

$$|(f^p)'(x_1)| \geq C^{-1} \mathcal{D}_1^{-1} \delta e^{\ell_c r} e^{-2\alpha p} \geq C^{-1} \mathcal{D}_1^{-1} \delta e^{(\ell_c - \frac{4\alpha\ell_c}{\Lambda})r}.$$

Since  $x \in \hat{I}_r$  we have  $|f'(x)| \geq C^{-1} |x - c|^{\ell_c - 1} \geq C^{-1} e^{-(r+2)(\ell_c - 1)}$  and therefore:

$$\begin{aligned} |(f^{p+1})'(x)| &= |(f^p)'(x_1)| \cdot |f'(x)| \\ &\geq C^{-2} e^{-2(\ell_c - 1)} \mathcal{D}_1^{-1} e^{(1 - \frac{4\alpha}{\Lambda})\ell_c r} e^{-r(\ell_c - 1)} \\ &\geq C^{-2} e^{-2(\ell_c - 1)} \mathcal{D}_1^{-1} e^{(1 - \frac{4\alpha\ell_c}{\Lambda})r} \\ &= C^{-2} e^{-2(\ell_c - 1)} \mathcal{D}_1^{-1} e^{\frac{\alpha\ell_c}{\Lambda} r} e^{(1 - \frac{5\alpha\ell_c}{\Lambda})r}. \end{aligned}$$

It therefore only remains to show that  $C^{-2} e^{-2(\ell_c - 1)} \mathcal{D}_1^{-1} e^{\frac{\alpha\ell_c}{\Lambda} r} \geq 1/\kappa$  for any  $r \geq r_\delta$  and any  $\ell_c$ . This can clearly be arranged by taking  $\delta$  sufficiently small (and thus  $r_\delta$  sufficiently large) since  $\ell_c \geq 1$  and the other constants do not depend on  $\delta$ .

*Derivative estimates in terms of the length of the binding period.* Finally, we prove the second inequality in the statement of the lemma. By (6) we have  $r \geq \Lambda p/2\ell$  and therefore

$$e^{\theta r} \geq e^{\frac{\theta\Lambda}{2\ell} p}.$$

Since the minimum binding period can be taken large by taking  $\delta$  small this clearly implies the statement for some  $\hat{\theta}$  between 0 and  $\theta\Lambda/2\ell$ .  $\square$

#### 4. DISTORTION ESTIMATES

In this section we show that our construction yields intervals for which some uniform distortion bounds hold. Let  $\omega \subset \hat{J}$  be an arbitrary interval,  $n \geq 1$  a positive integer such that  $\omega$  has a sequence  $t_0, \dots, t_q \leq n$  of *free* returns to  $\Delta$  (with respective return depth sequence  $r_{t_0}, \dots, r_{t_q}$ ) followed by corresponding binding periods  $[t_m + 1, t_m + p_m]$ , as described above. In particular, for  $m = 1, \dots, q$ , the interval  $\omega_{t_m}$  is contained in the union of three adjacent elements of the form  $I_{r,j}$  of the critical partition  $\mathcal{I}$  of  $\Delta$ .

**Proposition 3.** *There exist constants  $\mathcal{D}_\delta$  and  $\tilde{\mathcal{D}}_\delta$  depending on  $\delta$ , and  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$  independent of  $\delta$ , such that for every interval  $\omega$  and integer  $n$  as described in the previous paragraph, for every  $k \leq n$  and  $x, y \in \omega$  we have*

$$\left| \frac{(f^k)'(x)}{(f^k)'(y)} \right| \leq \mathcal{D}_\delta \quad \text{and} \quad \left| \frac{(f^k)'(x)}{(f^k)'(y)} - 1 \right| \leq \frac{\tilde{\mathcal{D}}_\delta}{|\omega_k|} |f^k(x) - f^k(y)|.$$

Moreover, the constants  $\mathcal{D}_\delta$  and  $\tilde{\mathcal{D}}_\delta$  can be replaced by  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$  under the following constraints:

- (1) either we allow every  $x, y \in \omega$  but restrict to values of  $k \leq t_q + p_q$ ;
- (2) or we allow all  $k \leq n$ , in particular,  $t_q + p_q \leq k \leq n$  but restrict to  $x, y \in \tilde{\omega}$  where  $\tilde{\omega} \subset \omega$  is such that  $\tilde{\omega}_j \cap \Delta = \emptyset$  for  $k > j \geq t_m + p_m$ , and  $\tilde{\omega}_k \subset \Delta$ ; in this case we replace  $\omega_k$  by  $\tilde{\omega}_k$  in the second inequality.

Most of the proof is devoted to proving the first inequality, then in Section 4.5 we show that the second follows almost immediately.

4.1. **Preliminary calculations.** First of all we prove the following

**Lemma 3.** *There exists a constant  $\mathcal{D}_2 > 0$  such that*

$$\log \left| \frac{(f^k)'(x_0)}{(f^k)'(y_0)} \right| \leq \mathcal{D}_2 \sum_{j=0}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)}.$$

*Proof.* We carry out some relatively standard algebraic manipulations and then use the nondegeneracy of the critical set together with our assumptions about the itinerary of  $\omega$ .

*Preliminary reductions.* We start by rewriting the expression for the distortion as follows. For  $j \geq 0$ , let  $x_j = f^j(x)$ ,  $y_j = f^j(y)$ , and  $\omega_j = f^j(\omega)$ . By the chain rule and the convexity of the log function we have

$$\log \left| \frac{(f^k)'(x_0)}{(f^k)'(y_0)} \right| = \sum_{j=0}^{k-1} \log \left| 1 + \frac{f'(x_j) - f'(y_j)}{f'(y_j)} \right| \leq \sum_{j=0}^{k-1} \frac{|f'(x_j) - f'(y_j)|}{|f'(y_j)|}. \quad (7)$$

Since  $f$  is  $C^2$  outside  $\mathcal{C}$ , by the Mean Value Theorem we can write  $|f'(x_j) - f'(y_j)| = |f''(\xi_j)||x_j - y_j|$  for some  $\xi_j \in (x_j, y_j) \subset \omega_j$  and so

$$\frac{|f'(x_j) - f'(y_j)|}{|f'(y_j)|} \leq \frac{|f''(\xi_j)|}{|f'(y_j)|} |\omega_j|. \quad (8)$$

*Using the nondegeneracy of the critical set.* Outside some fixed neighbourhood of the critical set  $\mathcal{C}$  the ratio  $|f''(\xi_j)|/|f'(y_j)|$  is uniformly bounded above (and below). Inside such a neighbourhood we have (2) and (3) and therefore, as long as the distance of  $\xi_j$  and  $y_j$  to the critical point  $c$  are comparable, we get

$$\frac{|f''(\xi_j)|}{|f'(y_j)|} \leq C^2 \frac{|\xi_j - c|^{\ell_c - 2}}{|y_j - c|^{\ell_c - 1}} \leq \frac{\mathcal{D}_2}{\mathfrak{D}(\omega_j)} \quad (9)$$

for some constant  $\mathcal{D}_2 > 0$ . The distances of  $\xi_j$  and  $y_j$  to the critical point are indeed comparable, and thus (9) holds in the situations we are considering. To see this, we distinguish two cases. If  $\mathfrak{D}(\omega_j) \geq \delta/2$  the distances clearly are comparable since  $|\omega_j| \leq \delta$ . On the other hand,  $\mathfrak{D}(\omega_j) \leq \delta/2$  implies that  $\omega_j$  is contained in at most three elements of the form  $I_{r,j}$  of the critical partition  $\mathcal{I}$  and therefore  $\mathfrak{D}(\omega_j) \gg |\omega_j|$ .

Now, substituting (9) into (8) and then into (7) gives the statement in the lemma.  $\square$

*Basic strategy.* By Lemma 3, we only need to get an upper bound, independent of  $k$ ,  $\omega$  and  $\delta$ , for the sum

$$\mathcal{S} = \sum_{j=0}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)}.$$

This relies on two main ideas.

- On the one hand we need at least  $\sum |\omega_j|$  uniformly bounded. This depends on the fact that  $\omega_j$  is growing exponentially in size, the sequence  $|\omega_j|$  is a geometric increasing sequence with a uniformly bounded last term, and thus has a uniformly bounded overall sum.

- On the other hand, this bound is not sufficient since  $\mathfrak{D}(\omega_j)$  is not uniformly bounded below. We therefore need to use the additional information that  $\mathfrak{D}(\omega_j)$  being small implies  $\omega_j \subset \Delta$  and in this case we have additional information on the size of  $\omega_j$  in relation to  $\mathfrak{D}(\omega_j)$  given by our assumptions that  $\omega_j$  must be contained in some  $\hat{I}_{r,i}$ .

We obtain this bound in two steps. First of all we divide all iterates into free and bound iterates and get estimates for the contribution to  $\mathcal{S}$  of the blocks of consecutive free iterates and bound iterates respectively. Secondly we add all these blocks together. This requires some care as we have no uniform bound on the number of such blocks, therefore the process of combining the estimates for each block need to be refined at this stage.

*Free and bound iterates.* We split the sum  $\mathcal{S}$  into free iterates and bound iterates to get

$$\sum_{m=1}^q \left( \sum_{j=t_{m-1}+p_{m-1}+1}^{t_m-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} + \sum_{j=t_m}^{t_m+p_m} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \right) + \sum_{j=t_q+p_q+1}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)}. \quad (10)$$

We include the return iterates, even though formally these are free, with the bound iterates. This is not necessary but simplifies slightly the calculations. For notational simplicity we define  $t_0+p_0+1=0$  so that in the general case the sum starts with the free iterates  $0, \dots, t_1-1$  preceding the first return at time  $t_1$ . In the special case in which the initial iterate is already an essential return (such as when we construct the escape partition  $\mathcal{P}(\Delta^*)$ ) we have  $t_1=0$  and thus the first sum inside the parenthesis is empty for  $k=1$ .

#### 4.2. Distortion during free iterates.

**Lemma 4.** *There exists a constant  $\mathcal{D}_3 > 0$  such that for  $m = 1, \dots, q$ ,*

$$\sum_{j=t_{m-1}+p_{m-1}+1}^{t_m-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_3 |\omega_{t_m}| e^{rt_m}.$$

*For the last free period we have*

$$\sum_{j=t_q+p_q+1}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \frac{\mathcal{D}_3}{\delta},$$

*and, restricting to a subinterval  $\tilde{\omega}$  as in the statement of Proposition 3 we have*

$$\sum_{j=t_q+p_q+1}^{k-1} \frac{|\tilde{\omega}_j|}{\mathfrak{D}(\tilde{\omega}_j)} \leq \mathcal{D}_3.$$

*Proof.* For the first  $q$  free periods, since  $\omega_{t_m} \subset \Delta$  the expansivity condition (H1) implies

$$|\omega_j| \leq \kappa^{-1} e^{-\lambda(t_{m+1}-j)} |\omega_{t_{m+1}}|,$$

and thus for  $m \leq q$ , using also the fact that  $\mathfrak{D}(\omega_j) \geq \delta \geq e^{-rt_{m+1}}$ ,

$$\sum_{j=t_{m-1}+p_{m-1}+1}^{t_m-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \frac{|\omega_{t_m}|}{e^{-rt_m}} \sum_{j=t_{m-1}+p_{m-1}+1}^{t_m-1} \kappa^{-1} e^{-\lambda(t_m-j)} \leq \mathcal{D}_3 |\omega_{t_m}| e^{rt_m},$$

for some constant  $\mathcal{D}_3 > 0$ . This proves the first inequality.

For the last free period, restricting to  $\tilde{\omega}$  and  $k$  as in the statement of Proposition 3, we get exactly the same estimates with  $\omega_{t_m}$  replaced by  $\tilde{\omega}_k$ , and so, using the fact that  $|\tilde{\omega}_k| \leq \delta$  we get the third inequality. Without restricting to  $\tilde{\omega}$  and without assuming that  $\omega_k \subset \Delta$ , (H1) only gives a weaker expansion estimate which implies  $|\omega_j| \leq \kappa^{-1} \delta^{-1} e^{-\lambda(k-j)} |\omega_k|$  and therefore a final estimate as in the second inequality.  $\square$

**Remark 1.** The dependence of  $\mathcal{D}_\delta$  and  $\tilde{\mathcal{D}}_\delta$  on  $\delta$  comes entirely from the contribution of the last term of the sum in (10) which is bounded by  $\mathcal{D}_3/\delta$  as shown in Lemma 4. For simplicity we shall now continue the proof of Proposition 3 under the assumptions which give bounds independent of  $\delta$ . The additional statement follows by making minimal and obvious modifications.

Notice also that restricting to the subinterval  $\tilde{\omega}$  for the last term of (10) does not affect the bounds obtained for the previous terms as  $\tilde{\omega}_j \subseteq \omega_j$  always implies  $|\tilde{\omega}_j|/\mathfrak{D}(\tilde{\omega}_j) \leq |\omega_j|/\mathfrak{D}(\omega_j)$ .

### 4.3. Distortion during binding periods.

**Lemma 5.** *There exists a constant  $\mathcal{D}_4 > 0$  such that*

$$\sum_{j=t_m}^{t_m+p_m} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_4 |\omega_{t_m}| e^{r t_m}.$$

*Proof.* For each  $j \geq t_m$  we let  $x_j, y_j$  be two arbitrary points in  $\omega_j$ , and let  $c(y_j)$  denote the critical point closest to  $y_j$ , so that  $|y_j - c(y_j)| = \mathfrak{D}(y_j)$  and, in particular,  $c(y_{t_m})$  is the critical point involved in the binding period. Then we can write

$$\frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \frac{|x_j - y_j|}{|y_j - c(y_j)|} = \frac{|x_j - y_j|}{|y_j - f^{j-t_m}(c(y_{t_m}))|} \cdot \frac{|y_j - f^{j-t_m}(c(y_{t_m}))|}{|y_j - c(y_j)|}. \quad (11)$$

We estimate the two ratios on the right hand side separately.

**Sublemma 5.1.** *There exists a constant  $\mathcal{D}_5 > 0$  such that for each  $j = t_m, \dots, t_m + p_m$  we have*

$$\frac{|x_j - y_j|}{|y_j - f^{j-t_m}(c(y_{t_m}))|} \leq \mathcal{D}_5 \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})}. \quad (12)$$

*Proof.* For  $j = t_m$  we have

$$\frac{|x_j - y_j|}{|y_j - f^{j-t_m}(c(y_{t_m}))|} = \frac{|x_{t_m} - y_{t_m}|}{|y_{t_m} - c(y_{t_m})|} \leq \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})}.$$

For  $t_m - 1 < j \leq t_m + p_m$  recall the bounded distortion property during binding periods, (5), which gives

$$\frac{|x_j - y_j|}{|y_j - f^{j-t_m}(c_{t_m})|} \leq \mathcal{D}_1 \frac{|x_{t_m+1} - y_{t_m+1}|}{|y_{t_m+1} - f(c(y_{t_m}))|}. \quad (13)$$

Since  $|\omega_{t_m}| \ll \mathfrak{D}(\omega_{t_m})$  we have, using the non-degeneracy conditions (2) and (1) on the order of the critical points,

$$|x_{t_m+1} - y_{t_m+1}| \lesssim |\omega_{t_m}| \mathfrak{D}(\omega_{t_m})^{\ell_c - 1}. \quad (14)$$

and

$$|y_{t_m+1} - f(c(y_{t_m}))| \approx |y_{t_m} - c(y_{t_m})|^{\ell_c} \approx \mathfrak{D}(\omega_{t_m})^{\ell_c}. \quad (15)$$

We use here the symbol  $\lesssim$ , respectively  $\approx$ , to mean that the left hand side is bounded above, respectively above and below, by constants that depend only on the map  $f$ .

Substituting (14) and (15) into (13) we obtain (12).  $\square$

**Sublemma 5.2.** *There exists a constant  $\mathcal{D}_6 > 0$  such that for each  $j = t_m, \dots, t_m + p_m$  we have*

$$\frac{|y_j - f^{j-t_m}(c(y_{t_m}))|}{|y_j - c(y_j)|} \leq \mathcal{D}_6 e^{-\alpha(j-t_m)},$$

*Proof.* For  $j = t_m$  we have

$$\frac{|y_j - f^{j-t_m}(c(y_{t_m}))|}{|y_j - c(y_j)|} = \frac{|y_{t_m} - c(y_{t_m})|}{|y_{t_m} - c(y_{t_m})|} = 1.$$

For  $t_m - 1 < j \leq t_m + p_m$ , the definition of binding period gives

$$|y_j - f^{j-t_m}(c(y_{t_m}))| \leq \delta e^{-2\alpha(j-t_m)}$$

and, in conjunction with the bounded recurrence condition (H2),

$$|y_j - c(y_j)| \geq \delta e^{-\alpha(j-t_m)} - \delta e^{-2\alpha(j-t_m)} = \delta e^{-\alpha(j-t_m)}(1 - e^{-\alpha(j-t_m)}) \geq \delta e^{-\alpha(j-t_m)}(1 - e^{-\alpha}).$$

Therefore we have

$$\frac{|y_j - f^{j-t_m}(c(y_{t_m}))|}{|y_j - c(y_j)|} \leq \frac{\delta e^{-2\alpha(j-t_m)}}{\delta e^{-\alpha(j-t_m)}(1 - e^{-\alpha})} \leq (1 - e^{-\alpha})e^{-\alpha(j-t_m)}. \quad (16)$$

□

Returning to the proof of Lemma 5, substituting the bounds obtained in Sublemmas 5.1 and 5.2 into (11) and letting  $\mathcal{D}_7 = \mathcal{D}_5 \mathcal{D}_6$  and  $\mathcal{D}_4 = \mathcal{D}_7 \sum_{i=0}^{\infty} e^{-\alpha i}$  we get

$$\sum_{j=t_m}^{t_m+p_m} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \sum_{j=t_m}^{t_m+p_m} \mathcal{D}_7 \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})} e^{-\alpha(j-t_m)} \leq \mathcal{D}_7 \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})} \sum_{i=0}^{\infty} e^{-\alpha i} = \mathcal{D}_4 \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})}$$

Finally, to get the statement in the Lemma recall that  $\mathfrak{D}(\omega_{t_m}) \geq e^{r t_m}$  by construction. □

**4.4. Combining free period and bound period estimates.** Combining the estimates of Lemmas 4 and 5 and letting  $\mathcal{D}_8 = \mathcal{D}_3 + \mathcal{D}_4$  gives

$$\mathcal{S} = \sum_{j=0}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_3 + \mathcal{D}_8 \sum_{m=1}^q \frac{|\omega_{t_m}|}{\mathfrak{D}(\omega_{t_m})} \leq \mathcal{D}_3 + \mathcal{D}_8 \sum_{m=1}^q |\omega_{t_m}| e^{r t_m}.$$

Therefore it is sufficient to show that  $\sum_{m=1}^q |\omega_{t_m}| e^{r t_m}$  is uniformly bounded above. Recall that by construction we have  $|\omega_{t_m}| \lesssim e^{-r t_m} / (r_{t_m})^2$  and so  $|\omega_{t_m}| e^{r t_m} \lesssim 1 / r_{t_m}^2$  and  $\sum_{m=1}^q |\omega_{t_m}| e^{r t_m} \lesssim \sum_{m=1}^q 1 / r_{t_m}^2$ . This does not however imply a uniform upper bound for the overall sum, since the sequence  $r_{t_m}$  is not generally monotone and might take on the same value with unbounded multiplicity. We therefore need to refine our estimates as follows. First of all we subdivide the  $q$  returns under consideration into returns with the same return depth:

$$\sum_{m=1}^q |\omega_{t_m}| e^{r t_m} = \sum_r \sum_{t_m: r_{t_m}=r} |\omega_{t_m}| e^{r t_m} = \sum_r e^r \sum_{t_m: r_{t_m}=r} |\omega_{t_m}|. \quad (17)$$

**Sublemma 5.3.** *There exists a constant  $\mathcal{D}_9$  such that for each  $r$  we have*

$$\sum_{t_m: r_{t_m}=r} |\omega_{t_m}| \leq \mathcal{D}_9 \frac{e^{-r}}{r^2}. \quad (18)$$



*Proof.* The statement follows from the fact that the interval  $\omega_j$  is growing exponentially fast between one return and the next. Indeed, by (H1) and Lemma 2, between any two consecutive returns we have:

$$|\omega_{t_{m+1}}| \geq \kappa |f^{p_m+1}(\omega_{t_m})| \geq e^{\theta r t_m} |\omega_{t_m}| \geq e^{\theta r \delta} |\omega_{t_m}|.$$

Iterating this process we get that the terms in the sum (18) form an exponentially decreasing sequence bounded above by the length of the last term, i.e. the term with the highest return time  $t_m$ . For this term, our assumptions imply  $|\omega_{t_m}| \leq e^{-r}/r^2$  thus proving the result.  $\square$

Substituting (18) into (17) gives

$$\sum_{m=1}^q |\omega_{t_m}| e^{r t_m} = \sum_r e^r \sum_{t_m: r t_m = r} |\omega_{t_m}| \leq \mathcal{D}_9 \sum_r \frac{1}{r^2}$$

which is bounded above by a uniform constant. This gives

$$\mathcal{S} = \sum_{j=0}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_3 + \mathcal{D}_8 \sum_{m=1}^q |\omega_{t_m}| e^{r t_m} \leq \mathcal{D}_3 + \mathcal{D}_8 \mathcal{D}_9 \sum_r \frac{1}{r^2} =: \mathcal{D}_{10}. \quad (19)$$

This together with Lemma 3 completes the proof of the first inequality in Proposition 3, with constant  $\mathcal{D} = e^{\mathcal{D}_2 \mathcal{D}_{10}}$ .

**4.5. Lipschitz distortion.** First of all, for any subinterval  $\bar{\omega} \subset \omega$  we have

$$\frac{|\bar{\omega}_j|}{\mathfrak{D}(\bar{\omega}_j)} \leq \frac{|\bar{\omega}_j|}{\mathfrak{D}(\omega_j)} = \frac{|\bar{\omega}_j|}{|\omega_j|} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)}.$$

For all  $x, y \in \bar{\omega}$ , from the first inequality of Proposition 3 we have

$$\mathcal{D}_2 \mathcal{D}_{10} \geq \log \left| \frac{(f^k)'(x)}{(f^k)'(y)} \right| \geq \frac{1}{\mathcal{D}_2 \mathcal{D}_{10}}.$$

Therefore, since  $x - 1$  and  $\log x$  are comparable on any interval bounded away from 0 and  $\infty$ , there exists a constant  $\mathcal{D}_{11}$ , depending on  $\mathcal{D}$ , such that

$$\left| \frac{(f^k)'(x)}{(f^k)'(y)} - 1 \right| \leq \mathcal{D}_{11} \log \left| \frac{(f^k)'(x)}{(f^k)'(y)} \right|.$$

Thus, by Lemma 3, and (19) we have

$$\left| \frac{(f^k)'(x)}{(f^k)'(y)} - 1 \right| \leq \mathcal{D}_{11} \mathcal{D}_2 \sum_{j=0}^{k-1} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_{11} \mathcal{D}_2 \sum_{j=0}^{k-1} \frac{|\bar{\omega}_j|}{|\omega_j|} \frac{|\omega_j|}{\mathfrak{D}(\omega_j)} \leq \mathcal{D}_{11} \mathcal{D}_2 \mathcal{D}_{10} \frac{|\bar{\omega}_j|}{|\omega_j|}.$$

Finally, using once again the first inequality of Proposition 3 and the Mean Value Theorem, the ratios  $|\bar{\omega}_j|/|\omega_j|$  are all uniformly comparable to the ratio  $|\bar{\omega}_k|/|\omega_k|$ . This implies the second inequality in Proposition 3 and thus completes the proof.

## 5. ESCAPE TIME ESTIMATES

In this section we give a complete proof of Proposition 1. We assume throughout that  $\omega = \Delta^*$  or  $\omega$  is an arbitrary interval with  $\delta \geq |\omega| \geq \delta/3$ . We fix from now on some  $n \geq 1$  and recall the definition of

$$\mathcal{E}_n(\omega) = \{\omega' \subseteq \omega \text{ which have not escaped by time } n\}.$$

Each  $\omega' \in \mathcal{E}_n(\omega)$  has an associated sequence

$$\nu_1, \nu_2, \dots, \nu_s$$

of *essential* return times occurring before time  $n$ , and a corresponding sequence

$$r_1, r_2, \dots, r_s$$

of *essential* return depths. (Notice that in contrast to Section 4 we use a different labelling of the subscripts for the return depths). If  $\omega = \Delta^*$  then  $\nu_1 = 0$  and the sequences  $\nu_i$  and  $r_i$  are non-empty. In the general case these sequences may be empty if  $\omega'$  escapes without intersecting  $\Delta$ . We let

$$\mathcal{E}_{n,R}(\omega) = \{\omega' \in \mathcal{E}_n(\omega) : r_1 + \dots + r_s = R\}$$

denote the union of elements of  $\mathcal{E}_n(\omega)$  with a given value  $R$  for the *accumulated return depth*. Notice that  $R$  can be 0 if the sequence of returns is empty, but if it is non-zero then we must have  $R \geq r_\delta$ . We split the proof of Proposition 1 into several lemmas. First of all we show that intervals  $\omega'$  in  $\mathcal{E}_{n,R}(\omega)$  are exponentially small in  $R$ .

**Lemma 6.** *For every  $n \geq 1, R \geq 1$  and  $\omega' \in \mathcal{E}_{n,R}(\omega)$  we have*

$$|\omega'| \leq \kappa^{-1} e^{-\theta R}. \quad (20)$$

Next, we show that the number of intervals  $\omega'$  with the same accumulated return depth can grow exponentially but at most with a very small exponential rate.

**Lemma 7.** *There exists a constant  $\tilde{\eta} > 0$  which can be made arbitrarily small if  $\delta$  is small, such that for every  $n \geq 1, R \geq 0$*

$$\#\{\mathcal{E}_{n,R}(\omega)\} \leq e^{\tilde{\eta}R}.$$

Lemmas 6 and 7 immediately give

$$|\mathcal{E}_{n,R}(\omega)| \leq \kappa^{-1} e^{-(\theta - \tilde{\eta})R} \quad (21)$$

where  $\theta - \tilde{\eta} > 0$  if  $\delta$  is sufficiently small. We can of course apply these estimates to get  $|\mathcal{E}_n(\omega)| = \sum_{R \geq r_\delta} |\mathcal{E}_{n,R}(\omega)| \lesssim \sum_{R \geq r_\delta} e^{-(\theta - \tilde{\eta})R} \lesssim e^{-(\theta - \tilde{\eta})r_\delta}$  but this bound is not good enough since it does not give an exponential bound in  $n$ . We need to show that there is a relation between  $n$  and possible values of  $R$ . This relation is given in the next two lemmas.

**Lemma 8.** *There exists  $n_\delta > 0$  depending on  $\delta$ , such that*

$$\omega' \in \mathcal{E}_{n,0}(\omega) \text{ implies } \omega' = \omega \text{ and } n \leq n_\delta.$$

**Lemma 9.** *For all  $n \geq 1$  and  $R \geq 1$  such that  $\mathcal{E}_{n,R} \neq \emptyset$  we have*

$$R \geq (n - n_\delta)/\tilde{\theta}.$$

Lemmas 8 and 9 and equation (21) give, for  $n > n_\delta$  and some constant  $\tilde{C}_1 > 0$ ,

$$|\mathcal{E}_n(\omega)| = \sum_{R \geq (n - n_\delta)/\tilde{\theta}} |\mathcal{E}_{n,R}(\omega)| \leq \sum_{R \geq (n - n_\delta)/\tilde{\theta}} e^{-(\theta - \tilde{\eta})R} \leq \tilde{C}_1 e^{\frac{(\theta - \tilde{\eta})n_\delta}{\tilde{\theta}}} e^{-\frac{\theta - \tilde{\eta}}{\tilde{\theta}}n}.$$

Finally we multiply the right hand side by  $|\Delta^*|/\delta^*$  or  $|\omega|/\delta$  to get the statement in Proposition 1 with  $\gamma_1 = (\theta - \tilde{\eta})/\tilde{\theta}$ . In the next four subsections we prove the four Lemmas above.

### 5.1. Escapes with essential returns: metric estimates.

*Proof of Lemma 6.* Let  $\omega' \in \mathcal{E}_{n,R}$  with an associated sequence  $\nu_1, \dots, \nu_s$  of essential returns corresponding return depths  $r_1, \dots, r_s$ . Then, by construction we have

$$|\omega'_{\nu_s}| \leq e^{-r_s}.$$

Moreover, using the expansion after the binding periods from Lemma 2 and the expansion during the free periods (H1), including the fact that each free period under consideration ends in a return to  $\Delta$ , we have

$$|(f^{\nu_1})'(x)| \geq \kappa e^{\lambda \nu_1} \geq \kappa$$

for all  $x \in \omega'$  and

$$|(f^{\nu_{i+1}-\nu_i})'(x_{\nu_i})| \geq e^{\theta r_i}$$

for all  $i = 1, \dots, s-1$  and  $x_{\nu_i} \in \omega'_{\nu_i}$ . This gives

$$|(f^{\nu_s})'(x)| \geq \kappa e^{\theta(r_1 + \dots + r_{s-1})}$$

for all  $x \in \omega'$ . Thus, by the mean value theorem we have

$$|\omega'| \leq \kappa^{-1} e^{-\theta(r_1 + \dots + r_{s-1})} |\omega'_{\nu_s}| \leq \kappa^{-1} e^{-\theta(r_1 + \dots + r_s)}.$$

□

### 5.2. Escapes with essential returns: combinatorial estimates.

*Proof of Lemma 7.* We divide the proof into two steps. The first one is purely combinatorial and bounds the number of theoretically possible combinatorially distinct elements. The second one relies on the construction and bounds the number of possible elements with the same combinatorics.

*Cardinality of possible sequences.* Let  $N_{R,s}$  denote the number of integer sequences  $(t_1, \dots, t_s)$ ,  $t_i \geq r_\delta$  for all  $i$ ,  $1 \leq i \leq s$ , such that  $\sum_{i=1}^s t_i = R$ . The number  $N_{R,s}$  of such sequences is the same as the number of ways to choose  $s$  balls from a row of  $k+s$  balls, thus partitioning the remaining  $k$  balls into at most  $s+1$  disjoint subsets. Therefore we have

$$N_{R,s} \leq \binom{R+s}{s} = \binom{R+s}{R}.$$

Since each term  $t_i \geq r_\delta$  we must always have  $s \leq R/r_\delta$ , and, since the right hand side above is monotonically increasing in  $s$ , writing  $\eta = 1/r_\delta$  for simplicity, we get

$$N_{R,s} \leq \binom{(1+\eta)R}{R} = \frac{[(1+\eta)R]!}{(\eta R)! R!}.$$

Using Stirling's formula  $k! \in [1, 1 + \frac{1}{4k}] \sqrt{2\pi k} k^k e^{-k}$  we obtain

$$\begin{aligned} N_{R,s} &\leq \frac{[(1+\eta)R]^{(1+\eta)R}}{(\eta R)^{\eta R} R^R} = (1+\eta)^{(1+\eta)R} \eta^{-\eta R} \\ &\leq \exp\{(1+\eta)R \log(1+\eta) - \eta R \log \eta\} \leq \exp\{((1+\eta)\eta - \eta \log \eta)R\} = e^{\eta_1 R} \end{aligned}$$

where  $\eta_1 = ((1+\eta)\eta - \eta \log \eta)$  can be made arbitrarily small with  $\delta$ . Now, letting  $N_R$  denote the number of all possible integer sequences  $(t_1, \dots, t_s)$ ,  $t_i \geq r_\delta$  for all  $i$ ,  $1 \leq i \leq s$ , such that  $\sum_{i=1}^s t_i = R$ , for all possible values of  $s$  we have

$$N_R = \sum_{s=1}^{\eta R} N_{R,s} \leq \sum_{s=1}^{\eta R} e^{\eta_1 R} = \eta R e^{\eta_1 R}. \quad (22)$$

*Multiplicity of intervals sharing the same sequence.* From the subdivision procedure described in Section 2, the only way in which several elements can share the same combinatorics is by having the same sequence of return depth  $r_1, r_2, \dots, r_s$  but being associated to different critical points and to different intervals  $I_{r_i, j}$  (recall that each element  $I_r$  in the critical partition is subdivided into  $r^2$  subintervals of equal length). Therefore, letting  $N_c$  denote the number of (one-sided) critical points and  $\mathcal{E}_{n, r_1, \dots, r_s}$  the set of escaping intervals having the same given sequence  $r_1, \dots, r_s$  of return depths, we have

$$\#\mathcal{E}_{n, r_1, \dots, r_s} \leq \prod_{j=1}^s N_c r_j^2.$$

To get an upper bound for the right hand side we take logs and obtain

$$\log \prod_{j=1}^s N_c r_j^2 = \sum_{j=1}^s \log N_c r_j^2 = s \log N_c + \sum_{j=1}^s 2 \log r_j. \quad (23)$$

Since  $R \geq sr_\delta$  we have

$$s \log N_c \leq \frac{\log N_c}{r_\delta} R = \eta_2 R \quad (24)$$

where  $\eta_2$  can be made arbitrarily small by taking  $\delta$  small. Moreover since  $r_j \geq r_\delta$  we also have

$$\frac{2 \log r_j}{r_j} \leq \frac{2 \log r_\delta}{r_\delta} = \eta_3 \quad (25)$$

and therefore

$$\sum_{j=1}^s 2 \log r_j \leq \eta_3 \sum_{j=1}^s 2r_j = \eta_3 R. \quad (26)$$

where  $\eta_3$  can be made arbitrarily small by taking  $\delta$  small. Substituting (24) and (26) into (23) gives

$$\log \prod_{j=1}^s N_c r_j^2 \leq (\eta_2 + \eta_3) R$$

or

$$\#\{\mathcal{E}_{n, r_1, \dots, r_s}\} \leq \prod_{j=1}^s N_c r_j^2 \leq e^{(\eta_2 + \eta_3)R} \quad (27)$$

*Final estimate.* The final upper bound for the number of possible intervals in  $\mathcal{E}_{n, R}(\omega)$  is therefore just given by multiplying the number of possible sequences with a given total return depth  $R$  by the number of possible partition elements which can potentially have exactly such a value as their total return depth. Thus, multiplying (22) by (27) we get

$$\#\{\mathcal{E}_{n, R}(\omega)\} \leq \eta R e^{(\eta_1 + \eta_2 + \eta_3)R} \leq e^{(\eta_1 + \eta_2 + \eta_3)R + \log \eta + \log R} \leq e^{(\eta_1 + \eta_2 + 2\eta_3)R}.$$

In the last inequality we have used the fact that  $\log \eta < 0$  since  $\eta = 1/r_\delta$  is small, and that  $\log R < \eta_3 R$  from (25) and the fact that  $R \geq r_\delta$ . Thus we get the statement in Lemma 7 with  $\tilde{\eta} = \eta_1 + \eta_2 + 2\eta_3$  where  $\tilde{\eta}$  can be made arbitrarily small if  $\delta$  is chosen small enough.

□

### 5.3. Escapes with no essential returns.

*Proof of Lemma 8.* First we show that  $\omega' = \omega$  and then estimate  $n_\delta$ . For convenience we introduce here a couple of constants which will be used here and in the next subsection. Let

$$\hat{\lambda} = \min\{\lambda, \hat{\theta}\} \quad \text{and} \quad \tilde{\lambda} = \min\{\lambda, \hat{\theta}, \Lambda\}.$$

*Claim:*  $\omega' = \omega$ . This follows directly from the construction. Indeed, suppose by contradiction that  $\omega' \subset \omega$ . This would mean that  $\omega$  had been chopped at some time  $\nu \leq n$  for which  $\omega_\nu \cap \Delta \neq \emptyset$ . However, at this time all subintervals of  $\omega$  arising from this chopping procedure would qualify as having had an essential return. This is true even if  $\omega_\nu$  is not strictly contained in  $\Delta$  since either the components of  $\omega_\nu \setminus \Delta$  are smaller than  $\delta$  and are therefore, by construction, attached to their adjacent elements which fall inside  $\Delta$ , or they are bigger than  $\Delta$  which implies that  $|\omega_\nu| \geq \delta$  implying that  $\nu$  is an escape time for  $\omega$ . This contradicts our assumption that  $\omega'$  does not escape before time  $n$ . Thus  $\omega' = \omega$ .

*No inessential returns.* To obtain a bound for  $n$  we suppose first of all that there are no (inessential) returns, i.e.  $\omega$  stays outside  $\Delta$  up to time  $n$ . Then, by conditions (H1) we have  $|(f^n)'(x)| \geq \kappa\delta e^{\lambda n}$  and so, by the mean value theorem and using the fact that  $|\omega| \geq \delta/3$  and that  $|\omega_n| \leq \delta$  we get

$$\delta \geq |\omega_n| \geq \kappa\delta e^{\lambda n} |\omega| \geq \kappa\delta^2 e^{\lambda n} / 3.$$

This gives  $e^{\lambda n} \leq \frac{3}{\kappa\delta}$  and, solving for  $n$  we obtain  $n \leq \frac{1}{\lambda} \log \frac{3}{\kappa\delta}$ .

*If  $n$  is a free iterate.* If there are any returns to  $\Delta$  we distinguish two further cases. If  $n$  is a free iterate, then we can combine the binding period expansion estimates and condition (H1) to get  $|(f^n)'(x)| \geq \kappa\delta e^{\hat{\lambda} n}$ . Then reasoning exactly as above we get  $n \leq \frac{1}{\hat{\lambda}} \log \frac{3}{\kappa\delta}$ .

*If  $n$  is a bound iterate.* If  $n$  is not a free iterate we have the additional minor complication of not being able to use the binding period estimate for the last incomplete binding period. Let  $\nu < n$  be the last inessential return before  $n$  such that  $n$  belongs to the binding period which follows the return at time  $\nu$ . Then, by the calculation above we have  $|(f^\nu)'(x)| \geq \kappa\delta e^{\hat{\lambda}\nu}$  and, in particular,

$$|\omega_\nu| \geq \kappa\delta^2 e^{\hat{\lambda}\nu} / 3 \geq \kappa\delta^2 / 3.$$

Therefore, for the return at time  $\nu$  to be inessential it cannot be too deep, i.e. it must have a return depth  $r \leq \log \kappa\delta^2 / 3$ , and therefore the derivative at points  $x_\nu \in \omega_\nu$  must be of the order of  $e^r \approx \kappa\delta^2 / 3$ . Therefore  $|(f^{\nu+1})'(x)| \gtrsim \delta^3 e^{\hat{\lambda}\nu} e^r \kappa / 3$ . During the remaining iterates, the bounded distortion during the binding periods implies that the derivative is growing exponentially fast at rate  $\Lambda$ , from condition (H2). We therefore have  $|(f^n)'(x)| \gtrsim \delta^3 e^{\tilde{\lambda} n}$ . Arguing once again as in the previous case we then have

$$\delta \geq |\omega_n| \gtrsim \delta^3 e^{\tilde{\lambda} n} |\omega| \gtrsim \delta^4 e^{\tilde{\lambda} n}.$$

Solving for  $n$  once again we get the result.  $\square$

#### 5.4. Escape times and return depths.

*Proof of Lemma 9.* Let  $\omega' \in \mathcal{E}_{n,R}$  and let  $\nu_1, \dots, \nu_s$  be the sequence of essential returns,  $r_1, \dots, r_s$  the corresponding sequence of return depths, and  $p_1, \dots, p_s$  the corresponding binding periods. To simplify the notation we let  $\nu_{s+1} := n$ .

Lemma 8 implies  $\nu_1 \leq n_\delta$  and therefore it is sufficient to prove that there exists a constant  $\tilde{\theta}$  such that for all  $i = 1, \dots, s$  we have

$$\nu_{i+1} - \nu_i \leq \tilde{\theta} r_i. \quad (28)$$

Indeed, this immediately gives  $n \leq n_\delta + \tilde{\theta} R$  which is equivalent to the statement in the Lemma.

**Sublemma 9.1.** *For all  $i = 1, \dots, s$  we have*

$$\inf_{x \in \omega'} |(f^{\nu_{i+1}-\nu_i})'(x)| \leq \delta e^{r_i} r_i^2. \quad (29)$$

*Proof.* Since  $\nu_i$  is an essential return there exists an interval  $\hat{\omega}'$  with  $\omega' \subseteq \hat{\omega}' \subseteq \omega$  such that the image  $\hat{\omega}'_{\nu_i}$  at time  $\nu_i$  contains a unique interval  $I_{r,j} = I_{r_i,j_i}$  and therefore  $|\hat{\omega}'_{\nu_i}| \geq e^{-r_i}/r_i^2$ . Since we are assuming here that  $\hat{\omega}'$  does not escape before or at time  $\nu_{i+1}$  we also have  $\delta \geq |\hat{\omega}'_{\nu_{i+1}}|$ . Therefore, by the mean value theorem we have

$$\delta \geq |\hat{\omega}'_{\nu_{i+1}}| \geq \inf_{x \in \hat{\omega}'} |(f^{\nu_{i+1}-\nu_i})'(x)| |\hat{\omega}'_{\nu_i}| \geq \inf_{x \in \hat{\omega}'} |(f^{\nu_{i+1}-\nu_i})'(x)| e^{-r_i}/r_i^2$$

which gives (29).  $\square$

**Sublemma 9.2.** *For all  $i = 1, \dots, s$  we have*

$$\inf_{x \in \omega'} |(f^{\nu_{i+1}-\nu_i})'(x)| \geq \delta e^{\hat{\lambda}(\nu_{i+1}-\nu_i)}. \quad (30)$$

*Proof.* Let  $\rho_1, \rho_2, \dots, \rho_{t(i)}$  denote the sequence of return depths corresponding to inessential returns occurring between  $\nu_i$  and  $\nu_{i+1}$  and  $p_1, \dots, p_{t(i)}$  the corresponding binding periods. Then, combining Lemma 2 and the expansion (H1) during free iterates, keeping in mind that every free iterate under consideration ends in  $\delta$ , we have  $|(f^{\nu_{i+1}-\nu_i})'(x)| \geq e^{\hat{\lambda}(\nu_{i+1}-\nu_i)}$  for all  $i = 1, \dots, s-1$  and  $|(f^{\nu_{i+1}-\nu_i})'(x)| \geq \delta e^{\hat{\lambda}(\nu_{i+1}-\nu_i)}$  for  $i = s$ . The factor  $\delta$  in the second inequality is due to the fact that  $n$  is not necessarily a return.  $\square$

Returning to the proof of the Lemma, we combine (30) into (29) to get

$$e^{r_i} r_i^2 \geq e^{\hat{\lambda}(\nu_{i+1}-\nu_i)}$$

and solving for  $\nu_{i+1} - \nu_i$  gives

$$\nu_{i+1} - \nu_i \leq \frac{1}{\hat{\lambda}} (r_i + 2 \log r_i) \leq \frac{1 + \eta_3}{\hat{\lambda}} r_i$$

where  $\eta_3$  is the constant in (25). Thus the Lemma follows with  $\tilde{\theta} = (1 + \eta_3)/\hat{\lambda}$ .  $\square$

## 6. RETURN TIME ESTIMATES

We are now ready to prove Proposition 2. We remark that from this point onwards we shall *consider  $\delta$  to be fixed once and for all*. In particular we shall make use of the distortion constant  $\mathcal{D}_\delta$  from Proposition 3; this constant depends on  $\delta$  but this will not cause any problems as we shall not impose any additional conditions on the size of  $\delta$ .

We proceed initially as in the proof of the analogous estimate for the escape time, although the argument here is more probabilistic because we do not have such sharp control of the combinatorics. We consider a fixed  $n \geq 1$  and recall the definition of

$$\mathcal{Q}^{(n)} = \{\omega : T(\omega) > n\}$$

as the set of intervals which have not yet had a full return at time  $n$ . By construction, each  $\omega \in \mathcal{Q}^{(n)}$  is contained in a nested sequence of intervals

$$\omega \subset \omega^{(s)} \subset \omega^{(s-1)} \subset \dots \subset \omega^{(1)} \subset \Delta^*$$

corresponding to escape times  $E_1, \dots, E_s$ , such that  $|f^{E_i}(\omega^{(i)})| \geq \delta$  for  $i = 1 \dots, s$ . This sequence is empty for those elements of  $\mathcal{Q}^{(n)}$  which have not had any escape before time  $n$  (such as those which start very close to the critical point). For  $s = 0, \dots, n$  (clearly there cannot be more than  $n$  escapes) we let  $\mathcal{Q}_s^{(n)}$  denote the collection of intervals in  $\mathcal{Q}^{(n)}$  which have exactly  $s$  escapes before time  $n$ . Then for a constant  $\zeta \in (0, 1)$  whose value will be determined below, we write

$$|\mathcal{Q}^{(n)}| = \sum_{s \leq n} |\mathcal{Q}_s^{(n)}| = \sum_{s \leq \zeta n} |\mathcal{Q}_s^{(n)}| + \sum_{\zeta n < s \leq n} |\mathcal{Q}_s^{(n)}|. \quad (31)$$

This corresponds to distinguishing those intervals which have had lots of escape times and those that have had only a few. The Proposition then follows immediately from the following two lemmas.

**Lemma 10.** *There exists a constant  $C_3 > 0$  depending on  $\delta$  and  $\delta^*$  and a constant  $\gamma_3 > 0$  independent of  $\delta$  and  $\delta^*$  such that, for all  $\zeta > 0$  sufficiently small and all  $n \geq 1$  we have*

$$\sum_{0 \leq s \leq \zeta n} |\mathcal{Q}_s^{(n)}| \leq C_3 e^{-\gamma_3 n} |\Delta^*|.$$

**Lemma 11.** *There exists a constant  $C_4 > 0$  depending on  $\delta$  and  $\delta^*$  and a constant  $\gamma_4 > 0$  independent of  $\delta$  and  $\delta^*$  such that, for all  $\zeta > 0$  sufficiently small and all  $n \geq 1$  we have*

$$\sum_{\zeta n < s \leq n} |\mathcal{Q}_s^{(n)}| \leq C_4 e^{-\gamma_4 n} |\Delta^*|.$$

### 6.1. Returns after few escapes.

*Proof of Lemma 10.* Starting with the escape partition on  $\Delta^*$ , letting  $t_1 \geq 1$  be an integer, we write

$$\{E_1 = t_1\}$$

for the set of points in  $\Delta^*$  which belong to intervals which have a first escape at time  $t_1$ . By Proposition 1, using the fact that  $|\Delta^*| = \delta^*$ , we have

$$|\{E_1 = t_1\}| \leq |\mathcal{E}_{t_1}(\Delta^*)| \leq C_1 e^{-\gamma_1 t_1} = \frac{C_1}{\delta^*} e^{-\gamma_1 t_1} |\Delta^*|.$$

We then repeat the escape partition construction on each component of  $\{E_1 = t_1\}$ . We let

$$\{E_2 = t_2 + t_1 : E_1 = t_1\}$$

denote the set of points which belong to an interval which has a first escape at time  $t_1$  and a second escape at time  $t_2 + t_1$ , i.e.  $t_2$  iterates after the first escape.

Proposition 1 gives an estimate for the size of  $\{E_2 = t_2 + t_1 : E_1 = t_1\}$  at time  $t_1$ : if  $\omega$  is an escape component belonging to  $\{E_1 = t_1\}$  then

$$|\mathcal{E}_{t_2}(f^{t_1}(\omega))| \leq C_1 e^{-\gamma t_2} \leq \frac{C_1}{\delta} e^{-\gamma t_2} |(f^{t_1}(\omega))|.$$

Using the bounded distortion property from Proposition 3 this ratio is preserved for the initial interval  $\omega$  up to the bounded distortion constant  $\mathcal{D}_\delta$  to get

$$|\{E_2 = t_2 + t_1 : E_1 = t_1\}| \leq \frac{C_1 \mathcal{D}_\delta}{\delta} e^{-\gamma t_2} |\{E_1 = t_1\}| \leq \frac{C_1^2 \mathcal{D}_\delta^2}{\delta \delta^*} e^{-\gamma(t_1+t_2)} |\Delta^*|.$$

Continuing in this way we get

$$|\{E_s = t_1 + \dots + t_s : E_1 = t_1, E_2 = t_2, \dots, E_{s-1} = t_{s-1}\}| \leq \frac{C_1^s \mathcal{D}_\delta^{s-1}}{\delta^{s-1} \delta^*} e^{-\gamma E_s} |\Delta^*|.$$

Letting  $t_{s+1} = n - E_s$  we then apply one more iteration of this formula to get

$$|\mathcal{Q}_s^{(n)}(t_1, \dots, t_{s+1})| \leq \frac{C_1^{s+1} \mathcal{D}_\delta^s}{\delta^s \delta^*} e^{-\gamma n} |\Delta^*|.$$

where

$$\mathcal{Q}_s^{(n)}(t_1, \dots, t_{s+1}) = \{\omega \in \mathcal{Q}_s^{(n)} : E_{i+1}(\omega) - E_i(\omega) = t_i, i = 1, \dots, s, n - E_s = t_{s+1}\}.$$

Thus for each  $s$  we have

$$|\mathcal{Q}_s^{(n)}| \leq \sum_{\substack{(t_1, \dots, t_{s+1}) \\ \sum t_j = n}} |\mathcal{Q}_s^{(n)}(t_1, \dots, t_{s+1})| \leq N_{n,s+1} \frac{C_1^{s+1} \mathcal{D}_\delta^s}{\delta^s \delta^*} e^{-\gamma n} |\Delta^*|$$

where  $N_{n,s+1}$  is the number of possible sequences  $(t_1, \dots, t_{s+1})$  such that  $\sum t_j = n$ , and therefore

$$\sum_{0 \leq s \leq \zeta n} |\mathcal{Q}_s^{(n)}| \leq \sum_{s \leq \zeta n} N_{n,s} \frac{C_1^{s+1} \mathcal{D}_\delta^s}{\delta^s \delta^*} e^{-\gamma n} |\Delta^*|$$

Using exactly the same counting argument used in the proof of Lemma 7 we can choose  $\hat{\zeta} > 0$  arbitrarily small as long as  $s \leq \zeta n$  and  $\zeta$  is sufficiently small, so that  $N_{n,s+1} \leq e^{\hat{\zeta} n}$ . This gives

$$\sum_{0 \leq s \leq \zeta n} |\mathcal{Q}_s^{(n)}| \leq \frac{C_1}{\delta^*} \left( \frac{C_1 \mathcal{D}_\delta}{\delta} \right)^{\zeta n} e^{\hat{\zeta} n} e^{-\gamma n} |\Delta^*|.$$

Since both  $\zeta$  and  $\hat{\zeta}$  can be taken arbitrarily small, the result follows.  $\square$

**Remark 2.** Notice that taking  $\zeta$  small (and thus obtaining  $\hat{\zeta}$  small) does not depend on choosing  $\delta$  or  $\delta^*$  sufficiently small as  $\zeta$  can be chosen arbitrarily in the decomposition of the sum in (31).

## 6.2. Returns after many escapes.

*Proof of Lemma 11.* To bound the set of points which have many escapes we use a softer argument. Clearly we have

$$|\{E_1 \leq n\}| \leq |\Delta^*|.$$

By Lemma 1 there is a fixed proportion  $\xi$  of the image  $f^{E_1(\omega)}$  of each escaping component which actually has a “full return” to  $\Delta^*$  within  $t^*$  iterates after this escape. Using the bounded distortion estimate in Proposition 3 again this translates



to a minimum proportion  $\xi/\mathcal{D}_\delta$  of the actual interval  $\omega$ . Therefore the maximum proportion of  $|\{E_1 \leq n\}|$  which is even allowed potentially to have a second escape (whether it be before or after time  $n$ ) is bounded by  $1 - \xi/\mathcal{D}_\delta < 1$  and so we have

$$|\{E_2 \leq n\}| \leq |\{E_2 \text{ defined} : E_1 \leq n\}| \leq \left(1 - \frac{\xi}{\mathcal{D}_\delta}\right) |\{E_1 \leq n\}| \leq \left(1 - \frac{\xi}{\mathcal{D}_\delta}\right) |\Delta^*|.$$

Iterating this formula we get

$$|\{E_s \leq n\}| \leq \left(1 - \frac{\xi}{\mathcal{D}_\delta}\right)^{s-1} |\Delta^*|$$

and so there exists some constant  $C_4 > 0$  such that

$$\sum_{\zeta n < s < n} |\mathcal{Q}_s^{(n)}| \leq \sum_{\zeta n < s < n} \left(1 - \frac{\xi}{\mathcal{D}_\delta}\right)^{s-1} |\Delta^*| \leq C_4 \left(1 - \frac{\xi}{\mathcal{D}_\delta}\right)^{\zeta n} |\Delta^*|. \quad (32)$$

The statement in the Lemma follows with  $\gamma_4 = -\zeta \log(1 - \xi/\mathcal{D}_\delta)$ . □

## 7. STATISTICAL PROPERTIES

By Proposition 2, we know the tail of the return times is decaying exponentially fast and thus  $\mathcal{Q}$  is indeed a partition of  $\Delta^*$ , while by Proposition 3, we have the required bounded distortion property. Thus together, these two Propositions imply Theorem 1. Moreover, the construction and estimates given above also yield the following additional properties which will be used below: there exists some  $\lambda' > 1$  such that

$$d(f^T x, f^T y) \geq \lambda' d(x, y) \quad (33)$$

and there exists a constant  $C > 0$  such that

$$d(f^k x, f^k y) \leq C d(f^T x, f^T y), \text{ for all } k < T. \quad (34)$$

Indeed, (33) follows by the same expansion estimates used repeatedly in the proof and (34) follows almost trivially from the bounded distortion property and the fact that  $\omega$  never gets bigger than  $\delta$ . It remains to show how these properties imply the required statistical properties.

**7.1. Absolutely continuous invariant probability measures.** The bounded distortion property of the induced map  $f^T : \Delta^* \rightarrow \Delta^*$  implies, by classical results, the existence of an ergodic, in fact mixing, absolutely continuous invariant measure  $\hat{\mu}$  for  $f^T$  with bounded density with respect to Lebesgue. Therefore, the exponential tail of the return time function  $T$  implies in that

$$\int T d\mu = \sum_{\omega \in \mathcal{Q}} T(\omega) \mu(\omega) < \infty.$$

This means that the measure obtained by *pushing forward*  $\hat{\mu}$  by iterates of the original map  $f$  is finite and thus, after normalization, yields an absolutely continuous probability measure  $\mu$  which, again by standard arguments, is also  $f$ -invariant and ergodic, and in fact mixing for some power of  $f$ . The expansivity estimates obtained above clearly imply that  $\mu$  has a positive Lyapunov exponent.

**7.2. The Markov extension.** First of all we define a *Markov extension* or *Markov tower*. Let  $\mathcal{T}_0$  be a copy  $\Delta^*$  and let  $\mathcal{T}_{j,0}$  be a copy of the element  $\omega_j$  of the partition  $\mathcal{Q}$ . The return time function  $T : \Delta^* \rightarrow \mathbb{N}$  is constant on partition elements  $\mathcal{T}_{j,0}$  with value  $T|_{\mathcal{T}_{j,0}} = T_j \geq 1$ . Then define

$$\mathcal{T} = \{(x, k) : x \in \mathcal{T}_0, k = 0, \dots, T(x) - 1\} = \bigcup_{j \geq 1} \bigcup_{k=0}^{T_j-1} \mathcal{T}_{j,k},$$

where

$$\mathcal{T}_{j,k} = \mathcal{T}_{j,0} \times \{k\}.$$

So  $\mathcal{T}$  is the disjoint union of  $T_j$  copies of each  $\mathcal{T}_{j,0}$ . Define the *tower map*

$$F : \mathcal{T} \rightarrow \mathcal{T}$$

by setting

$$F(x, k) = (x, k + 1)$$

for  $0 \leq k < T(x) - 1$  and

$$F(x, T(x) - 1) = (f^T x, 0).$$

$F$  is Markov with respect to the partition  $\{\mathcal{T}_{j,k}\}$  and the return map  $F^T : \mathcal{T}_{j,0} \rightarrow \mathcal{T}_{j,0}$ , or equivalently  $f^T : \Delta^* \rightarrow \Delta^*$ , is full branched Markov with respect to the partition  $\{\mathcal{T}_{j,0}\} = \mathcal{Q}$  on  $\mathcal{T}_0 = \Delta^*$ .

**7.3. Projections.** Define the projection

$$\pi : \mathcal{T} \rightarrow \hat{J}$$

by  $\pi(x, k) = f^k(x)$ . Clearly,  $\pi$  is a semi-conjugacy between  $F$  and  $f$ :

$$f \circ \pi = \pi \circ F. \tag{35}$$

Using the projection  $\pi$ , any measure  $\tilde{\mu}$  on  $\mathcal{T}$  can be projected to a measure  $\mu = \pi_* \tilde{\mu}$  on  $\hat{J}$  with the property that  $\mu(A) = \tilde{\mu}(\pi^{-1}(A))$  for every measurable set  $A \subset \hat{J}$ . Moreover any observable

$$\varphi : \hat{J} \rightarrow \mathbb{R}$$

can be lifted to an observable

$$\tilde{\varphi} : \mathcal{T} \rightarrow \mathbb{R} \quad \text{given by} \quad \tilde{\varphi} = \varphi \circ \pi.$$

Notice that (35) implies

$$\tilde{\varphi} \circ F^n = \varphi \circ f^n \circ \pi \quad \text{and} \quad \int \tilde{\varphi} d\tilde{\mu} = \int \varphi \circ \pi d\tilde{\mu} = \int \varphi d(\pi_* \tilde{\mu}).$$

Therefore, for any two observables  $\varphi, \psi : \hat{J} \rightarrow \mathbb{R}$  and their corresponding lifts  $\tilde{\varphi}, \tilde{\psi} : \mathcal{T} \rightarrow \mathbb{R}$ , any measure  $\tilde{\mu}$  and its corresponding projection  $\mu = \pi_* \tilde{\mu}$  it is sufficient to prove statistical properties for the lifts on the tower to obtain similar results for the original map using the fact that  $\pi$  is a “measure-preserving” (by

definition, since one measure is the projection of the other by  $\pi$ ) semi-conjugacy. For example, for the correlation function we have

$$\begin{aligned} \int (\tilde{\varphi} \circ F^n) \tilde{\psi} d\tilde{\mu} - \int \tilde{\varphi} d\tilde{\mu} \int \tilde{\psi} d\tilde{\mu} &= \int (\varphi \circ (f^n \circ \pi)) (\psi \circ \pi) d\tilde{\mu} - \int (\varphi \circ \pi) d\tilde{\mu} \int (\psi \circ \pi) d\tilde{\mu} \\ &= \int (\varphi \circ f^n) \psi d(\pi_* \tilde{\mu}) - \int \varphi d(\pi_* \tilde{\mu}) \int \psi d(\pi_* \tilde{\mu}) \\ &= \int (\varphi \circ f^n) \psi d\mu - \int \varphi d\mu \int \psi d\mu. \end{aligned}$$

An important observation however is that the properties must be proved on the tower for all observables which can be obtained as lifts of suitable (Hölder continuous) observables on the manifold.

**7.4. Statistical properties on the tower.** In this section we state precisely the results of Young [52] which give conditions for the exponential decay of certain classes of observables for the map  $F : \mathcal{T} \rightarrow \mathcal{T}$  and for the Central Limit Theorem. Young's setting is very general and only requires  $\Delta$  to be a measure space with some reference measure  $m$  and the Markov tower structure described above. The Markov structure naturally induces a symbolic metric on  $\mathcal{T}$  which can be used to define notions of regularity both for the Jacobian of  $F$  and for observables on  $\mathcal{T}$ .

**7.4.1. Symbolic metric.** First of all we define a *separation time*: if  $x, y \in \mathcal{T}_{j,0}$  for some  $j$ , then

$$s(x, y) = \max \{n \geq 0 \text{ s.t. } (F^T)^n x \text{ and } (F^T)^n y \text{ lie in the same partition element of } \mathcal{T}_0\}$$

If  $x, y$  lie in distinct partition elements, then  $s(x, y) = 0$ . If  $x, y \in \mathcal{T}_{j,k}$  then write  $x = F^k x_0, y = F^k y_0$  where  $x_0, y_0 \in \mathcal{T}_{j,0}$  and define  $s(x, y) = s(x_0, y_0)$ . Then, for a fixed constant  $\sigma \in (0, 1)$ , we define

$$d_\sigma(x, y) = \sigma^{s(x,y)}.$$

By the expansivity property of  $F^T$  any two distinct points eventually separate and therefore this defines a metric on  $\mathcal{T}$ .

**7.4.2. Regularity of observables in the symbolic metric.** Using the metric  $d_\sigma$  we define

$$\mathcal{H}_\sigma(\mathbb{R}) = \{\phi : \mathcal{T} \rightarrow \mathbb{R} : \exists C_\phi : |\phi(x) - \phi(y)| \leq C_\phi \sigma^{s(x,y)} \forall x, y \in \mathcal{T}\}$$

to be the space of ‘‘Hölder continuous’’ functions on  $\mathcal{T}$ .

**7.4.3. Young's Theorem.** Let  $F : \mathcal{T} \rightarrow \mathcal{T}$  be a Markov map on a tower  $\mathcal{T}$  with a reference measure  $m$ , a Markov return map  $F^T : \mathcal{T}_0 \rightarrow \mathcal{T}_0$ , an exponential tail of the return time function  $T : \mathcal{T}_0 \rightarrow \mathbb{N}$ , and such that the greatest common divisor of all values taken by the function  $T$  is 1. Let  $JF^T$  denote the Jacobian of  $F^T$  with respect to the reference measure  $m$ .

**Theorem** ([52]). *Suppose that the following bounded distortion condition holds: there exists a constant  $\tilde{C}$  such that for all  $x, y \in \mathcal{T}_0$  we have*

$$\left| \frac{|JF^T(x)|}{|JF^T(y)|} - 1 \right| \leq \tilde{C} \sigma^{s(x,y)}. \quad (36)$$

*Then the correlation function  $C_n(\tilde{\varphi}, \tilde{\psi}, \tilde{\mu})$  decays exponentially fast (at a uniform rate) for every observable  $\tilde{\varphi} \in L^\infty(\tilde{\mu})$  and  $\tilde{\psi} \in \mathcal{H}_\sigma$ . Moreover, the Central Limit Theorem holds for every  $\tilde{\psi} \in \mathcal{H}_\sigma$ .*

Notice that the bounded distortion condition (36) is the *only* assumption of the theorem over and above the Markov tower structure, the exponential decay of the return time function, and the assumption on the greatest common divisor of  $T$ . Moreover, both the exponential tail and the assumption on the gcd can be relaxed to some extent. If the gcd of  $T$  is  $k > 1$  we can just consider  $\tilde{F} = F^k : \mathcal{T} \rightarrow \mathcal{T}$  which will continue to have the same Markov structure and properties and a return time function  $\tilde{T}$  with  $\text{gcd}(\tilde{T}) = 1$ . The exponential decay of the tail of the return time function can also be relaxed to assume only subexponential or even polynomial decay (in fact including these cases is one of the main motivations for [52]), although in this case the rate of decay of the correlation function is correspondingly slower.

**7.5. Decay of correlations for the original map.** To obtain our desired results for  $f$  we therefore just need to show that the bounded distortion condition of Proposition 3 implies the required bounded distortion condition (36) in the symbolic metric, and that observables on  $\hat{J}$  which are Hölder continuous with respect to the usual Euclidean metric lift to bounded observables on  $\mathcal{T}$  which are Hölder continuous in the symbolic metric.

**Lemma 12.** *The bounded distortion condition (36) holds.*

*Proof.* For the bounded distortion condition, notice first of all that  $F^T$  is differentiable in our setting and therefore the left hand side of (36) formulated in terms of the Jacobian of  $F$  is exactly the same as the left hand side of the second inequality in Proposition 3 formulated in terms of the derivative of  $f$ , with  $k = T$ . Thus all we need to show is that

$$\tilde{\mathcal{D}}|f^T(x) - f^T(y)| \leq C\sigma^{s(x,y)}|\Delta^*|$$

for any two points  $x, y$  belonging to the same element of the partition  $\mathcal{Q} = \mathcal{T}_0$ . To see this, let  $\lambda'$  be the expansion constant in (33), then cylinder sets are shrinking exponentially at rate  $\lambda'$  and so we have

$$|x - y| \leq |\Delta^*|\lambda'^{-s(x,y)} \quad \text{and so} \quad |f^T(x) - f^T(y)| \leq |\Delta^*|\lambda'^{-s(x,y)+1}. \quad (37)$$

Thus the result follows as long as  $1 > \sigma > \lambda'^{-1}$ .  $\square$

**Lemma 13.** *If  $\varphi$  is Hölder continuous with exponent  $\gamma \geq \tilde{\gamma}$  then  $\tilde{\varphi} = \pi \circ \varphi \in \mathcal{H}_\sigma$ .*

*Proof.* Let  $\varphi : \hat{J} \rightarrow \mathbb{R}$  be Hölder continuous: there exists  $C_\varphi$  such that

$$|\varphi(x) - \varphi(y)| \leq C_\varphi|x - y|^\alpha$$

for any two  $x, y \in \hat{J}$ . Now, for any  $\tilde{x}, \tilde{y}$  belonging to some partition element of  $\mathcal{T}$  we have

$$|\tilde{\varphi}(\tilde{x}) - \tilde{\varphi}(\tilde{y})| = |\varphi(\pi(\tilde{x})) - \varphi(\pi(\tilde{y}))| \leq C_\varphi|\pi(x) - \pi(y)|^\alpha.$$

Since  $\pi(x)$  and  $\pi(y)$  belong to the image  $\omega_k$  of some element  $\omega \in \mathcal{Q}$  for  $k < T$ , then (34) and (37) give

$$|\pi(x) - \pi(y)|^\alpha \leq C^\alpha|f^{T-k}(\pi(x)) - f^{T-k}(\pi(y))|^\alpha \leq |C^\alpha\Delta^*|^\alpha\lambda'^{\alpha(-s(x,y)+1)}.$$

The result follows if  $1 > \sigma > \lambda'^{-\alpha}$ .  $\square$

## REFERENCES

- [1] J. Aaronson, *An Introduction to Infinite Ergodic Theory*. Math. Surveys and Monographs **50**. Amer. Math. Soc, (1997).
- [2] J. Aaronson, M. Denker, and M. Urbanski, *Ergodic theory for Markov fibred systems and parabolic rational maps*, Trans. Am. Math. Soc **337** (1993), no. 2, 495–548.
- [3] J. F. Alves, C. Bonatti, and M. Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), no. 2, 351–398.
- [4] J. F. Alves, S. Luzzatto, and V. Pinheiro, *Lyapunov exponents and rates of mixing for one-dimensional maps.*, Ergodic Th. & Dyn. Syst **24** (2004), 637–657.
- [5] J. F. Alves, Stefano Luzzatto, and Vilton Pinheiro, *Markov structures and decay of correlations for non-uniformly expanding maps.*, Ann. Inst. H. Poin. Anal. Non-Lin. **22** (2005), 817–839.
- [6] M. Benedicks and L. Carleson, *On iterations of  $1 - ax^2$  on  $(-1, 1)$* , Ann. of Math. **122** (1985), 1–25.
- [7] H. Bruin, S. Luzzatto, and S. van Strien, *Decay of correlations in one-dimensional dynamics*, Ann. Sci. Éc. Norm. Sup. **36** (2003), no. 4, 621–646.
- [8] H. Bruin, W. Shen, and S. van Strien, *Invariant measures exist without a growth condition*, Comm. Math. Phys. **241** (2003), no. 2-3, 287–306.
- [9] J. Buzzi and O. Sarig, *Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps*, Ergodic Theory Dynam. Systems **23** (2003), no. 5, 1383–1400.
- [10] J.R. Chazottes and S. Gouëzel, *On almost-sure versions of classical limit theorems for dynamical systems*, Preprint (2006).
- [11] P. Collet and J.-P. Eckmann, *Positive Lyapunov exponents and absolute continuity for maps of the interval*, Ergodic Theory Dynam. Systems **3** (1983), no. 1, 13–46.
- [12] M. Denker, Z. Nitecki, and M. Urbański, *Conformal measures and  $S$ -unimodal maps*, Dynamical systems and applications, 1995, pp. 169–212.
- [13] M. Denker, F. Przytycki, and M. Urbanski, *On the transfer operator for rational functions on the Riemann sphere*, Ergodic Theory Dynam. Systems **16** (1996), 155–266.
- [14] K. Díaz-Ordaz, *Decay of correlations for non-Hölder observables for one-dimensional expanding Lorenz-like maps*, Discrete Contin. Dyn. Syst. **15** (2006), no. 1, 159–176.
- [15] K. F. Gauss, *Vol. X-1* **371** (1870), 552–552.
- [16] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math. **139** (2004), 29–65.
- [17] ———, *Decay of correlations for nonuniformly expanding systems*, Bulletin de la Société Mathématique de France, to appear (2004).
- [18] F. Hofbauer and G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. **180** (1982), no. 1, 119–140.
- [19] M. Holland, *Slowly mixing systems and intermittency maps*, Erg. Th. & Dyn. Syst. **25:1** (2005), 133–159.
- [20] M. Holland and S. Luzzatto, *Dynamics of two dimensional maps with criticalities and singularities* (2003), In progress.
- [21] M. V. Jakobson, *Topological and metric properties of one-dimensional endomorphisms*, Sov. Math. Dokl. **19** (1978), 1452–1456.
- [22] ———, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), 39–88.
- [23] M. V Jakobson, *Piecewise smooth maps with absolutely continuous invariant measures and uniformly scaled Markov partitions*, Proceedings of Symposia in Pure Mathematics **69** (2001), 825–881.
- [24] O. Jenkinson, R. D. Mauldin, and M. Urbański, *Zero temperature limits of Gibbs-equilibrium states for countable alphabet subshifts of finite type*, J. Stat. Phys. **119** (2005), no. 3-4, 765–776.
- [25] G. Keller, *Generalized bounded variation and applications to piecewise monotonic transformations.*, Z. Wahrsch. Verw. Gebiete **69** (1985), no. 3, 461–478.
- [26] G. Keller and T. Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps*, Comm. Math. Phys. **149** (1992), no. 1, 31–69.
- [27] A. Lasota and J. A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **186** (1973), 481–488 (1974).

- [28] C. Liverani, *Decay of correlations*, Ann. of Math. (2) **142** (1995), no. 2, 239–301.
- [29] S. Luzzatto, *Stochastic behaviour in non-uniformly expanding maps*, Handbook of Dynamical Systems (2005).
- [30] S. Luzzatto and W. Tucker, *Non-uniformly expanding dynamics in maps with singularities and criticalities*, Inst. Hautes Études Sci. Publ. Math. (1999), no. 89, 179–226.
- [31] S. Luzzatto and M. Viana, *Lorenz-like attractors without continuous invariant foliations*, In progress.
- [32] R. D. Mauldin and M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet*, Israel J. Math. **125** (2001), 93–130.
- [33] I. Melbourne and M. Nicol, *Statistical properties of endomorphisms and compact group extensions*, J. London Math. Soc. (2004), no. 70, 427–446.
- [34] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Inst. Hautes Études Sci. Publ. Math. (1981), no. 53, 17–51.
- [35] T. Nowicki and S. van Strien, *Absolutely continuous invariant measures for  $C^2$  unimodal maps satisfying the Collet-Eckmann conditions*, Invent. Math. **93** (1988), no. 3, 619–635.
- [36] ———, *Invariant measures exist under a summability condition for unimodal maps*, Invent. Math. **105** (1991), no. 1, 123–136.
- [37] W. Parry, *On the  $\beta$ -expansions of real numbers*, Acta Math. Acad. Sci. Hungar. **11** (1960), 401–416.
- [38] Y. Pesin and S. Senti, *Equilibrium measures for some one-dimensional maps*, Preprint (2005).
- [39] Y. Pesin and K. Zhang, *Phase transitions for uniformly expanding maps*, J. Stat. Phys. (2005).
- [40] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar **8** (1957), 477–493.
- [41] D. Ruelle, *Applications conservant une mesure absolument continue par rapport à  $dx$  sur  $[0, 1]$* , Comm. Math. Phys. **55** (1977), no. 1, 47–51 (French, with English summary).
- [42] M. Rychlik, *Bounded variation and invariant measures*, Studia Math. **76** (1983), no. 1, 69–80.
- [43] O. Sarig, *Thermodynamic formalism for countable Markov shifts*, Ergodic Theory Dynam. Systems **19** (1999), no. 6, 1565–1593.
- [44] ———, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653.
- [45] ———, *Existence of Gibbs measures for countable Markov shifts*, Proc. Amer. Math. Soc. **131** (2003), no. 6, 1751–1758 (electronic).
- [46] F. Schweiger, *Number theoretical endomorphisms with  $\sigma$ -finite invariant measures*, Israel Journal of Math **21** (1975), 308–318.
- [47] M. Thaler, *Estimates of the invariant densities of endomorphisms with indifferent fixed points*, Israel Journal of Math **37** (1980), 303–314.
- [48] S. Ulam and J. von Neumann, *On combination of stochastic and deterministic processes*, Bull. AMS **53** (1947), 1120.
- [49] M. Viana, *Stochastic dynamics of deterministic systems*, Lecture Notes XXI Braz. Math. Colloq., IMPA, Rio de Janeiro, 1997.
- [50] L.-S. Young, *Decay of correlations for certain quadratic maps*, Comm. Math. Phys. **146** (1992), no. 1, 123–138.
- [51] ———, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650.
- [52] ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.
- [53] M. Yuri, *Thermodynamic formalism for certain nonhyperbolic maps*, Ergodic Theory Dynam. Systems **19** (1999), no. 5, 1365–1378.
- [54] R. Zweimüller, *Invariant measures for general(ized) induced transformations*, Proc. Amer. Math. Soc. **133** (2005), no. 8, 2283–2295 (electronic).

MATHEMATICS DEPARTMENT, IMPERIAL COLLEGE, 180 QUEEN'S GATE, LONDON SW7 2AZ

*E-mail address:* [karla.diaz-ordaz@imperial.ac.uk](mailto:karla.diaz-ordaz@imperial.ac.uk)

*URL:* <http://www.ma.ic.ac.uk/~kd2>

MATHEMATICS DEPARTMENT, UNIVERSITY OF EXETER, EXETER EX4 4QF, UK

*E-mail address:* [M.P.Holland@exeter.ac.uk](mailto:M.P.Holland@exeter.ac.uk)

*URL:* <http://www.secam.ex.ac.uk/people/staff/mp204>

MATHEMATICS DEPARTMENT, IMPERIAL COLLEGE, 180 QUEEN'S GATE, LONDON SW7 2AZ

*E-mail address:* [Stefano.Luzzatto@imperial.ac.uk](mailto:Stefano.Luzzatto@imperial.ac.uk)

*URL:* <http://www.ma.ic.ac.uk/~luzzatto>