INVARIANT MEASURES FOR INTERVAL MAPS WITH CRITICAL POINTS AND SINGULARITIES

VÍTOR ARAÚJO, STEFANO LUZZATTO, AND MARCELO VIANA

Abstract. We prove that, under a mild summability condition on the growth of the derivative on critical orbits any piecewise monotone interval map possibly containing discontinuities and singularities with infinite derivative (cusp map) admits an ergodic invariant probability measures which is absolutely continuous with respect to Lebesgue measure.

1. Introduction and statement of results

1.1. **Introduction.** The existence of absolutely continuous invariant probability measures (*acip*'s) for dynamical systems is a problem with a history going back more than 70 years, see for example pioneering papers by Hopf [9] and Ulam and von Neumann [15]. Notwithstanding an extensive amount of research in this direction in the last two or three decades, the problem is still not completely solved even in the one-dimensional setting which is the focus of this paper. Quite general conditions are known which guarantee the existence of *acip*'s for uniformly expanding maps in the smooth case or possibly admitting singularities, i.e. discontinuities with possibly unbounded derivatives (see [16][10] for additional remarks and references), and for smooth maps with a finite number of critical points (see [4] for first and strongest results including decay of correlations, and [5] for the most recent and possibly the most general conditions for the existence of absolutely continuous invariant measures in this setting) and even for smooth maps with a countable number of critical points [2]. We are interested here in a general class of maps which contain critical points *and* singularities.

A natural family of maps belonging to this class was introduced in [11, 12] and motivated by the study of the return map of the Lorenz equations near classical parameter values, see Figure 1. It is clear from the arguments in these papers, that the presence of both critical points and singularities and their interaction can give rise to significant technical as well as fundamental issues. In particular, as we shall see in the present setting, it is not enough to have just some expansivity conditions in order to obtain the existence of an

---

Date: 20 February 2009.

This research was partly supported by a Royal Society International Joint Project Grant and EPSRC grant number GR T09699 01. V.A. was also supported by FAPERJ, CNPq (Brazil) and CMUP-FCT (Portugal). M.V. was also supported by CNPq (Brazil), FAPERJ, and PRONEX-Dynamical Systems. The authors acknowledge the hospitality of Imperial College London and IMPA Rio de Janeiro where most of this work was carried out. Thanks also to Colin Little and to the anonymous referee for their careful reading of a preliminary version of the paper and for their useful remarks.

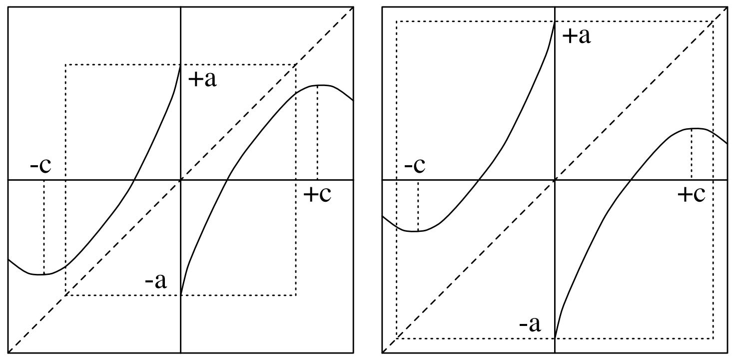

FIGURE 1. Interval maps with critical points and singularities

acip, as expansivity might occur due to the regions of unbounded derivative even when the deeper dynamical structure of the map is very pathological. Moreover, it is possible that the interaction of critical points and singularities could give rise to new phenomena which are still unexplored.

1.1.1. *Exponential growth and subexponential recurrence.* Some general results for the existence of *acip*'s and their properties in maps with critical points and singularities were obtained in [1] under the assumption that *Lebesgue almost every point* satisfy some exponential derivative growth and subexponential recurrence conditions. These conditions provide an interesting conceptual picture but may be hard to verify in practice. On the other hand, it was proved in [11] [12] that with positive probability in the parameter space of Lorenz-like families, the orbits of the *critical points* satisfy such exponential derivative growth and subexponential recurrence conditions. In [8] it was shown, within a more general setting of maps with multiple critical points and singularities, that these conditions are in fact sufficient to guarantee the existence of an ergodic *acip* (from which it can in fact be proved that Lebesgue almost every point also satisfies such conditions).

1.1.2. *Summability conditions.* Our aim in this paper is to obtain the same conclusion but relax as much as possible the conditions on the orbit of the critical points, to include in particular cases in which the derivative growth may be subexponential and/or the recurrence of the critical points exponential. A crucial observation concerning the difference between the smooth case and the case with singularities discussed here is that in the smooth case, for which in particular the derivative is bounded, any condition on the growth of the derivative is also implicitly a condition on the recurrence to the critical set. Indeed sufficiently strong recurrence to the critical set will always kill off any required derivative growth. On the other hand, this is not the case in our setting. Derivative growth may be exponential but arise as a consequence of very strong recurrence to the singularities even if we have at the same time very strong recurrence to the critical set. Strong recurrence to either the singular or the critical set brings its own deep structural problems and can be an intrinsic obstruction to the existence of an *acip*. We shall formulate below a condition which simultaneously keeps track of the growth of the derivative along critical orbits and of the recurrence of such orbits to the critical set within a single summability condition. This optimizes the result to include a larger class of maps than would be possible by having to independent conditions both of whichneed to be satisfied. We conjecture that it is not possible to obtain a general result on the existence of acip's in the presence of both critical points and singularities by assuming only conditions on the derivative growth of critical points.

1.2. **Statement of results.** We now give the precise statement of our result. We let $\mathcal{F}$ denote the class of interval map satisfying the conditions formulated in Sections 1.2.1, 1.2.2 and 1.2.3 below. Then we have the following

Theorem. *Every map $f \in \mathcal{F}$ admits a finite number of absolutely continuous invariant (physical) probability measures whose basins cover I up to a set of measure 0.*

1.2.1. *Nondegenerate critical/singular set.* Let M be an interval and $f : M \to M$ be a piecewise C^2 map: By this we mean that there exists a finite set $\mathcal{C}'$ such that f is C^2 and monotone on each connected component of $M \setminus \mathcal{C}'$ and admits a continuous extension to the boundary so that $f(c) := \lim_{x\to c^\pm} f(x)$ exists. We denote by $\mathcal{C}$ the set of all "one-sided critical points" c^+ and c^- and define corresponding one-sided neighbourhoods

$$\Delta(c^+,\delta) = (c^+, c^+ + \delta) \quad \text{and} \quad \Delta(c^-,\delta) = (c^- - \delta, c^-),$$

for each $\delta > 0$. For simplicity, from now on we use c to represent the generic element of $\mathcal{C}$ and write Δ for $\cup_{c\in\mathcal{C}}\Delta(c,\delta)$. We assume that each $c \in \mathcal{C}$ has a well-defined (one-sided) critical order $\ell = \ell(c) > 0$ in the sense that

$$(1) \quad |f(x) - f(c)| \approx d(x,c)^\ell \quad \text{and} \quad |Df(x)| \approx d(x,c)^{\ell-1} \quad \text{and} \quad |D^2f(x)| \approx d(x,c)^{\ell-2}$$

for all x in some $\Delta(c,\delta)$. Note that we say that $f \approx g$ if the ratio f/g is bounded above and below uniformly in the stated domain. If $\ell(c) < 1$ we say that c is a *singular* point as this implies unbounded derivative near c; if $1 < \ell(c)$ we say that c is a *critical* point as this implies that the derivative tends to 0 near c. We shall assume also that $\ell(c) \neq 1$ for every c as othis would be a degenerate case which is not hard to deal with but would require having to introduce special notation and special cases, whereas the other cases can all be dealt with in a unified formalism.

Remark 1.1. For future reference we point out that this immediately implies

$$(2) \qquad \frac{|D^2f(x)|}{|Df(x)|} \approx \frac{1}{d(x)}$$

for all x, where $d(x)$ denotes the distance of the point x to the critical/singular set $\mathcal{C}$ (indeed this is the actual property of which we will make use).

1.2.2. *Uniform expansion outside the critical neighbourhood.* We suppose that f is "uniformly expanding away from the critical points", meaning that the following two conditions are satisfied: there exists a constant $\kappa > 0$, independent of δ, such that for every point x

and every integer $n \geq 1$ such that $d(f^j(x), \mathcal{C}) > \delta$ for all $0 \leq j \leq n-1$ and $d(f^n(x), \mathcal{C}) \leq \delta$ we have

$$|Df^n(x)| \geq \kappa \tag{3}$$

and, for every $\delta > 0$ there exist constants $c(\delta) > 0$ and $\lambda(\delta) > 0$ such that

$$|Df^n(x)| \geq c(\delta) e^{\lambda(\delta) n} \tag{4}$$

for every x and $n \geq 1$ such that $d(f^j(x), \mathcal{C}) > \delta$ for all $0 \leq j \leq n-1$.

We remark that both these conditions are quite natural and are often satisfied for smooth maps without discontinuities. More specifically, the first one is satisfied if f is C^3, has negative Scharzian derivative and satisfies the property that the the derivative along all critical orbits tends to infinity, see Theorem 1.3 in [6]. The second is satisfied in even greater generality, namely when f is C^2 and all periodic points are repelling [13].

1.2.3. *Summability condition along the critical orbit.* For each $c \in \mathcal{C}$ we write

$$D_n(c) = |(f^n)'(f(c))| \quad \text{and} \quad d(c_n) = d(c_n, \mathcal{C})$$

to denote the derivative along the orbit of c and the distance of c from the critical set respectively. We then assume that for every critical point c with $\ell = \ell(c) > 1$ we have

$$\sum_n \frac{-n \log d(c_n)}{d(c_n) D_{n-1}^{1/(2\ell-1)}} < \infty. \tag{$\star$}$$

Remark 1.2. This condition plays off the derivative against the recurrence in such a way as to optimize to some extent the class of maps to which it applies. As mentioned in Section 1.1.2 above, we cannot expect to obtain the conclusions of our main theorem in this setting using a condition which only takes into account the growth of the derivative. Notice that condition ($\star$) is satisfied if the derivative is growing exponentially fast and the recurrence is not faster than exponential in the sense that

$$D_{n-1} \gtrsim e^{\lambda n} \quad \text{and} \quad d(c_n) \gtrsim e^{-\alpha n} \quad \text{with} \quad \alpha < \frac{\lambda}{2\ell - 1}.$$

Here and in the rest of the paper, the symbol $\gtrsim$ means that the inequality holds up to some multiplicatice constant, i.e. there exists a constant $C > 0$ indpendent of n or any other constants, such that $D_{n-1} \geq e^{\lambda n}$ and $d(c_n) \geq Ce^{-\alpha n}$.

2. The main technical theorem

2.1. **Inducing.** Our strategy for the proof is to construct a countable partition into intervals $\mathcal{I}$ of M (mod 0), define an inducing time function $\tau : M \to \mathbb{N}$ which is constant on elements of $\mathcal{I}$, and let $\hat{f} : M \to M$ denote the induced map defined by

$$\hat{f}(x) = f^{\tau(x)}(x).$$

This induced map is uniformly expanding on each element of $\mathcal{I}$ but does *not* have many desirable properties such as uniformly bounded distortion or long branches. Nevertheless it has the two key properties we shall require which are summable inducing times and summable variation. We recall that the *variation* of a function $\varphi : M \to \mathbb{R}$ over a subinterval $I = [a, b]$ of M is defined by

$$\operatorname*{var}_I \varphi = \sup \sum_{i=1}^{N} |\varphi(c_i) - \varphi(c_{i-1})|$$

where the supremum is taken over all $N \geq 1$ and all choices of points $a = c_0 < c_1 < \cdots < c_{N-1} < c_N = b$. For each $I \in \mathcal{I}$ we define the function $\omega_I : M \to M$ by

$$\omega_I(x) = \begin{cases} |D\hat{f}(\hat{f}^{-1}(x) \cap I)|^{-1} & \text{if } x \in \hat{f}(I) \\ 0 & \text{otherwise} \end{cases}$$

Our main technical result in this paper is the following

Theorem 1. *There exists a countable partition $\mathcal{I}$ of M (mod 0) and an inducing time function $\tau : M \to \mathbb{N}$, constant on elements of $\mathcal{I}$, such that the induced map $\hat{f} = f^{\tau(x)}(x)$ is uniformly expanding and satisfies the the following properties.*

(1) *(Summable variation)*

$$\sum_{I \in \mathcal{I}} \operatorname*{var}_I \omega_I < \infty$$

(2) *(Summable inducing times)*

$$\sum_{I \in \mathcal{I}} \tau(I)|I| < \infty$$

Theorem 1 implies the Main Theorem by known arguments. Indeed, by a result of Rychlik the summable variation property together with uniform expansion implies that $\hat{f}$ admits a finite number of ergodic absolutely continuous invariant measure whose basins cover I up to a set of measure zero [14, 16, 3]. By standard arguments the summable inducing time property implies that these measures can be pulled back to a absolutely continuous invariant probability measure for the original map f satisfying the same properties [7].

Remark 2.1. The arguments used in [4, 1, 8] also involve the construction of an induced map with summable return times, but in those papers the induced map has some very strong properties such as uniformly bounded distortion and the Gibbs-Markov property (the image of each partition element maps diffeomorphically to the entire domain of definition of the induced map). To achieve these properties a quite complicated construction is required, involving the inductive definition of an infinite number of finer and finer partitions together with a combinatorial and probabilistic argument showing that the procedure eventually converges. Besides the fact that we deal here with a significantly larger class of systems, a major difference is the construction of an induced map satisfying a different set of conditions as formalized in the summable variation property stated in the theorem. These induced maps do not necessarily have bounded distortion and there is no uniform lower bound for the size of the images. For this reason the construction of these induced maps is *much* simpler, and in fact can be fully achieved in less than two pages of text in the following section. The rest of the paper is just devoted to checking the required properties.

2.2. **Definition of the induced map.** The induced map $\hat{f}$ can in fact be defined in complete generality with essentially no assumptions on the map f. We will only require our assumptions to show that this induced map has the desired properties.

2.2.1. *Notation.* For a point x in the neighbourhood $\Delta(c,\delta)$ of one of the critical points c, we let

$$\hat{I} = \hat{I}_0 = (x, c) \quad \text{and} \quad \hat{I}_j = (x_j, c_j) = (f^j(x), f^j(c)).$$

For an arbitrary interval I we let $|I|$ denote the length of I and $d(I)$ denote it's distance to the critical set $\mathcal{C}$, i.e. the minimum distance of all point in I to $\mathcal{C}$. For each critical point c with $\ell = \ell(c) > 1$, and every integer $n \geq 1$ we let

$$\gamma_n(c) = \min\left\{\frac{1}{2},\ \frac{1}{d(c_n) D_{n-1}^{1/(2\ell-1)}}\right\} \tag{5}$$

It follows immediately from the summability condition ($\star$) that

$$\sum_n \gamma_n < \infty.$$

2.2.2. *Binding.* Given $c \in \mathcal{C}$,we define the *binding period* of a point $x \in \Delta(c,\delta)$ as follows. If $\ell(c) < 1$ we just define the binding period as $p = 1$. Otherwise we define the binding period as the *smallest* $p = p(x) \in \mathbb{N}$ such that

$$|\hat{I}_j| \leq \gamma_j\, d(c_j) \text{ for } 1 \leq j \leq p-1 \quad \text{and} \quad |\hat{I}_p| > \gamma_p\, d(c_p).$$

For each $c \in \mathcal{C}$ and $p \geq 1$, define $I(c,p)$ to be the interval of points $x \in \Delta(c,\delta)$ such that $p(x) = p$. Observe that from the definition of binding it follows immediately that

$$h(\delta) := \inf\{p(x) : x \in \Delta(c,\delta),\ c \in \mathcal{C}\} \to \infty$$

monotonically when $\delta \to 0$. Notice also that the interval $I(c,p)$ may be empty and indeed that is the case, for instance, for all $p < h(\delta)$.

2.2.3. *Fixing δ.* Using the monotonicity of $h(\delta)$ we can fix at this moment and for the rest of the paper δ sufficiently small so that

(1) the critical neighbourhood of size δ of all critical/singular points are disjoint and the images of the critical/singular neighbourhoods are also disjoint from the critical/singular neighbourhoods themselves;
(2) $\gamma_n < 1/2$ for all $n \geq h(\delta)$;
(3) $D_{n-1}^{\frac{1}{2\ell-1}} \gg 2/\kappa$ for all $n \geq h(\delta)$. The symbol $\gg$ here means that $D_{n-1}^{\frac{1}{2\ell-1}}$ must be larger than some constant factor of $2/\kappa$ for a constant which depends only on the map itself and which is determined in the course of the proof but which could in principle we specified explicitly at this point.

2.2.4. *Fixing q_0.* We now fix an integer $q_0 = q_0(\delta) \geq 1$ sufficiently large so that

$$C(\delta)e^{\lambda(\delta)q_0(\delta)} \geq 2.$$

Notice that the constants $C(\delta)$ and $\lambda(\delta)$ come from the expansion outside the critical neighbourhoods given in Section 1.2.2. The choice of q_0 is motivated by the fact that any finite piece of orbit longer than q_0 iterations staying outside a δ neighbourhood of the critical points has an accumulated derivative of at least 2.

2.2.5. *The inducing time.* Let

$$M_f = \{x \in M : f^i(x) \notin \Delta \text{ for all } 0 \leq i < q_0\} \quad \text{and} \quad M_b = M \setminus M_f$$

so that M_f denotes the set of points of M which remain outside Δ for the first $q_0 - 1$ iterations, and M_b denotes those which enter Δ at some time before q_0. For $x \in M_b$ let

$$l_0 = l_0(x) = \min\{0 \leq l < q_0 : f^l(x) \in \Delta\} \quad \text{and} \quad p_0 = p_0(f^{l_0}(x))$$

so that l_0 is the first time the orbit of x enters Δ and p_0 denotes the binding period corresponding to the point $f^{l_0}(x)$. Then we define the inducing time by

$$\tau(x) = \begin{cases} q_0 & \text{if } x \in M_f \\ l_0 + p_0 & \text{if } x \in M_b. \end{cases} \tag{6}$$

2.2.6. *The induced map.* We define the induced map as

$$\hat{f}(x) = f^{\tau(x)}(x)$$

and let $\mathcal{I}$ denote the partition of M into the maximal intervals restricted to which the induced map $\hat{f}$ is smooth, and write $\mathcal{I}_f = \mathcal{I}|M_f$ and $\mathcal{I}_b = \mathcal{I}|M_b$. This completes the definitions of the induced map.

3. Variation, Distortion and Expansion

In this section we prove a general formula relating the variation, the distortion and the expansion. First of all we define the notion of *generalized distortion*. This is a very natural notion which is no more difficult to compute than standard distortion and which appears in variation calculations. Strangely it does not seem to us to have been defined before in the literature. For any interval I and integer $n \geq 1$ we let $I_j = f^j(I)$ for $j = 0, ..., n$ and define the *(generalized) distortion*

$$\mathcal{D}(f^n, I) = \prod_{j=0}^{n-1} \sup_{x_j, y_j \in I_j} \frac{|Df(x_j)|}{|Df(y_j)|}.$$

We remark here that we are taking the supremum over all choices of sequences $x_j, y_j \in I_j$. If these sequences are chosen so that $x_j = f^j(x), y_j = f^j(y)$ for some $x, y \in I$ then we recover the more standard notion of distortion. In particular, by choosing the sequence x_j arbitrary and the sequence $y_j = f^j(y)$ as the actual orbit of a point, we can compare the two products and,in this case, the definition given above of generalized distortion immediately implies

$$\prod_{j=0}^{n-1} \sup_{I_j} \frac{1}{|Df|} \leq \frac{\mathcal{D}(f^n, I)}{|Df^n(x)|} \tag{7}$$

for any $x \in I$. For future reference we remark also that by the mean value theorem, there exists some $\xi_j \in I_j$ such that

$$\frac{Df(x_j)}{Df(y_j)} = 1 + \frac{Df(x_j) - Df(y_j)}{Df(y_j)} = 1 + \frac{D^2 f(\xi_j)}{Df(y_j)} |x_j - y_j|.$$

Therefore we have

$$\sup_{x_j, y_j \in I_j} \frac{|Df(x_j)|}{|Df(y_j)|} \leq 1 + \frac{\sup_{I_j} |D^2 f|}{\inf_{I_j} |Df|} |I_j| \quad \text{and} \quad \mathcal{D}(f^n, I) \leq \prod_{j=0}^{n-1} \left(1 + \frac{\sup_{I_j} |D^2 f|}{\inf_{I_j} |Df|} |I_j| \right). \tag{8}$$

We are now ready to state the main result of this section.

Lemma 3.1. *For any interval I and integer $l \geq 1$ such that $f^l : I \to f^l(I)$ is a diffeomorphism, we have*

$$\operatorname*{var}_I \frac{1}{|Df^l|} \lesssim \frac{\mathcal{D}(f^l, I)}{\inf_I |Df^l|} \cdot \sum_{j=0}^{l-1} \int_{I_j} \frac{dx}{d(x)}.$$

Before starting the proof we recall a few elementary properties of functions with bounded variation which will be used here and later on. Proofs can be found, for instance, in [16] or [3]. For any interval $I \subset M$, $a, b \in \mathbb{R}$, and $\varphi, \psi : M \to \mathbb{R}$,

- (V1) $\operatorname{var}_I |\varphi| \leq \operatorname{var}_I \varphi$;
- (V2) $\operatorname{var}_I(a\varphi + b\psi) \leq |a| \operatorname{var}_I \varphi + |b| \operatorname{var}_I \psi$;
- (V3) $\operatorname{var}_I(\varphi\psi) \leq \sup_I |\varphi| \operatorname{var}_I \psi + \operatorname{var}_I |\varphi| \sup_I \psi$;
- (V4) $\operatorname{var}_J \varphi = \operatorname{var}_I(\varphi \circ h)$ if $h : I \to J$ is a homeomorphism;
- (V5) if φ is of class C^1 then $\operatorname{var}_I \varphi = \int_I |D\varphi(x)| \, dx$.
- (V6) for any interval I, any bounded variation function φ, and any probability ν on I,

$$\int_I \varphi \, d\nu - \operatorname*{var}_I \varphi \leq \inf_I \varphi \leq \sup_I \varphi \leq \int_I \varphi \, d\nu + \operatorname*{var}_I \varphi. \tag{9}$$

In particular, this holds when $\nu =$ normalized Lebesgue measure on I.

Proof. We start by writing

$$\operatorname*{var}_I \frac{1}{Df^l} = \operatorname*{var}_I \left[\prod_{j=0}^{l-1} \frac{1}{Df} \circ f^j \right] = \operatorname*{var}_I \left[\left(\frac{1}{Df} \circ f^{l-1} \right) \left(\prod_{j=0}^{l-2} \frac{1}{Df} \circ f^j \right) \right]$$

Thus, from (V3) we have

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\sup_I \frac{1}{|Df|} \circ f^{l-1}\right)\left(\operatorname*{var}_I \prod_{j=0}^{l-2} \frac{1}{Df} \circ f^j\right) + \left(\operatorname*{var}_I \frac{1}{Df} \circ f^{l-1}\right)\left(\sup_I \prod_{j=0}^{l-2} \frac{1}{|Df|} \circ f^j\right)$$

Since the supremum of the product is clearly less than or equal to the product of the supremums this gives

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\sup_I \frac{1}{|Df|} \circ f^{l-1}\right)\left(\operatorname*{var}_I \prod_{j=0}^{l-2} \frac{1}{Df} \circ f^j\right) + \left(\operatorname*{var}_I \frac{1}{Df} \circ f^{l-1}\right)\left(\prod_{j=0}^{l-2} \sup_I \frac{1}{|Df|} \circ f^j\right)$$

Thus, multiplying and dividing through by both the first and last term of the right hand side of this expression, we get

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\prod_{j=0}^{l-1} \sup_I \frac{1}{|Df(f^j)|}\right)\left[\frac{\operatorname{var}_I \prod_{j=0}^{l-2} \frac{1}{Df(f^j)}}{\prod_{j=0}^{l-2} \sup_I \frac{1}{|Df(f^j)|}} + \frac{\operatorname{var}_I \frac{1}{Df(f^{l-1})}}{\sup_I \frac{1}{|Df(f^{l-1})|}}\right] \tag{10}$$

We have used here the simplified notation $[Df(f^j)]^{-1}$ to denote $[Df]^{-1} \circ f^j$. Using this bound recursively we get

$$\operatorname*{var}_I \prod_{j=0}^{l-2} \frac{1}{Df(f^j)} \le \left(\prod_{j=0}^{l-2} \sup_I \frac{1}{|Df(f^j)|}\right)\left[\frac{\operatorname{var}_I \prod_{j=0}^{l-3} \frac{1}{Df(f^j)}}{\prod_{j=0}^{l-3} \sup_I \frac{1}{|Df(f^j)|}} + \frac{\operatorname{var}_I \frac{1}{Df(f^{l-2})}}{\sup_I \frac{1}{|Df(f^{l-2})|}}\right] \tag{11}$$

and therefore, substituting (11) into (10) we get

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\prod_{j=0}^{l-1} \sup_I \frac{1}{|Df(f^j)|}\right)\left[\frac{\operatorname{var}_I \prod_{j=0}^{l-3} \frac{1}{Df(f^j)}}{\prod_{j=0}^{l-3} \sup_I \frac{1}{|Df(f^j)|}} + \frac{\operatorname{var}_I \frac{1}{Df(f^{l-2})}}{\sup_I \frac{1}{|Df(f^{l-2})|}} + \frac{\operatorname{var}_I \frac{1}{Df(f^{l-1})}}{\sup_I \frac{1}{|Df(f^{l-1})|}}\right]$$

Continuing in this way and and then using (V4) we arrive at

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\prod_{j=0}^{l-1} \sup_I \frac{1}{|Df(f^j)|}\right)\left[\sum_{j=0}^{l-1} \frac{\operatorname{var}_I \frac{1}{Df(f^j)}}{\sup_I \frac{1}{|Df(f^j)|}}\right] = \left(\prod_{j=0}^{l-1} \sup_{I_j} \frac{1}{|Df|}\right)\left[\sum_{j=0}^{l-1} \frac{\operatorname{var}_{I_j} \frac{1}{Df}}{\sup_{I_J} \frac{1}{|Df|}}\right]$$

From the definition of generalized distortion, in particular (7), this gives

$$\operatorname*{var}_I \frac{1}{Df^l} \le \left(\prod_{j=0}^{l-1} \sup_{I_j} \frac{1}{|Df|}\right)\left[\sum_{j=0}^{l-1} \frac{\operatorname{var}_{I_j} \frac{1}{Df}}{\sup_{I_J} \frac{1}{|Df|}}\right] \le \frac{\mathcal{D}(f^l, I)}{\inf_I |Df^l|}\left[\sum_{j=0}^{l-1} \frac{\operatorname{var}_{I_j} \frac{1}{Df}}{\sup_{I_j} \frac{1}{|Df|}}\right].$$

Finally from (V5) and (2) we get

$$\operatorname*{var}_{I_j} \frac{1}{Df} = \int_{I_j} \left|\frac{D^2 f}{(Df)^2}\right| \le \sup_{I_j} \frac{1}{|Df|} \int_{I_j} \left|\frac{D^2 f}{Df}\right| dx \lesssim \sup_{I_j} \frac{1}{|Df|} \int_{I_j} \frac{dx}{d(x, \mathcal{C})}.$$

$\square$

4. Binding

4.1. Distortion during binding periods.

Lemma 4.1. *For any $x \in \Delta$, $c \in \mathcal{C}$, the critical point closest to x, $\hat{I}_0 = (x, c)$, and any $1 \le j \le p(x) - 1$ we have*

$$\frac{|\hat{I}_j|}{d(\hat{I}_j)} \le 2\gamma_j \quad \text{and} \quad \sup_{x_j, y_j \in \hat{I}_j} \frac{|D^2 f(x_j)|}{|Df(y_j)|} \lesssim \frac{1}{d(\hat{I}_j)} \tag{12}$$

In particular there exists $\Gamma > 0$ independent of x such that for all $1 \le k \le p(x) - 1$ we have

$$\mathcal{D}(f^k, \hat{I}_1) \le \Gamma \quad \text{and} \quad \int_{\hat{I}_j} \frac{1}{d(x)} dx \le 2\gamma_j$$

and for all $y, z \in [x, c]$ we have

$$|Df^k(f(y))| \approx |Df^k(f(z))|.$$

Proof. The definition of binding period is designed to guarantee that the length $|\hat{I}_j|$ of the interval $\hat{I}_j = (f^j(x), f^j(c))$ is small compared to its distance $d(\hat{I}_j)$ to the critical set. Indeed, from the definition we have $d(\hat{I}_j) \ge d(f^j(c), \mathcal{C}) - d(f^j(c), f^j(x)) \ge (1 - \gamma_j) d(f^j(c), \mathcal{C})$ and therefore, for every $1 \le j \le p - 1$ we have

$$\frac{|\hat{I}_j|}{d(\hat{I}_j)} \le \frac{d(f^j(x), f^j(c))}{(1 - \gamma_j) d(f^j(c), \mathcal{C})} \le \frac{\gamma_j}{1 - \gamma_j} \le 2\gamma_j.$$

In particular this also implies, from the order of the critical points, that $\sup_{\hat{I}_j} |Df^2| \lesssim d(\hat{I}_j)^{\ell - 2}$ and $\inf_{\hat{I}_j} |Df| \gtrsim d(\hat{I}_j)^{\ell - 1}$ and therefore

$$\sup_{x_j, y_j \in \hat{I}_j} \frac{|D^2 f(x_j)|}{|Df(y_j)|} = \frac{\sup_{\hat{I}_j} |D^2 f|}{\inf_{\hat{I}_j} |Df|} \lesssim \frac{1}{d(\hat{I}_j)}$$

where $\lesssim$ means that the bound holds up to a multiplicative constant independent of δ, I or j. Now, from (8) and (12) we have

$$\mathcal{D}(f^k, \hat{I}_1) \le \prod_{j=1}^{k} \left(1 + \sup_{x_j, y_j \in I_j} \frac{|D^2 f(x_j)|}{|Df(y_j)|} |\hat{I}_j| \right) \le \prod_{j=1}^{k} \left(1 + C \frac{|\hat{I}_j|}{d(\hat{I}_j)} \right) \le \prod_{j=1}^{k} (1 + 2C\gamma_j)$$

The right hand side is uniformly bounded by the summability of the γ_j's. Indeed, taking logs and using the inequality $\log(1 + x) \le x$ for all $x \ge 0$ we get $\log \prod (1 + C\gamma_j) = \sum \log(1 + C\gamma_j) \le \sum C\gamma_j$. This proves the uniform bound on the distortion $\mathcal{D}(f^k, \hat{I}_1)$. The fact that $|Df^k(f(x))| \approx |Df^k(f(c))|$ then follows directly from the definition of $\mathcal{D}(f^k, \hat{I}_1)$ and the fact that it is uniformly bounded. Finally notice that $\int_{\hat{I}_j} 1/d(x) \le |\hat{I}_j| / d(\hat{I}_j)$ and therefore the required bound follows from (12). $\square$

4.2. **The binding period partition.** The partition $\mathcal{I}$ is defined quite abstractly and we do not have direct information about the sizes of the partition elements and in particular the relation between their sizes and their distances to the critical set. However, using the distortion bounds obtained above, we can prove the following

Lemma 4.2. *Let $I \in \mathcal{I}$ with $p(I) = p$ and I in the neighbourhood of a critical point with order ℓ. Then*

$$D_{p-1}^{-2/(2\ell-1)} \lesssim \inf_{x\in I} d(x) \le \sup_{x\in I} d(x) \lesssim D_{p-2}^{-2/(2\ell-1)}. \tag{13}$$

In particular, letting $\ell_k = \ell_k(c)$ denote the order of the critical/singular point closest to c_k we have

$$\mathcal{D}(f, I) \lesssim \left[\frac{D_{p-1}}{D_{p-2}}\right]^{\frac{2(\ell-1)}{2\ell-1}} \lesssim d(c_{p-1})^{\frac{2(\ell-1)(\ell_{p-1}-1)}{2\ell-1}} \tag{14}$$

and

$$\int_I \frac{1}{d(x)} dx \lesssim \log \left[\frac{D_{p-1}}{D_{p-2}}\right]^{\frac{2(\ell-1)}{2\ell-1}} \lesssim \log d(c_{p-1})^{-1}. \tag{15}$$

Remark 4.3. We remark that the distortion not uniformly bounded in p implying that the induced map does not have uniformly bounded distortion. Notice also that for some values of p it may happen that $D_{p-2}^{-2/(2\ell-1)} \ll D_{p-1}^{-2/(2\ell-1)}$; in this case the corresponding interval I would necessarily be empty, i.e. there is no x with binding period p.

Proof. From Lemma 4.1 and the definition of binding period we have, for any $x \in I$,

$$d(x) = |\hat{I}_0| \approx |\hat{I}_1|^{1/\ell} \approx [D_{p-1}^{-1}|\hat{I}_p|]^{1/\ell} \ge [D_{p-1}^{-1}\gamma_p d(c_p)]^{1/\ell}$$

and

$$d(x) = |\hat{I}_0| \approx |\hat{I}_1|^{1/\ell} \approx [D_{p-2}^{-1}|\hat{I}_{p-1}|]^{1/\ell} \le [D_{p-2}^{-1}\gamma_{p-1} d(c_{p-1})]^{1/\ell}$$

By taking a sufficiently small δ we can assume that p is sufficiently large so that $\gamma_{p-1}, \gamma_p < 1/2$ and therefore, from the definition of the sequence $\{\gamma_n\}$ we get

$$\gamma_n d(c_n) = D_{n-1}^{-1/(2\ell-1)}.$$

Thus, substituting into the expressions above gives

$$d(x) \gtrsim [D_{p-1}^{-1}\gamma_p d(c_p)]^{1/\ell} = [D_{p-1} D_{p-1}^{-1/(2\ell-1)}]^{1\ell} = [D_{p-1}^{-2\ell/(2\ell-1)}]^{1/\ell} = D_{p-1}^{-2/(2\ell-1)}$$

and, similarly,

$$d(x) \lesssim D_{p-2}^{-2/(2\ell-1)}.$$

This gives the first set of inequalities. As a consequence we immediately get

$$D_{p-2}^{-2(\ell-1)/(2\ell-1)} \gtrsim \sup_I |Df(x)| \ge \inf_I |Df(x)| \gtrsim D_{p-1}^{-2(\ell-1)/(2\ell-1)}$$

and therefore,

$$\mathcal{D}(f, I) = \sup_{x,y\in I} \frac{|Df(x)|}{|Df(y)|} \lesssim \left[\frac{D_{p-1}}{D_{p-2}}\right]^{\frac{2(\ell-1)}{2\ell-1}}$$

This gives the first inequality in (14). To get the second inequality we simply use the fact that $D_{p-1} \approx D_{p-2} d(c_{p-1})^{\ell_{p-1}-1}$. To get the last inequality we simply integrate $1/d(z)$ over the interval $I = (x, y)$ to get

$$\int_I \frac{1}{d(z)} dx = |\log d(x) - \log d(y)| \lesssim \log \left[\frac{D_{p-1}}{D_{p-2}} \right]^{2/(2\ell-1)}$$

and then argue as above. □

4.3. Expansion during binding periods.

Lemma 4.4. *For all $c \in \mathcal{C}$, $x \in \Delta(c, \delta)$ and $p = p(x)$, we have*

$$|Df^p(x)| \gtrsim D_{p-1}^{\frac{1}{2\ell-1}} \tag{16}$$

In particular we can choose δ small enough so that

$$|Df^p(x)| \geq 2/\kappa.$$

Proof. Using the chain rule, bounded distortion in binding periods and Lemma 4.2 we have

$$|Df^p(x)| = |Df^{p-1}(f(x)) \cdot Df(x)| \gtrsim D_{p-1} D_{p-1}^{-2(\ell-1)/(2\ell-1)} = D_{p-1}^{\frac{1}{2\ell-1}}$$

This gives (16). The inequality $|Df^p(x)| \geq 2/\kappa$ then just follows from the choice of δ in Section 2.2.3. □

5. Inducing

5.1. Expansion of the induced map.

Lemma 5.1. *For every $x \in M$ we have*

$$|D\hat{f}(x)| \geq 2.$$

Proof. This follows immediately from the definition of the induced map and the expansion estimates during binding periods obtained in Lemma 4.4 together with conditions 3 and 4, the choice of δ and the corresponding choice of q_0. □

5.2. Distortion of the induced map.

We now study the distortion of the induced map $\hat{f}$ on each of its branches.

Lemma 5.2. *There exists a constant $D = D(\delta) > 0$ such that*

$$\mathcal{D}(f^\tau, I) \leq D \quad \text{and} \quad \mathcal{D}(f^\tau, I) \lesssim d(c_{p-1})^{\frac{2(\ell-1)(\ell_{p-1}-1)}{2\ell-1}} \tag{17}$$

for all $I \in \mathcal{M}_f$ (in which case $\tau \equiv q_0$) and $I \in \mathcal{M}_b$ (in which case $\tau = l + p$) respectively, where ℓ is the order of the critical point associated to I_l. Also, we have

$$\sum_{j=0}^{\tau-1} \int_{I_j} \frac{1}{d(x)} dx \leq D$$

and

$$\sum_{j=0}^{\tau-1}\int_{I_j}\frac{1}{d(x)}dx \le D+\log d(c_{p-1})^{-1} \tag{18}$$

respectively for $I\in\mathcal{M}_f$ *and* $I\in\mathcal{M}_b$.

Proof. For $I\in\mathcal{I}_f$ we have standard distortion estimates for uniformly expanding maps which give a uniform distortion bound D depending on the size of Δ. For $I\in\mathcal{I}_b$ on the other hand we write

$$\mathcal{D}(f^\tau,I)=\mathcal{D}(f^l,I)\cdot\mathcal{D}(f,I_l)\cdot\mathcal{D}(f^{p-1},I_{l+1}).$$

The first term consists of iterates for which I_j lies always outside Δ and therefore is bounded above by the same constant D as above. The second and third term have already been estimated above in Lemmas 4.1 and 4.2. Combining these estimates we complete the first set of estimates.

For $I\in\mathcal{M}_f$, using the uniform expansion outside Δ we have $|I_j|\le c(\delta)^{-1}e^{-\lambda(\delta)(\tau-j)}$ and therefore

$$\sum_{j=0}^{\tau-1}\int_{I_j}\frac{1}{d(x)}dx\le\sum_{j=0}^{\tau-1}\frac{|I_j|}{d(I_j)}\le\sum_{j=0}^{\tau-1}\frac{c(\delta)^{-1}e^{-\lambda(\delta)(\tau-j)}}{\delta}\le D.$$

For $I\in\mathcal{M}_b$ we again split the sum into three parts corresponding to the initial iterates outside Δ, the first iterate in Δ, and the following binding period. The fist part of the sum is bounded by the same constant D as above. The second and third have already been estimated above. Thus, from Lemmas 4.1 and 4.2 and in particular (15) we get the statement.

□

6. Summability

We are now ready to prove the summable variation and the summable inducing time properties.

6.1. **Summable variation.** From the definition of ω_I that we have

$$\operatorname*{var}_M\omega_I=\operatorname*{var}_I\omega_I+2\sup_I\omega_I=\operatorname*{var}_I\frac{1}{|Df^\tau|}+2\sup_I\frac{1}{|Df^\tau|}$$

For the supremum we have, from Lemma 5.1,

$$\sup_I\frac{1}{|Df^\tau|}\le\frac{1}{D_{p-1}^{1/(2\ell-1)}} \tag{19}$$

and for the variation, we have, substituting the estimates in (19), (17) and (18) into the formula obtained in Lemma 3.1,

$$\operatorname*{var}_I\frac{1}{|Df^\tau|}\lesssim\frac{\mathcal{D}(f^\tau,I)}{\inf_I|Df^\tau|}\cdot\sum_{j=0}^{\tau-1}\int_{I_j}\frac{dx}{d(x)}\lesssim\frac{D+\log d(c_{p-1})^{-1}}{d(c_{p-1})^{-\frac{2(\ell-1)(\ell_{p-1}-1)}{2\ell-1}}D_{p-1}^{1/(2\ell-1)}}.$$

We can write

$$D_{p-1}^{\frac{1}{2\ell-1}} \approx D_{p-2}^{\frac{1}{2\ell-1}} d(c_{p-1})^{\frac{\ell_{p-1}-1}{2\ell-1}}$$

and

$$d(c_{p-1})^{-\frac{2(\ell-1)(\ell_{p-1}-1)}{2\ell-1}} d(c_{p-1})^{\frac{\ell_{p-1}-1}{2\ell-1}} = d(c_{p-1})^{-\frac{(2\ell-1)(\ell_{p-1}-1)}{2\ell-1}} = d(c_{p-1})^{(1-\ell_{p-1})}$$

and so, substituting above, gives

$$\operatorname*{var}_I \frac{1}{|Df^\tau|} \lesssim \frac{D+\log d(c_{p-1})^{-1}}{d(c_{p-1})^{(1-\ell_{p-1})} D_{p-2}^{1/(2\ell-1)}} \leq \frac{D+\log d(c_{p-1})^{-1}}{d(c_{p-1}) D_{p-2}^{1/(2\ell-1)}}. \tag{20}$$

The summability then follows immediately from ($\star$).

6.2. **Summable inducing times.** To prove the summability of the inducing time notice first of all that the number of intervals of a given inducing time is uniformly bounded. Therefore it is sufficient to prove the summability with respect to the binding time. For this we give a basic upper bound for the size of each element $I \in \mathcal{I}$ using the mean value theorem and Lemma 5.1. This gives

$$\sum \tau(I)|I| \lesssim \sum_p p|I| \lesssim \sum_p \frac{p}{D_{p-1}^{1/(2\ell-1)}}.$$

Again, the summability follows directly from ($\star$). This completes the proof of the Theorem.

6.3. **Final remarks.** Notice that there is a signficant gap between the first bound and the second bound in (20), particularly in the special case in which there are no singularities and where therefore $\ell > 1$ for every critical point. In this case we get

$$\operatorname*{var}_I \frac{1}{|Df^\tau|} \lesssim \frac{\log d(c_{p-1})^{-1}}{d(c_{p-1})^{(1-\ell_{p-1})} D_{p-2}^{1/(2\ell-1)}} \leq \frac{\log d(c_{p-1})^{-1}}{D_{p-2}^{1/(2\ell-1)}}.$$

This leaves only an extremely mild condition on the recurrence of the critical points and therefore the summability conditions reduces almost to the condition $\sum 1/D_n^{1/(2\ell-1)}$ assumed for smooth maps in [4]. Ideally we would therefore like to replace condition ($\star$) by the summability condition

$$\sum_n \frac{n \log d(c_n)^{-1}}{d(c_n)^{(1-\ell_n)} D_{n-1}^{1/(2\ell-1)}} < \infty \tag{$\star\star$}$$

which would automatically reduce to the condition

$$\sum_n \frac{n \log d(c_n)^{-1}}{D_{n-1}^{1/(2\ell-1)}} < \infty$$

in the smooth case. This however gives rise to technical difficulties that we have not been able to overcome, mainly in the definition of the sequence γ_n, recall (5). Condition ($\star\star$) does not imply the summability of the γ_n with the definition given in (5) and on the other hand, changing the definition of γ_n to something more natural in terms of ($\star\star$), such as for example $1/(d(c_n)^{(1-\ell_n)} D_{n-1}^{1/(2\ell-1)})$ gives rise to additional complications in the calculations and estimates related to the binding period in Section 4. It is not clear to us whether these are superficial difficulties which can be overcome or whether they reflect deeper issues.

References

[1] José F. Alves, Stefano Luzzatto, and Vilton Pinheiro. Lyapunov exponents and rates of mixing for one-dimensional maps. *Ergodic Theory Dynam. Systems*, 24(3):637–657, 2004.

[2] Vitor Araujo and Maria Jose Pacifico. Physical measures for infinite-modal maps. *Preprint*, 2007.

[3] Abraham Boyarsky and Paweł Góra. *Laws of chaos*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1997. Invariant measures and dynamical systems in one dimension.

[4] H. Bruin, S. Luzzatto, and S. Van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup. (4)*, 36(4):621–646, 2003.

[5] Henk Bruin, Juan Rivera-Letelier, Weixao Shen, and Sebastian van Strien. Large derivatives, backward contraction and invariant densities for interval maps. *preprint*.

[6] H. Bruin and S. van Strien. Expansion of derivatives in one-dimensional dynamics. *Israel J. Math.*, 137:223–263, 2003.

[7] Welington de Melo and Sebastian van Strien. *One-dimensional dynamics*, volume 25 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1993.

[8] K. Díaz-Ordaz, M. P. Holland, and S. Luzzatto. Statistical properties of one-dimensional maps with critical points and singularities. *Stoch. Dyn.*, 6(4):423–458, 2006.

[9] Eberhard Hopf. Theory of measure and invariant integrals. *Trans. Amer. Math. Soc.*, 34(2):373–393, 1932.

[10] Stefano Luzzatto. Stochastic-like behaviour in nonuniformly expanding maps. In *Handbook of dynamical systems. Vol. 1B*, pages 265–326. Elsevier B. V., Amsterdam, 2006.

[11] Stefano Luzzatto and Warwick Tucker. Non-uniformly expanding dynamics in maps with singularities and criticalities. *Inst. Hautes Études Sci. Publ. Math.*, (89):179–226 (2000), 1999.

[12] Stefano Luzzatto and Marcelo Viana. Positive Lyapunov exponents for Lorenz-like families with criticalities. *Astérisque*, (261):xiii, 201–237, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[13] Ricardo Mañé. Hyperbolicity, sinks and measure in one-dimensional dynamics. *Comm. Math. Phys.*, 100(4):495–524, 1985.

[14] Marek Rychlik. Bounded variation and invariant measures. *Studia Math.*, 76(1):69–80, 1983.

[15] S Ulam and J von Neumann. On combination of stochastic and deterministic processes. *Bull AMS*, 53:1120, 1947.

[16] Marcelo Viana. *Stochastic dynamics of deterministic systems.* Lecture Notes XXI Braz. Math. Colloq. IMPA, Rio de Janeiro, 1997.

Instituto de Matematica, UFRJ, CP 68.530, Rio de Janeiro 21.945-970 Brazil and Centro de Matematica, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

E-mail address: `vitor.araujo@im.ufrj.br; vdaraujo@fc.up.pt`

Mathematics Department, Imperial College London, SW7 2AZ, UK

URL: `http://www.ma.ic.ac.uk/~luzzatto`

E-mail address: `stefano.luzzatto@imperial.ac.uk`

Instituto de Matematica Pura e Aplicada, Est. D. Castorina 110, Rio de Janeiro, Brazil

URL: `http://www.impa.br/~viana`

E-mail address: `viana@impa.br`