ON THE LEBESGUE MEASURE OF LI-YORKE PAIRS FOR INTERVAL MAPS

HENK BRUIN AND VÍCTOR JIMÉNEZ LÓPEZ

ABSTRACT. We investigate the prevalence of Li-Yorke pairs for C^2 and C^3 multimodal maps f with non-flat critical points. We show that every measurable scrambled set has zero Lebesgue measure and that all strongly wandering sets have zero Lebesgue measure, as does the set of pairs of asymptotic (but not asymptotically periodic) points.

If f is topologically mixing and has no Cantor attractor, then typical (w.r.t. two-dimensional Lebesgue measure) pairs are Li-Yorke; if additionally f admits an absolutely continuous invariant probability measure (acip), then typical pairs have a dense orbit for $f \times f$. These results make use of so-called nice neighborhoods of the critical set of general multimodal maps, and hence uniformly expanding Markov induced maps, the existence of either is proved in this paper as well.

For the setting where f has a Cantor attractor, we present a trichotomy explaining when the set of Li-Yorke pairs and distal pairs have positive two-dimensional Lebesgue measure.

1. INTRODUCTION

In interval dynamics there are many ways to deal with the notion of asymptotic complexity ("chaos"). Probably it is pointless to try and decide which is the best of them, but in applications there are two of them which are by far the most popular, see *e.g.* [D, MMN]. One is topological chaos, that is, the existence of an uncountable scrambled set in the sense of the famous Li and Yorke paper [LY]. The other one is ergodic chaos, that is, the existence of an invariant probability measure absolutely continuous with respect to Lebesgue measure (acip). Neither of them is without drawbacks. (To keep this introduction at the expository level, we have deferred most of definitions to the subsequent sections.)

There are easy conditions implying the existence of Li-Yorke chaos and its stability under small perturbations. One such condition is the existence of a periodic orbit of period not a power of two [Bl]. Nevertheless, this chaos need not be "observable": for instance, the

orbits of almost all points (in the sense of Lebesgue measure) may be attracted by this periodic orbit [Gu].

On the other hand, ergodic chaos ensures complicated dynamics for a set of points with positive measure. For instance, if a smooth multimodal map (with non-flat critical points) f has an acip, then, as a simple consequence of the zero measure of Cantor metric attractors [SV] and Proposition 9 below, there is a positive measure set of points whose orbit under f is dense in some interval. However, the converse is not true, even for the family of logistic maps [Jo, Ly2], so acips can only exist under additional conditions (for instance, hyperbolic repelling periodic points and $|Df^n(f(c))| \to \infty$ for any critical point c of the map f [BRSS]).

A map f is infinitely renormalizable if there is an infinite collection of nested cycles of periodic intervals; the intersection of these cycles is a Cantor set, called a solenoidal attractor (because the suspension over this attractor is a topological solenoid). The Feigenbaum map, or more correctly Coullet-Tresser-Feigenbaum, (see *e.g.* [MS, pp. 151-152]) is the best known example of this. A solenoidal attractor is Lyapunov stable and as shown in [BrJ] (see Proposition 31 below), points in the basin of such an attractor are approximately periodic:

Definition 1. *A point x is* approximately periodic *if for every $\varepsilon > 0$ there is a periodic point p such that* $\limsup_{n\to\infty} |f^n(x) - f^n(p)| < \varepsilon$.

Hence, up to small errors, almost all points eventually behave as periodic points.

Remark 2. *An intrinsic characterization of adding machines (namely any system in which every point is regularly recurrent) is presented in* [BK]. *In our setting, it applies to the solenoidal attractor itself, whereas approximate periodicity gives information on a neighborhood of the solenoidal attractor.*

However, a multimodal (even polynomial) map may have a dense orbit while, simultaneously, almost all orbits are attracted by a Cantor set (a so-called wild attractor) [BKNS]. As we will explain below, in such a case it is still possible but not necessary that a.e. point is approximately periodic.

We see that there is a variety of smooth multimodal maps featuring a certain degree of "observable" dynamical complexity which is, however, not strong enough to be realized by an acip. It is natural to return to the Li-Yorke notion of chaos and investigate to what extent it can be used to measure this complexity. This is what we intend in the present paper.

Definition 3. *Let $f : I = [0, 1] \to I$ be a continuous map. A pair of points (x, y) is called:*

- distal *if* $\liminf_{n\to\infty} |f^n(y) - f^n(x)| > 0$;

- asymptotic *if* $\lim_{n\to\infty} |f^n(y) - f^n(x)| = 0$;

- Li-Yorke *if it is neither asymptotic nor distal, that is,*
$$0 = \liminf_{n\to\infty} |f^n(y) - f^n(x)| < \limsup_{n\to\infty} |f^n(y) - f^n(x)|.$$

We denote the set of distal, asymptotic and Li-Yorke pairs by Dis, Asymp *and* LY, *respectively. A pair that is not distal (hence asymptotic or Li-Yorke) is called* proximal. *The set of Li-Yorke pairs with* $\limsup_{n\to\infty}|f^n(y) - f^n(x)| \geq \varepsilon$ *is denoted as* LY_ε.

Note that $f^n(x) \neq f^n(y)$ for all $n \geq 0$ whenever (x, y) is a Li-Yorke or distal pair.

Definition 4. *Let $f : I \to I$ be a continuous map.*

- *A set $S \subset X$ is called* scrambled *if any two distinct points in S form a Li-Yorke pair.*
- *The map f is called* chaotic (in the sense of Li-Yorke) *if it has an uncountable scrambled set.*

The above definitions can be strengthened by assuming that there is a positive lower bound of $\limsup|f^n(x) - f^n(y)|$ *independently of x, y:*

- *If there is $\varepsilon > 0$ such that for every $x \neq y \in S$, $\limsup|f^n(x) - f^n(y)| \geq \varepsilon$ independent of $x, y \in S$, then S is called ε-scrambled.*
- *If there is $\varepsilon > 0$ such that for every $x \in X$ and every neighborhood $U \ni x$, there is $y \in U$ such that $(x, y) \in LY_\varepsilon$, then f is* Li-Yorke sensitive.

We emphasize that ε-scrambled is a stronger property than just scrambled. For example, [BHS, Proposition 5] shows the possibility of having scrambled sets that are not ε-scrambled for any $\varepsilon > 0$.

Our concrete aim is to investigate the Lebesgue measure of the above sets for C^2 (or sometimes C^3) multimodal maps from the interval I into itself with non-flat critical points (denoted by $C^2_{\mathrm{nf}}(I)$ and $C^3_{\mathrm{nf}}(I)$, respectively). The advantage of working in this setting is that there are many tools at hand to deal with measure-theoretic properties. Remarkably the most important of these tools is purely topological: maps from $C^2_{\mathrm{nf}}(I)$ have no wandering intervals, see [MS, Theorem A, p. 267].

As it happens, some of the results in the paper are based on a generalization of this property which is of interest in itself.

Definition 5. *A point x is* asymptotically periodic, *written $x \in$ AsPer, if there is a periodic point p such that $\lim_{n\to\infty}|f^n(x) - f^n(p)| = 0$. A measurable set W is* strongly wandering *if $f^n(W) \cap f^m(W) = \emptyset$ for all $n > m \geq 0$ and W contains no asymptotically periodic points. A* wandering interval *is just an interval which is strongly wandering.*

The notion of strongly wandering set was introduced by Blokh and Lyubich in [BL] (in a slightly different way). Under the assumption of negative Schwarzian derivative, the non-existence of strongly wandering sets of positive measure was proved in the unimodal case (and stated in the multimodal case without inflection points) in [BL]. Here we prove it for maps from $C^2_{\mathrm{nf}}(I)$, for which inflection points are now allowed.

Theorem A. *Let $f \in C^2_{\mathrm{nf}}(I)$. Then every strongly wandering set has zero Lebesgue measure.*

Concerning the size of Li-Yorke chaos, the first natural question is whether smooth multimodal maps may have scrambled sets of positive Lebesgue measure. There is extensive

literature on the subject. Examples of continuous maps possessing scrambled sets of positive or even full measure are well known: [K, S2, Mi, BH]. In fact, if f is chaotic (respectively, has a dense orbit), then it is topologically conjugate to a map having a positive (respectively, full) measure scrambled set [JS, S3] (respectively, [SS]).

Of course, it is not possible that the whole interval is scrambled (in fact, a scrambled set cannot be residual on any subinterval of I, see [G]), although there are maps with scrambled forward invariant Cantor sets [HY]. However such maps cannot be multimodal (see Proposition 26).

It is worth emphasizing that maps of type 2^∞ in the Sharkovskiy ordering (that is, those having periodic points of periods all powers of 2, but no other periods) can possess scrambled sets of positive measure, but not of full measure because any Li-Yorke chaotic map of type 2^∞ has a wandering interval. Indeed, if a map of type 2^∞ has no wandering intervals, then all points are approximately periodic [S3]. However the following result is well known (see *e.g.* [BC, p. 144]):

Proposition 6. *If $f : I \to I$ is continuous and x, y are approximately periodic points, then (x, y) is either asymptotic or distal.*

Hence a scrambled set can contain at most one approximately periodic point. Finally, positive measure scrambled sets may also exist for C^∞ maps (with flat critical points) or C^1 maps with non-flat critical points, but a C^1 map cannot have a full measure scrambled set [J1, J2, BJ2].

Nevertheless, it is a widely held view that these are rather pathological examples. For instance, it is known that neither maps from $C^2_{\mathrm{nf}}(I)$ with hyperbolic periodic points and whose critical points satisfy the Misiurewicz condition, nor maps in $C^3_{\mathrm{nf}}(I)$ with negative Schwarzian derivative and having no wild attractors, may possess scrambled sets of positive measure [BJ1, BrJ]. Our next result confirms these expectations:

Theorem B. *If $f \in C^3_{\mathrm{nf}}(I)$, then it has no measurable scrambled sets of positive Lebesgue measure.*

It seems rather paradoxical that scrambled sets have zero measure even in the case when there is an acip, but this is not really so. The key point is *measurability*. For instance, one can easily derive from [S1] that the full logistic map $f(x) = 4x(1 - x)$ possesses a non-measurable scrambled set with full *exterior* Lebesgue measure. Since the sets $\mathrm{Dis}, \mathrm{Asymp}, \mathrm{LY}, \mathrm{LY}_\varepsilon, \mathrm{AsPer}$ are all Borel, hence measurable sets [J3, BrJ], the moral is that we should measure these sets rather than scrambled sets. The idea of passing to the square $I \times I$ to study topological or measure-theoretic properties of Li-Yorke chaos is due to Lasota and was first used by Piórek [P].

To begin with, we prove that there almost no "non-trivial" asymptotic pairs.

Theorem C. *If $f \in C^2_{\mathrm{nf}}(I)$, then $\mathrm{Asymp} \setminus (\mathrm{AsPer} \times \mathrm{AsPer})$ has zero (two-dimensional) Lebesgue measure.*

We can consider Theorem C as a weak form of sensitivity to initial conditions in the absence of periodic attractors. It is really quite weak: it applies in particular to the

infinitely renormalizable case where almost all points are attracted by a solenoidal set, so there is no sensitivity to initial conditions in the standard Guckenheimer sense [Gu]. What Theorem C emphasizes is that there are no "privileged" routes (that is, with positive measure) to measure-theoretic attractors.

The Li-Yorke property describes how chaotically pairs of points behave with respect to each other, and is hence a property of the Cartesian product (I^2, f_2), for $I^2 = I \times I$ and $f_2(x, y) := (f \times f)(x, y) = (f(x), f(y))$. If (x, y) is Li-Yorke, then $\mathrm{orb}((x, y))$ accumulates on, but does not converge to, the diagonal of I^2. Hence LY is a weaker property than (x, y) having a dense orbit in I^2. We want to find conditions ensuring that LY has positive (or full) mass w.r.t. two-dimensional Lebesgue measure λ_2. Recall the a map g is called *topologically mixing* if for every pair of non-empty open sets A, B there is $N \in \mathbb{N}$ such that $f^n(A) \cap B \neq \emptyset$ for all $n \geq N$. According to Proposition 9, the orbit of Lebesgue a.e. point is either attracted by a periodic point or a solenoidal set, or eventually falls into an interval K such that $f^r(K) = K$ for some r and the restriction $g = f^r|_K$ is topologically mixing. As we explained earlier, approximately periodic points cannot be used to produce Li-Yorke pairs. On the other hand, all sets from Definition 3 are the same for f as for any of its iterates f^r (except that the ε in LY_ε may change, but not its positivity). Thus we can restrict ourselves to the case where the map f itself is topologically mixing.

Now, if f is topologically mixing, then either almost all points have a dense orbit, or the orbits of almost all points are attracted by finitely many pairwise disjoint wild Cantor attractors. It turns out that in the first case λ_2-a.e. pair (x, y) is Li-Yorke.

Theorem D. *Let $f \in C^3_{\mathrm{nf}}(I)$ be a topologically mixing map having no Cantor attractors. Then the Cartesian product (I^2, λ_2, f_2) is ergodic and for every $x \in I$ there is a full measure set $C_x \subset I$ such that*

$$\liminf_{n \to \infty} |f^n(y) - f^n(x)| = 0, \quad \limsup_{n \to \infty} |f^n(y) - f^n(x)| \geq \mathrm{diam}(I)/2,$$

for every $y \in C_x$. In particular, $\mathrm{LY}_{\mathrm{diam}(I)/2}$ has full measure and f is Li-Yorke sensitive.

Weaker versions of this result under additional Misiurewicz or negative Schwarzian conditions were proved earlier [BJ1, BrJ]. Ergodicity of (I^2, λ_2, f_2) is parallel to Lebesgue measure λ being weak mixing, although in its standard definition, weak mixing applies to invariant measures only, see Subsection 2.4. If f admits an acip, then we can say more: λ_2-a.e. pair (x, y) has a dense orbit in I^2, see Corollary 28. However, it seems possible that there are cases where f admits no acip, λ_2-a.e. pair (x, y) is Li-Yorke, but has no dense orbit in I^2.

It remains to consider the case where $f \in C^2_{\mathrm{nf}}(I)$ is a topologically mixing map having a wild attractor $\mathcal{A}$. Let the *basin* $\mathrm{Bas}(\mathcal{A})$ be the set of points whose orbit is attracted by $\mathcal{A}$. We already know (Theorem C) that Asymp has zero measure. Also, if some $x \in \mathrm{Bas}(\mathcal{A})$ is approximately periodic, then by Proposition 31, $\mathcal{A}$ is conjugate to an adding machine, so all points in $\mathrm{Bas}(\mathcal{A})$ are approximately periodic. Finally, recall that every pair of approximately periodic points is either asymptotic or distal. Our result in this area classifies which behaviors can (and indeed do) occur:

Theorem E. *Let $f \in C^2_{\mathrm{nf}}(I)$ be a topologically mixing map with a wild attractor $\mathcal{A}$. Then one of the following alternatives must occur:*

(a) *Lebesgue a.e. pair of points in* $\mathrm{Bas}(\mathcal{A})$ *is distal and every point in* $\mathrm{Bas}(\mathcal{A})$ *is approximately periodic;*
(b) *Lebesgue a.e. pair of points in* $\mathrm{Bas}(\mathcal{A})$ *is distal and no point in* $\mathrm{Bas}(\mathcal{A})$ *is approximately periodic;*
(c) *Lebesgue a.e. pair of points in* $\mathrm{Bas}(\mathcal{A})$ *is Li-Yorke;*
(d) *Both* Dis *and* LY *have positive Lebesgue measure in* $\mathrm{Bas}(\mathcal{A}) \times \mathrm{Bas}(\mathcal{A})$.

There are examples of polynomial unimodal maps of all above types (a)-(d) so that additionally, in cases (b)-(d), $\mathrm{Bas}(\mathcal{A})$ *contains* ε-*scrambled sets for a fixed* $\varepsilon > 0$ *and* f *is Li-Yorke sensitive on* $\mathrm{Bas}(\mathcal{A})$.

2. Preliminaries

2.1. **Interval maps.** A continuous map $f : I \to I$ (for $I = [0,1]$) is called *multimodal* if $[0,1]$ can be decomposed into finitely many subintervals on which f is (strictly) monotone. A point c is *critical* if $f'(c) = 0$; the set of critical points is denoted by Crit. Critical points can be *turning* points (if f assumes a local extremum at c) or *inflection* points, and hence there can be more critical points than maximal intervals of monotonicity. A differentiable map having exactly one critical (turning) point is called *unimodal*. Near a turning point c, there is a largest interval $[a,b]$ such that $f(a) = f(b)$ and f is monotone on each of the intervals $[a,c]$ and $[c,b]$. Then there is a continuous involution $\tau_c : [a,b] \to [a,b]$ such that $f(\tau_c(x)) = f(x)$ and $\tau_c(x) \neq x$ for every $x \neq c$.

We say that $f \in C_{\mathrm{nf}}^k(I)$ if $f : I \to I$ has a finite critical set Crit, is of class C^k and each critical point is *non-flat*, *i.e.*, for each $c \in$ Crit, there exist a C^k diffeomorphism φ_c with $\varphi_c(c) = 0$ and an $\ell_c \in (1, \infty)$, called the *critical order* of c, such that $f(x) = \pm |\varphi_c(x)|^{\ell_c} + f(c)$ for x close to c. It follows that if n is the smallest integer $\geq \ell_c$, then $D^m f(c) = 0$ for $1 \leq m < n$, but $D^n f(c) \neq 0$. Conversely, if f is C^{k+1} near c and $D^n f(c) \neq 0$ for some $2 \leq n \leq k$, then c is non-flat.

We can always enlarge the domain of f (without adding new critical points) and rescale such that

$$f(\partial I) \subset \partial I \quad \text{and} \quad \mathrm{Crit} \cap \partial I = \emptyset. \tag{1}$$

Observe that this operation does not change the zero or positive measure quality of the sets from Definition 3. Hence, except when f is topologically mixing, we will always assume that (1) holds. Also, we will assume without loss of generality that

$$\text{Crit does not contain any periodic point,} \tag{2}$$

because if c is a periodic critical point, then we can modify slightly f near the orbit of c so that the resulting map has a periodic orbit containing no critical points and attracting the same points as previously attracted by the orbit of c.

In a metric space, topological mixing (see definition on page 5) implies that every iterate f^n has a dense orbit. This excludes the case that there is a proper compact subinterval J of I such that $f^r(J) \subset J$ for some $r \geq 1$ (*i.e.*, f is *non-renormalizable*). The map f is *topologically exact* (also called *locally eventually onto*) if for every non-degenerate interval $J \subset I$, there is n such that $f^n(J) = I$. For multimodal interval maps, topologically

mixing, the existence of a dense orbit for f^n for each iterate $n \in \mathbb{N}$, and topologically exactness are all equivalent, see *e.g.* [BC, pp. 157–158].

2.2. **Attractors of interval maps.** The *orbit* $\{f^n(x)\}_{n=0}^{\infty}$ of a point x is denoted by $\mathrm{orb}(x)$. More generally, the orbit of a set A is $\mathrm{orb}(A) := \bigcup_{n=0}^{\infty} f^n(A)$. The ω-*limit set* $\omega(x) := \bigcap_{n \in \mathbb{N}} \mathrm{Cl} \bigcup_{m \geq n} f^n(x)$ of x is the set of limit points of the orbit of x. If $\mathcal{A}$ is a subset of I, then we call $\mathrm{Bas}(\mathcal{A}) = \{x \in I : \omega(x) \subset \mathcal{A}\}$ the *basin (of attraction)* of $\mathcal{A}$. A periodic orbit $\mathcal{O}$ is called *attracting* if its basin contains an open set. The union of the components of $\mathrm{Bas}(\mathcal{O})$ intersecting $\mathcal{O}$ is called the *immediate basin* of $\mathcal{O}$. If $\mathrm{Bas}(\mathcal{O})$ contains a neighborhood of $\mathcal{O}$, then $\mathcal{O}$ is called a *two-sided* attracting periodic orbit; otherwise it is called a *one-sided* attracting periodic orbit. If p is a periodic point of period r and $|Df^r(p)|$ is less than, equal to, or greater than 1, then the orbit of p is called *hyperbolic attracting*, *parabolic*, or *hyperbolic repelling* respectively. Of course, only hyperbolic attracting and parabolic orbits can be attracting. If $f \in C_{\mathrm{nf}}^2(I)$, then it is well known that all periodic orbits of sufficiently high period are hyperbolic repelling [MMS]. Furthermore, maps in $C_{\mathrm{nf}}^2(I)$ have no wandering intervals. One of the consequences of this is the following useful result. Below $\mathrm{dist}(A, B)$ denotes the distance between the sets A and B (by convention $\mathrm{dist}(A, \emptyset) = \infty$).

Proposition 7 (The non-Contraction Principle). *Let $f : I \to I$ be a multimodal map without wandering intervals. Then for every $\varepsilon > 0$ there is $\delta > 0$ such that if J is an interval such that $|J| < \delta$ and $\mathrm{dist}(J, p) > \varepsilon$ for each attracting periodic point p, then every component of every preimage $f^{-n}(J)$ of J has length less than ε.*

Proof. This follows easily from the Contraction Principle as stated in [MS, p. 305]. The principle is a bit of a misnomer, so we added non- in our version. $\square$

Definition 8. *We call a closed invariant set $\mathcal{A}$ a* (measure-theoretic) attractor *if its basin has positive Lebesgue measure, and there is no proper subset $\mathcal{A}' \subset \mathcal{A}$ with the same properties.*

Note that $\mathcal{A}$ need not be an attractor in any topological sense: the basin $\mathrm{Bas}(\mathcal{A})$ need not contain a neighborhood of $\mathcal{A}$, nor be of second Baire category.

In order to understand the nature of measure-theoretic attractors of maps from $C_{\mathrm{nf}}^2(I)$, certain types of interval cycles are of particular interest. We say that a compact interval K is *periodic (of period r)* if $K, \ldots, f^{r-1}(K)$ have disjoint interiors and $f^r(K) \subset K$. We call the union $\bigcup_{i=0}^{r-1} f^i(K)$ a *cycle of intervals* and denote it by $\mathrm{cyc}(K)$. If $K_0 \supset K_1 \supset \cdots$ is a nested sequence of periodic intervals of periods $r_0 < r_1 < \cdots$, then $S = \bigcap_{i=0}^{\infty} \mathrm{cyc}(K_i)$ is called a *solenoidal set* and all points from S are called *solenoidal points*. If $r_{i+1} = 2r_i$ for every i sufficiently large, then we say that both S and its points are of *Feigenbaum type*. If i_0 is such that $\mathrm{cyc}(K_{i_0}) \cap \mathrm{Crit} = S \cap \mathrm{Crit}$ and all periodic points in $\mathrm{cyc}(K_{i_0})$ are hyperbolic repelling, then we say that $\mathrm{cyc}(K_{i_0})$ is *solenoidal*. Thus all solenoidal sets are of Cantor type and all cycles $\mathrm{cyc}(K_i)$ are solenoidal if i is sufficiently large.

A compact invariant set A is *Lyapunov stable* if for every neighborhood U of A, there exists a neighborhood V of A such that $f^n(V) \subset U$ for every n. We have the following classification of measure-theoretic attractors, see [M, Ke, BL, Ly1, MS, SV].

Proposition 9. *If $f \in C_{nf}^2(I)$, then f has countably many attractors $\mathcal{A}$, which are of the following types:*

(1) *$\mathcal{A}$ is an attracting periodic orbit;*
(2) *$\mathcal{A}$ is a cycle of intervals on which Lebesgue a.e. orbit is dense;.*
(3) *$\mathcal{A}$ is a solenoidal set (the infinitely renormalizable case). Then $\mathcal{A}$ is Lyapunov stable and the basin $\mathrm{Bas}(\mathcal{A})$ is of second Baire category.*
(4) *$\mathcal{A}$ is a minimal Cantor set, but not of the above type. In particular, $\mathcal{A}$ is not Lyapunov stable, and $\mathrm{Bas}(\mathcal{A})$ is of first Baire category.*

For Lebesgue a.e. $x \in I$, either x has a finite orbit (that is, x is eventually periodic) or $\omega(x)$ is one of the attractors above, and the number of attractors of type (2)-(4) together is no more than the number of critical points (because each of them must contain at least one critical point).

More can be said: there can be countably many disjoint cycles, but all but finitely many of them must be disjoint from the basins of periodic attractors. If a cycle is solenoidal, then almost all its orbits are attracted by the solenoidal set contained in the cycle. If a cycle $\mathrm{cyc}(K)$ contains a dense orbit, then one of the following two holds:

- The whole cycle is an attractor and almost all its points are dense in $\mathrm{cyc}(K)$. If K has period r this can still mean that K consists of two intervals J and J' with a common boundary point such that $f^r(J) = J'$ and $f^r(J') = J$. In this case f^{2r} is topologically mixing on J. Otherwise f^r is topologically mixing on K.
- The cycle contains finitely many attractors of type (4) attracting the orbits of almost all points in the cycle (hence almost no point has a dense orbit in the cycle).

An attractor of type (3) or (4) is called a *solenoidal* and a *wild attractor* respectively. Proposition 9 implies that the orbit of λ-a.e. $x \notin \mathrm{AsPer}$ accumulates on Crit, so every solenoidal set is in fact a solenoidal attractor. It is well known that attractors of type (1) and (3) are uniquely ergodic, and it follows from [BSS, Theorem 4] that this is also true for attractors of type (4). The existence of wild attractors has been proved for unimodal maps only if the combinatorial properties of the map are very specific, and the critical order ℓ_c is sufficiently large ($\ell_c \gg 2$). The prototype is the unimodal Fibonacci map [BKNS], but there are other Fibonacci-like combinatorics that allow wild attractors, see [Br3]. For multimodal maps, there are combinatorial types that allow wild attractors also if all critical orders are $\ell_c = 2$ [vS].

2.3. **Distortion results.** In what follows we denote by $|A|$ or $\lambda(A)$ the Lebesgue measure of a measurable set $A \subset I$ (also, λ_2 will denote the two-dimensional Lebesgue measure). The *density* of a set X in J is $|X \cap J|/|J|$. A point x is a *(Lebesgue) density point* of X if $\lim_{\varepsilon \to 0} |X \cap (x - \varepsilon, x + \varepsilon)|/2\varepsilon = 1$.

Many of the arguments in this paper rely on measuring images under f^n of neighborhoods U of density points of certain sets. If $f^n|_U$ is diffeomorphic then the Koebe Principle (see Proposition 13) is used to estimate how densities change, but in general, U can visit several critical points in its first n iterates. In this case, we need more advanced techniques and

results (relying on work in [BM1, MS, SV]), which we summarize below in Theorems 20 and 21.

We call a sequence $(G_i)_{i=0}^{l}$ of intervals a *chain* if G_i is a maximal interval such that $f(G_i) \subset G_{i+1}$, $i = 0, \ldots, l-1$. If $(H_i)_{i=0}^{l}$ and $(G_i)_{i=0}^{l}$ are chains and $H_i \subset G_i$ for every i, then we will write $(H_i)_{i=0}^{l} \subset (G_i)_{i=0}^{l}$. If $x \in G_0$ (or J is a subinterval of G_0), then we call $(G_i)_{i=0}^{l}$, or sometimes just the interval G_0, *the pullback (chain) of G_l along $x, \ldots, f^l(x)$* (or *along $J, \ldots, f^l(J)$*). The *order* of a chain is the number of intervals G_i, $0 \le i < l$, intersecting Crit.

Remark 10. *Under the hypothesis $f \in C_{\mathrm{nf}}^2(I)$, if G_l is a small interval not too close to any attracting periodic orbit, then all intervals G_i, $i < l$, are also very small by Proposition 7. (Notice that the closure of the set of attracting periodic points only contains periodic points because the periods of attracting orbits are bounded and recall that* Crit *only contain non-periodic points, see assumption (2).)*

Given intervals $J \subset K$, we say that J is ξ-*well inside K* if the components L and R of $K \setminus J$ satisfy $|L|, |R| \ge \xi|J|$. If in addition $\xi|L| \le |R|$ and $\xi|R| \le |L|$, then we say that J is ξ-*well centered in K*.

A differentiable map without critical points $f : J \to \mathbb{R}$ has *distortion bounded by $\kappa > 0$* if

$$\sup_{x,y \in J} \frac{|f'(x)|}{|f'(y)|} \le \kappa.$$

We emphasize that if the density of a subset X of J is very close to 1, say $> 1 - \varepsilon$, then the density of $f(X)$ in $f(J)$ is $> 1 - \kappa\varepsilon$, so it is very close to 1 as well.

An open subset V of $\mathbb{R}$ is called *nice* if $\mathrm{orb}(\partial V) \cap V = \emptyset$. The *first entry map* to a nice set V is the map $\phi_V : D(V) \to V$ defined on the domain $D(V) = \cup_{n \ge 1} f^{-n}(V)$ by $\phi_V(x) = f^{r_V}(x)$, where $r_V = \min\{k > 0 : f^k(x) \in V\}$ is the *first entry time*. The maximal intervals J on which the first entry time r_V is constant are called *entry domains*. By convention we assume that J does not intersect V: if J is a subset of a component of V with first entry time $r_V(J) \equiv r$, then we prefer to call J a *return domain* with *return time r*. The main reason why nice sets are "nice" is that the return and entry domains are all disjoint. Furthermore, all components in the backward orbit of a nice set are nice, and two such intervals are either nested or disjoint.

Lemma 11. *Let $f \in C_{\mathrm{nf}}^2(I)$. Then for every $\xi > 0$ there exists $\xi' = \xi'(\xi, f) > 0$ such that if T is a component of the preimage of an interval V and U is an interval ξ-well inside V (respectively, ξ-well centered in V), then the preimage J of U in T is ξ'-well inside T (respectively, ξ'-well centered in T).*

Proof. This follows easily from [BM1, Lemmas 3.2 and 3.3]. $\qquad\square$

The *Schwarzian derivative* of a C^3 map f is defined for every $x \notin$ Crit by

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2.$$

A C^1 version reads: f has negative Schwarzian derivative if $1/\sqrt{|f'|}$ is convex on every interval where it is defined. If f has negative Schwarzian derivative, then we can frequently estimate distortion using the Koebe Principle, see [MS, Section IV.1] or [BM1, "Koebe lemma"]:

Proposition 12 (Koebe Principle for negative Schwarzian derivative maps)**.** *If $f : G \to f(G)$ is a diffeomorphism with negative Schwarzian derivative, $H \subset G$ and $f(H)$ is ξ-well inside $f(G)$, then H is $\xi^3/(2(3\xi + 2)^2)$-well inside G and $f|_G$ has distortion bounded by $((1 + \xi)/\xi)^2$.*

However, in this paper we will use also use a C^2-version of these classical results.

Proposition 13 (C^2 Koebe Principle)**.** *Given $f \in C^2_{\mathrm{nf}}(I)$, there is a function $Q : (0, \infty) \to (0, \infty)$ with $\lim_{\varepsilon \to 0} Q(\varepsilon) = 0$ such that the following holds. Suppose that $H \subset G$ are intervals such that $f^l|_G$ is a diffeomorphism and $f^l(H)$ is ξ-well-inside $f^l(G)$ for some $\xi > 0$. Then there exists $\xi' = \xi'(\xi, f) > 0$ (with $\xi' \to \infty$ as $\xi \to \infty$), and*

$$\kappa = \exp\left(Q(\max_{0 \leq i < l} |f^i(G)|) \cdot \sum_{i=0}^{l-1} |f^i(H)| \right) \cdot \left(\frac{1 + \xi}{\xi} \right)^2 \tag{3}$$

such that the distortion of $f^l|_H$ is bounded by κ and H is ξ'-well inside G.

Proof. See [SV, Proposition 2]. $\qquad\square$

We have an obvious bound $\sum_{i=0}^{l-1} |f^i(H)| \leq 1$ when the $f^i(H)$, $0 \leq i < l$, are pairwise disjoint. This is what we will use in the following corollary, which uses Lemma 11 and Proposition 13 as well.

Corollary 14. *Let $f \in C^2_{\mathrm{nf}}(I)$. Then for any $\xi > 0$ and $k \geq 0$, there are $\xi' = \xi'(\xi, k, f) > 0$, $\sigma = \sigma(\xi, k, f) > 0$ such that the following statement holds: Let $(H_i)_{i=0}^l \subset (G_i)_{i=0}^l$ be chains such that (G_i) has order at most k, G_l is a small interval close enough to Crit, and the intervals H_i are pairwise disjoint. If H_l is ξ-well inside G_l, then H_0 is ξ'-well inside G_0. If in addition $k = 0$, then there is $\kappa = \kappa(\xi, f) > 0$ such that $f^l|_{H_0}$ has distortion bounded by κ.*

Proof. By saying that "G_l is a small interval close enough to Crit" we mean that there is $\varepsilon_0 = \varepsilon_0(f)$ such that $|G_l| < \varepsilon_0$ and $\mathrm{dist}(G_l, c) < \varepsilon_0$ for some non-periodic critical point c, where ε_0 is chosen so that, if ε_1 satisfies $Q(\varepsilon) < 1$ for every $\varepsilon \leq \varepsilon_1$ in Proposition 13, then $|G_i| < \varepsilon_1$ for every $i < l$ (see Remark 10).

The case $k = 0$ is just Proposition 13. Let us give the proof for $k = 1$; the idea is the same for $k > 1$. Let G_t be the interval from the chain containing the critical point. Now we construct three subchains, $(H_i)_{i=0}^t \subset (G_i)_{i=0}^t$, $(H_i)_{i=t}^{t+1} \subset (G_i)_{i=t}^{t+1}$ and $(H_i)_{i=t+1}^l \subset (G_i)_{i=t+1}^l$. Applying Proposition 13 to the third chain, we find $\xi_1 = \xi_1(\xi, f)$ and $\kappa_1 = \kappa_1(\xi, f)$ such that H_{t+1} is ξ_1-well inside G_{t+1}. Applying Lemma 11 to the middle chain, we find $\xi_2 = \xi_2(\xi_1, f)$ such that H_t is ξ_2-well inside G_t. Finally, applying again Proposition 13 to the first chain, we find $\xi = \xi(\xi_2, f)$ such that H_0 is ξ-well inside G_0. $\qquad\square$

The Koebe property refers to distortion control in the presence of Koebe space. Slightly weaker is the Macroscopic Koebe property, which refers to the preservation of Koebe space

under pullback. Hence the fact that H is ξ'-well inside G in Proposition 13 is basically a Macroscopic Koebe statement.

In order to use the above results, we need conditions guaranteeing the existence of Koebe space at the end of chains. The following propositions are particularly useful in this regard.

Proposition 15. *Let $f \in C^2_{\mathrm{nf}}(I)$. Then for every $\xi > 0$ there exists $\xi' = \xi'(\xi, f) > 0$ such that if V and U are nice intervals, U is ξ-well inside V, $x \in V$ and $f^k(x) \in U$ (with $k \geq 1$ not necessarily minimal), then the pullback of U along $x, \ldots, f^k(x)$ is ξ'-well inside the return domain to V containing x.*

In particular, if U is a return domain to V which is ξ-well inside V, then all return domains to U are ξ'-well inside U.

Proof. This is Theorem C(1) from [SV] (see also the remark below Theorem C(1) and the erratum to that paper). The second statement follows easily from the first one, by fixing an arbitrary return domain K to U and $x \in K$, and choosing k as the return time of K. The interval K is then the pullback of U along $x, \ldots, f^k(x)$ and U is the return domain to V containing x. $\qquad\square$

Proposition 16. *Let $f \in C^2_{\mathrm{nf}}(I)$ and let x be a recurrent point of f which is neither periodic nor of Feigenbaum type. Then there are $\xi_0 = \xi_0(f) > 0$ and an arbitrarily small nice neighborhood J of x such that the return domain to J containing x is ξ_0-well inside J. Assume in addition that x is not solenoidal, and that I_0 is a nice neighborhood of x so small that it contains no periodic neighborhood of x. Let $(I_m)_{m=0}^{\infty}$ be the sequence of nice intervals such that I_m the return domain to I_{m-1} containing x. In this case, there are infinitely many m such that I_{m+1} is ξ_0-well inside I_m.*

Proof. This is a mixture of Theorems A(1) and A'(2) from [SV]. $\qquad\square$

2.4. **Notions from ergodic theory.** Let X be a topological space with Borel σ-algebra $\mathcal{B}$ and let $f : X \to X$ be a Borel measurable map. Recall that a probability measure μ on $\mathcal{B}$ is called *invariant* (respectively, *non-singular*) if $\mu(f^{-1}(A)) = \mu(A)$ (respectively, $\mu(A) = 0$ if and only if $\mu(f^{-1}(A)) = 0$) for any $A \in \mathcal{B}$. In what follows we assume that μ is non-singular but not necessarily invariant.

Definition 17. *Let (X, μ, f) be defined as above. Write $X^2 := X \times X$, $\mu_2 := \mu \times \mu$, $f_2 := f \times f$ and let $\mathcal{B}_2$ be the Borel σ-algebra in $X \times X$. We call the system (X, μ, f)*

- conservative *if for every set $A \in \mathcal{B}$ with $\mu(A) > 0$, there is $n \geq 0$ such that $\mu(A \cap f^n(A)) > 0$;*
- ergodic *if $f^{-1}(A) = A \in \mathcal{B}$ implies $\mu(A) = 0$ or 1. If f is ergodic and conservative, then μ-a.e. orbit is dense in $\mathrm{supp}(\mu)$, see Lemma 19 below.*
- exact *if $f^{-n}(f^n(A)) = A \in \mathcal{B}$ for every $n \geq 0$ implies $\mu(A) = 0$ or 1;*
- mixing *if μ is invariant and*

$$\lim_{n \to \infty} \mu(A \cap f^{-n}(B)) = \mu(A)\mu(B)$$

for all sets $A, B \in \mathcal{B}$;

- weak mixing *if μ is invariant, but 1 is the only eigenvalue corresponding to a measurable eigenfunction of the operator $P : L^1(\mu) \to L^1(\mu)$ defined by $P\psi = \psi \circ f^{-1}$. An equivalent definition of weak mixing is that the Cartesian product system (X^2, μ_2, f_2) is ergodic.*
- *For non-invariant measures, we can still speak of mildly mixing: A non-singular probability measure is* mildly mixing *if for every set $A \in \mathcal{B}$ of positive measure*

$$\liminf_{n \to \infty} \mu(A \cap f^{-n}(A)) > 0.$$

Mild mixing implies that (X^2, μ_2, f_2) is ergodic. (If f is invertible, then mild mixing is equivalent to (X^2, μ_2, f_2) being ergodic. In this case f also preserves a probability measure equivalent to μ, but neither of these stronger statements holds in general if f is non-invertible. See [HS] for more results.)

Lemma 18. *If (X, μ, f) is exact, then (X^2, μ_2, f_2) is ergodic.*

Proof. Assume by contradiction that (X^2, μ_2, f_2) is not ergodic, so there is $U \in \mathcal{B}_2$ such that $f_2^{-1}(U) = U$ and $0 < \mu_2(U) < 1$. Then there is $a \in X$ such that $0 < \mu(U_a) < 1$ for $U_a = \{y \in X : (a, y) \in U\}$. Let $V_a = X \setminus U_a$. Then $f^n(U_a) \cap f^n(V_a) = \emptyset$ for all $n \geq 0$, so $(f^{-n} \circ f^n)(U_a) = U_a$ for all n, contradicting that μ is exact. $\square$

Lemma 19. *If X is separable and (X, μ, f) is ergodic and conservative, then μ-a.e. x has a dense orbit in $\mathrm{supp}(\mu)$.*

Proof. Let $\{U_n\}_{n \in \mathbb{N}} \subset \mathcal{B} \cap \mathrm{supp}(\mu)$ be a countable basis for the relative topology of $\mathrm{supp}(\mu)$, *i.e.*, $U_n = \mathrm{supp}(\mu) \cap U_n'$ where U_n' are those elements of a countable basis of X that intersect $\mathrm{supp}(\mu)$. Since $\mathrm{supp}(\mu)$ is by definition the smallest closed set of full measure, $\mu(U_n) > 0$ for every n.
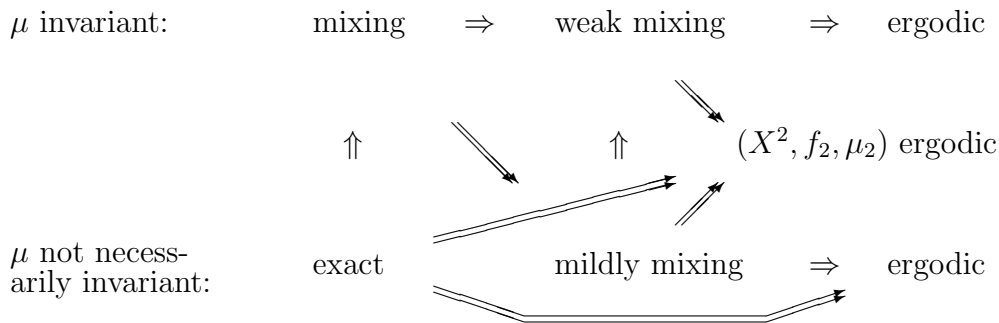
Set $Y_n := \{x \in X : f^k(x) \in U_n \text{ infinitely often}\}$. Then $f^{-1}(Y_n) = Y_n$, so $\mu(Y_n) = 0$ or 1 for each n. If $\mu(Y_n) = 1$ for each n, then $Y := \bigcap_n Y_n$ has full measure, and every $x \in Y$ has a dense orbit in $\mathrm{supp}(\mu)$.

So assume that n is such that $\mu(Y_n) = 0$. For $Z = U_n \setminus Y_n$, we can write $Z = \bigcup_{k \geq 0} Z_k$ where $Z_k = \{x \in U_n : k = \max\{i \geq 0 : f^i(x) \in U_n\}\}$. Since $\mu(Z) = \mu(U_n) > 0$, there is k such that $\mu(Z_k) > 0$, and hence $\mu(f^k(Z_k)) > 0$. But then $f^k(Z_k) \subset U_n$ is a set of positive measure that never visits U_n again, and this contradicts that μ is conservative. $\square$

The following diagram summarizes the implications between these various notions of mixing. The notions on the bottom line can be defined for non-invariant measures μ, but any implication to the top line requires that μ is f-invariant and in particular conservative.

3. INDUCING TO CRITICAL NEIGHBORHOODS AND STRONGLY WANDERING SETS

In this section, we prove that strongly wandering sets have zero measure as a consequence of Theorem 20 below. Theorem 20 is a refinement of Theorem 1 of [CL] (where the slightly stronger hypothesis $f \in C_{\mathrm{nf}}^3(I)$ is used), which in turns improves Theorem D of [SV].

μ invariant: mixing $\Rightarrow$ weak mixing $\Rightarrow$ ergodic

$\Uparrow$ $\Uparrow$ (X^2, f_2, μ_2) ergodic

μ not necess-arily invariant: exact   mildly mixing $\Rightarrow$ ergodic

Theorem 20. *Let $f \in C^2_{\mathrm{nf}}(I)$. Then there are positive constants $\xi = \xi(f)$, $\kappa = \kappa(f)$, $\delta = \delta(f)$ and, for every $\varepsilon > 0$ and $c \in \mathrm{Crit}$, open intervals $c \in U_c \subset V_c$, with $|V_c| < \varepsilon$, such that the following conditions hold:*

  (i) *The set $U = \bigcup_{c \in \mathrm{Crit}} U_c$ is nice and U_c is ξ-well inside V_c for every $c \in \mathrm{Crit}$.*

  (ii) *If J is an entry domain of the first entry map ϕ to U, say $\phi|_J = f^j|_J$ and $\phi(J) = U_c$, then there is $K \supset J$ such that $f^j|_K$ is a diffeomorphism and $f^j(K) = V_c$. Moreover, $\phi|_J$ has distortion bounded by κ.*

  (iii) *If $c \in \mathrm{Crit}$, then there are an open interval $W_c \subset U_c$ and $k_c \in \mathbb{N}$ such that $|W_c| > \delta|U_c|$, $f^{k_c}|_{W_c}$ is a diffeomorphism with distortion bounded by κ, and*
     – *either $W_c \subset f^{k_c}(W_c) \subset U_c$ (if c is of Feigenbaum type);*
     – *or $k_c = 1$ and $\partial W_c \cap D(U) = \emptyset$ (if c is not of Feigenbaum type).*

The existence of the sets $U_c \subset V_c$ from Theorem 20 suggests the construction of an induced Markov map $F : U \to U$ with all branches mapping monotonically onto a component of U. The first, very detailed, constructions of such induced maps go back to Jakobson [Ja], but the abstract statement for unimodal maps having no periodic or Cantor attractors comes from Martens' PhD. thesis [Ma]. During the writing of this paper, we learnt about a multimodal C^3 version by Cai and Li [CL], see Proposition 45. Their construction precludes the existence of parabolic points, hence we improve it slightly by showing the version below, where Crit' denotes the set of critical points interior to metric attractors of type (2) in Proposition 9.

Let $E := \bigcup_{n \geq 0} f^{-n}(A)$ where A is the union of all attractors of type (2) in Proposition 9, that is, the union of cycles of intervals in which λ-a.e. orbit is dense.

Theorem 21. *Let $f \in C^3_{\mathrm{nf}}(I)$ and let U_c, V_c, $c \in \mathrm{Crit}'$, be defined as in Theorem 20 for ε sufficiently small and let $U' = \bigcup_{c \in \mathrm{Crit}'} U_c$. Then for Lebesgue a.e. $x \in E$ we can find $k_x \in \mathbb{N}$ and intervals $G_x \supset H_x \ni x$ (with either $H_x = H_y$ or $H_x \cap H_y = \emptyset$) such that $f^{k_x} : G_x \to V_c$ is diffeomorphic for some $c \in \mathrm{Crit}'$ and $f^{k_x}(H_x) = U_c$.*

Let $F : U' \to U'$ be defined by $F|_{H_x} = f^{k_x} : H_x \to U_c$ for the appropriate $c \in \mathrm{Crit}'$. Then all iterates of F is well defined λ-a.e and its branches have uniformly bounded distortion, i.e., there is $\kappa = \kappa(f) > 0$ such that for every $n \in \mathbb{N}$ and every interval H on which $F^n|_H : H \subset U \to U_c$ is a diffeomorphism, the distortion of $F^n|_H$ is bounded by κ.

We postpone the somewhat technical proofs of Theorems 20 and 21 to the appendix.

Proof of Theorem A. Let $f \in C^2_{\mathrm{nf}}(I)$ and assume that W is a strongly wandering set of positive measure. Let x be a density point of W. It is not restrictive to assume that x belongs to the interior of a cycle $\mathrm{cyc}(K)$ which either contains a dense orbit or is of solenoidal type. We can also assume that $\mathrm{orb}(x)$ accumulates on Crit.

Let $\tau > 0$ be such that $(x - \tau, x + \tau) \subset \mathrm{cyc}(K)$. We can assume that the density of W in any subinterval of $(x - \tau, x + \tau)$ containing x is larger than $1 - \varepsilon_0$, with $\varepsilon_0 > 0$ so small that if $d = \#\,\mathrm{Crit}$, then $1 - \delta^{-d}\kappa^{2d+1}\varepsilon_0 > 1/2$, where $\delta = \delta(f)$ and $\kappa = \kappa(f)$ are the numbers from Theorem 20. Finally we fix $\varepsilon > 0$ such that if U is defined as in Theorem 20, then all components of U and entry domains to U have length less than τ (by the non-Contraction Principle).

Let J_1 be either the entry domain to U containing x, say $\phi|_{J_1} = f^{j_1}|_{J_1}$, $f^{j_1}(J_1) = U_{c_1}$, or the component U_{c_1} of U containing x (then we take $j_1 = 0$). Since $J_1 \subset (x - \tau, x + \tau) \subset \mathrm{cyc}(K)$ and $\mathrm{cyc}(K)$ is invariant, the density of W in J_1 is $> 1 - \varepsilon_0$ and $U_{c_1} \subset \mathrm{cyc}(K)$.

Let W_{c_1} be the interval from Theorem 20, part (iii). Since $|W_{c_1}| > \delta|U_{c_1}|$ and the density of $f^{j_1}(W)$ in U_{c_1} is $> 1 - \kappa\varepsilon_0$, the density of $f^{j_1}(W)$ in W_{c_1} is $> 1 - \delta^{-1}\kappa\varepsilon_0$ and the density of $f^{j_1+k_1}(W)$ in $f^{k_1}(W_{c_1})$ is $> 1 - \delta^{-1}\kappa^2\varepsilon_0$, where $k_1 = k_{c_1}$ is as in Theorem 20(iii). If c_1 is a Feigenbaum critical point, then $W_{c_1} \subset f^{k_{c_1}}(W_{c_1}) \subset U_{c_1}$, so the density of $f^{j_1}(W)$ in $f^{k_{c_1}}(W_{c_1})$ is also $> 1 - \delta^{-1}\kappa\varepsilon_0$. Therefore the densities of both $f^{j_1}(W)$ and $f^{j_1+k_1}(W)$ in $f^{k_{c_1}}(W_{c_1})$ are $> 1/2$, hence $f^{j_1}(W) \cap f^{j_1+k_1}(W) \neq \emptyset$, contradicting that W is strongly wandering.

Now we deal with the non-Feigenbaum case. Here $k_1 = 1$, $f|_{W_{c_1}}$ is a diffeomorphism and $\partial W_{c_1} \cap D(U) = \emptyset$. On the other hand, $U_{c_1} \subset \mathrm{cyc}(K)$, so $f(W_{c_1}) \subset \mathrm{cyc}(K)$ as well. Moreover, the choice of $\mathrm{cyc}(K)$ guarantees that the orbits of almost all its points accumulate on Crit. Hence $f(W_{c_1})$ is the pairwise disjoint union (up to a measure zero set) of entry domains to U and components of U. Since the density of $f^{j_1+1}(W)$ in $f(W_{c_1})$ is $> 1 - \delta^{-1}\kappa^2\varepsilon_0$, it is also $> 1 - \delta^{-1}\kappa^2\varepsilon_0$ in one of these entry domains or components, call it J_2.

Now we repeat the argument. Say that $\phi|_{J_2} = f^{j_2}|_{J_2}$, $f^{j_2}(J_2) = U_{c_2}$, or J_2 is a component U_{c_2} of U (when we take $j_2 = 0$). Then $f^{j_1+j_2+1}(W)$ has density $> 1 - \delta^{-2}\kappa^3\varepsilon_0$ in U_{c_2}. If c_2 is a critical point of Feigenbaum type, then we get a contradiction as before. If not, then we find an entry domain to U or a component of U, call it J_3, such that $f^{j_1+j_2+2}(W)$ has density $> 1 - \delta^{-2}\kappa^4\varepsilon_0$ in J_3. Proceeding in this way, we find intervals $J_1, J_2, \ldots, J_{d+1}$, each of them either an entry domain to U or a component of U, say $f^{j_i}(J_i) = U_{c_i}$, such that the density of $f^{t_i}(W)$ in U_{c_i} is $> 1 - \delta^{-(i-1)}\kappa^{2i-1}\varepsilon_0$ for every $i = 1, 2, \ldots, d+1$, with $t_i = j_1 + \cdots + j_i + i - 1$. Find $i < i'$ such that $c_i = c_{i'}$. Then the densities of $f^{t_i}(W)$ and $f^{t_{i'}}(W)$ in U_{c_i} are $> 1/2$, and we have the required contradiction. $\square$

4. lim sup FULLNESS OF LEBESGUE MEASURE

The following definition was first used in Barnes [Ba].

Definition 22. *The system (X, μ, f) is called* lim sup full *if $\limsup_{n \to \infty} \mu(f^n(A)) = 1$ for any $A \in \mathcal{B}$ with $\mu(A) > 0$.*

Theorem 23. *Let $f \in C_{\mathrm{nf}}^3(I)$ be a topologically mixing map having no Cantor attractors. Then f is* $\limsup$ *full with respect to Lebesgue measure.*

Proof. It suffices to show that if $f \in C_{\mathrm{nf}}^3(I)$ (and satisfying (1)) has an invariant interval I' such that $f|_{I'}$ is topologically mixing and has no Cantor attractors, then we have $\limsup_n \lambda(f^n(A)) = \lambda(I')$ for any measurable set $A \subset I'$ of positive measure.

Fix a critical point c interior to I' and let $\varepsilon > 0$ be small enough so that Theorem 21 holds and the corresponding interval U_c lies in I'. Let x be a density point of A. Since f has no Cantor attractors in I', there is no loss of generality in assuming $\omega(x) = I'$ and (after replacing if necessary A by some of its iterates) $x \in U_c$. Since all iterates of the induced map F from Theorem 21 are well defined for λ-a.e. point in U_c, we can assume that this is the case for x. This implies that there are intervals $x \in H_n \subset U_c$ with $\bigcap_n H_n = \{x\}$ and integers k_n such that $f^{k_n}|_{H_n} : H_n \to U_{c'}$ are diffeomorphisms with uniform distortion bound for some $c' \in \mathrm{Crit}'$. Hence $\lambda(f^{k_n}(A \cap H_n)) \to \lambda(U_{c'})$ as $n \to \infty$. Observe that $U_{c'} \subset I'$, so there is j such that $I' = f^j(U_{c'})$ (because $f|_{I'}$ is topologically mixing). Then also $\lambda(f^{k_n+j}(A \cap H_n)) \to \lambda(I')$, as required. $\square$

Proposition 24. *Let $f \in C_{\mathrm{nf}}^3(I)$ be topologically mixing. Then the following statements are equivalent:*

  (i) *f has an acip μ;*
  (ii) *$\liminf_n \lambda(f^n(A)) > 0$ for every measurable set A of positive Lebesgue measure;*
  (iii) *$\liminf_n \lambda(f^n(A)) = 1$ for every measurable set A of positive Lebesgue measure.*

In this case μ is equivalent to λ (that is, $\mu(A)$ if and only if $\lambda(A) = 0$) and λ_2 is ergodic and conservative.

Remark 25. *Under C^3 assumptions, no transitive Cantor set can have positive measure, [SV, Theorem E(1), cf. also Remark 1], so in particular a Cantor attractor has Lebesgue measure 0, and cannot support an acip. We expect this to be true in the C^2 setting as well, but we have no proof.*

Proof. The implication (iii)$\Rightarrow$(ii) is trivial.

We prove (ii)$\Rightarrow$(i). According to [St], the existence of an acip for f is equivalent to the existence of $\delta > 0$ and $0 < \alpha < 1$ such that $\lambda(A) < \delta$ implies $\lambda(f^{-n}(A)) < \alpha$ for every n and A. Hence, if f does not admit an acip, then there are sets $\{A_k\}$ and integers $(n_k)_{k \geq 1}$ with $n_k \to \infty$, $\lambda(A_k) \to 0$ and $\lambda(f^{-n_k}(A_k)) \geq 1 - 2^{-k-1}$. Take $A = \bigcap_k f^{-n_k}(A_k)$, then $\lambda(A) \geq \frac{1}{2}$ and $\liminf_n \lambda(f^n(A)) = 0$, contrary to condition (ii).

For the implication (i)$\Rightarrow$(iii), assume that f admits an acip μ. Since (wild) Cantor attractors have zero Lebesgue measure, μ cannot be supported on them. Then Theorem 23 applies.

We claim that μ is equivalent is to λ. Assume the contrary to find a measurable set A such that $\mu(A) = 0 < \lambda(A)$. Let $B = \bigcup_{n=0}^\infty f^{-n}(A)$. Then $\mu(B) = 0$ but, by Theorem 23, $\lambda(B) = 1$. This is impossible.

The equivalence of μ and λ (or just the absolute continuity of μ) and Theorem 23 imply that if $\mu(A) > 0$, then $\limsup_n \mu(f^n(A)) = 1$. Since μ is invariant, the sequence

$\{\mu(f^n(A))\}_{n\in\mathbb{N}}$ is non-decreasing. Therefore $\lim_n \mu(f^n(A)) = 1$, and also $\lim_n \lambda(f^n(A)) = 1$ by the equivalence of μ and λ. Moreover, λ is exact, hence λ_2 is ergodic (Lemma 18), and μ_2 is conservative (because of Poincaré recurrence), so λ_2 is conservative as well (because μ_2 and λ_2 are equivalent). $\qquad\square$

5. Li-Yorke pairs and scrambled sets

In this section we prove our results on scrambled sets and the λ_2-measure of the sets Dis, Asymp and LY.

Proof of Theorem B. Assume that a scrambled set S has positive measure. We can assume that all its points are attracted by the same attractor, which must be either a cycle of intervals containing a dense orbit, or a minimal Cantor set. Since f^n is one-to-one on S for every n, the first possibility can be immediately discarded by Theorem 23. Hence we may assume that all points of S are attracted by a minimal Cantor set W. (In fact W must be a wild attractor, because points attracted by solenoidal sets are approximately periodic and a scrambled set can contain at most one approximately periodic point by Proposition 6.)

By Theorem A there are integers $n < m$ such that $f^n(S) \cap f^m(S) \neq \emptyset$. Let $x \in f^n(S) \cap f^m(S)$. Then there is $y \in f^n(S)$ such that $f^{m-n}(y) = x$. Since (x,y) is a Li-Yorke pair, so is $(f^{m-n}(x), f^{m-n}(y)) = (f^{m-n}(x), x)$. Find a sequence $\{l_k\}$ such that $|f^{m-n+l_k}(x) - f^{l_k}(x)| \to 0$. We may assume that $(f^{l_k}(x))_{k\in\mathbb{N}}$ accumulates at $p \in W$. Then $f^{m-n}(p) = p$, which is impossible because W is infinite and minimal. $\qquad\square$

Proposition 26. *A multimodal map f has no closed invariant scrambled set (apart from a singleton set).*

Proof. Suppose by contradiction that the closed non-singleton S is invariant and scrambled. Then clearly it can contain at most one fixed point. Let $y \in S$ be a non-fixed point. Then because $(y, f(y))$ is Li-Yorke, there is a sequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_k |f^{n_k}(y) - f^{n_k}(f(y))| = 0$ and $\{f^{n_k}(y)\}$ converges. By continuity $\lim_k f^{n_k}(y) = \lim_k f^{n_k+1}(y) = f(\lim_k f^{n_k}(y))$, so the limit is a fixed point $p \in S$. Since $(y, f(y))$ is Li-Yorke, $\{f^n(y)\}$ does not converge to p.

Since f is multimodal, there are a sequence (m_k) and a number $\varepsilon > 0$ such that $f^{m_k}(y) \to p$ but $|f^{m_k-1}(y) - p| > \varepsilon$. By taking a subsequence, we can assume that $\{f^{m_k-1}(y)\}_{k\in\mathbb{N}}$ converges, say to q. But then $p = \lim_k f^{m_k}(y) = f(\lim_k f^{m_k}(y))) = f(q)$, so (p, q) is not Li-Yorke. This contradiction proves the proposition. $\qquad\square$

Proof of Theorem C. Assume that $\lambda_2(\mathrm{Asymp} \setminus (\mathrm{AsPer} \times \mathrm{AsPer})) > 0$. Then there are a point $x \in I \setminus \mathrm{AsPer}$ and a Borel set $Y \subset I \setminus \mathrm{AsPer}$ of positive measure such that $\lim_{n\to\infty} |f^n(x) - f^n(y)| = 0$ for every $y \in Y$. We can assume that $\mathrm{orb}(x)$ accumulates on a non-periodic point u.

For every $n \geq 0$, let $d_n : Y \to \mathbb{R}$ be defined by $d_n(y) = \sup_{m\geq n} |f^m(x) - f^m(y)|$. Then (d_n) is a sequence of Borel measurable maps converging pointwise to zero. According to Egorov's theorem we can remove from Y a small set (so that the remaining set Z

has positive measure) in such a way that $\{d_n|_Z\}_{n \in \mathbb{N}}$ converges uniformly to zero, that is, $\operatorname{diam}(f^n(Z)) \to 0$ as $n \to \infty$.

Use Theorem A to find integers $k > m$ such that $f^k(Z) \cap f^m(Z) \neq \emptyset$. Recall that $\operatorname{orb}(Z)$ accumulates at a point u with $|f^j(u) - u| = \varepsilon > 0$ for $j = k - m$. Find $\delta \in (0, \varepsilon/4)$ such that $|f^j(v) - f^j(w)| < \varepsilon/2$ whenever $|v - w| < 2\delta$. Next take $l > m$ so that $\operatorname{dist}(f^l(Z), u) < \delta$ and $\operatorname{diam}(f^l(Z)) < \delta$. Then $|f^l(z) - u| < 2\delta < \varepsilon/2$, hence $|f^{l+j}(z) - f^j(u)| < \varepsilon/2$ for every $z \in Z$. Thus $f^{l+j}(Z) \cap f^l(Z) = \emptyset$, contradicting $f^{m+j}(Z) \cap f^m(Z) \neq \emptyset$ and $l > m$. $\qquad\square$

Proposition 27. *If (I, μ, f) is exact, then $\mu_2(\operatorname{Dis}) = 0$.*

Proof. Assume by contradiction that $\mu_2(\operatorname{Dis}) > 0$. Write $\operatorname{Dis}_x = \{y \in I : (x, y) \text{ is distal}\}$ and $G_{\operatorname{Dis}} := \{x : \mu(\operatorname{Dis}_x) > 0\}$. Then by Fubini's Theorem, $\mu(G_{\operatorname{Dis}}) > 0$. For $x \in G_{\operatorname{Dis}}$, take
$$Y_\varepsilon := \{y \in \operatorname{Dis}_x : \liminf_{n \to \infty} |f^n(x) - f^n(y)| \geq \varepsilon\}.$$
Clearly $Y_\delta \subset Y_\varepsilon$ if $\delta > \varepsilon$ and $f^n(Y_\varepsilon) \cap f^n(I \setminus Y_\varepsilon) = \emptyset$ for all $n \geq 0$. If for some $\varepsilon > 0$, both Y_ε and $I \setminus Y_\varepsilon$ have positive measure, then we have a contradiction to exactness. The remaining possibility is that there is $\varepsilon = \varepsilon(x) > 0$ such that Y_ε has full measure, whereas $\lambda(Y_\delta) = 0$ for all $\delta > \varepsilon$. Now take $\eta > 0$ such that $G_\eta := \{x \in G_{\operatorname{Dis}} : \varepsilon(x) > \eta\}$ has positive measure. Let $R > 1/\eta$, and take distinct points $x_0, x_1, \ldots, x_R \in G_\eta$. Clearly these points can be chosen so that $\liminf_n |f^n(x_i) - f^n(x_j)| > \eta$ for all $i \neq j$. Take N minimal such that $|f^n(x_i) - f^n(x_j)| > \eta$ for all $i \neq j$ and $n \geq N$. However, by the choice of R, there is no space in I to fit the points $f^N(x_i)$, $i = 0, \ldots, R$, so that they have pairwise distance greater than η. This contradiction proves that $\mu_2(\operatorname{Dis}) = 0$. $\qquad\square$

Next, for the proof of Theorem D, recall that we assumed that f has no wild attractor.

Proof of Theorem D*.* Theorem 23 implies that f is $\limsup$ full w.r.t. Lebesgue measure. By [Ba, Theorem A], if $f : I \to I$ is a multimodal surjective map that is $\limsup$ full with respect to a measure μ, then f is exact. (The proof is stated for d-to-1 maps, but applies with minor changes to the "at most d-to-1" setting as well.) Hence λ is exact. By Lemma 18, λ_2 is ergodic for (I^2, f_2).

Take $x \in I$. We prove that there is a full measure set A_x such that $\limsup_{n \to \infty} |f^n(y) - f^n(x)| \geq \operatorname{diam}(I)/2$ for every $y \in A_x$. If the opposite is true, then there are a set A of positive measure and an $m \in \mathbb{N}$ such that $|f^n(y) - f^n(x)| < \operatorname{diam}(I)/2$ for every $y \in A$ and every $n \geq m$. But this contradicts the conclusion of Theorem 23 applied to A, *i.e.*, that $\limsup_n \lambda(f^n(A)) = 1$. Similarly, we can prove that there is a full measure B_x such that $\limsup_{n \to \infty} |f^n(y) - f^n(x)| = 0$ for every $y \in B_x$. Indeed, if this were false then there are $m \in \mathbb{N}$, $\varepsilon > 0$ and a set B of positive measure such that $|f^n(y) - f^n(x)| > \varepsilon$ for every $y \in A$ and every $n \geq m$. Again, this contradicts that $\limsup_n \lambda(f^n(B)) = 1$. Hence
$$\liminf_{n \to \infty} |f^n(y) - f^n(x)| = 0, \quad \limsup_{n \to \infty} |f^n(y) - f^n(x)| \geq \operatorname{diam}(I)/2,$$
for every $y \in U_x = A_x \cap B_x$. $\qquad\square$

By Proposition 24 and Lemma 19 we immediately recover a well-known result on weak-mixing.

Corollary 28. *If $f \in C^3_{\mathrm{nf}}(I)$ is topologically mixing and has an acip, then λ_2-a.e. (x, y) has a dense orbit in I^2.*

Remark 29. *Probably the earliest result in this direction dates back to Ledrappier who proves in [Le, Theorem 1] that a certain class of interval maps is weak Bernoulli. Keller [Ke] proved (weak-)mixing for acips μ of multimodal maps assuming negative Schwarzian derivative. he also showed that $d\mu/d\lambda$ is bounded away from zero on $\mathrm{supp}(\mu)$. One can prove that μ is also mixing by means of an induced map $F : U \to U$ as in Theorem 21 which possesses an acip ν, cf. [Y]. In fact, by taking an appropriate power of F^N, U decomposes into a finite number of F^N-invariant parts on which ν is invariant and mixing. Pulling back ν to the original system, we recover μ and by the topologically mixing condition, μ can have only one mixing component.*

Conservativity of λ_2 is crucial in Lemma 19, and even if λ itself is conservative for (I, f), this does not guarantee that λ_2 is conservative for the Cartesian product. It is for this reason that Proposition 27 and Theorem D are not just direct consequences of ergodicity of λ_2 from Lemma 18. The following conjecture suggests conditions under which λ_2-a.e. pair is Li-Yorke, but has no dense orbit.

Conjecture 30. *We think that if (I, f, λ) is conservative and has no acip, but instead the induced time k_x associated to the induced map in Theorem 21 is non-integrable w.r.t. Lebesgue, and in fact the tail $\lambda(\{x : k_x > s\}) \geq 1/\log s$, then the product system (I^2, f_2, λ_2) is dissipative.*

6. Li-Yorke chaos in the presence of Cantor attractors

In this section, we concentrate on C^2 unimodal maps with Cantor attractors $\mathcal{A} = \omega(c)$ for the unique critical point c. If f is infinitely renormalizable (*i.e.*, type (3) in Proposition 9), then the situation regarding Li-Yorke pairs is well-known: there are none. Instead, the attractor is Lyapunov stable and conjugate to an *adding machine* (Ω, g). In other words, $\Omega = \{(\omega_j)_{j \geq 1} : 0 \leq \omega_j < p_j\}$, for some sequence $(p_i)_{i \geq 1}$ of integers $p_i \geq 2$ (where $p_1 \cdots p_i$ are the periods of the periodic intervals) such that is equipped with product topology and the map g of "adding 1 and carry":

$$g(\omega_1, \omega_2, \dots) = \begin{cases} 0, 0, , \dots, 0, \omega_k + 1, \omega_{k+1}, \omega_{k+2}, \dots & \text{if } k = \min\{i : \omega_i < p_i - 1\}; \\ 0, 0, 0, 0, \dots & \text{if } \omega_i = p_i - 1 \text{ for all } i \geq 1. \end{cases} \quad (4)$$

The following classification is due to [BrJ, Proposition 5.1].

Proposition 31. *Let $f : I \to I$ be a continuous map and $x \in I$. The system $(\omega(x), f)$ is conjugate to some adding machine if and only if x is approximately but not asymptotically periodic, see Definition 1.*

However, there are several constructions leading to *strange adding machines*, *i.e.*, (critical) omega-limit sets that are conjugate to adding machines, but not involving periodic intervals, see [BKM, BrJ, Br4]. In [Br4] it is shown that a strange adding machine can still be an attractor, but in this case it is a wild and not a solenoidal attractor.

The dynamics on such attractors can frequently be understood in terms of generalized adding machines as done in [BKS] and in the proof of Theorem 34 below. Such generalized adding machines are based on the sequence of cutting time, which we will define now.

Let $Z_n(x)$ be the *n-cylinder*, *i.e.,* maximal interval containing x on which f^n is monotone. Unless $x \in \cup_{m<n} f^{-m}(c)$, $Z_n(x)$ is well-defined, but the critical point itself has two sets $Z_n^{\pm}(c)$ on either side of it, with $\mathcal{D}_n := f^n(Z_n^{\pm}(c))$ independent of $\pm$, if n is sufficiently large. One can show that for every x and $n \geq 1$, there is $m \leq n$ such that $f^n(Z_n(x)) = \mathcal{D}_m$; if $x \in f^{-m}(c)$, then x is the common boundary point of two n-cylinder sets $Z_n^{\pm}(x)$, but $f^n(Z_n^{\pm}(x)) = \mathcal{D}_m$ independently of $\pm$, again if n is sufficiently large. We say that n is a *cutting time* if $\mathcal{D}_n \ni c$, and we list them in increasing order $1 = S_0 < S_1 < \dots$. Write $c_n := f^n(c)$. By a short induction proof one can show that $\mathcal{D}_1 = [0, c_1]$ and for $n \geq 2$

$$\mathcal{D}_n = [c_n, c_{\beta(n)}] \text{ and } \mathcal{D}_n \subset \mathcal{D}_{\beta(n)}, \tag{5}$$

where the map $\beta : \mathbb{N} \to \mathbb{N}$ is defined as $\beta(n) = n - \max\{S_k : S_k < n\}$. In the special case that $n = S_k$ is a cutting time, this means that $S_k - S_{k-1}$ is again a cutting time, so one can define the *kneading map* $Q : \mathbb{N} \to \mathbb{N} \cup \{0, \infty\}$ by

$$S_k = S_{k-1} + S_{Q(k)},$$

see [H, Br2]. Here $Q(k) = \infty$ means that S_k does not exist; in this case c is attracted to an orbit of period S_{k-1} or $2S_{k-1}$. Since we assumed that f has a Cantor attractor, $Q(k) < \infty$ for all $k \in \mathbb{N}$. Some properties of the kneading map related to $\omega(c)$ are as follows.

Proposition 32. *(a) The minimal n such that $Z_n^{\pm}(c) \subset [c, c_{S_k}]$ is $n = S_{Q(k+1)}$.*
(b) If f is a unimodal map with kneading map $Q(k) \to \infty$, then the length $|\mathcal{D}_n| \to 0$ and $\omega(c)$ is a minimal Cantor set.

(c) If $\max_k k - Q(k) \leq B$, then $\omega(c) \subset \cup_{n=1+S_k}^{S_k+B} \mathcal{D}_n$ for every k.

(d) The map is renormalizable if and only if there is $k \geq 1$ such that $k = Q(k+1) \leq Q(k+j)$ for all $j > 1$. In this case S_k is the period of renormalisation.

Proof. By definition of the kneading map, $n = S_{Q(k+1)}$ is the smallest positive iterate such that $f^n([c, c_{S_k}]) \ni c$, implying statement (a).
Under the assumption that $Q(k) \to \infty$, this means that $|c_{S_k} - c| \to 0$ and also $|c_{S_{Q(k)}} - c| \to 0$ as $k \to \infty$. Since $\mathcal{D}_{S_k} = [c_{S_k}, c_{S_{Q(k)}}]$, the non-Contraction Principle shows that $|\mathcal{D}_n| \to 0$ as well. Minimality of $\omega(c)$ follows as in *i.e.,* [Br2, Proposition 2] (which proves that c is persistently recurrent) and [Br2, Lemma 8]. This proves statement (b).
For statement (c), observe that $\mathcal{D}_{1+S_l} \subset \mathcal{D}_{1+S_k}$ for all $l \geq k+B$ because of the assumption $\max_k\{k - Q(k)\} \leq B$ and statement (a). It follows that $\mathcal{D}_{m+S_{k+B}} \subset \mathcal{D}_{m+S_k}$ for all $m \geq 1$, and $\cup_{n \geq 1+S_k} \mathcal{D}_n \subset \cup_{n=1+S_k}^{S_k+B} \mathcal{D}_n$. By definition, $\omega(c) \subset \text{Cl} \cup_{n \geq 1+S_k} \mathcal{D}_n \subset \text{Cl} \cup_{n=1+S_k}^{S_k+B} \mathcal{D}_n = \cup_{n=1+S_k}^{S_k+B} \mathcal{D}_n$ as asserted.
Part (d) is [Br2, Proposition 1(iii)]. $\qquad\square$

Special types of unimodal maps are the Feigenbaum map ($S_k = 2S_{k-1}$) and the Fibonacci map ($S_k = S_{k-1} + S_{k-2}$). We call f *Fibonacci-like* if $\{k - Q(k)\}_k$ is bounded.

The proof of the existence of wild attractors was first established in [BKNS] for Fibonacci maps with sufficiently large critical order ℓ. In [Br3], this result was extended to C^3

unimodal maps with negative Schwarzian derivative, sufficiently large critical order and eventually non-decreasing kneading map Q such that $\limsup_k k - Q(k) \leq B$. If the cutting times increase more slowly than Fibonacci-like, then no unimodal map with finite critical order can have a wild attractor.

Before we continue, let us give a short exposition how the existence of wild attractor is proved for unimodal maps. There are two approaches, both based on a random walk on a Markov graph. In [BKNS, Br3, BHa], this Markov graph of I is based on preimages of the orientation reversing fixed point p. The states of that Markov system are pairs of intervals $U_k \subset (Z^+_{S_k}(c) \cup Z^-_{S_k}(c)) \setminus (Z^+_{S_{k+B}}(c) \cup Z^-_{S_{k+B}}(c))$ and ∂U_k belongs to the backward orbit of p. The induced Markov map defined by $G|_{U_k} = f^{S_k}$ preserves the partition $\{U_k\}$ of $[\hat{p}, p]$, where $f^{-1}(p) = \{p, \hat{p}\}$. Viewing the dynamics of G as a random walk, we define "random variables" χ_n by

$$\chi_n(x) = k \text{ if } G^n(x) \in U_k. \tag{6}$$

It is then shown that this process has *positive drift*, *i.e.*, the expectations (measured with respect to Lebesgue measure)

$$\mathbb{E}(\chi_{n+1} - k | \chi_n = k) \geq \eta > 0$$

uniformly in n and l. The second moments $\mathbb{E}((\chi_{n+1} - k)^2 | \chi_n = k)$ are shown to be bounded as well. It follows that $\chi_n(x) \to \infty$ for λ-a.e. x and hence $G^n(x) \to c$. Since $U_k \subset (Z^+_{S_k}(c) \cup Z^-_{S_k}(c))$, $f^{S_k}(U_k) \subset \mathcal{D}_{S_k}$ and since $|\mathcal{D}_n| \to 0$ as $n \to \infty$, this means for the original map that $f^n(x) \to \omega(c)$ for λ-a.e. x, so $\mathcal{A} = \omega(c)$ is an attractor.

In [BHa, Theorem 5.2], a further conclusion is drawn from the positive drift, namely a Borel-Cantelli Lemma argument shows that for λ-a.e. $x \in \text{Bas}(\mathcal{A})$, there is $k_0 = k_0(x)$ such that such that for all $k \geq k_0$

$$\text{if } G^m(x) \in U_k, \text{ then } G^{m+j}(x) \notin U_k \text{ for } j > k. \tag{7}$$

In this paper, we will use a second approach from [Br1, BKS], where the Markov system is a disjoint union $\hat{I} = \sqcup_{n \geq 2} \mathcal{D}_n$ (the *Hofbauer tower*) for intervals $\mathcal{D}_n$ defined above. Let $\hat{f} : \hat{I} \to \hat{I}$ be defined by

$$\hat{f}(\mathcal{D}_n) = \begin{cases} \mathcal{D}_{n+1} & \text{if } n \text{ is not a cutting time;} \\ \mathcal{D}_{1+S_k} \sqcup (\mathcal{D}_{1+S_{Q(k)}} \setminus \{c_1\}) & \text{if } n = S_k \text{ is a cutting time.} \end{cases}$$

Then $i \circ \hat{f} = f \circ i$, where $i : \hat{I} \to I$ is the inclusion map, and $\hat{f}$ is continuous, except at the points $c \in \mathcal{D}_{S_k}$, $k \geq 1$, which are mapped to $c_1 \in \mathcal{D}_{1+S_k}$. Since these are only countable many points, this has no effect on the Lebesgue typical behavior. The collection $\{\mathcal{D}_n\}_{n \geq 2}$ is a *Markov partition* of $\hat{I}$, because $\hat{f}$ maps each $\mathcal{D}_n$ to the union of partition elements (ignoring again the point $c_1 \in \mathcal{D}_{1+S_{Q(k)}}$). Starting in some $\mathcal{D}_n$, the subsequence intervals visited are unique determined up to the moment we reach some $\mathcal{D}_{S_k}$, where we have a choice between $\mathcal{D}_{1+S_k}$ and $\mathcal{D}_{1+S_{Q(k)}}$. Therefore it suffices to consider the transitions from interval $E_l := \mathcal{D}_{1+S_l}$ realized by the $S_{Q(l+1)}$-th image of f, see Figure 1. It follows that $\{E_l\}_{l \geq 0}$ is a Markov partition for the induced map

$$F : \sqcup_{l \geq 0} E_l \to \sqcup_{l \geq 0} E_l, \quad F|_{E_l} = \hat{f}^{S_{Q(l+1)}}.$$
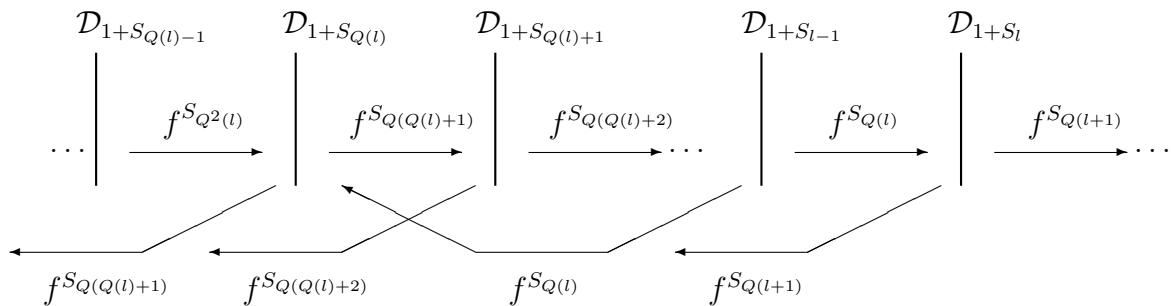
FIGURE 1. The transitions from $\mathcal{D}_{1+S_l}$, $l \geq 1$. Backward transitions go from $\mathcal{D}_{1+S_l}$ to $\mathcal{D}_{1+S_{Q(l+1)}}$.

One can show that if f is renormalizable of period S_k, then $\sqcup_{l \geq k-1} E_l$ is a trapping region for F, see part (d) of Proposition 32. Now we need to translate the results on positive drift to the current set-up. Write $\hat{\chi}_n(y) = l$ if $F^n(y) \in E_l$.

Lemma 33. *If $\{\chi_n\}_{n \geq 0}$ has positive drift as in* (6) *(and bounded second moments), then for λ-a.e. $y \in \sqcup_{l \geq 0} E_l$, we have $\hat{\chi}_n(y) \to \infty$. Furthermore, there is $\hat{k}_0 = \hat{k}_0(y)$ and $C > 0$ such that for all $\hat{k} \geq \hat{k}_0$, if $\hat{\chi}_n(y) = \hat{k}$, then $\hat{\chi}_{n+j}(y) > \hat{k}$ for all $j > C\hat{k}$.*

Proof. The positive drift of $\{\chi_n\}$ ensures that for λ-a.e. x and $k \in \mathbb{N}$, there is n_k such that $G^n(x) \in \cup_{j \geq k} U_j$ for all $n \geq n_k$. Since $U_k \subset Z_{S_k}^+(c) \cup Z_{S_k}^-(c)$, the image $f^{S_k}(U_k) \subset \mathcal{D}_{S_k}$ and $f^{1+S_k}(U_k) \subset \mathcal{D}_{1+S_k} \cup \mathcal{D}_{1+S_{Q(k)}} = E_k \cup E_{Q(k)}$. This means that $y := f^{1+S_k} \circ G^n(x) \in E_k \cup E_{Q(k)}$. If consequently $G^{n+1}(x) \in U_l$, $l > k$, then $f^{S_l}(y) \in E_l \cup E_{Q(l)}$ but the move from E_k or $E_{Q(k)}$ up to E_l or $E_{Q(l)}$ in $\sqcup_{j \geq 0} E_j$ involves passages through the intermediate states E_i as well, but "lower" states E_i, $i < k - B \leq \min_{j \geq k} Q(j)$ are avoided. Therefore $\hat{\chi}_n(f(x)) \to \infty$ for all $n \to \infty$.

For the second statement, observe that the passage from E_k or $E_{Q(k)}$ up to E_l or $E_{Q(l)}$ requires u_n iterates of F for some $l - k - B \leq u_n \leq l - k + B$. Suppose $G^m(x) \in U_k$, $k \geq k_0(x)$ as in (7), then $G^{m+j}(x) \notin U_k$ for $j \geq k$. Take $\hat{k} \in \{k, Q(k)\}$ such that $y = f^{1+S_k}(G^m(x)) \in E_{\hat{k}}$. The iterates $m+1, \ldots, m+k$ of G correspond to $\sum_{j=m+1}^{m+k} u_j$ iterates of F. Each $u_j \geq 1$, and if $\chi_{m+j}(x) < \chi_{m+j-1}(x)$, then this single iterate of G corresponds to a single iterate of F, reducing the index of the state by at most B. If $\chi_{m+j}(x) \gg \chi_{m+j-1}(x)$, then one iterate of G corresponds to many iterates of F, but if some of these iterates brings y above state $E_{\hat{k}+kB}$, then it would take more than k steps of G to return, hence this will not occur. Thus, if the $\sum_{j=m+1}^{m+\hat{k}} u_j$ iterates of F (corresponding k iterates of G) keep y close to state $E_{\hat{k}}$, then $\sum_{j=m+1}^{m+\hat{k}} u_j \leq 2Bk \leq 3B\hat{k}$, where the last inequality follows because $k - B \leq \hat{k} \leq k$. This proves the second statement for $C = 3B$. $\qquad\square$

The dynamics of wild attractors has been investigated in [BKS]. In that paper, various combinatorial types are presented for which $(\mathcal{A}, f)$ is semi-conjugate to a monothetic group (G, g), where $g : G \to G$ is an isometry for which every orbit is dense. The best known example goes back to Lyubich and Milnor [LM]; it is the Fibonacci map and its

omega-limit set $\omega(c)$ factorizes over the golden mean circle rotation. In [BKS], similar examples are shown factorising onto tori of any dimension, and even onto a solenoid. On the other hand, [BKS] gives examples for which $(\mathcal{A}, f)$ is weak mixing with respect to the unique invariant probability measure μ supported on $\mathcal{A}$. In [BHa], the simplest such example is shown to be Lebesgue exact as well.

Let $\mathfrak{fp}(x) := x - \mathrm{round}(x) \in [-\frac{1}{2}, \frac{1}{2})$ be the signed distance of x to the nearest integer.

Theorem 34. *Assume that a unimodal map has a wild attractor with positive drift. If there exists ρ such that the cutting times satisfy*

$$\sum_k k \max_{i \geq k - B} |\mathfrak{fp}(\rho S_k)| < \infty,$$

for $B = \limsup_k k - Q(k)$, then $\lambda_2(\mathrm{Dis}) = 1$.

Proof. **Step 1: Construction of the factor map.**
An *enumeration scale* is a symbolic system resembling an adding machine as in (4) based on, in this case, the sequence of cutting times. Any non-negative integer n can be written in a canonical way as a sum of cutting times: $n = \sum_j e_j S_j$, where

$$e_j := \begin{cases} 1 & \text{if } j = \max\{k; S_k \leq n - \sum_{m>k} e_m S_m\}, \\ 0 & \text{otherwise.} \end{cases}$$

In particular $e_j = 0$ if $S_j > n$. In this way we can code the non-negative integers as zero-one sequences with a finite number of ones: $n \mapsto \langle n \rangle \in \{0,1\}^{\mathbb{N}}$. Let $E_0 = \langle \mathbb{N} \cup \{0\} \rangle$ be the set of such sequences, and let E be the closure of E_0 in the product topology. This results in

$$E := \{e \in \{0,1\}^{\mathbb{N}} \; ; \; e_i = 1 \Rightarrow e_j = 0 \text{ for } Q(i+1) \leq j < i\}.$$

The condition in this set follows because if $e_i = e_{Q(i+1)} = 1$, then this should be rewritten to $e_i = e_{Q(i+1)} = 0$ and $e_{i+1} = 1$. It follows immediately that for each $e \in E$ and $j \geq 0$,

$$e_0 S_0 + e_1 S_1 + \cdots + e_j S_j < S_{j+1}. \tag{8}$$

We denote by g the standard addition of 1 by means of "add and carry", cf. (4). Let $\langle n \rangle$ be the representation of $n \in \mathbb{N} \cup \{0\}$ in the enumeration scale based on $\{S_k\}_{k \geq 0}$. Obviously $g(\langle n \rangle) = \langle n+1 \rangle$. Under the condition that $Q(k) \to \infty$, $g : E \to E$ is continuous, and is invertible on $E \setminus \{\langle 0 \rangle\}$, see [BKS, GLT]. Since $\sum_k |\mathfrak{fp}(\rho S_k)| < \infty$, we can define a continuous projection $\pi_\rho : E \to \mathbb{S}^1$ by

$$\pi_\rho(e) = \sum_k e_k \, \mathfrak{fp}(\rho S_k) \bmod 1,$$

and $\pi_\rho \circ g = R_\rho \circ \pi_\rho$ for the circle rotation $R_\rho : x \mapsto x + \rho \bmod 1$. At the same time, the map $P : E \to \mathcal{A}$ defined as the continuous extension of $P(\langle n \rangle) = f^n(c)$ satisfies $P \circ g = f \circ P$. We know from [BKS] that there is semi-conjugacy $\pi = \pi_\rho \circ P^{-1} : \mathcal{A} \to \mathbb{S}^1$ such that $\pi \circ f = R_\rho \circ \pi$, provided $\sum_k |\mathfrak{fp}(\rho S_k)| < \infty$. A more direct way to construct $\pi : \mathcal{A} \to \mathbb{S}^1$ is by setting

$$\pi(x) = \begin{cases} \rho n \bmod 1 & \text{if } x = c_n, \\ \lim_{j \to \infty} \rho n_j \bmod 1 & \text{if } x \in \mathcal{A} \setminus \mathrm{orb}(c) \text{ and } (n_j)_{j \in \mathbb{N}} \text{ is such that } x \in \cap_j \mathcal{D}_{n_j}. \end{cases}$$

$$(E, g)$$

$$P \swarrow \qquad \searrow \pi_\rho$$

$$(\mathcal{A}, f) \quad \xrightarrow{\pi = \pi_\rho \circ P^{-1}} \quad (\mathbb{S}^1, R_\rho)$$

$$\cap \qquad\qquad \tilde{\pi} \nearrow$$

$$(\mathrm{Bas}(\mathcal{A}), f)$$

In [BHa] it was shown how to extend the map $\pi_\rho \circ P^{-1}$ to a measurable factor map $\tilde{\pi} : \mathrm{Bas}(\mathcal{A}) \to \mathbb{S}^1$. Here we will give a construction of $\tilde{\pi}$ which is more closely connected to [BKS].

For any x with $\hat{\chi}_n(x) \to \infty$, the number $b_n(x)$ defined as

$$b_n(x) := \max_j \{ j : f^j(Z_j(x)) = \mathcal{D}_n \}$$

exists. If $x \in f^{-k}(c)$ for some $k \geq 0$, then we need to write $Z_n^\pm(x)$ for $n > k$, because x is the common boundary point of two cylinder sets, and $f^n(x) = c_{n-k} \in \mathcal{D}_{n-k}$ for n sufficiently large. So $b_n(x)$ is well-defined in this case too. Set

$$\tilde{\pi}_n(x) := -\sum_k \mathfrak{fp}(\rho \langle b_n(x) - n \rangle_k S_k) \bmod 1 = -\rho(b_n(x) - n) \bmod 1. \qquad (9)$$

If $n + 1$ is not a cutting time, then $b_{n+1}(x) = b_n(x) + 1$; in this case $\tilde{\pi}_{n+1}(x) = \tilde{\pi}_n(x)$. If $n + 1 = S_k$ is a cutting time, $f^{b_n(x)}(Z_{b_n(x)}(x)) = \mathcal{D}_{S_k - 1}$ and $f^{b_n(x)+1}(Z_{b_n(x)}(x)) = \mathcal{D}_{S_k}$, but $b_{n+1}(x)$ can be strictly larger than $b_n(x) + 1$. In this case, however,

$$b_{n+1}(x) = b_n(x) + 1 + \sum_{j \geq k} d_j S_{Q(j)}$$

for some non-negative integers d_j. Recall from Lemma 33 that for λ-a.e. $x \in \mathrm{Bas}(\mathcal{A})$, there is $\hat{k}_0 = \hat{k}_0(x)$ such that for all $k > \hat{k}_0$

$$\text{if } F^m(x) \in E_k, \text{ then } F^{m+j}(x) \notin E_k \text{ for all } j > Ck. \qquad (10)$$

This means that $\sum_{j \geq k} d_j \leq k$, and therefore $\langle b_n(x) - n \rangle$ and $\langle b_{n+1}(x) - (n+1) \rangle$ are sequences which differ by at most Ck entries and the indices of these entries are $\geq k - B$. This means that by the definition of (9)

$$|\tilde{\pi}_n(x) - \tilde{\pi}_{n+1}(x)| \leq Ck \max_{i \geq k - B} |\mathfrak{fp}(\rho S_i)|, \qquad (11)$$

which is summable over k by assumption. It follows that $\{\tilde{\pi}_n(x)\}_n$ is a Cauchy sequence in $\mathbb{S}^1$ for λ-a.e. $x \in \mathrm{Bas}(\mathcal{A})$, and hence $\tilde{\pi}(x) := \lim_{n \to \infty} \tilde{\pi}_n(x)$ exists.

Let us complete Step 1 by showing that $\tilde{\pi} \circ f = R_\rho \circ \tilde{\pi}$ for λ-a.e. $x \in \mathrm{Bas}(\mathcal{A})$. Assume that $x \notin \cup_j f^{-j}(c)$, then $f(Z_j(x)) = Z_{j-1}(f(x))$ for all sufficiently large j. Therefore $b_n(f(x)) = b_n(x) - 1$, so substituting into (9) gives $\tilde{\pi}_n(f(x)) = \tilde{\pi}_n(x) + \rho$ for each n. In the limit, $\tilde{\pi} \circ f = R_\rho \circ \tilde{\pi}$.

Remark 35. *It can be shown that $\tilde{\pi}$ is well-defined on $\mathcal{A}$ and coincides with π, but since it plays no role in Theorem 34, we will omit the proof.*

Step 2: The measure of Dis.
We will show that if $\tilde\pi(x) \neq \tilde\pi(y)$, then (x,y) form a distal pair. This is more involved than in Proposition 38 below, because $\tilde\pi|_{\mathrm{Bas}(\mathcal{A})}$ is not continuous. With the exception of a set of measure zero, we can assume that x and y satisfy (10); let $k_1 = \max\{\hat{k}_0(x), \hat{k}_0(y)\}$. Suppose that $\tilde\pi(x) \neq \tilde\pi(y)$ and take $N \in \mathbb{N}$ and $\eta > 0$ such that $|\tilde\pi_n(x) - \tilde\pi_n(y)| > 2\eta$ for all $n \geq N$. Take $k_2 \geq k_1$ so large that $Ck \max_{i \geq k-B} |\mathfrak{fp}(\rho S_i)| < \eta$ for all $k \geq k_2$ and C as in Lemma 33.

We know that $\mathcal{A} \subset \cup_{n \geq S_{k_2}} \mathcal{D}_n$, but by Proposition 32, part (c), $\omega(c) \subset \cup_{n=1+S_{k_2}}^{S_{k_2}+B} \mathcal{D}_n$. Take $\varepsilon > 0$ so small that every two intervals $\mathcal{D}_n, \mathcal{D}_{n'}$, $S_{k_2} < n, n' \leq S_{k_2+B}$ either are at least ε apart or their intersection has length at most ε.

For λ-a.e. $x \in \mathrm{Bas}(\mathcal{A})$, we have that $\hat\chi_n(x) \to \infty$. Hence $f^i(x) \in \mathcal{D}$ for i sufficiently large, even though $\mathcal{D}$ only contains a "one-sided" open neighbourhood of $\omega(c)$. (In fact it is possible to prove that the same statement is true for all $x \in \mathrm{Bas}(\mathcal{A})$, but this is enough for our purposes.)

Suppose now by contradiction that (x,y) is not distal. Then there is $i \geq N$ such that $f^i(x), f^i(y) \in \mathcal{D}$ and $|f^i(x) - f^i(y)| < \varepsilon$. So $f^i(x)$ and $f^i(y)$ belong to the same interval $\mathcal{D}_n$ for some $n \geq S_{k_2}$, and taking i larger if necessary, we can assume that n is a cutting time. By (10), $\langle b_n(x) - i \rangle$ and $\langle b_n(y) - i \rangle$ are sequences which differ by at most Ck entries and the indices of these entries are $\geq k - B$. Similar to (11), we have

$$|\tilde\pi_n(x) - \tilde\pi_n(y)| \leq Ck \max\{|\mathfrak{fp}(\rho S_i)| : k - B \leq i \leq 2k - B\} < \eta.$$

This contradiction to the choice of η and N proves that (x,y) is distal. Therefore proximal pairs (x,y) can only exist within fibers of $\tilde\pi$.

Finally, if $W = \tilde\pi^{-1}(s)$ with $\lambda(W) > 0$ for some $s \in \mathbb{S}^1$, then since $\tilde\pi \circ f = R_\rho \circ \tilde\pi$ and R_ρ is invertible, it follows that $f^m(W) \cap f^n(W) = \emptyset$ for all $0 \leq m < n$, and this contradicts the non-existence of strongly wandering sets. Therefore each fiber has measure zero. This completes the proof. $\qquad\square$

Corollary 36. *If $S_k = S_{k-1} + S_{k-d}$ for $d = 2, 3, 4$, and f is a map with cutting times $\{S_k\}_{k \geq 0}$ and sufficiently large critical order so that $\mathcal{A}$ is a wild attractor with positive drift, then $\lambda_2(\mathrm{Dis}) = 1$.*

Proof. We know from [BKS] that the dynamics $(\mathcal{A}, f)$ is semi-conjugate to a minimal rotation on a $d-1$-dimensional torus. If fact, the characteristic equation $\lambda^d = \lambda^{d-1} + 1$ of the recursive relation $S_k = S_{k-1} + S_{k-d}$ has a leading root ρ which is a Pisot number, *i.e.*, all its algebraic conjugates lie within the unit disk. It follows easily that $\mathfrak{fp}(\rho S_k)$ is exponentially small in k so that the condition $\sum_k k |\mathfrak{fp}(\rho S_k)| < \infty$ is obviously satisfied, and Theorem 34 shows that $\lambda_2(\mathrm{Dis}) = 1$. By defining $\Pi : \mathrm{Bas}(\mathcal{A}) \to \mathbb{T}^{d-1}$ as $\Pi(x) = (\tilde\pi_\rho(x), \ldots, \tilde\pi_{\rho^{d-1}}(x))$ (which is well-defined λ-a.e.), we obtain a factor map onto $(\mathbb{T}^{d-1}, R_{\rho,\ldots,\rho^{d-1}})$ with Haar measure, which is the maximal automorphic factor of $(\mathrm{Bas}(\mathcal{A}), \lambda, f)$. $\qquad\square$

Proposition 37. *Under the conditions of Theorem 34 and if there is a continuous factor of $(\mathcal{A}, f)$ onto a d-dimensional torus $\mathbb{T}^d$ (with $d \geq 1$), then the fiber $\tilde\pi^{-1}(\tau) \subset \mathrm{Bas}(\mathcal{A})$ for each $\tau \in \mathbb{T}^d$ contains an uncountable ε-scrambled set for some $\varepsilon > 0$.*

Proof. Assume for simplicity that $d = 1$ and take $\rho \in \mathbb{R}$ such that $\tilde{\pi} \circ f = \tilde{\pi} + \rho \mod 1$ for λ-e.a. $x \in \mathcal{A}$. Because a zero-dimension set cannot be mapped injectively onto a set of higher dimension, see [E], the continuity of the factor map $\pi : \mathcal{A} \to \mathbb{T}^d$ implies that π cannot be injective. Therefore we can find $a \neq \hat{a} \in \mathcal{A}$ such that $f(a) = f(\hat{a})$. By Proposition 32, part (c), we can find, for every k, integers κ and $\hat{\kappa}$ with $k < \kappa, \hat{\kappa} \leq k + B$ and $N \leq \min(S_\kappa, S_{\hat{\kappa}})$, such that $\mathcal{D}_{S_\kappa - N} \ni a$ and $\mathcal{D}_{S_{\hat{\kappa}} - N} \ni \hat{a}$.

Hence we can find two sequences $(\kappa_j)_{j \in \mathbb{N}}$ and $(\hat{\kappa}_j)_{j \in \mathbb{N}}$ with $|\kappa_j - \hat{\kappa}_j| \leq B$, but possibly $\kappa_{j+1} \gg \kappa_j$, and another sequence $(N_j)_{j \in \mathbb{N}}$ such that $\mathcal{D}_{S_{\kappa_j} - N_j} \ni a$ and $\mathcal{D}_{S_{\hat{\kappa}_j} - N_j} \ni \hat{a}$.

Recall the Markov map was defined as $F : \sqcup_{l \geq 0} E_l \to \sqcup_{l \geq 0} E_l$. For each j, we will create loops from E_{κ_j} to itself and from $E_{\hat{\kappa}_j}$ to itself, as indicated in Figure 2. Both loops require the same steps under F, only arranged in a different order, hence they involve the same number s_j of iterates of f. Because $\limsup_k k - Q(k) \leq B$, both loops involve no more than $2B$ steps, and the width (highest vertex minus smallest vertex) is less than $2B$ as well.
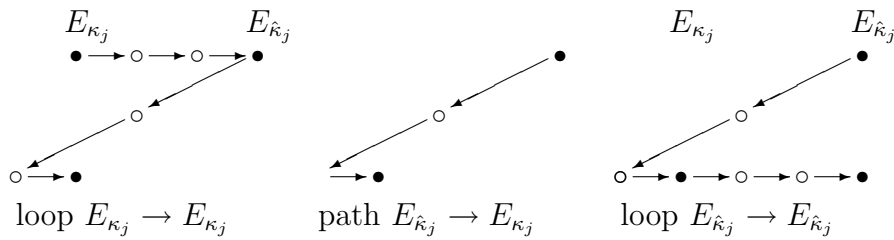

FIGURE 2. Different loops ending at E_{κ_j} and $E_{\hat{\kappa}_j}$ respectively. Both loops contain the same path from $E_{\hat{\kappa}_j}$ to E_{κ_j}, depicted in the middle. Backward arrows go from E_l to $E_{Q(l+1)}$.

Assume that $\kappa_j < \hat{\kappa}_j$ and $E_{\kappa_j} \supset E_{\hat{\kappa}_j}$. (The other three cases can be treated similarly.) Since $\{E_l\}_{l \geq 0}$ is a Markov partition for F, there are intervals $J_j \subset E_{\kappa_j}$ and $\hat{J}_j \subset E_{\hat{\kappa}_j}$ such that $f^{s_j} : J_j \to E_{\kappa_j}$ and $f^{s_j} : \hat{J}_j \to E_{\hat{\kappa}_j}$ are diffeomorphic and onto. There is a similar interval $\hat{K}_j \subset E_{\hat{\kappa}_j}$ representing the path from $E_{\hat{\kappa}_j}$ to E_{κ_j}, that is if the path from $E_{\hat{\kappa}_j}$ to E_{κ_j} requires t_j iterates of f, then $f^{t_j} : \hat{K}_j \to E_{\kappa_j}$ is diffeomorphic and onto.

Combining the two, we find intervals $H_j \subset J_j$ such that

$$
\begin{cases}
(a) & f^{s_j}(H_j) = \hat{K}_j \subset E_{\hat{\kappa}_j} \subset E_{\kappa_j}, \\
(b) & f^{s_j + t_j}(H_j) = E_{\kappa_j}, \text{ and} \\
(c) & f^{s_j - (N_j + 1)}(H_j) \text{ contains, or is close to, } a.
\end{cases}
$$

Similarly, we can find $\hat{H}_j \subset \hat{J}_j$ such that

$$
\begin{cases}
(\hat{a}) & f^{s_j}(\hat{H}_j) = \hat{K}_j \subset E_{\hat{\kappa}_j}, \\
(\hat{b}) & f^{s_j + t_j}(\hat{H}_j) = E_{\kappa_j}, \text{ and} \\
(\hat{c}) & f^{s_j - (N_j + 1)}(\hat{H}_j) \text{ contains, or is close to, } \hat{a}.
\end{cases}
$$

Let $\Sigma \subset \{0,1\}^{\mathbb{N}}$ be an uncountable scrambled subset of the full shift. The idea is now, for each $\tau \in \mathbb{T}$ and each $\sigma \in \Sigma$, to find a point $x \in \pi^{-1}(\tau)$ such that

$$f^{r_j}(x) \in \begin{cases} \mathcal{D}_{S_{\kappa_j}-N_j} \ni a & \text{if } \sigma_j = 0, \\ \mathcal{D}_{S_{\hat{\kappa}_j}-N_j} \ni \hat{a} & \text{if } \sigma_j = 1, \end{cases} \tag{12}$$

where the sequence r_j depends of t but not on σ.

Start with some $y \in \mathcal{D}_2 = E_0$ with $\tilde{\pi}(y) = \tau + \varepsilon$ where ε will be determined later. Then, when the orbit of y under iteration of F goes from E_{κ_j} to $E_{\hat{\kappa}_j}$, we insert one of the loops as in Figure 2 according to whether $\sigma_j = 0$ or 1, and the extra path from E_{κ_j} to $E_{\hat{\kappa}_j}$. That is, when q_1 is such that $f^{q_1}(y) \in E_{\kappa_j}$, we insert one of the extended loops, both taking $s_1 + t_1$ iterates, and iterate $r_1 := q_1 + s_1 - (N_1 + 1)$ brings the path close to a or $\hat{a}$, depending on whether $\sigma_1 = 0$ or 1. Then, when y visit E_{κ_2}, the new extended loop takes $s_1 + t_1$ iterates more to reach it; call this number q_2, insert the appropriate extended loop of $s_2 + t_2$ iterates, and find that after $r_2 = q_2 + s_2 - (N_2 + 1)$ iterates, the path will be close to a or $\hat{a}$, etc.

Due to the Markov property, there is some $x \in E_0$ whose infinite path under F is precisely the path we have created, so x satisfies (12). Furthermore, the values $b_n(x) = b_n(y) + \sum_{j,q_j+B<n} s_j + t_j$. When we compute $\tilde{\pi}(x)$, the contribution ε of all the inserted extended loops bounded by

$$\sum_j 3B \max_{\kappa_j-B \le i \le \kappa_j+B} |\mathfrak{fp}(\rho S_i)| < \infty.$$

Here we used that each loop requires at most $2B$ steps and each path from E_{κ_j} to $E_{\hat{\kappa}_j}$ (or vice versa) has at most B steps, and that all these steps are taken within a distance B from the step $E_{\kappa_j} \to E_{\kappa_j+1}$.

In particular, $\tilde{\pi}(x)$ is well-defined, and $\tilde{\pi}(x) = \tilde{\pi}(y) - (\tilde{\pi}(y) - \pi(x)) = \tilde{\pi}(y) - \varepsilon = \tau$, so $x \in \tilde{\pi}^{-1}(\tau)$.

Finally, if $\sigma \ne \sigma'$ both belong to Σ, there are infinitely many js such that $\sigma_j \ne \sigma'_j$, and consequently $\limsup_{j\to\infty} |f^{r_j}(x) - f^{r_j}(x')| = |a - \hat{a}| > 0$. On the other hand, $\lim_{j\to\infty} |f^{q_j}(x) - f^{q_j}(x')| = 0$, because $f^{q_j}(x)$ and $f^{q_j}(x')$ both belong to E_{κ_j} or to $E_{\hat{\kappa}_j}$ Therefore $\{x(\sigma) : \sigma \in \Sigma\}$ is an $|a - \hat{a}|$-scrambled subset of $\pi^{-1}(\tau)$, as required. □

Proposition 38. *Let f be a unimodal map with kneading map $Q(k) = \max\{k - d, 0\}$ for $d = 2, 3, 4$, and let μ denote the unique invariant probability measure supported by $\omega(c)$. Then*

- *μ_2-a.e. pair of points is distal;*
- *if $d = 2$ (the Fibonacci map) then $\omega(c)$ contains no Li-Yorke pair, and the only asymptotic pairs (x, y) are such that $f^n(x) = f^n(y) = c$ (only one such pair for each $n \ge 1$);*
- *if $d = 3, 4$ then there are uncountably many Li-Yorke pairs in $\omega(c)$.*

Proof. As in Corollary 36, there is a continuous map $\pi : \omega(c) \to \mathbb{T}^{d-1}$ onto the $d - 1$-dimensional torus and an irrational rotation $R := R_{\rho,\dots,\rho^d} : \mathbb{T}^d \to \mathbb{T}^d$ such that $\pi \circ f = R \circ \pi$. If $\pi(x) \ne \pi(y)$ then (x, y) is distal, as in Corollary 36. It follows from [BKS] that this happens for μ_2-a.e. (x, y).

Furthermore, $\pi^{-1}(b)$ consists of at most d points $a_1, \ldots, a_d$ for any $b \in \mathbb{T}^{d-1}$. If there are indeed d distinct points, then there is $n \geq 0$ such that $f^n(a_1) = f^n(a_2) = \cdots = f^n(a_d) = c$, cf. [Br4]. For the Fibonacci map, this accounts for all non-distal pairs. For $d = 3, 4$, other non-singleton fibers $\pi^{-1}(b)$ are possible, and Li-Yorke pairs exists within such fibers. They are related to incidences in the substitution shift description, as described implicitly in [BD]. $\square$

The question is whether the situation is the same for the "next" Fibonacci-like map with $Q(k) = \max\{k - 5, 0\}$. In this case, the system of $(\omega(c), f)$ with its unique probability measure is weak mixing, so there is no continuous (or even measurable) factor map onto a group rotation. The difference with the previous cases is that the characteristic equation of the recursive relation

$$0 = \lambda^5 - \lambda^4 - 1 = (\lambda^2 - \lambda + 1)(\lambda^3 - \lambda - 1)$$

is reducible, and more decisively, its leading root is not a Pisot number. The following curiosity about the cutting times holds in this case:

$$S_k = S_{k-2} + S_{k-3} + \begin{cases} +1 & \text{if } k \equiv 2, 3 \bmod 6; \\ -1 & \text{if } k \equiv 5, 0 \bmod 6; \\ 0 & \text{if } k \equiv 1, 4 \bmod 6. \end{cases} \tag{13}$$

Note that the same algebraic curiosity holds for any characteristic equation $\lambda^{6m-1} - \lambda^{6m-2} - 1 = 0$, because in each such case, $\lambda^2 - \lambda + 1$ (with solutions $\lambda = \frac{1 \pm i\sqrt{3}}{2}$ on the unit circle) divides the equation. As an example, the case $m = 2$ gives:

$$\lambda^{11} - \lambda^{10} - 1 = (\lambda^2 - \lambda + 1)(\lambda^9 - \lambda^7 - \lambda^6 + \lambda^4 + \lambda^3 - \lambda - 1),$$

and one can indeed check that

$$S_k = S_{k-2} + S_{k-3} - S_{k-5} - S_{k-6} + S_{k-8} + S_{k-9} + \begin{cases} +1 & \text{if } k \equiv 2, 3 \bmod 6; \\ -1 & \text{if } k \equiv 5, 0 \bmod 6; \\ 0 & \text{if } k \equiv 1, 4 \bmod 6. \end{cases}$$

Proposition 39. *If $S_k = S_{k-1} + S_{Q(k)}$ for $Q(k) = \max\{0, k - 5\}$ and f is a map with cutting times $\{S_k\}_{k \geq 0}$ and a wild attractor with positive drift, then $\lambda_2(\mathrm{LY}_\varepsilon) = 1$ for some $\varepsilon > 0$.*

Proof. As pointed out before, it was shown in [Br3] that $\mathcal{A} = \omega(c)$ is a wild attractor, and from [BHa] it follows that dynamics on the basin of the attractor is Lebesgue exact. Thus Proposition 27 implies that $\lambda_2(\mathrm{LY}) = 1$. In fact, the construction of Proposition 37 can also be used here to show that there is $\varepsilon > 0$ such that λ_2-a.e. pair belongs to LY_ε. $\square$

Example 40. *Let f be a unimodal map with cutting times satisfying*

$$S_0 = 1, \ S_1 = 2, \ S_2 = 3, \ S_3 = 4, \ S_4 = 6, \ S_5 = 8, \ S_6 = 10, \ S_7 = 12$$

and

$$S_k = S_{k-1} + S_{k-5} \text{ for } k \geq 8.$$

This means that the cutting times S_k are even for $k \geq 3$ and eventually are twice the numbers occurring in the example of Proposition 39. Assume also that the critical order of ℓ is so large that f has a wild attractor $\mathcal{A}$. Then $\mathcal{A}$ decomposes into two disjoint Cantor

sets $\mathcal{A}_0$ and $\mathcal{A}_1$ which are permuted by f. Note, however, that f is not renormalizable, see Proposition 32, and therefore f is topologically mixing on $[c_2, c_1]$.

Let B_0 and B_1 be disjoint neighborhoods of $\mathcal{A}_0$ and $\mathcal{A}_1$; for example we can take $B_0 = [c_2, c_{14}] \cup [c_4, c_6]$ and $B_1 = [c_3, c_{15}] \cup [c_5, c_1]$. Every point in the basin of $\mathcal{A}$ will eventually be trapped in $B_0 \cup B_1$. But every pair (x, y) with $x \in B_0$ and $y \in B_1$ such that $\mathrm{orb}(x), \mathrm{orb}(y) \subset B_0 \cup B_1$ is distal. On the other hand $f^2|_{\mathcal{A}_0}$ and $f^2|_{\mathcal{A}_1}$ behave like the example of Proposition 39, so λ_2-a.e. every pair $(x, y) \in B_0 \times B_0$ (or $(x, y) \in B_1 \times B_1$) such that $\mathrm{orb}(x), \mathrm{orb}(y) \subset B_0 \cup B_1$ is Li-Yorke.

Proof of Theorem E. From Theorem C we know that $\lambda_2(\mathrm{Asymp}) = 0$. Also, if some $x \in \mathrm{Bas}(\mathcal{A})$ is approximately periodic, then by Proposition 31, $\mathcal{A}$ is conjugate to an adding machine, so all points in $\mathrm{Bas}(\mathcal{A})$ are approximately periodic. By Proposition 6, $\mathrm{Bas}(\mathcal{A})$ contains no Li-Yorke pairs. Therefore (a)-(d) are the only possibilities, and they all occur:
(a) The strange adding machine case as wild attractor, see [Br4].
(b) The Fibonacci-like map with kneading map $Q(k) = \max\{k - d, 0\}$, $d = 2, 3, 4$, see Theorem 34 and Corollary 36.
(c) The Fibonacci-like map with kneading map $Q(k) = \max\{k - 5, 0\}$, see Proposition 39.
(d) See Example 40.
Proposition 37 implies the existence of ε-scrambled sets in the fibers $\tilde{\pi}^{-1}(\tau)$ for all $\tau \in \mathbb{S}^1$ and factor maps $\tilde{\pi}$ in case (b), and for cases (c) and (d), the ε-scrambled set is also immediate. Li-Yorke sensitivity follows as well. □

Remark 41. *In case (a) and (b) of Theorem E, $(\mathcal{A}, \mu_{\mathcal{A}}, f)$ is not weakly mixing. Instead, there is an f_2-invariant set of positive μ_2-measure that is bounded away from the diagonal of $\mathcal{A}_2$. We expect that case (c) always corresponds to the weakly mixing case, cf. Proposition 39, and in particular, $\mu_{\mathcal{A}} \times \mu_{\mathcal{A}}$-a.e. pair (x, y) has a dense orbit in $\mathcal{A} \times \mathcal{A}$, and hence is Li-Yorke.*

Remark 42. *In case (a), points x in the basin of $\mathcal{A}$ have a distinct target point $t_x \in \mathcal{A}$ such that $\mathrm{dist}(f^n(x), f^n(t_x)) \to 0$, see Proposition 6. In case (b), such target points $t_x \in \mathcal{A}$ do not exist in general, cf. the proof of Proposition 37 and also [BHa, Remark 2].*

7. Appendix

In this appendix, we prove the more technical results of the paper, divided into four parts. First we give a result on the structure of solenoidal sets of Feigenbaum type, next we prove Theorem 20, then we formulate an improved C^3 Koebe distortion lemma, and finally we prove Theorem 21.

7.1. **Neigborhoods of Solenoidal Sets.** Assume that S is a solenoidal set and $\mathrm{cyc}(K)$ is a solenoidal cycle of period r containing S. Let $S_i = S \cap f^i(K)$, $0 \leq i < r$. Since the critical points in $\mathrm{cyc}(K)$ belong to S, the convex hulls J_i of S_i form a cycle of intervals, that is, $f(J_i) = J_{i+1}$ for every i (with this we also mean $f(J_{r-1}) = J_0$). We call it the *r-minimal solenoidal cycle covering S.* We emphasize that the intervals J_i are pairwise disjoint (because a solenoid contains no periodic points). Moreover, if they are ordered in

the real line as $J_{i_1} < J_{i_2} < \cdots < J_{i_r}$, then there is a union of periodic orbits $P = \{p_i\}_{i=1}^{r-1}$ such that $J_{i_1} < p_1 < J_{i_2} < p_2 < \cdots < p_{r-1} < J_{i_r}$ (see [MT] or [AJS]).

The existence of these periodic orbits of smaller period interlaced among the intervals J_i allows us to prove immediately that the union L_i of all periodic intervals of period r containing J_i is also periodic of period r. Clearly, $f(L_i) \subset L_{i+1}$ and $f(\partial L_i) \subset \partial L_{i+1}$ for every i. However, although the intervals L_i have pairwise disjoint interiors, they do not form a cycle of intervals because $f(L_i) = L_{i+1}$ need not hold. In fact $(L_i)_{i=0}^r$ is the pullback chain of $L_r := L_0$ along $J_0, J_1, \ldots, J_r = J_0$. Let $M_i = \operatorname{Int} L_i$. We call $M = \bigcup_{i=0}^{r-1} M_i$ the r-shell covering S. Then M is nice and $J_i \subset M_i$ for every i.

Proposition 43. *Let $f \in C^2_{\mathrm{nf}}(I)$ and let S be a solenoidal set of Feigenbaum type. Let T be the r-minimal solenoidal cycle covering S for some r, $T = \bigcup_{i=0}^{r-1} f^i(J)$, and let $M = \bigcup_{i=0}^{r-1} M_i$ be the r-shell covering S, with $f^i(J) \subset M_i$ for every $0 \le i < r$. Then there is $\xi = \xi(f) > 0$ such that $f^i(J)$ is ξ-well centered in M_i for every i.*

Proof. Since S is a solenoid, there is a turning point $c \in S$. There is no loss of generality in assuming that $c \in J$. Also, we can assume that r is large enough so that each interval M_i contains at most one critical point of f and there are no critical points in $\operatorname{Cl} M$ outside T. Hence $(M_i)_{i=0}^r$ is the pullback chain of $M_r := M_0$ along $J, f(J), \ldots, f^r(J) = J$. Let $\operatorname{cyc}(R)$ be the $2r$-minimal solenoidal cycle covering S, with $c \in R$. We can for example assume that $[u, v] = R$, $f^r(R) = [w, z]$ and $c \in R$ is a local maximum, Note that J is the convex hull of $R \cup f^r(R)$. Observe that there is a periodic point p of period r such that $f^i(p)$ lies between $f^i(R)$ and $f^{i+r}(R)$ for every i. Since T is a solenoidal cycle, $\operatorname{orb}(p)$ is hyperbolic repelling. In fact, since f is monotone on each on the intervals connecting $f^i(R)$ and $f^{i+r}(R)$, it is the only r-periodic orbit in T and $f^r([u, p)) = (p, z]$, $f^r((p, z]) = [u, p)$. Then $(\tau_c(p), p)$ is clearly the only nice periodic interval of period $2r$ containing c.
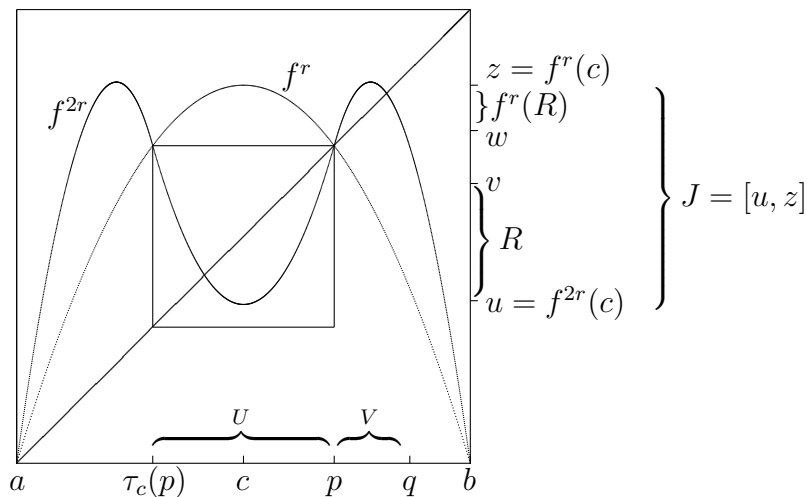

FIGURE 3. The graph of $f^r : M_0 = (a, b) \to M_0$ with J and R and relevant points

By [SV, Theorem A'(1)] there are an integer $s > 0$, a number $\xi_0 = \xi_0(f) > 0$ and nice periodic intervals $N^m \ni c$ of periods $s2^m$, $m \ge 0$, such that N^{m+1} is ξ_0-well inside N^m

for every m. We have shown in the previous paragraph that if $s2^m$ is large enough, then there is just one nice periodic interval of period $s2^m$ containing c. Thus we can assume that $M_0 = N^m$ and $(\tau_c(p), p) = N^{m+1}$ for some m, with $r = s2^m$.

Let $q > p$ be the closest point to p satisfying $f^r(q) = \tau_c(p)$. Then $V = (p, q)$ is just the pullback of $U = (\tau_c(p), p)$ along $f^r(R), f^{r+1}(R), \ldots, f^{2r}(R) = R$. Since U is ξ_0-well inside M_0, there is $\xi_1 = \xi_1(f)$ such that V is ξ_1-well inside M_0, due to Corollary 14.

Let $W = (q, b]$ with b the right endpoint of M_0. Then $\xi_0|U| \leq |V| + |W|$ and $\xi_1|V| \leq |W|$, hence

$$\frac{\xi_0 \xi_1}{1 + \xi_0 + \xi_1}(|U| + |V|) \leq |W|.$$

Then $(\tau_c(q), q)$ is well centered in M_0 and so is any subinterval of $(\tau_c(q), q)$. In particular, J is well centered in M_0.

To finish the proof we must show that *every* interval $f^i(J)$ is well centered in the corresponding interval M_i. Note that we cannot directly use Corollary 14 because it only guarantees that $f^i(J)$ is well inside M_i. Instead we proceed as follows. According to [SV, Lemma 2], $M_r := M_0$ is well inside an interval G_r which contains at most $e = 2^{d+1} + 3$ of the intervals $f^i(M_0)$, $0 \leq i < r$, with $d = \# \text{Crit}$. Therefore it also contains at most e of the intervals M_i. Now [SV, Lemma 3] implies that the pullback chain $(G_i)_{i=0}^r$ of G_r along $M_0, \ldots, f^r(M_0)$ has order bounded by $2(e + d(e + 2)) + 1$. Hence, by Corollary 14, M_i is well inside G_i for every i and, additionally, if some iterate f^l maps diffeomorphically M_i onto M_{i+l}, then this diffeomorphism has bounded distortion. Thus, if $f^{i+l}(J)$ is well centered in M_{i+l}, $f^i(J)$ is well centered in M_i. Using now Lemma 11 and recalling that $J = f^r(J)$ is well centered in $M_0 = M_r$, we conclude that every interval $f^i(J)$ is well centered in M_i as we set out to show. $\qquad\square$

7.2. **Proof of Theorem 20.** Later on we will apply the lemma below to the set $Q = \text{Crit}$ of critical points of our map $f \in C^2_{\text{nf}}(I)$; recall that we are assuming that f has no periodic critical points, see (2).

Lemma 44. *Let $f : I \to I$ be a multimodal map without wandering intervals and let $Q \subset I \setminus \partial I$ be a finite set containing no periodic points. Let $x \in Q$. Then there is an arbitrarily small nice interval $J \ni x$ such that $\text{orb}(Q) \cap \partial J = \emptyset$ and $\text{dist}(\text{orb}(\partial J), Q) > 0$.*

Proof. Let $Q' = \bigcup_{n=-\infty}^\infty f^n(Q)$. We claim that the set $P = \text{AsPer} \setminus Q'$ is dense in I. This implies the lemma. Indeed, let $\varepsilon > 0$ be small. Since x is not periodic, there is no loss of generality in assuming that $\text{orb}(Q) \cap (x - \varepsilon, x + \varepsilon)$ does not contain any periodic point. Take points $\hat{a} \in (x - \varepsilon, x) \cap P$, $\hat{b} \in (x, x + \varepsilon) \cap P$, and let $a < x < b$ be the points from $\text{Cl}(\text{orb}(\hat{a}) \cup \text{orb}(\hat{b}))$ closest to x from both sides. We emphasize that both $\hat{a}$ and $\hat{b}$ are asymptotically periodic and x is not periodic, so a and b are well defined. Moreover, (a, b) is nice. Also, $a \notin Q'$. If $a \in \text{orb}(\hat{a}) \cup \text{orb}(\hat{b})$ this is obvious because $\hat{a}, \hat{b} \notin Q'$. If $a \notin \text{orb}(\hat{a}) \cup \text{orb}(\hat{b})$, then a must belong to a periodic orbit attracting either $\text{orb}(\hat{a})$ or $\text{orb}(\hat{b})$, and again $a \notin Q'$ because Q contains no periodic points and neither does $\text{orb}(Q) \cap (x - \varepsilon, x + \varepsilon)$. Similarly, $b \notin Q'$. We have shown that $\text{orb}(Q) \cap \{a, b\} = \emptyset$ and $(\text{orb}(a) \cup \text{orb}(b)) \cap Q = \emptyset$. Since a and b are asymptotically periodic and Q

contains no periodic points, the property $(\mathrm{orb}(a) \cup \mathrm{orb}(b)) \cap Q = \emptyset$ implies in fact that $\mathrm{dist}(\mathrm{orb}(a) \cup \mathrm{orb}(b), Q) > 0$. Thus $J = (a, b)$ is the small nice interval we are looking for.

We prove that every interval K intersects P. If the intervals $\{f^n(K)\}_n$ are pairwise disjoint, then the absence of wandering intervals for f implies that these intervals are attracted by a periodic orbit. Since f is multimodal, the set of points in $K \cap Q'$ is countable. Thus K intersects P.

Now assume that $f^n(K) \cap f^m(K) \neq \emptyset$ for some $n < m$. Let $k = m - n$. Using again that f is multimodal, we get that $T = \mathrm{Cl}(\bigcup_{r=0}^{\infty} f^{n+rk}(K))$ is a nondegenerate interval. Moreover, it is invariant for f^k. If $f^k|_T$ has finitely many periodic points, then all points from T are asymptotically (or eventually) periodic, see *e.g.* [BC, p. 127], so all points from K are asymptotically periodic as well and $K \cap P \neq \emptyset$ as before. If $f^k|_T$ has infinitely many periodic points, there is a family of disjoint periodic orbits $\{\mathcal{O}_j\}_{j=1}^{\infty}$ such that $\mathrm{orb}(K) \cap \mathcal{O}_j \neq \emptyset$ for every j. If j is large enough, then $\mathrm{orb}(Q) \cap \mathcal{O}_j = \emptyset$. Let $y \in K$ be a preimage of such $\mathcal{O}_j$. Then $y \notin Q'$, which finishes the proof. $\qquad\square$

The next proposition strengthens [CL, Proposition 5]. In what follows, we say that a nice set V is ξ-*nice* if all return domains to V are ξ-well inside the components of V containing them. Sometimes we say that V is *uniformly nice* if it is ξ-nice for some constant ξ depending only on f. We denote

$$Z = \{x \in I : x \notin \mathrm{orb}(\mathrm{Crit}), \mathrm{dist}(\mathrm{orb}(x), \mathrm{Crit}) > 0\}.$$

Observe that $f^{-1}(Z) = Z$ and that (reasoning as in the proof of Lemma 44) Z is dense for $f \in C_{\mathrm{nf}}^2(I)$. We denote by Feig the critical points of Feigenbaum type.

Proposition 45. *Let $f \in C_{\mathrm{nf}}^2(I)$. Then there are $\xi_0 = \xi_0(f) > 0$ and, for every $\varepsilon > 0$ and $c \in \mathrm{Crit} \setminus \mathrm{Feig}$, open intervals $c \in V_c$ with $|V_c| < \varepsilon$, such that $V = \bigcup_{c \in \mathrm{Crit} \setminus \mathrm{Feig}} V_c$ is ξ_0-nice and $\partial V \subset Z$.*

Proof. In Cai and Li's version of this proposition, all critical points of solenoidal type (not only those of Feigenbaum type) are excluded and no additional properties on ∂V are obtained. Nevertheless, the proof remains very much the same. We sketch it below, emphasizing the specific points where it must be modified.

First of all, we prove:

Claim 1. There is $\xi_1 = \xi_1(f)$ such that if $c \in \mathrm{Crit} \setminus \mathrm{Feig}$, then there are arbitrarily small ξ_1-nice intervals J containing c such that $\partial J \subset Z$.

This is [CL, Corollary 3], but including also solenoidal critical points of non-Feigenbaum type. If c is recurrent, then Claim 1 follows immediately from Propositions 15 and 16. In fact, if c is not solenoidal, and we take as the starting interval I_0 in Proposition 16 a small nice interval containing c with $\partial I_0 \subset Z$ (which is possible by Lemma 44), then the central return domain I_m to I_{m-1} satisfies $\partial I_m \subset Z$ as well for all m. Thus if m is large enough and I_{m+1} is well inside I_m, then, by Proposition 15, I_{m+1} is the uniformly nice interval we need. If c is solenoidal, then Propositions 15 and 16 again imply that there is an arbitrarily small uniformly nice interval J containing c. Observe that if J is sufficiently small, then it is contained in a solenoidal cycle very close to $\omega(c)$, which in particular implies that the orbits of its endpoints cannot accumulate on any critical point

outside $\omega(c)$. Also, since J is nice and $\omega(c)$ is minimal, they cannot accumulate on $\omega(c)$ either. Thus $\partial J \subset Z$.

If c is not recurrent then the argument from [CL, Corollary 3] applies without any significant changes. Namely, let $I' \ni c$ be a small nice interval with $\partial I' \subset Z$. We can assume that $c \notin D(I')$. Let δ be the minimal length of the components of $I' \setminus \{c\}$ and take an interval $(a', b') \subset (c - \delta/2, c + \delta/2)$ with $a', b' \in Z$. This is possible by Lemma 44. If $a' \notin D(I')$, define $a = a'$. Otherwise, let K be the return domain to I' containing a' and let a be the endpoint of K' in (a', c). The point b is defined similarly. Then $J = (a, b)$ is $1/2$-well inside I', $\partial J \subset Z$ and $\partial J \cap D(I') = \emptyset$. If $x \in J \cap D(J)$, then the return domain L to J containing x is well inside the return domain L' to I' containing x by Proposition 15. Since $\partial J \cap D(I') = \emptyset$, we obtain $L' \subset J$. Then L is well inside J and J is uniformly nice.

Claim 2. Let $c_1, c_2, \ldots, c_k \in \mathrm{Crit}$ and $V_i \ni c_i$ be nice intervals such that $V = \bigcup_{i=1}^{k} V_i$ is a ξ-nice set and $\partial V \subset Z$. Then there is $\xi' = \xi'(\xi) > 0$ such that the following hold.

(1) For each $1 \le i \le k$, there exist nice intervals $W_i \supset \tilde{V}_i \ni c_i$ such that
  $-$ $\tilde{V}_i$ is ξ'-well inside W_i and W_i is ξ'-well inside V_i;
  $-$ $\partial \tilde{V}_i \cap D(V) = \emptyset$ and $\partial W_i \cap D(V) = \emptyset$. In particular, $\bigcup_{i=1}^{k} \tilde{V}_i$ is a nice set;
(2) For each $x \in V_i \setminus \mathrm{Cl}\, \tilde{V}_i$, there is a nice interval $J_x \ni x$ such that J_x is ξ'-well inside V_i and $J_x \cap \tilde{V}_i = \emptyset$, $\partial J_x \cap D(V) = \emptyset$.
(3) The endpoints of all intervals above belong to Z.

Claim 2 is exactly [CL, Lemma 3], except that we additionally request $\partial V \subset Z$ and get the extra property (3) in return. The proof requires no changes: only, instead of defining the auxiliary interval $(p', q') = (a - \xi'(b-a)/4, b + \xi'(b-a)/4)$ for $\tilde{V}_i = (a, b)$, we choose $p', q' \in Z$ with $p' \in (a - \xi'(b-a)/3, a - \xi'(b-a)/4)$ and $q' \in (b + \xi'(b-a)/4, b + \xi'(b-a)/3)$.

We are now in position to prove Proposition 45. This is done inductively. Let $\mathrm{Crit} \setminus \mathrm{Feig} = \{c_1, \ldots, c_m\}$. If $m = 1$, then this is just Claim 1. Assume that we have constructed intervals $V_i \ni c_i$ with $|V_i| < \varepsilon$ and $\partial V_i \subset Z$, $1 \le i \le k$, such that $\bigcup_{i=1}^{k} V_i$ is ξ_k-nice for some constant $\xi_k > 0$. We will show that there are smaller intervals $\tilde{V}_i \ni c_i$ with $\partial \tilde{V}_i \subset Z$, and a constant ξ_{k+1} depending only on ξ_k, such that $\bigcup_{i=1}^{k+1} \tilde{V}_i$ is ξ_{k+1}-nice.

The intervals $\tilde{V}_i$, $1 \le i \le k$, are those from Claim 2. To define $\tilde{V}_{k+1}$ and conclude the proof, two cases must be considered. If $c_{k+1} \in D(\bigcup_{i=1}^{k} V_i)$, then Cai and Li's proof works without any changes (it uses Claim 2 in its full extension).

If $c_{k+1} \notin D(\bigcup_{i=1}^{k} V_i)$, then we need to find intervals $c_{k+1} \subset \tilde{V}_{k+1} \subset V_{k+1}$ with $|V_{k+1}| < \varepsilon$, $\tilde{V}_{k+1}$ well inside V_{k+1} and $\partial \tilde{V}_{k+1} \subset Z$, such that $\bigcup_{i=1}^{k+1} V_i$ is nice and $\partial \tilde{V}_i \notin D(\bigcup_{i=1}^{k+1} V_i)$ for every $1 \le i \le k+1$.

We define V_{k+1} and $\tilde{V}_{k+1}$. Since $\tilde{V}_{k+1}$ is well inside V_{k+1} and $\partial \tilde{V}_{k+1} \subset Z$, we only need to show that $\partial V_{k+1}, \partial \tilde{V}_{k+1} \notin D(\bigcup_{i=1}^{k+1} V_i)$. (Recall that the endpoints of the intervals $V_i, \tilde{V}_i$, $i \le k$, belong to Z, hence their orbits cannot visit V_{k+1} if it is sufficiently small.)

As in the proof of Claim 1, three possibilities arise for c_{k+1}. The simplest case is when c_{k+1} is solenoidal. Then V_{k+1} is defined as in Claim 1 and $\tilde{V}_{k+1}$ is the return domain to V_{k+1} containing c_{k+1}; everything works because $c_{k+1} \notin D(\bigcup_{i=1}^{k} V_i)$.

Now assume that c_{k+1} is not solenoidal. Starting from a small interval $(a', b') \ni c_{k+1}$ with $a', b' \in Z$ and repeating the reasoning in Case 2 of Cai and Li's proof, we find an interval $c_{k+1} \in (a, b) \subset (a', b')$ such that $a, b \in Z$ and $(a, b) \cup \bigcup_{i=1}^{k} V_i$ is nice. If c_{k+1} is recurrent, we take $I_0 = (a, b)$ in Proposition 16, and define accordingly the intervals I_m. Then $I_m \cup \bigcup_{i=1}^{k} V_i$ is nice for every m because $c_{k+1} \notin D(\bigcup_{i=1}^{k} V_i)$. Now it suffices to fix m' such that $I_{m'}$ is uniformly nice and define $V_{k+1} = I_{m'}$ and $\tilde{V}_{k+1} = I_{m'+1}$. If c_{k+1} is non-recurrent, then we find an interval $J \ni c_{k+1}$, similarly as we did in Claim 1, such that $\partial J \subset Z$, $\partial J \cap D((a, b) \cup \bigcup_{i=1}^{k} V_i) = \emptyset$, and J is 1/2-well inside (a, b). Then $V_{k+1} = (a, b)$ and $\tilde{V}_{k+1} = J$ are adequate to our purposes.

Let us finally show that with this choice, $\tilde{V} = \bigcup_{i=1}^{k+1} \tilde{V}_i$ is uniformly nice. Let $x \in \tilde{V}_n \cap D(\tilde{V})$, say $\phi_{\tilde{V}}(x) \in \tilde{V}_l$. Let $0 \le s \le t$ be the entry time of x to $\tilde{V}_l$ and the return time of x to $\tilde{V}$, respectively, and denote the return domain to $\tilde{V}$ and the entry domain to V_l containing x by J and K respectively, so $x \in J \subset K$. By Proposition 15, the pullback H of $\tilde{V}_l$ along $f^s(x), \ldots, f^t(x)$ is well inside V_l. Note that J is the pullback of H along $x, \ldots, f^s(x)$ and K is the pullback of V_l along $x, \ldots, f^s(x)$. Hence J is well inside K by Corollary 14. Since $\partial \tilde{V}_n \notin D(\bigcup_{i=1}^{k+1} V_i)$, K is contained in $\tilde{V}_n$, so J is well inside $\tilde{V}_n$. This proves that $\tilde{V}$ is uniformly nice. $\qquad\square$

Proof of Theorem 20. Fix $\varepsilon > 0$. We define the sets $c \in U_c \subset V_c$ and W_c as follows. If $c \in \mathrm{Crit} \setminus \mathrm{Feig}$, then V_c is the component of the set V from Proposition 45 containing c. Write $V_c = (a, b)$ and say, for instance, that (c, b) is the smallest component of $V_c \setminus \{c\}$. Define the convex combinations

$$v_0 := \frac{c + \xi_0' \, b}{1 + \xi_0'} \quad \text{and} \quad v_i := \frac{v_{i-1} + \xi_0' \, b}{1 + \xi_0'}, \quad i = 1, 2, 3,$$

with $\xi_0' = \xi_0(f)/2$ and $\xi_0(f)$ the constant from Proposition 45. Take $v' \in (v_0, v_1) \cap Z$ close to v_0. If $v' \notin D(V)$, then set $v = v'$; otherwise let v be the right endpoint of the return domain to V containing v', and by having chosen v' close to v_0, we obtain $v_0 < v' \le v < v_1$. This gives $v \in ((v_0, v_1) \cap Z) \setminus D(V)$, and, similarly, there is $w \in ((v_2, v_3) \cap Z) \setminus D(V)$. If $c \in D(V)$, let u denote the left endpoint of the return domain to V containing c; if $c \notin D(V)$, take $u_0 = c - (v_0 - c) = 2c - v_0$ and find as before $u \in (u_0, c)$ belonging to $Z \setminus D(V)$. Finally we define $U_c = (u, w)$ and $W_c = (v, w)$.

For critical points of Feigenbaum type we rely on Proposition 43. More precisely, for every Feigenbaum solenoidal set S, we find a minimal solenoidal cycle $T = \mathrm{cyc}(\hat{J}) = \bigcup_{i=0}^{s-1} f^i(\hat{J})$ containing S whose constituting intervals are small enough to ensure that $|V_c| < \varepsilon$ for the components of the corresponding set V. Then the intervals U_c are those from the shell $M = \bigcup_{i=0}^{s-1} M_i$ of T containing points from Crit (that is, from Feig). To define the sets V_c assume, after reordering, that $M_0 = M_s$ is the smallest of all intervals M_i. Since the intervals $f^i(\hat{J})$ are well centered in the intervals M_i by Proposition 43, there is an interval G_s such that M_s is well inside G_s and G_s intersects no other interval from the solenoidal cycle T than $\hat{J} = f^s(\hat{J})$. Pulling back G_s along $f(\hat{J}), \ldots, f^s(\hat{J})$, we construct similarly intervals G_i, $i = 1, \ldots, s$, such that M_i is well inside G_i and $G_i \cap T = f^i(\hat{J})$. The intervals V_c are those intervals G_i containing points from Crit.

To define the sets W_c we take a boundary point q of M_0 with $f^s(q) = q$ (in fact the proof of Proposition 43 shows that it is possible that $f^{s/2}(q) = q$). For every $0 \le i < s$, let u_i denote

the middle point between $f^i(q)$ and the endpoint of $f^i(\hat{J})$ closest to $f^i(q)$. If M_i contains a critical point, let L_i be the interval with endpoints $f^i(q)$ and u_i; if not, let $L_i = M_i$. Finally, let $L'_i \subset L_i$ be the largest interval with endpoint $f^i(q)$ such that $f^j(L'_i) \subset L_{i+j}$ for all $0 \leq j < s$ with indices taken $\mathrm{mod}\, s$, and hence $f^s|_{L'_i}$ is a diffeomorphism. Now the intervals W_c are those intervals L'_i contained in intervals $U_c = M_i$ intersecting Crit.

We show that (i)-(iii) in Theorem 20 hold. Clearly, the construction implies the existence of a number $\xi = \xi(f) > 0$, thus not depending on ε, such that U_c is ξ-well inside V_c for every $c \in$ Crit. Moreover, since the endpoints of all sets $W_{c'}, U_{c'}, V_{c'}$, $c' \in$ Crit $\setminus$ Feig, belong to $Z \setminus D(V)$, we can assume that they do not belong to $D(V_c)$, $c \in$ Feig, either. If $c' \in$ Feig, then the niceness and invariance of shells still guarantees $\partial U_{c'} \cap D(U_c) = \emptyset$ for every $c \in$ Crit and $\partial U_{c'} \cap D(V_c) = \emptyset$ for every $c \in$ Crit not belonging to the same solenoidal set as c'. In particular, $U = \bigcup_{c \in \mathrm{Crit}} U_c$ is nice. This proves Theorem 20(i).

Now let J be an entry domain to U, say $\phi|_J = f^j|_J$ and $\phi(J) = U_c$. Then $f^j|_J$ is a diffeomorphism. We want to show that $f^j|_J$ extends to a diffeomorphism $f^j|_K : K \to V_c$. Assume by contradiction that this is not the case. Then there is $K' \supset J$ such that $f^j|_{K'}$ is a diffeomorphism, one of the endpoints a of K' satisfies $c' = f^n(a) \in$ Crit for some $1 \leq n < j$, and $f^j(\{a\} \cup K') \subset V_c$.

Let $b \in K'$ such that $f^n(b) \in \partial U_{c'}$ (here we use that $f^n(J)$ does not intersect U). We can assume that both c and c' belong to the same Feigenbaum solenoidal set S, because otherwise $\partial U_{c'} \cap D(V_c) = \emptyset$, which contradicts $f^j(b) \in V_c$. Let $T = \mathrm{cyc}(\hat{J})$ be the minimal solenoidal cycle for S we used earlier to construct the sets U_c, V_c. Then there is $i \in \mathbb{N}$ such that $f^j(J) = U_c \supset f^i(\hat{J})$. Since $f^j(a) = f^{j-n}(c') \in S$, there is $i' \in \mathbb{N}$ such that $f^j(a) \in f^{i'}(\hat{J})$. But $f^j|_{\{a\} \cup K'}$ is a homeomorphism, so $f^i(\hat{J})$ and $f^{i'}(\hat{J})$ are different. This is impossible, because by its definition V_c intersects exactly one interval from T.

We have shown that if J is an entry domain to U, then $f^j|_J : J \to U_c$ extends to a diffeomorphism $f^j|_K : K \to V_c$. Since U_c is well inside V_c, there is $\kappa = \kappa(f)$ such that $f^j|_J$ has distortion bounded by κ by the C^2 Koebe Principle (Proposition 15). This finishes the proof of Theorem 20(ii).

It remains to prove Theorem 20(iii). If $c \in$ Crit $\setminus$ Feig, then the definition of W_c easily implies that it is not too small compared to U_c. On the other hand, it is not too large compared to the subinterval of U_c between c and W_c, so $f|_{W_c}$ has bounded distortion because c is non-flat. Since $\partial W_c \cap D(U) = \emptyset$, Theorem 20(iii) holds in this case.

Assume now that $c \in$ Feig and U_c is one of the above intervals M_{i_1}, with $W_c = L'_{i_1}$. Put $k_c = s$ and let $i_1 < i_2 < \cdots < i_t < i_1 + s$ be the indices i such that M_i contains some critical point. As shown in the proof of Proposition 43, each map $f^{i_{r+1} - i_r - 1}|_{M_{i_r+1}}$ has bounded distortion. Also, each map $f|_{L_{i_r}}$ has bounded distortion because $|L_{i_r}|$ is less than half the distance from $f^{i_r}(q)$ to the critical point in M_{i_r} and non-flatness applies. Hence $f^{k_c}|_{W_c} = f^s|_{L'_{i_1}}$ has bounded distortion. Recall that all periodic orbits in the intervals M_i are repelling, and the same is true for $\mathrm{orb}(q)$. Then $f^{k_c}(W_c) \supset W_c$.

Finally, let us show that W_c is not too small compared to U_c. From its definition, L_{i_1} is not too small compared to U_c. Let A_j denote the largest interval with endpoint $f^j(q)$ such that $f^k(A_j) \subset L_{j+k}$ for all $0 \leq k \leq i_j - i_{j-1}$; hence $A_1 = L_{i_1}$ and $A_t = W_c$. We claim that A_{j+1}

is not too small compared with A_j. Indeed, $L_{i_{j+1}}$ is not too small compared to $M_{i_{j+1}}$, so *a fortiori* is not too small compared to $f^{i_{j+1}-i_1}(A_j)$. Since $f^{i_{j+1}-i_1}|_{A_j}$ has bounded distortion and $A_{j+1} = A_j \cap f^{-(i_{j+1}-i_1)}(L_{i_{j+1}})$, A_{j+1} cannot be too small compared to A_j either.

This concludes the proof of Theorem 20. $\square$

7.3. A C^3 Koebe Distortion Lemma.

Lemma 46. *Let $f \in C^3_{\mathrm{nf}}(I)$. Then for any $\xi > 0$ and $k \geq 0$, there is $\xi' = \xi'(\xi, k, f) > 0$ such that the following statement holds: Let $(H_i)^l_{i=0} \subset (G_i)^l_{i=0}$ be chains such that $(G_i)^l_{i=0}$ has order at most k and G_l is a small nice interval close enough to* Crit. *If H_l is ξ-well inside G_l, then H_0 is ξ'-well inside G_0. Moreover, if $k = 0$, then there is $\kappa = \kappa(\xi, f) > 0$ such that $f^l|_{H_0}$ has distortion bounded by κ.*

A slightly weaker version of this lemma (requiring that the intervals G_i are not too close to parabolic periodic points) appears in [LS], who in turns refer to [SV, Theorem C(2)]. Our proof is based on [SV, Proposition 3] and Theorem 20.

Proof of Lemma 46. Assuming that the components of U in Theorem 20 are sufficiently small, we can conclude that iterates of f one beyond those mapping into U have negative Schwarzian derivative. The idea behind obtaining negative Schwarzian derivative goes back to Kozlovski [Ko], see also [GSS], and the precise statement is as follows.

> Let $f \in C^3_{\mathrm{nf}}(I)$ and U be as in Theorem 20. If the components U_c of U are sufficiently small, and x is such that $f^n(x) \in U$ and $f^i(x) \notin$ Crit for every $0 \leq i \leq n$, then the Schwarzian derivative of f^{n+1} at x is negative.

"Sufficiently small" in this statement should be interpreted as that for each component U_c of U and each $i \geq 0$, each component of $f^{-i}(U_c)$ has length $\leq \tau$, where $\tau = \tau(f)$ is taken from [SV, Proposition 3] (choosing, with the notation in [SV, Proposition 3], $S = 1$, $N = 0$ and δ the number $\xi_0 = \xi(f)$ from Theorem 20). If U_c is sufficiently small, then this holds by Proposition 7. To prove the statement, let $(J_i)^n_{i=0}$ be the pullback chain of U_c along $x, f(x), \ldots, f^n(x) \in U_c$. Let $t_1 < t_2 < \cdots < t_m$ be the iterates such that $f^{t_j}(x) \in U$, and hence the intervals J_i, $t_j + 1 \leq i \leq t_{j+1}$ are pairwise disjoint and $\sum_{i=t_j+1}^{t_{j+1}} |J_i| \leq 1$. Fix $j > 1$ for the moment. Then J_{t_j+1} is contained in an entry domain J to U, say $f^{t_{j+1}-t_j-1}(J) = U_{c'}$. By Theorem 20(ii), there is $K \supset J$ such that $f^{t_{j+1}-t_j-1}$ maps diffeomorphically K onto $V_{c'}$, hence the pullback chain of $V_{c'}$ along $J_{t_j+1}, \ldots, J_{t_{j+1}}$ has order 0. Moreover, $J_{t_{j+1}}$ is contained in $U_{c'}$, so it is ξ_0-well inside $V_{c'}$. Therefore [SV, Proposition 3] implies that $f^{t_{j+1}-t_j}$ has negative Schwarzian derivative at $f^{t_j+1}(x)$. For $j = 1$ and $t_1 > 0$ (so $x \notin U$), f^{t_1+1} has negative Schwarzian derivative at x by the same reasoning. If $t_1 = 0$ (so $x \in U$), then f has negative Schwarzian derivative at x by non-flatness. Since compositions of maps with negative Schwarzian derivatives have negative Schwarzian derivative, the statement follows.

Now we continue with the proof of the lemma. Assume first $k = 0$. Fix $\xi > 0$ and let $(H_i)^l_{i=0} \subset (G_i)^l_{i=0}$ be chains such that $(G_i)^l_{i=0}$ has order 0 and G_l is contained in a component U_c of the set U above. If G_0 is an entry domain to G_l, then the statement is just

Proposition 13 because the intervals G_i are pairwise disjoint (G_l is nice). If not, again due to the niceness of G_l, there is an interval $G_t \subset G_l$ such that the intervals $G_{t+1}, \ldots, G_l$ are pairwise disjoint, so we can apply Proposition 13 to the subchains $(H_i)_{i=t+1}^l \subset (G_i)_{i=t+1}^l$. Also, $f^{t+1}|_{G_0}$ has negative Schwarzian derivative by the above statement, so we can apply Proposition 12 to the subchains $(H_i)_{i=0}^{t+1} \subset (G_i)_{i=0}^{t+1}$.

The general case $k > 0$ follows easily from this one (take also Lemma 11 into account). Now we need G_l to be small enough so that every component of every preimage of G_l is contained in U whenever it intersects Crit (again by Proposition 7).   □

7.4. **Proof of Theorem 21.** Let $x \in I$ be such that orb(x) is disjoint from ∂I. For every $n \in \mathbb{N}$ and every $0 < \varepsilon < d(f^n, \partial I)$ we construct the pullback chain of $(f^n(x)-\varepsilon, f^n(x)+\varepsilon)$ along $x, f(x), \ldots, f^n(x)$. We define

$$r_n^k(x) = \sup\{\varepsilon > 0 : \text{ order of the pullback chain of } (f^n(x) - \varepsilon, f^n(x) + \varepsilon) \le k\}.$$

If for every k we have $r_n^k(x) \to 0$ as $n \to \infty$, then we call x a *super-persistent* point. In [BM1, Theorem 2.7] (see also [BM2]) it is shown that for $f \in C_{\mathrm{nf}}^2(I)$, the ω-limit set of any super-persistent recurrent point of f is minimal.

Proof of Theorem 21. By definition of E and type (2) attractors, λ-a.e. $x \in E$ has a dense orbit in some cycle cyc(K) and therefore cannot be super-persistently recurrent, nor map into ∂I. We can then find a sequence $n_j \to \infty$ and $N_x \in \mathbb{N}$ and $\delta_x > 0$ such that the pullback chain of $(f^{n_j}(x) - \delta_x, f^{n_j}(x) + \delta_x)$ along $x, \ldots, f^{n_j}(x)$ has order at most N_x. Clearly we can take $\delta_{f(x)} \ge \delta_x$ and $N_{f(x)} \le N_x$ (in fact, we have equality for large j whenever x is not a critical point). Since Lebesgue measure has only finitely many ergodic components [Ly1, Theorem 2], it follows that there are a single $\delta > 0$ and $N \in \mathbb{N}$ such that $\delta_x \ge \delta$ and $N_x \le N$ for λ-a.e. $x \in E$.

Next choose $\varepsilon > 0$ in Theorem 20 so small that the intervals $U_c \subset V_c$, $c \in$ Crit$'$, satisfy:

- if c is in the interior of the metric attractor cyc(K)r, then dist$(V_c, \partial \text{cyc}(K)) > 2\delta$; moreover, if $d \in$ Crit $\setminus$ Crit$'$, then either $d \in \partial \text{cyc}(K)$ or dist$(d, \text{cyc}(K)) > \delta$;
- every component of every preimage of V_c, and every image $f^n(V_c)$, $n \le 2r$ (with r the period of cyc(K)), has length less than δ.

Recall that $U' = \bigcup_{c \in \text{Crit}'} U_c$, $V' = \bigcup_{c \in \text{Crit}'} V_c$. From their definition, both sets are nice. Moreover, $\partial U' \cap D(V') = \emptyset$ where as before $D(V') = \cup_{n>0} f^{-n}(V')$.

Since orb(x) is dense in cyc(K), we can assume that the numbers n_j are large enough so that $f^{n_j}(x) \in \text{cyc}(K)$ for every j. Let J be the entry domain to U (or the component of U) containing $f^{n_j}(x)$, say $f^{m_j}(J) = U_c$. Two possibilities arise. If $J \subset \text{cyc}(K)$, then $c \in$ Crit$'$. Moreover, if $K \supset J$ is such that f^{m_j} maps K diffeomorphically onto V_c, then $K \subset (f^{n_j}(x) - \delta, f^{n_j}(x) + \delta)$, so the pullback chain of V_c along $x, \ldots, f^{n_j+m_j}(x)$ has order at most N. The second possibility is that J contains some point from $\partial \text{cyc}(K)$, call it a. It is not possible that a belongs to a periodic orbit contained in $\partial \text{cyc}(K)$, because this would imply that one of the points of this periodic orbit belongs to Crit $\setminus$ Crit$'$, which contradicts that cyc(K) contains dense orbits. Hence there is $r_j \le 2r$ (with r the period of cyc(K)) such that $|f^{n_j-r_j}(x) - c'| < \delta$ for some $c' \in$ Crit$'$. In fact, $f^{n_j-r_j}(x) \in U_{c'}$, because otherwise $f^{n_j-r_j}(x)$ would belong to an entry domain J' to U contained in cyc(K), which

leads to the contradiction $f^{r_j}(J') = J \subset \mathrm{cyc}(K)$. Since $f^{r_j}(V_{c'}) \subset (f^{n_j}(x) - \delta, f^{n_j}(x) + \delta)$, the pullback chain of $V_{c'}$ along $x, \ldots, f^{n_j - r_j}(x)$ has order at most N.

We have proved that there is a number $N = N(f)$ such that, for a.e $x \in D$, there are $c \in \mathrm{Crit}'$ and a sequence $s_j \to \infty$ such that $f^{s_j}(x) \in U_c$ and the pullback of V_c along $x, \ldots, f^{s_j}(x)$ has order at most N. However, we need critical order 0, not N, for this theorem, so a further argument is required.

Let $k_x \geq 1$ be the smallest integer for which there are $G_x \supset H_x \ni x$ and $c \in \mathrm{Crit}'$ such that f^{k_x} maps G_x diffeomorphically onto V_c and $f^{k_x}(H_x) = U_c$. If $x \notin U'$, $\omega(x) = \mathrm{cyc}(K)$ and the entry domain $H_x \ni x$ to U' is contained in $\mathrm{cyc}(K)$, then k_x exists; it is the first entry time $k_x = r_{U'}(x)$. But if $x \in U'$, then k_x is more difficult to find.

Claim 1. The set $B := \{x \in E : k_x \text{ exists}\}$ has full Lebesgue measure in E.

Assume by contradiction that this claim fails, and that x is a density point of $E \setminus B$. We can find the sequence $(s_j)_{j \in \mathbb{N}}$ and $c \in \mathrm{Crit}'$ as above, and let $(G_i^j)_{i=0}^{s_j}$ and $(H_i^j)_{i=0}^{s_j}$ be the pullback chains of $G_{s_j}^j = V_c$ and $H_{s_j}^j = U_c$ along $x, \ldots, f^{s_j}(x)$, respectively. The chains $(G_i^j)_{i=0}^{s_j}$ have order $\leq N$.

Let $(W_t)_{t \in \mathbb{N}}$ be an enumeration of the return domains to V' within $V_c \setminus U_c$ and also take $W_0 := U_c$. Recall that $\partial U_c \cap D(V') = \emptyset$, so $(W_t)_{t \geq 0}$ is a family of pairwise disjoint intervals whose union has full measure in V'. Since $f^{s_j}|_{G_0^j}$ has at most N critical values, we can arrange the enumeration of the W_ts such that no critical value of $f^{s_j}|_{G_0^j}$ is contained in W_t if $t > N$. By Proposition 45 there is $\xi = \xi(f)$ such that each W_t is ξ-well inside V_c. By Lemma 46, there is $\zeta = \zeta(\xi, N, f)$ such that each component of $G_0^j \cap f^{-s_j}(\bigcup_{t=0}^N W_t)$ is ζ-well inside G_0^j. This holds in particular for the at most $M \leq 2^{N+1} - 1$ components of $f^{-s_j}(\bigcup_{t=0}^N W_t)$ in G_0^j. Let us denote these components as Y_i, numbered in order of decreasing size.

Claim 2. There are $b = b(\zeta, N) > 0$ and $B^j \subset G_0^j$ such that $\lambda(B^j) \geq b\lambda(G_0^j)$ and for every $y \in B^j$, there are $k_y \geq s_j$, $G_y \supset H_y \ni y$ and $c(y) \in \mathrm{Crit}'$ such that f^{k_y} maps G_y diffeomorphically onto $V_{c(y)}$ and $f^{k_y}(H_y) = U_{c(y)}$.

Take $b = \zeta^M/(M+1)!$. The components Y_i are pairwise disjoint, and have ζ-collars around them in G_0^j, which may intersect other components $Y_{i'}$ or their collars. We will show that $\lambda(\bigcup_i Y_i) \leq (1-b)\lambda(G_0^j)$. If $\lambda(Y_1) < \lambda(G_0^j)/(M+1)$, then we get directly

$$\lambda(\bigcup_i Y_i) < \lambda(G_0^j)M/(M+1) = (1 - 1/(M+1))\lambda(G_0^j) < (1-b)\lambda(G_0^j).$$

Thus we can assume $\lambda(Y_1) \geq \lambda(G_0^j)/(M+1)$ and the ζ-collar of Y_1 has two components of length $\geq \zeta\lambda(G_0^j)/(M+1)$. If the second largest Y_2 satisfies $Y_2 < \zeta\lambda(G_0^j)/((M+1)M)$, then there is a set in the ζ-collar of Y_1 of measure $> \zeta\lambda(G_0^j)/((M+1)M)$ which do not intersect $\bigcup_i Y_i$. Hence

$$\lambda(\bigcup_i Y_i) < (1 - \zeta/((M+1)M))\lambda(G_0^j) < (1-b)\lambda(G_0^j)$$

again. Continuing this way, we find that at least one component Y_k has a ζ-collar disjoint from $\bigcup_i Y_i$ and its size is $\geq \zeta^M\lambda(G_0^j)/(M+1)! = b\lambda(G_0^j)$.

Now take $B^j := E \cap (G_0^j \setminus \bigcup_i Y_i)$. Then clearly $\lambda(B^j) \geq b\lambda(G_0^j)$, and if $y \in B^j$, we find k_y as follows. We have $z := f^{s_j}(y) \in W_t$ for some $t > N$. There is r_t such that f^{r_t} maps W_t diffeomorphically onto a component of V'. If $f^{r_t}(z) \in U'$, then $k_y = s_j + r_t$. Otherwise $f^{r_t}(z) \in V' \setminus U'$ and in fact $f^{r_t}(z)$ belong to another return domain to V'. We continue iterating diffeomorphically until finally z falls into U', and we choose k_y accordingly. This proves the Claim 2. Obviously $B^j \subset B$, so $\lambda(B \cap G_0^j) \geq b\lambda(G_0^j)$ independently of j. Since $\lambda(G_0^j) \to 0$ as $j \to \infty$, this contradicts that x is a density point of $I \setminus B$, proving Claim 1.

From Claim 1 the first part of Theorem 21 easily follows: if $x, y \in B$ and $k_x \leq k_y$, then either $H_x \cap H_y = \emptyset$ or $H_y \subset H_x$. Now an easy maximality argument allows us to redefine these sets if necessary so that either $H_x \cap H_y = \emptyset$ or $H_y = H_x$.

It remains to prove the second part of Theorem 21. Since a.e. point from U' belongs to B, $F : U' \to U'$ is well defined a.e. It is important to note that if F is defined on x, then $H_x \subset G_x \subset U'$ because $\partial(U') \cap D(V') \neq \emptyset$. Now it is immediate to show by induction on n that every branch $F^n|_H = f^j|_H : H \subset U \to U_c$ of F^n admits a diffeomorphic extension to an interval $H \subset G \subset U$ with $f^j(G) = V_c$. Therefore $F^n|_H$ has distortion bounded by a constant κ depending neither on n nor on H by Lemma 46. $\square$

References

[AJS]   L. Alsedà, V. Jiménez López, L. Snoha, *All solenoids of piecewise smooth maps are period doubling,* Fund. Math. **157** (1998) 121–138.

[BJ1]   F. Balibrea, V. Jiménez López, *A structure theorem for C^2 functions verifying the Misiurewicz condition,* in: Proceedings of the European Conference on Iteration Theory (Lisbon, 1991), 12–21, World Sci. Publishing, Singapur, 1992.

[BJ2]   F. Balibrea, V. Jiménez López, *The measure of scrambled sets: a survey,* Acta Univ. M. Belii Ser. Math. No. **7** (1999) 3–11.

[BD]   M. Barge, B. Diamond, *Proximality in Pisot tiling spaces,* Fund. Math. **194** (2007) 191–238.

[Ba]   J. A. Barnes, *Conservative exact rational maps of the sphere,* J. Math. Anal. Appl. **230** (1999) 350–374.

[BrJ]   A. Barrio Blaya, V. Jiménez López, *An almost everywhere version of Smítal's order-chaos dichotomy for interval maps,* J. Austral. Math. Soc. **85** (2008) 29–50.

[BHS]   F. Blanchard, W. Huang, L. Snoha, *Topological size of scrambled sets,* Colloq. Math. **110** (2008) 293–361.

[Bl]   L. S. Block, *Stability of periodic orbits in the theorem of Šarkovskii,* Proc. Amer. Math. Soc. **81** (1981) 333–336.

[BC]   L. S. Block, W. A. Coppel, *Dynamics in one dimension,* Lecture Notes in Mathematics, 1513, Springer-Verlag, Berlin, 1992.

[BK]   L. Block, J. Keesling, *A characterization of adding machine maps,* Topology Appl. **140** (2004) 151–161.

[BKM]   L. Block, J. Keesling, M. Misurewicz, *Strange adding machines,* Ergodic Theory Dynam. Systems **26** (2006) 673–682.

[BL]   A. Blokh, M. Lyubich, *Measurable dynamics of S-unimodal maps of the interval,* Ann. Sci. École Norm. Sup. (4) **24** (1991) 545-573.

[BM1]   A. Blokh, M. Misiurewicz, *Wild attractors of polymodal negative Schwarzian maps,* Comm. Math. Phys. **199** (1998) 397–416.

[BM2]   A. Blokh, M. Misiurewicz, *Typical limit sets of critical points for smooth interval maps,* Ergodic Theory Dynam. Systems **20** (2000) 15–45. Erratum in Ergodic Theory Dynam. Systems **23** (2003) 661.

[BH]   A. M. Bruckner, T. Hu, *On scrambled sets for chaotic functions,* Trans. Amer. Math. Soc. **301** (1987) 289–297.

[Br1]   H. Bruin. *Invariant measures of interval maps,* Ph.D. Thesis, Delft University of Technology, 1994.

[Br2]   H. Bruin, *Combinatorics of the kneading map,* Internat. J. Bifur. Chaos Appl. Sci. Engrg. **5** (1995) 1339–1349.

[Br3]   H. Bruin, *Topological conditions for the existence of absorbing Cantor sets,* Trans. Amer. Math. Soc. **350** (1998) 2229–2263.

[Br4]   H. Bruin, *(Non)invertibility of Fibonacci-like unimodal maps restricted to their critical omega-limit sets,* Preprint 2008.

[BHa]   H. Bruin, J. Hawkins, *Exactness and maximal automorphic factors of unimodal interval maps,* Ergodic Theory Dynam. Systems **21** (2001) 1009–1034.

[BKNS]   H. Bruin, G. Keller, T. Nowicki, S. van Strien, *Wild Cantor Attractors exist,* Ann. of Math. (2) **143** (1996) 97–130.

[BKS]   H. Bruin, G. Keller, M. St.-Pierre, *Adding machines and wild attractors,* Ergodic Theory Dynam. Systems **17** (1997) 1267–1287.

[BSS]   H. Bruin, W. Shen, S. van Strien, *Existence of unique SRB-measures is typical for real unicritical polynomial families,* Ann. Sci. École Norm. Sup. (4) **39** (2006) 381–414.

[BRSS]   H. Bruin, J. Rivera-Letelier, W. Shen, S. van Strien, *Large derivatives, backward contraction and invariant densities for interval maps,* Invent. Math. **172** (2008) 509–533.

[CL]   H. Cai, S. Li, *Distortion of interval maps and applications,* Nonlinearity **22** (2009) 2353–2363.

[D]   R. H. Day, *Complex economic dynamics. Vol. I. An introduction to dynamical systems and market mechanisms,* MIT Press, Cambridge, 1994.

[E]   R. Engelking, *Dimension theory,* PWN-Polish Scientific Publishers, Warsaw, 1978.

[G]   T. Gedeon, *There are no chaotic mappings with residual scrambled sets,* Bull. Austral. Math. Soc. **36** (1987) 411–416.

[GKR]   C. Good, R. Knight, B. Raines, *Nonhyperbolic one-dimensional invariant sets with a countably infinite collection of inhomogeneities,* Fund. Math. **192** (2006) 267–289.

[GLT]   P. J. Grabner, P. Liardet, R. F. Tichy, *Odometers and systems of enumeration,* Acta Arith. **70** (1995) 103–125.

[GSS]   J. Graczyk, D. Sands, G. Świątek, *Metric attractors for smooth unimodal maps,* Ann. of Math.(2) **159** (2004) 725–740.

[Gu]   J. Guckenheimer, *Sensitive dependence to initial conditions for one-dimensional maps,* Comm. Math. Phys. **70** (1979) 133–160.

[HS]   J. Hawkins, C. Silva, *Characterizing mildly mixing actions by orbit equivalence of products,* New York J. Math, **3A** (1998) 99-115.

[H]   F. Hofbauer, *The topological entropy of a transformation $x \mapsto ax(1-x)$,* Monatsh. Math. **90** (1980) 117–141.

[HK]   F. Hofbauer, G. Keller. *Quadratic maps without asymptotic measures,* Comm. Math. Phys. **127** (1990) 319–337.

[HY]   W. Huang, X. Ye, *Homeomorphisms with the whole compacta being scrambled sets,* Ergodic Theory Dynam. Systems **21** (2001) 77–91.

[Ja]   M. V. Jakobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps,* Comm. Math. Phys. **81** (1981) 39-88.

[JS]   K. Janková, J. Smítal, *A characterization of chaos,* Bull. Austral. Math. Soc. 34 (1986) 283–292.

[J1]   V. Jiménez López, C^1 *weakly chaotic functions with zero topological entropy and non-flat critical points,* Acta Math. Univ. Comenian. (N.S.) **60** (1991) 195–209.

[J2]   V. Jiménez López, *Large chaos in smooth functions of zero topological entropy,* Bull. Austral. Math. Soc. **46** (1992) 271–285.

[J3]   V. Jiménez López, *Order and chaos for a class of piecewise linear maps,* Internat. J. Bifur. Chaos Appl. Sci. Engrg. **5** (1995) 1379–1394.

[Jo]   S. D. Johnson, *Singular measures without restrictive intervals,* Comm. Math. Phys. **110** (1987) 185–190.

[K]   I. Kan, *A chaotic function possessing a scrambled set with positive Lebesgue measure,* Proc. Amer. Math. Soc. **92** (1984) 45–49.

[Ke]   G. Keller. *Exponents, attractors and Hopf decompositions for interval maps,* Ergodic Theory
        Dynam. Systems **10** (1990) 717–744.
[Ko]   O. S. Kozlovski, *Getting rid of the negative Schwarzian derivative condition,* Ann. of Math. (2)
        152 (2000) 743–762.
[Le]   F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval,* Er-
        godic Theory Dynam. Systems **1** (1981) 77–93.
[LS]   S. Li, W. Shen, *Hausdorff dimension of Cantor attractors in one-dimensional dynamics,* Invent.
        Math. **171** (2008) 1629–1643.
[LY]   T. Li, J. A. Yorke, *Period three implies chaos,* Amer. Math. Monthly **82** (1975) 985-992.
[Ly1]   M. Lyubich, *Ergodic theory for smooth one dimensional dynamical systems,* Stony Brook preprint
        1991/11.
[Ly2]   M. Lyubich, *Combinatorics, geometry and attractors of quasi-quadratic maps,* Ann. of Math. (2)
        140 (1994) 347–404.
[LM]   M. Lyubich, J. Milnor, *The Fibonacci unimodal map,* J. Amer. Math. Soc. **6** (1993) 425–457.
[MMN]   M. Majumdar, T. Mitra, K. Nishimura (eds.), *Optimization and chaos,* Studies in Economic
        Theory **11**, Springer-Verlag, Berlin, 2000.
[M]   R. Mañé, *Hyperbolicity, sinks and measure in one-dimensional dynamics,* Comm. Math. Phys.
        100 (1985) 495–524. Erratum in Comm. Math. Phys. **112** (1987) 721–724.
[Ma]   M. Martens, *Interval Dynamics,* Ph.D. Thesis, Delft University of Technology, 1990.
[MMS]   M. Martens, W. de Melo, S. van Strien, *Julia-Fatou-Sullivan theory for real one-dimensional
        dynamics,* Acta Math. **168** (1992) 273–318.
[MT]   M. Martens, C. Tresser, *Forcing of periodic orbits for interval maps and renormalization of
        piecewise affine maps,* Proc. Amer. Math. Soc. **124** (1996) 2863–2870.
[MS]   W. de Melo, S. van Strien, *One dimensional dynamics,* Ergebnisse Series **25**, Springer-Verlag,
        Berlin, 1993.
[Mi]   M. Misiurewicz, *Chaos almost everywhere,* in: Iteration theory and its functional equations
        (Lochau, 1984), 125–130, Lecture Notes in Math., 1163, Springer, Berlin, 1985.
[P]   J. Piórek, *On the generic chaos in dynamical systems,* Univ. Iagel. Acta Math. **25** (1985) 293–
        298.
[S1]   J. Smítal, *A chaotic function with some extremal properties,* Proc. Amer. Math. Soc. **87** (1983)
        54–56.
[S2]   J. Smítal, *A chaotic function with a scrambled set of positive Lebesgue measure,* Proc. Amer.
        Math. Soc. **92** (1984) 50–54.
[S3]   J. Smítal, *Chaotic functions with zero topological entropy,* Trans. Amer. Math. Soc. **297** (1986)
        269–282.
[SS]   J. Smítal, M. Štefánková, *Omega-chaos almost everywhere,* Discrete Contin. Dyn. Syst. **9** (2003)
        1323–1327.
[St]   E. Straube, *On the existence of invariant, absolutely continuous measures,* Comm. Math. Phys.
        81 (1981) 27–30.
[vS]   S. van Strien, *Transitive maps which are not ergodic with respect to Lebesgue measure,* Ergodic
        Theory Dynam. Systems **16** (1996) 833–848.
[SV]   S. van Strien, E. Vargas, *Real bounds, ergodicity and negative Schwarzian for multimodal maps,*
        J. Amer. Math. Soc. **17** (2004) 749–782. Erratum in J. Amer. Math. Soc. **20** (2007) 267–268.
[Y]   L.-S. Young, *Recurrence times and rates of mixing,* Israel. J. Math. **110** (1999) 153–188.

Department of Mathematics
University of Surrey
Guildford, Surrey, GU2 7XH
UK
h.bruin@surrey.ac.uk
http://personal.maths.surrey.ac.uk/st/H.Bruin/

Departamento de Matemáticas
Universidad de Murcia
Campus de Espinardo, 30100 Murcia
Spain
vjimenez@um.es
http://www.um.es/docencia/vjimenez