FOR ALMOST EVERY TENT-MAP, THE TURNING POINT IS TYPICAL

Henk Bruin

KTH, Department of Mathematics
100 44 Stockholm, Sweden
bruin@math.kth.se

Abstract. Let T_a be the tent-map with slope a. Let c be its turning point, and μ_a the absolutely continuous invariant probability measure. For an arbitrary, bounded, almost everywhere continuous function g, it is shown that for almost every a, $\int g d\mu_a = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} g(T_a^i(c))$. As a corollary, we obtain that the critical point of a quadratic map is generically not typical for its absolutely continuous invariant probability measure, if it exists.

1. Introduction

Let $T_a : I \to I$ be the tent-map with slope a. Brucks and Misiurewicz [BM] showed that for a.e. $a \in [\sqrt{2}, 2]$, the orbit of the turning point is dense in the dynamical core. It is well-known that for $a > 1$, the tent-map T_a has an absolutely continuous invariant probability measure (*acip*), μ_a, and that μ_a is ergodic. By Birkhoff's Ergodic Theorem,

$$(1) \qquad \int g d\mu_a = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(T_a^i(x)) \quad \mu_a\text{-a.e.}$$

Here we take $g \in \mathcal{G} = \{h : I \to \mathbb{R} \mid h \text{ is bounded and continuous a.e.}\}$. Because μ_a is absolutely continuous with respect to Lebesgue measure, (1) holds Lebesgue a.e. If (1) holds for a point x, then x is called *typical* with respect to g. Although most points are typical, it is very difficult to identify a typical point. It is natural to ask if the turning point c of T_a is typical. We will prove

1991 *Mathematics Subject Classification.* 58F11, 58F03, 28D20.

Supported by the Deutsche Forschungsgemeinschaft (DFG). The research was carried out at the University of Erlangen-Nürnberg, Germany.

Theorem 1 (Main Theorem). *Let $g \in \mathcal{G}$. Then*

$$\int g d\mu_a = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(T_a^i(c)), \tag{2}$$

for a.e. $a \in [1,2]$.

It follows that for a.e. $a \in [1,2]$, (2) holds for every bounded Riemann integrable function simultaneously. This answers a question of Brucks and Misiurewicz [BM]. Schmeling [Sc] recently obtained similar results for β-transformations. In our proof, as well as in [BM], the properties of the turning point are used in a few arguments. We think, however, that Theorem 1 is true not only for c, but also for an arbitrary point $y \in I$.

The tent-map T_a has topological entropy $\log a$. Hence one can state Theorem 1 as: For a.e. value of the topological entropy, the turning point of T_a is typical. Because the measure μ_a actually maximizes metric entropy [M], this has a striking consequence for unimodal maps in general:

Corollary 1. *For a.e. $h \in [0, \log 2]$, if f is a unimodal map with $h_{top}(f) = h$, then the turning point of f is typical for the measure of maximal entropy.*

A result by Sands [Sa] states that for a.e. $h \in [0, \log 2]$, every S-unimodal map f with $h_{top}(f) = h$ satisfies the Collet-Eckmann condition, and therefore has an acip. For an S-unimodal map, however, the acip in general doesn't maximize entropy, because if it did, and if f is conjugate to a tent-map, the conjugacy ψ would be absolutely continuous. But then ψ has to be $C^{1+\alpha}$ too in a large neighbourhood of the critical point, as [MS, Exercise 3.1 page 375] indicates. (In [M] an argument is given for unimodal maps with a nonrecurrent critical point.) As a consequence, all periodic points have to have the same Lyapunov exponent, which is very unlikely. The only exception we are aware of is the full quadratic map $x \mapsto 4x(1-x)$. Hence combining the Corollary 1 with Sands' result, we obtain a large class of S-unimodal maps satisfying the Collet-Eckmann condition, but for which c is not typical for the acip. In contrast, Benedicks and Carleson [BC, Theorem 3] show that for the quadratic family $f_a(x) = ax(1-x)$ there is a set of parameters of positive Lebesgue measure for which f_a is Collet-Eckmann and c is typical for the acip.[1] Thus we are led to the conclusion that the entropy map $a \mapsto h_{top}(f_a)$, even when we disregard its flat pieces, has very bad absolute continuity properties.

[1] Thunberg [T] showed another kind of typicality: for a positive measured set of parameters, f_a has an acip which can be approximated weakly by Dirac-measures on super-stable orbits of nearby maps.

The proof of the Main Theorem goes in short as follows. First we introduce some induced map of the tent-map. We show that if a point is typical in some strong sense for this induced map, it is also typical for the original tent-map, Proposition 1. In sections 4 and 5 we prove certain properties of the induced map. Finally we show, using a version of the Law of Large Numbers (Lemma 8), that the turning point is indeed typical in this strong sense for a.e. parameter value (sections 7 to 9).

Acknowledgments. I want to thank Karen Brucks for many discussions and Gerhard Keller for his help with Lemma 6. I am also grateful to the referee for the attentive comments.

2. Preliminaries

The tent-map $T_a : I = [0,1] \to I$ is defined as $T_a(x) = \min(ax, a(1-x))$. For $a \leq 1$, the dynamics is uninteresting, and for $a \in (1, \sqrt{2}]$, T_a is finitely renormalizable. By considering last renormalization instead of T_a, we reduce to the case $a \in (\sqrt{2}, 2]$. Let us only deal with $a \in (\sqrt{2}, 2]$.

The point $c = \frac{1}{2}$ is the turning point. We write $c_n = c_n(a) = T_a^n(c)$. Another notation is $\varphi_n(a) = T_a^n(c)$. The core $[c_2(a), c_1(a)]$ will be denoted as $J(a)$.

For $a \in [\sqrt{2}, 2]$, T_a has an absolutely continuous invariant measure μ_a (*acip* for short). Its precise form can be found in [DGP], although we will not use that paper here. $\mu_a|_{J(a)}$ is equivalent to Lebesgue measure.

In the Main Theorem we considered $g \in \mathcal{G}$. Using a well-known fact from measure theory (e.g. [P, page 40]), it suffices to prove the following: Let $\mathcal{B}$ be the algebra of subsets of I whose boundaries have zero Lebesgue measure (or equivalently μ_a-measure), and let $B \in \mathcal{B}$. Then for a.e. $a \in (\sqrt{2}, 2]$,

$$\mu_a(B) = \lim_{n \to \infty} \frac{1}{n} \#\{0 \leq i < n \mid T_a^i(c) \in B\}.$$

It is this statement that we are going to prove.

The induced map that we will use is closely related to the *Hofbauer-tower* (Markov extension) of the tent-map. This object was introduced by Hofbauer (e.g. [H]). It is the disjoint union of intervals $\{D_n\}_{n \geq 2}$, where $D_2 = [c_2, c_1]$ and for $n \geq 1$,

$$D_{n+1} = \begin{cases} T_a(D_n) & \text{if } D_n \not\ni c \\ [c_{n+1}, c_1] & \text{if } D_n \ni c \end{cases}$$

Hence the boundary points of D_n are forward images of c, one of which is c_n. If $D_n \ni c$, then we call n a *cutting time*. We enumerate them by S_k: $S_1 = 2$, and by abuse of notation $S_0 = 1$. In this way we get $D_{S_k+1} = [c_{S_k+1}, c_1]$ and an inductive argument shows that $D_n = [c_n, c_{n-S_k}]$ if $S_k < n \leq S_{k+1}$.

The action $\check{T}_a$ on the tower is as follows. If $x \in D_n$, then

$$\check{T}_a(x) = T_a(x) \in \begin{cases} D_{n+1} & \text{if } c \notin (c_n, x] \text{ or } x = c = c_n \\ D_{r+1} & \text{if } c \in (c_n, x], \end{cases}$$

where r is determined as follows: Clearly $c \in (c_n, x]$ implies that $c \in D_n$. So n is a cutting time, say S_k. Then we set $r = S_k - S_{k-1}$. In fact, it is not hard to show that r itself is a cutting time. One can define a function $Q : \mathbb{N} \to \mathbb{N}$ by

$$r = S_{Q(k)} = S_k - S_{k-1}.$$

The function Q is called the *kneading map*. For more details see [B2].

The tower can be viewed as a countable Markov chain with the intervals D_n as states. There is a transition from D_n to D_{n+1} for each n and a transition from D_{S_k} to $D_{1+S_{Q(k)}}$ for each k. This will be used in section 5 to estimate the number of branches of our induced map.

One other property of the tower is that if U is an interval in the tower, then $\check{T}_a^n|_U$ is continuous if and only if $T_a^n|_U$ is monotone.

3. The induced map F_a

Definition. *Let $\check{F}_a$ be the first return map to D_2 in the Hofbauer-tower. The induced map F_a is the unique map such that $\pi \circ \check{F}_a = F_a \circ \pi$.*

For a.e. x we can define the *transfer-time* $s(x)$ as the integer such that $F_a(x) = T_a^{s(x)}$, F_a has the following properties:

- Each branch of F_a is linear.
- The image closure of each branch is $D_2 = [c_2, c_1] = J(a)$. If $a < 2$, then D_2 is the only level in the tower that equals $J(a)$. Hence $s(x)$ is the smallest positive integer n such that there exists an interval H, $x \in H \subset J(a)$, such that $T_a^n(H) = J(a)$ and $T_a^n|_H$ is monotone.
- F_a has countably many branches. The branch-domain will be denoted by $J_i(a)$. They form a partition of $J(a)$. Lemma 1 below shows that $|J(a) \setminus \bigcup_i J_i(a)| = 0$.
- $s|_{J_i}$ is constant. Let us denote this number by s_i.

Let also

$$\Phi_n(a) = F_a^n(c_3(a)).$$

The third iterate of c is chosen here, because F_a^n is well-defined in it for most parameter values, see Lemma 3.

Lemma 1. *For every $a \in [\sqrt{2}, 2]$ and every $n \in \mathbb{N}$, F_a^n is well-defined for a.e. $x \in J(a)$.*

Proof. The tent-map T_a admits an acip μ_a with positive metric entropy $\log a$. According to [K], μ can be lifted to an acip $\breve{\mu}$ on the tower. Furthermore, $\breve{\mu}(D_2) > 0$, and due to Birkhoff's Ergodic Theorem, a.e. x in the tower visits D_2 infinitely often. Hence for every $n \in \mathbb{N}$, F_a^n is defined a.e. □

Lemma 2. *For each $a_0 \in (\sqrt{2}, 2]$ there exists a neighbourhood $U \ni a$ and a constant C_1 such that for all $a \in U$*

$$\sum_i s_i |J_i| = \int_J s(x)dx \leq C_1.$$

Proof. $\sum_i s_i|J_i| = \int_J s(x)dx < \infty$ follows from the existence of the acip [B]. In our case, the uniform bound follows because there exist $U \ni a_0$, $C_2 > 0$ and $r \in (0,1)$ such that for every $a \in U$,

$$\sum_{s_i = n} |J_i| \leq C_2 r^n. \tag{3}$$

We will prove this in Lemma 7. □

The induced map F_a preserves Lebesgue measure, because every branch of F_a is linear and surjective. The invariant measure μ of T_a can be written as:

$$\mu(B) = C \sum_i \sum_{j=0}^{s_i - 1} |T_a^{-j}(B) \cap J_i|,$$

where C is the normalizing factor. By Lemma 2, $\mu(I) = C\sum_i s_i|J_i| < \infty$, and the measure can indeed be normalized:

$$\sum_i s_i |J_i| = \frac{1}{C}.$$

Fix $B \in \mathcal{B}$. We call x *very typical* with respect to B if

i) For all $i \in \mathbb{N}$ and $0 \leq j < s_i$,

$$\lim_{n\to\infty} \frac{1}{n} \#\{0 \leq k < n \mid F_a^k(x) \in T_a^{-j}(B) \cap J_i\} = \frac{1}{|c_2 - c_1|} |T_a^{-j}(B) \cap J_i|.$$

In particular, this limit exists.

ii) For every branch-domain J_i of F_a,

$$\frac{1}{|c_2 - c_1|}|J_i| = \lim_{n\to\infty} \frac{1}{n} \#\{0 \leq j < n \mid F_a^j(x) \in J_i\},$$

and

iii)

$$\frac{1}{C} = \sum_i s_i |J_i| = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(F_a^i(x)).$$

Proposition 1. *If x is very typical with respect to B, then*

$$\mu(B) = \lim_{n\to\infty} \frac{1}{n}\#\{0 \le k < n \mid T_a^k(x) \in B\}.$$

In other words, x is typical with respect to B for the original map.

Proof of Proposition 1. Choose $\varepsilon > 0$ arbitrary. Let x be very typical. Because of (3), there exists L such that $\sum_{s_j \ge L} s_j |J_j| \le \varepsilon$. Define $N_k(x) = \sum_{i=0}^{k-1} s(F_a^i(x))$. By condition iii), $\lim_{n\to\infty} \frac{N_n(x)}{n} = \frac{1}{C}$. Abbreviate $v(n,i) = \#\{(k,j) \mid 0 \le k < n,\, 0 \le j < s_i,\, F_a^k(x) \in J_i \text{ and } T_a^j \circ F_a^k(x) \in B\}$.

$$\begin{aligned}
\mu(B) &= C \cdot \sum_i \sum_{j=0}^{s_i-1} |T_a^{-j}(B) \cap J_i| \\
&= C \cdot \sum_i \sum_{j=0}^{s_i-1} \lim_{n\to\infty} \frac{1}{n}\#\{0 \le k < n \mid F_a^k(x) \in T_a^{-j}(B) \cap J_i\} \\
&= C \cdot \sum_i \lim_{n\to\infty} \frac{1}{n} v(n,i) \\
&\le C \cdot \sum_{s_i < L} \lim_{n\to\infty} \frac{1}{n} v(n,i) + C \cdot \sum_{s_i \ge L} s_i \lim_{n\to\infty} \#\{0 \le k < n \mid F_a^k(x) \in J_i\} \\
&\le C \cdot \lim_{n\to\infty} \frac{1}{n} \sum_{s_i < L} v(n,i) + C \sum_{s_i \ge L} s_i |J_i| \\
&\le C \cdot \limsup_{n\to\infty} \frac{1}{n} \sum_i v(n,i) + C\varepsilon \\
&\le C \cdot \limsup_{n\to\infty} \frac{1}{n}\#\{0 \le k < N_n(x) \mid T_a^k(x) \in B\} + C\varepsilon \\
&= C \cdot \lim_{n\to\infty} \frac{N_n(x)}{n} \limsup_{n\to\infty} \frac{1}{N_n(x)}\#\{0 \le k < N_n(x) \mid T_a^k(x) \in B\} + C\varepsilon \\
&= \limsup_{n\to\infty} \frac{1}{N_n(x)}\#\{0 \le k < N_n(x) \mid T_a^k(x) \in B\} + C\varepsilon.
\end{aligned}$$

Because ε is arbitrary, and also $\lim_{n\to\infty} \frac{N_{n+1}(x) - N_n(x)}{n} = 0$, we obtain

$$\mu(B) \le \limsup_{N\to\infty} \frac{1}{N}\#\{0 \le k < N \mid T_a^k(x) \in B\}.$$

Combining properties i) and ii) gives

$$\lim_{n\to\infty} \frac{1}{n}\#\{0 \le k < n \mid F_a^k(x) \in T_a^{-j}(I \setminus B) \cap J_i\} = |T_a^{-j}(I \setminus B) \cap J_i|.$$

Therefore we can carry out the above computation for the complement $I \setminus B$ as well. Because $\frac{1}{N}\#\{0 \le k < N \mid T_a^k(x) \in B \cup (I \setminus B)\} = 1$, it follows that $\mu(B) = \lim_{N\to\infty} \frac{1}{N}\#\{0 \le k < N, | T_a^k(x) \in B\}$, as asserted. $\square$

Remark: Since $\check{F}_a$ is also an induced (in fact first return) map over $\check{T}_a$, we can use the same argument to show that $x \in D_2$ is typical with respect to $B \subset \bigsqcup_n D_n$ and lifted measure $\check{\mu}_a$ on the tower. It was shown in [B], that many induced maps over (T_a, I) correspond to first return maps to some subset A in the tower. As x is typical with respect to $\frac{\check{\mu}|_A}{\check{\mu}(A)}$ and the first return map to A, it immediately follows that x is typical for these induced maps.

In order to prove the Main Theorem, we need to show that c, or rather c_3, satisfies conditions i) to iii) for a.e. a. This will be done in Propositions 2 and 3.

4. Some more properties of J_i, φ_n and Φ_n.

Lemma 3. *If $orb(c(a))$ is dense in $J(a)$, then $\Phi_n(a)$ is defined for every $n \in \mathbb{N}$.*

It immediately follows by [BM] that

Corollary 2. *$\Phi_n(a)$ is defined for all n for a.e. $a \in [\sqrt{2}, 2]$.*

Proof of Lemma 3. Let k be such that there exists H, $c_3 \in H \subset J(a)$, such that $T_a^k|_H$ is monotone and $T_a^k(H) = J(a)$. Let p be the nonzero fixed point of T_a. Let

$$c_{-v} < c_{-v-2} < \ldots < p < \ldots < c_{-v-3} < c_{-v-1}$$

be pre-turning points closest to p, where $v > k$. As $orb(c(a))$ is dense in $J(a)$, there exists m such that $c_m \in (c_{-v}, c_{-v-1})$. Take m minimal. Let $H' \ni c_3$ be the maximal interval such that $T_a^{m-3}|_{H'}$ is monotone. Because $\partial T_a^{m-3}(H') \subset orb(c(a))$ and m is minimal, $T_a^{n-3}(H') \supset [c_{-v}, c_{-v-1}]$. Because $T_a^{v+2}([c_{-v}, c_{-v-1}]) = [c_2, c_1]$, we have for $k' = n-3+v+2 > k$ that $T_a^{k'}|_{H'}$ is monotone and $T_a^{k'}(H') = J(a)$. It follows that $\Phi^n(a)$ is defined for all $n \in \mathbb{N}$. $\square$

The previous lemmas showed that there exists a full-measured set $\mathcal{A} \subset [\sqrt{2}, 2]$ of parameters for which $\Phi_n(a)$ is defined for every n. In particular, c is not periodic for every $a \in \mathcal{A}$. Let us assume from now on that a is always taken from $\mathcal{A}$. The next lemma shows that all branches of $\Phi_n : (\sqrt{2}, 2] \to J(a)$ are onto.

Lemma 4. *Let $a \in \mathcal{A}$, and suppose $\Phi_n(a) = T_a^m(c_3(a))$. Then there exists an interval $U = [a_1, a_2] \ni a$, such that φ_{m+3} maps U monotonically onto $[c_1(a_1), c_2(a_2)]$ or $[c_2(a_1), c_1(a_2)]$.*

Proof. By definition $\pi^{-1} \circ \Phi_n(a) \cap D_2$ is the n-th return in the tower of $c_3 \in D_2$ to D_2. Suppose $\Phi_n(a) = \varphi_{m+3}(a) \in int\, J(a)$. Because any point in $\pi^{-1}(c)$ is mapped by $\check{T}_a$ to a boundary point of some level in the tower, and because boundary points are mapped to boundary points, it follows that $\varphi_j(a) \neq c$ for $j < m+3$. Hence φ_{m+3} is a diffeomorphism in a neighbourhood of a. Since this is true for every point a' such that $\Phi_n(a') \in int\, J(a')$, the existence of the interval U follows. □

For any C^1 function f, let $dis(f, J) = \sup_{x,y\in J} \frac{|Df(x)|}{|Df(y)|}$ be the *distortion* of f on J.

Lemma 5. *Let $U_n \subset [\sqrt{2}, 2]$ be an interval on which φ_n is monotone. Then*

$$\sup_{U_n \subset [\sqrt{2},2]} dis(\varphi_n, U_n) \to 1 \text{ as } n \to \infty.$$

Moreover $\frac{d}{da}\varphi_n(a) = \mathcal{O}(a^n)$.

Proof. See [BM]. □

Corollary 3. *There exists $K > 0$ with the following property. Let $x = x(a) \in I$ be such that $T_a^n(x) = c(a)$ for some n and $T_a^j(x) \neq c(a)$ for $j < n$. Moreover fix the itinerary of x up to entry n. Then $|\frac{dx(a)}{da}| \leq K$.*

Proof. Write $G(a, x) = T_a^n(x) - c$, then

$$0 = \frac{d}{da}G(a,x) = \frac{\partial}{\partial a}T_a^n(x) + \frac{\partial}{\partial x}T_a^n(x)\frac{dx}{da} = \frac{\partial}{\partial a}T_a^n(x) + a^n\frac{dx}{da}.$$

As T_a^n is an degree n polynomial with coefficients in $[0, 1]$, $|\frac{\partial}{\partial a}T_a^n| \leq Ka^n$. The result follows. □

The boundary points of $J_i(a)$ are preimages of c. As long as $J_i(a)$ persists, $|J_i(a)| = a^{-s_i}|J(a)|$ and $J_i(a)$ moves with speed $\mathcal{O}(1)$ as a varies. Take n large and let U_n be such that $\varphi_n|_{U_n}$ is monotone. By Lemma 5, the $dis(\varphi_n, U_n)$ is close to 1. There exists K ($K \to 1$ as $n \to \infty$) such that

$$\frac{|\varphi_n^{-1}(J_i(a)) \cap U_n|}{|U_n|} \leq K|J_i(a)| = Ka^{-s_i}|J(a)|.$$

Let us now try to analyze how the branch-domains $J_i(a)$ are born and die if the parameter varies. As $|J_i(a)| = a^{-s_i}|c_1(a) - c_2(a)|$,

$$\frac{d}{da}|J_i(a)| = \frac{1}{2}a^{-s_i}(2a - 1 - s_i(a-1)).$$

It is easy to see that for $s_i \geq 5$ and $a \in [\sqrt{2}, 2]$, $\frac{d}{da}|J_i(a)| < 0$. These branch-domains shrink as a increases, and therefore cannot be born in a point. The only way a branch-domain can be created is by merging (countably) many smaller branch-domains, with larger transfer-times, into a new one. This happens whenever c is n-periodic, and the central branch of T_a^n covers a point of $T_a^{-1}(c)$. This is the same moment on which the central branch of T_a^{n+2} covers (c_2, c_1).

As the kneading invariant (and topological entropy) of T_a increases with a, branch-domains cannot disappear either, except in this merging process.

5. The proof of statement (3)

Lemma 6. *For every $a_0 \in (\sqrt{2}, 2]$ for which c is not periodic under T_{a_0}, there exists $C_2, \delta > 0$ such that for every $a \in (a_0 - \delta/2, a_0 + \delta/2)$ and every $n \geq 1$,*

$$\#\{j \mid s_j(a) = n\} \leq C_2(a_0 - \delta)^n.$$

Proof. It is shown in [H] that $a = \exp h_{top}(T_a)$ is the exponential growth-rate of the number of paths in the tower starting from D_2. Let $G(a, n) = \#\{j \mid s_j(a) = n\}$ be the number of n-loops from D_2 to D_2 that do not visit D_2 in between. We will choose $\delta > 0$ below such that the combinatorics of the tower up to some level remains the same for all $a \in (a_0 - \delta, a_0 + \delta)$. Then we argue that the exponential growth-rate $\limsup_n \frac{1}{n} \log G(a, n)$ for all $a \in (a_0 - \delta/2, a_0 + \delta/2)$ is smaller than $h_{top}(T_{a_0 - \delta}) = \log(a_0 - \delta)$. ¿From this the lemma follows. We will compute these exponential growth-rates by means of the characteristic polynomials of well-chosen submatrices of the transition matrix corresponding to the tower.

Choice of δ: The assumption $a_0 > \sqrt{2}$ implies that c_3 lies to the left of the non-zero fixed point of T_{a_0}. It is easy to verify that for some integer $u \geq 0$, $c_3, \dots, c_{2u+2}$ lie to the right of c while c_{2u+3} lies to the left again. This corresponds to the fact that T_{a_0} is not renormalizable. In terms of the kneading map renormalizability is equivalent to the statement ([B2, Proposition 1]): There exists $k \geq 1$ such that

$$Q(k) = k - 1 \text{ and } Q(k + j) \geq k - 1 \text{ for all } j \geq 1.$$

Here S_k is the period of renormalization. In our case, this formula is false for $S_k = S_1 = 2$. Therefore there exists $u \geq 0$ such that

$$Q(1) = 0,\ Q(j) = 1 \text{ for } 2 \leq j \leq u + 1,\ Q(u + 2) = 0.$$

Take δ maximal such that the cutting times $S_0, \dots, S_{u+2}$ are the same for all $a \in (a_0 - \delta, a_0 + \delta)$. As c is not periodic under T_{a_0}, δ is positive.

A lower bound for the entropy: The tower $\sqcup_{n \geq 2} D_n$ gives rise to a countable transition matrix $M = (m_{i,j})_{i,j=2}^{\infty}$, where $m_{i,j} = 1$ if and only if a transition $D_i \to D_j$ is possible. Therefore $m_{i,i+1} = 1$ and $m_{S_k,1+S_{Q(k)}} = 1$ for all i, k, and all other entries are zero. For $a \in (a_0 - \delta, a_0 + \delta)$ let $M(u)$ be the $(2u+2) \times (2u+2)$ left upper submatrix of M. Denote the spectral radius of this matrix by $\rho_0(u)$. For example,

$$M(2) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Because $M(u)$ is the transition matrix of $\sqcup_{n=2}^{2u+3} D_n$, $\log \rho_0(u)$, the exponential growth-rate of the paths from D_2 in $\sqcup_{n=2}^{2u+3} D_n$ is less than or equal to the exponential growth-rate of the paths from D_2 in the whole tower. Therefore $\log \rho_0(u) \le \inf\{h_{top}(T_a) \mid a \in (a_0 - \delta, a_0 + \delta)\} = \log(a_0 - \delta)$.

An upper bound for $\mathbf{G(a,n)}$: In order to estimate $G(a,n)$, we use a larger submatrix of M. Assume that $S_{u+3} = S_{u+2} + v = 2u+3+v$. Let $\tilde{M}(u,v)$ be the $(2u+2+v) \times (2u+2+v)$ left upper submatrix of M in which we set $\tilde{m}_{2,2} = \tilde{m}_{2,3} = 0$ and $\tilde{m}_{2u+3+v,2u+4} = 1 + m_{2u+3+v,2u+4}$. Denote the spectral radius by $\rho_1(u,v)$. For example,

$$\tilde{M}(2,4) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

We claim that for $a \in (a_0 - \delta/2, a_0 + \delta_2)$, i.e. u fixed,

$$\limsup \frac{1}{n} \log G(a,n) \le \max\{\log \rho_1(u,v) \mid v = 1, 2, 4, \dots, 2u, 2u+2, 2u+3\}.$$

Clearly $G(a,1) = 1$ and for $n \ge 2$, $G(a,n)$ is the number of paths of length $n-1$ from D_3 to D_2 that do not visit D_2 in between. The total number of paths of length $n-1$ from D_3 to D_2 is $m_{3,2}^{n-1}$, the appropriate entry of the matrix M^{n-1}. By putting $\tilde{m}_{2,2} = \tilde{m}_{2,3} = 0$ we avoid counting the paths that visit D_2 in between. The tower $\sqcup_{n \ge 2} D_n$ can be pictured as a graph; the branchpoints are the cutting levels D_{S_k}. ¿From $D_{S_{u+2}}$ there is a path $D_{S_{u+2}} \to D_2$ and a path upwards in the tower. This path splits again at $D_{S_{u+3}}$ into a path to $D_{1+S_{Q(u+3)}}$ and another to $D_{1+S_{u+3}}$. This gives two paths $D_{S_{u+2}} \to D_{1+S_{u+3}}$ and $D_{S_{u+2}} \to D_{1+S_{Q(u+3)}}$, both of length $v = S_{Q(u+3)} \in \{1,2,4,6,\dots,2u,2u+2,2u+3\}$. At the branchpoint $D_{S_{u+3}}$ the same situation occurs: there are paths $D_{S_{u+3}} \to D_{1+S_{u+4}}$ and $D_{S_{u+3}} \to D_{1+S_{Q(u+4)}}$, both of length $v' = S_{Q(u+4)} \in \{1,2,4,6,\dots,2u,2u+2,2u+3,2u+3+v\}$. The number of paths of length n from D_2 increases if the path-lengths between branchpoints decrease. Therefore the choice $v' = 2u+3+v$ will give smaller values of $G(a,n)$ for n large than the choice $v' = 2u+3$. And if v is chosen such that $G(a,n)$ is maximized (i.e. the largest values for $G(a,n)$ are obtain for those a for which $S_{Q(u+3)} = v$), then choosing $v' = v$ (i.e. choosing a such that $S_{Q(u+4)} = v$) will also maximize $G(a,n)$. By induction we should take the same value for $S_{Q(k)}$ for each $k \geq u+3$. Therefore we can identify all branchpoints D_{S_k}, $k \geq u+3$. This gives rise to the transition matrix $\tilde{M}(u,v)$ and hence proves the claim.

The rome-technique: To prove the lemma, it suffices to show that $\rho_1(u,v) \leq \rho_0(u)$. The spectral radius is the leading root of the characteristic polynomial. We will compute the characteristic polynomials of $M(u)$ and $\tilde{M}(u,v)$ (denoted as cp_0 and cp_1 respectively) by means of the *rome*-technique from [BGMY, Theorem 1.7]. Let M be some $n \times n$ matrix with nonnegative integer entries. A *path* p is a sequence $p_0 \dots p_\ell$ of states such that $m_{p_{i-1},p_i} > 0$ for all $1 \leq i \leq \ell$. The *length* of the path is $\ell(p) = \ell$ and $w(p) = \prod_{i=1}^{\ell(p)} m_{p_{i-1},p_i}$ is the *width*. A *rome* $R = \{r_1 \dots r_k\}$, i.e. $\#(R) = k$, is a subset of the states with the property that every closed path (i.e. $p_0 = p_\ell$) contains at least one state from R. A path $p = p_0 \dots p_\ell$ is *simple* if $p_0, p_\ell \in R$ but $p_i \notin R$ for $1 \leq i < \ell$.

Theorem (Rome Theorem). *The characteristic polynomial of M equals*

$$(-1)^{n-k} x^n \det(A_R(x) - I),$$

where I is the identity on $\mathbb{R}^k$ and $A = (a_{i,j})_{i,j=1}^k$ is the matrix with entries $a_{i,j} = \sum_p w(p) x^{-\ell(p)}$. Here the sum runs over all simple paths from r_i to r_j.

The characteristic polynomials: Let $D_i \to_k D_j$ stand for a path of length k from D_i to D_j. For $M(u)$, the states D_2 and D_3 form a rome. The corresponding simple paths are $D_2 \to_1 D_2$, $D_2 \to_1 D_3$, $D_3 \to_{2u+1} D_2$ and $D_3 \to_j D_3$ for $j = 2,4,..,2u$. Therefore the characteristic polynomial of $M(u)$ is

$$cp_0(u) = x^{2u+2} \det \begin{pmatrix} \frac{1}{x} - 1 & \frac{1}{x} \\ \frac{1}{x^{2u+1}} & \frac{1}{x^2} + \dots + \frac{1}{x^{2u}} - 1 \end{pmatrix} = \frac{x^{2u+3} - 2x^{2u+1} - 1}{x+1}.$$

For $\tilde{M}(u,v)$ we distinguish four cases.

a) $v = 1$. In this case $\{D_2, D_3, D_{2u+4}\}$ forms a rome and the simple paths are $D_3 \to_{2u+1} D_2$, $D_3 \to_{2u+1} D_{2u+4}$, $D_3 \to_j D_3$ for

$j = 2, 4, \ldots, 2u$, $D_{2u+4} \to_1 D_2$ and $D_{2u+4} \to_1 D_{2u+4}$. We give the characteristic polynomial the sign that makes the leading coefficient positive.

$$\begin{aligned} -cp_1(u,1) &= -x^{2u+3} \cdot \det \begin{pmatrix} -1 & 0 & 0 \\ \frac{1}{x^{2u+1}} & \frac{1}{x^2} + \ldots + \frac{1}{x^{2u}} - 1 & \frac{1}{x^{2u+1}} \\ \frac{1}{x} & 0 & \frac{1}{x} - 1 \end{pmatrix} \\ &= \frac{x^2(x^{2u+2} - 2x^{2u} + 1)}{(x+1)}. \end{aligned}$$

Hence $-\frac{1}{x}cp_1(u,v) - cp_0(u) = 1$. Because $\frac{1}{x}$ is positive on $(1,\infty)$, it follows that $\rho_0(u) > \rho_1(u,1)$.

b) $v = 2$. In this case $\{D_2, D_3, D_{2u+4}\}$ forms a rome and the simple paths are $D_3 \to_{2u+1} D_2$, $D_3 \to_{2u+1} D_{2u+4}$, $D_3 \to_j D_3$ for $j = 2, 4, \ldots, 2u$, $D_{2u+4} \to_2 D_3$ and $D_{2u+4} \to_2 D_{2u+4}$. The characteristic polynomial is

$$\begin{aligned} cp_1(u,2) &= -x^{2u+4} \cdot \det \begin{pmatrix} -1 & 0 & 0 \\ \frac{1}{x^{2u+1}} & \frac{1}{x^2} + \ldots + \frac{1}{x^{2u}} - 1 & \frac{1}{x^{2u+1}} \\ 0 & \frac{1}{x^2} & \frac{1}{x^2} - 1 \end{pmatrix} \\ &= x(x^{2u+3} - 2x^{2u+1} + x - 1). \end{aligned}$$

Therefore $\frac{1}{x}cp_1(u,v) - (x+1)cp_0(u) = x$, which is positive on $(1,\infty)$. Because also $\frac{1}{x}$ and $x+1$ are positive on $(1,\infty)$, $\rho_0(u) > \rho_1(u,2)$.

c) $v = 4, 6, \ldots, 2u$. Here $\{D_2, D_3, D_{v+1}, D_{2u+4}\}$ forms a rome and the paths are $D_3 \to_{v-2} D_{v+1}$, $D_3 \to_j D_3$ for $j = 2, 4, \ldots, v-2$, $D_{u+1} \to_j D_3$ for $j = 2, \ldots, 2u-v+2$, $D_{u+1} \to_{2u-v+3} D_{2u+4}$, $D_{u+1} \to_{2u-v+3} D_2$, $D_{2u+4} \to_v D_{2u+4}$, $D_{2u+4} \to_v D_{u+1}$ and $D_{2u+4} \to_2 D_{2u+4}$. The characteristic polynomial is

$$\begin{aligned} cp_1(u,v) &= x^{2u+v+2}\cdot \\ &\det \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1}{x^2} + \ldots + \frac{1}{x^{v-2}} - 1 & \frac{1}{x^{v-2}} & 0 \\ \frac{1}{x^{2u+3-v}} & \frac{1}{x^2} + \ldots + \frac{1}{x^{2u+2-v}} & -1 & \frac{1}{x^{2u+3-v}} \\ 0 & 0 & \frac{1}{x^v} & \frac{1}{x^v} - 1 \end{pmatrix} \\ &= \frac{x(x^v - 1)(x^{2u+3} - 2x^{2u+1} + x + 1)}{(x-1)(x^2-1)}. \end{aligned}$$

It follows that $\frac{(x-1)(x^2-1)}{x(x^v-1)}cp_1(u,v) - (x+1)cp_0(u) = (x+2)$, which is positive on $(1,\infty)$. Because $\frac{(x-1)(x^2-1)}{x(x^v-1)}$ and $x+1$ are also positive in $(1,\infty)$, $\rho_0(u) > \rho_1(u,v)$.

d) $v = 2u+3$. Again $\{D_2, D_3, D_{2u+4}\}$ forms a rome. The paths are $D_3 \to_{2u+1} D_2$, $D_3 \to_{2u+1} D_{2u+4}$, $D_3 \to_j D_3$ for $j = 2, 4, \ldots, 2u$

and $D_{2u+4} \to_{2u+3} D_{2u+4}$. This last path has width 2. We obtain

$$-cp_1(u, 2u+3) = -x^{4u+5} \cdot \det \begin{pmatrix} -1 & 0 & 0 \\ \frac{1}{x^{2u+1}} & \frac{1}{x^2} + \ldots + \frac{1}{x^{2u}} - 1 & \frac{1}{x^{2u+1}} \\ 0 & 0 & \frac{2}{x^{2u+3}} - 1 \end{pmatrix}$$
$$= \frac{x(x^{2u+3} - 2)(x^{2u+2} - 2x^{2u} + 1)}{(x^2 - 1)}.$$

Therefore $-\frac{x-1}{x^{2u+3}-2} cp_1(u,v) - cp_0(u) = 1$. Because $\frac{1}{x^{2u+3}-2}$ and $x - 1$ are positive on $(2^{\frac{1}{2u+3}}, \infty)$ and $cp_0(2^{\frac{1}{2u+3}}) < 0$ it follows that $\rho_0(u) > \rho_1(u, 2u+3)$.

Hence in all cases $\rho_0(u) > \rho_1(u,v)$. Therefore $\limsup \frac{1}{n} \log G(a,n) \leq \max\{\rho_1(u,v) \mid v = 1, 2, 4, 6, \ldots, 2u, 2u+2, 2u+3\} < \rho_0(u) \leq a_0 - \delta$, proving the lemma. □

Lemma 7. *For every $a_0 \in [\sqrt{2}, 2]$, there exists $C_2, \delta > 0$ and $r \in (0,1)$ such that for every $a \in (a_0 - \delta/2, a_0 + \delta/2)$,*

$$\sum_{s_j = n} |J_i(a)| \leq C_2 r^n. \tag{3}$$

Proof. Because $|J_i(a)| = |c_2(a) - c_1(a)| a^{-s_i} \leq a^{-s_i}$, the statement follows immediately from Lemma 6. We can take δ and C_2 as in Lemma 6 and $r = \frac{a_0 - \delta}{a_0 - \delta/2} < 1$. □

6. Probabilistic lemmas

For each $n \in \mathbb{N}$ we consider the set of branch-domains of the map Φ_n as partition $\mathcal{Z}_n$ of the parameter space $[\sqrt{2}, 2]$. For $m < n$, $\mathcal{Z}_n$ is finer than $\mathcal{Z}_m$, and $\bigvee_n \mathcal{Z}_n$ contains no nondegenerate intervals. An element of $\mathcal{Z}_n$ will be denoted by $Z_{e_1 e_2 \ldots e_n}$, where $e_j = i$ if $\Phi_{j-1}(Z_{e_1 e_2 \ldots e_n}) \subset J_i(a)$.

Lemma 8. *Let $\{X_m\}$ be a sequence of random variables with the following properties:*

a) There exists $V < \infty$ such that for every $m \in \mathbb{N}$, $Var(X_m \mid Z_{e_1 e_2 \ldots e_m}) < V$ for every branch-domain $Z_{e_1 e_2 \ldots e_m}$.
b) X_{m-1} is constant on each interval $Z_{e_1 \ldots e_m}$.
c) There exist $M \in \mathbb{R}$, $N \in \mathbb{N}$ and $\varepsilon > 0$ such that for every $m > N$,

$$|M - \mathbb{E}(X_m | Z_{e_1 e_2 \ldots e_m})| < \varepsilon.$$

Then

$$\limsup_{m \to \infty} |M - \frac{1}{m} \sum_{i=0}^{m-1} X_i| \leq \varepsilon \quad a.s.$$

Notice that the random variables X_m are not independent, but only "eventually almost independent". We will use this lemma twice in the next two sections. In the next section however, we will consider only a subsequence of the branch-domain partitions $\{Z_{e_1\dots e_n}\}$. This does not effect the validity of the lemma.

Proof. Define $Y_m = X_m - \mathbb{E}(X_m|Z_{e_1e_2\dots e_m})$. Then $\mathbb{E}(Y_m|Z_{e_1e_2\dots e_m}) = 0$ and $Var(Y_m \mid Z_{e_1e_2\dots e_m}) = \mathbb{E}(Y_m^2 \mid Z_{e_1e_2\dots e_m}) < V$ for all m and all branch-domains $Z_{e_1e_2\dots e_m}$. Let $S_n = \sum_{m=1}^n Y_m$, so $\mathbb{E}(S_1^2) = \mathbb{E}(Y_1^2) \leq V$. By property b), S_{n-1} is constant on each set $Z_{e_1\dots e_n}$. Suppose by induction that $\mathbb{E}(S_{n-1}^2) \leq (n-1)V$, then

$$\begin{aligned}\mathbb{E}(S_n^2) &= \mathbb{E}(S_{n-1}^2) + \mathbb{E}(Y_n^2) + 2\mathbb{E}(Y_n S_{n-1}) \\ &\leq (n-1)V + V + 2 \sum_{Z_{e_1\dots e_n}} \mathbb{E}(Y_n S_{n-1} \mid Z_{e_1\dots e_n}) \\ &\leq nV + 2 \sum_{Z_{e_1e_2\dots e_n}} S_{n-1} \cdot \mathbb{E}(Y_n \mid Z_{e_1\dots e_n}) = nV.\end{aligned}$$

By the Chebyshev inequality $P(S_n > n\delta) \leq \frac{nV}{n^2\delta^2} = \frac{V}{n\delta^2}$. In particular $P(S_{n^2} > n^2\delta) \leq \frac{V}{n^2\delta^2}$. Therefore $\sum_n P(S_{n^2} > n^2\delta^2) < \infty$ and by the Borel-Cantelli Lemma, $P(S_{n^2} > n^2\delta^2 \text{ i.o.}) = 0$. As δ is arbitrary, $\frac{S_{n^2}}{n^2} \to 0$ a.s. For the intermediate values of n, let $D_n = \max_{n^2<k<(n+1)^2} |S_k - S_{n^2}|$. Because $|S_k - S_{n^2}| = |\sum_{j=n^2+1}^k X_j|$, $\mathbb{E}(|S_k - S_{n^2}|^2) \leq (k-n^2)V \leq 2nV$. Hence

$$\mathbb{E}(D_n^2) \leq \mathbb{E}(\sum_{k=n^2+1}^{(n+1)^2-1} |S_k - S_{n^2}|^2) \leq \sum_{k=n^2+1}^{(n+1)^2-1} 2nV = 4n^2V.$$

Using Chebyshev's inequality again we obtain $P(D_n \geq n^2\delta) \leq \frac{4n^2V}{n^4\delta^2} = \frac{4V}{n^2\delta^2}$. By the Borel-Cantelli Lemma, $P(D_n \geq n^2\delta \text{ i.o.}) = 0$, and $\frac{D_n}{n^2} \to 0$ a.s. Combining things and taking $n^2 \leq k < (n+1)^2$, we get

$$\frac{S_k}{k} \leq \frac{S_{n^2} + D_n}{n^2} \to 0 \text{ a.s.}$$

Because $X_m \in Y_m + [M-\varepsilon, M+\varepsilon]$ for $m > N$,

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^n X_i &= \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^N X_i + \limsup_{n\to\infty} \frac{1}{n} \sum_{i=N+1}^n X_i \\ &\leq \limsup_{n\to\infty} \frac{1}{n} S_N + \limsup_{n\to\infty} \frac{1}{n} \sum_{i=N+1}^n (Y_i + M + \varepsilon) \\ &\leq \limsup_{n\to\infty} \frac{1}{n} S_N + \limsup_{n\to\infty} \frac{n-N}{n}(M+\varepsilon) \leq M + \varepsilon.\end{aligned}$$

The other inequality is proved similarly. □

An additional lemma is needed to deal with the a-dependence of the acip.

Lemma 9. *Let A be an interval, and let $M : A \to \mathbb{R}$ and $g_n : A \to \mathbb{R}$ be functions with the following properties:*
a) M is continuous a.e. on A.
Let $A(a_0, \varepsilon) = \{a \in A \mid \limsup_{n\to\infty} |g_n(a) - M(a_0)| \leq \varepsilon\}$.
b) If $\varepsilon > 0$, then a.e. $a_0 \in A$ is a density point of $A(a_0, \varepsilon)$.
Then $\lim_{n\to\infty} g_n(a) = M(a)$ a.e.

Proof. Set $B_k = \{a \in A \mid \limsup_{n\to\infty} |g_n(a) - M(a)| \geq \frac{1}{k}\}$. Assume by contradiction that there exists k such that $|B_k| > 0$. Take $\varepsilon < \frac{1}{3k}$ and let $a_0 \in B_k$ be a density point, both of B_k and of $A(a_0, \varepsilon)$. Assume also that M is continuous in a_0. Let A' be a neighbourhood of a_0 which is so small that

- $|M(a) - M(a_0)| \leq \varepsilon$ for all $a \in A'$,
- $|A' \cap A(a_0, \varepsilon)| \geq \frac{3}{4}|A'|$ and
- $|A' \cap B_k| \geq \frac{3}{4}|A'|$.

Then $a \in A' \cap A(a_0, \varepsilon) \cap B_k \neq \emptyset$ and for all $a \in A' \cap A(a_0, \varepsilon) \cap B_k$,

$$\limsup_{n\to\infty} |g_n(a) - M(a)| \leq \limsup_{n\to\infty} |g_n(a) - M(a_0)| + |M(a) - M(a_0)| \leq 2\varepsilon < \frac{1}{k}.$$

This contradicts that $a \in B_k$, proving the lemma. □

7. Condition i)

Choose $B \in \mathcal{B}$. Hence ∂B is a closed zero-measured set.

Lemma 10. *Choose $\varepsilon > 0$, $a_0 \in \mathcal{A}$, $k_1 \in \mathbb{N}$ and $0 \leq k_2 < s_{k_1}(a_0)$ arbitrary. Let for a close or equal to a_0, $B'(a) = T_a^{-k_2}(B) \cap J_{k_1}(a)$. Then there exists a neighbourhood $A \ni a_0$ such that*

$$\limsup_{n\to\infty} \left| \frac{1}{n} \#\{0 \leq i < n \mid \Phi_i(a) \in B'(a)\} - \frac{|B'(a_0)|}{|J(a_0)|} \right| \leq \varepsilon.$$

Proof of Lemma 10. Suppose we have chosen $a_0 \in \mathcal{A}$ and $\varepsilon > 0$. Let $\mathcal{J} = \{J_i\}_i$ be the partition of $J(a_0)$ into branch-domains of F_{a_0}. The partition $\mathcal{J} \vee F_{a_0}^{-1}\mathcal{J} \vee F_{a_0}^{-2}\mathcal{J} \vee \ldots$ contains no nondegenerate intervals. Furthermore, as $B \in \mathcal{B}$, also $\partial B'(a_0)$ is a closed set of zero Lebesgue measure. Therefore we can find N and an neighbourhood U of $\partial B'(a_0)$ with the following properties:

- $|U| \leq \frac{\varepsilon}{8}$.
- U consists of a finite number of intervals, say U_i, $i = 1, ..., L$.
- The boundary-points of each U_i are boundary points of cylinder-sets in $\mathcal{J} \vee F_{a_0}^{-1}\mathcal{J} \vee ... \vee F_{a_0}^{-N}\mathcal{J}$.

In this way, we have chosen at most $2L$ cylinder sets, say K_i, $i = 1, ..., 2L$ which determine the neighbourhood U in a topological way: U can be defined persistently under small changes of the parameter. Let us write $U = U(a)$.

Let $Z_{e_1 e_2 \dots e_n} \subset A$ denote a branch-domain of Φ_n. Fix $R \in \mathbb{N}$ and an interval $A \ni a_0$ such that

- $J_{k_1}(a)$ persists as a varies in A.
- $dis(\Phi_r, Z_{e_1 \dots e_r}) \leq 1 + \frac{\varepsilon}{4}$ for every $r \geq R$ and every branch-domain $Z_{e_1 \dots e_r}$ such that $Z_{e_1 \dots e_r} \cap A \neq \emptyset$.
- The intervals K_i, $i = 1, \dots, 2L$, persist as a varies in A, and $|U(a)| \leq \frac{\varepsilon}{4}$ for all $a \in A$.
- $|\frac{|B'_a|}{|J(a)|} - \frac{|B'_{a_0}|}{|J(a_0)|}| \leq \frac{\varepsilon}{4}$ for all $a \in A$.

Let

$$\tilde{X}_r^+ = \begin{cases} 1 & \text{if } \Phi_r(a) \in B'(a) \cup U(a), \\ 0 & \text{otherwise.} \end{cases}$$

and

$$\tilde{X}_r^- = \begin{cases} 1 & \text{if } \Phi_r(a) \in B'(a) \setminus U(a), \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\tilde{X}_r^\pm$ are constant on $Z_{e_1 \dots e_{r+N}}$. We claim that for any set $Z_{e_1 e_2 \dots e_r} \subset A$,

$$\mathbb{E}(\tilde{X}_r^+ \mid Z_{e_1 e_2 \dots e_r}) \leq \frac{|B'_{a_0}|}{|J(a_0)|} + \varepsilon.$$

Here the expectation is taken with respect to normalized Lebesgue measure on A. Indeed, we have

$$\begin{aligned} \mathbb{E}(\tilde{X}_r^+ \mid Z_{e_1 \dots e_r}) &\leq (1 + \frac{\varepsilon}{4}) \frac{|B'(a) \cup U(a)|}{|J(a)|} \\ &\leq (1 + \frac{\varepsilon}{4})(\frac{|B'_a|}{|J(a)|} + \frac{\varepsilon}{4}) \\ &\leq (1 + \frac{\varepsilon}{4})(\frac{|B'_{a_0}|}{|J(a_0)|} + \frac{\varepsilon}{2}) \leq \frac{|B'_{a_0}|}{|J(a_0)|} + \varepsilon. \end{aligned}$$

Similarly one shows that

$$\mathbb{E}(\tilde{X}_r^- \mid Z_{e_1 \dots e_r}) \geq \frac{|B'_{a_0}|}{|J(a_0)|} - \varepsilon.$$

The variances of $\tilde{X}_r^+$ and $\tilde{X}_r^-$ are clearly bounded. We can use Lemma 8 for $M = \frac{|B'_{a_0}|}{|J(a_0)|}$, $X_i^\pm = \tilde{X}_{iN+j}^\pm$ and the corresponding partitions $\{Z_{e_1 \dots e_{iN+j}}\}$. It follows that

$$M - \varepsilon \leq \liminf_{i \to \infty} \frac{1}{m} \sum_{i=0}^{m-1} X_i^- \leq \limsup_{m \to \infty} \frac{1}{m} \sum_{i=0}^{m-1} X_i^+ \leq M + \varepsilon.$$

Since this is true for $j = 1, 2, \dots, N$, also

$$M - \varepsilon \leq \liminf_{m \to \infty} \frac{1}{m} \sum_{i=0}^{m-1} \tilde{X}_i^- \leq \limsup_{m \to \infty} \frac{1}{m} \sum_{i=0}^{m-1} \tilde{X}_i^+ \leq M + \varepsilon.$$

Because

$$\sum_{i=0}^{m-1} \tilde{X}_i^- \leq \#\{0 \leq i < m \mid \Phi_i(a) \in B'_a\} \leq \sum_{i=0}^{m-1} \tilde{X}_i^+,$$

the lemma follows. $\square$

Proposition 2. *Let B, J_{k_1}, B' and A be as above. Then for a.e. $a \in A$,*

$$\lim_{n\to\infty} \#\{0 \le k < n \mid \Phi_k(a) \in B'_a\} = \frac{|B'_a|}{|J(a)|}.$$

Proof. Combine the previous lemma and Lemma 9. Clearly $a \mapsto \frac{|B'_a|}{|J(a)|}$ is continuous in A and we can indeed use Lemma 9, with $M(a) = \frac{|B'(a)|}{|J(a)|}$ and $g_n = \frac{1}{n}\#\{0 \le i \le n \mid \Phi_i(a) \in B'(a)\}$. □

8. Condition ii)

Condition ii) can be proved exactly as condition i). In fact, we recover it by taking $B = I$, $i = i$ and $j = 0$ in condition i).

9. Condition iii)

Let for $a \in \mathcal{A}$, $M(a) = \sum_i s_i(a)|J_j(a)|$. Let as before $Z_{e_1\dots e_m}$ be the set of parameters a such that $\Phi_{j-1}(a) \in J_{e_j}(a)$ for $1 \le j \le m$.

Lemma 11. *Let $a_0 \in \mathcal{A}$. For every $\varepsilon > 0$ there exists N, a neighbourhood $A \ni a_0$ and sets $W_n \subset A$ such that*

- *For every $n \ge N$, $|W_n| \le \mathcal{O}(a_0^{-n})|A|$,*
- *For every $n \ge N$ and $Z_{e_1\dots e_n} \subset A$*

$$|\mathbb{E}(s \circ \Phi_n \mid Z_{e_1\dots e_n} \setminus W_n) - M(a_0)| \le \varepsilon.$$

- *Moreover, there exists V, independently of ε, such that*

$$Var(s \circ \Phi_n \mid Z_{e_1\dots e_n} \setminus W_n) \le V.$$

Proof. Let $a_0 \in \mathcal{A}$. Choose ε arbitrary. By Lemma 7, one can find $C_2, \delta > 0$, such that for every $a \in (a_0 - \delta/2, a_0 + \delta/2)$ we have $|\bigcup_{s_j(a)=n} J_i(a)| \le C_2 r^n$, where $r = \frac{2-\delta}{2-\delta/2} < 1$. Choose t_0 so that

$$\sum_{t>t_0} \sum_{s\ge t} s r^s \le \frac{\varepsilon}{8C_2}. \tag{4}$$

Next choose N so large that $\frac{\varepsilon}{2C_1} \gg a^{-N/2}$ and also so large that for every $n \ge N$ and every $Z_{e_1\dots e_n}$ satisfying $Z_{e_1\dots e_n} \cap (a_0 - \delta/2, a_0 + \delta/2) \ne \emptyset$,

$$dis(\Phi \mid Z_{e_1\dots e_n}) \le 1 + \frac{\varepsilon}{8C_1}.$$

Here C_1 is taken from Lemma 2, so it is an upper bound for $\sum_i s_i|J_i(a)|$ for each $a \in (a_0 - \delta/2, a_0 + \delta/2)$. Finally choose a neighbourhood $a_0 \in A \subset (a_0 - \delta/2, a_0 + \delta/2)$ so small that for every $a \in A$, and every

j such that $s_j < t_0$, $J_j(a)$ persists in A, no new branch-domain of transfer-time $s_j < t_0$ is created, and

$$(5) \qquad 1 - \frac{\varepsilon}{8s_j 2^j} \leq \frac{|J_j(a)|}{|J_j(a_0)|} \leq 1 + \frac{\varepsilon}{8s_j 2^j}.$$

Take from now on $n \geq N$ and $a \in A$. Let $J_j(a) \ni \Phi_n(a)$. If $s_j < t_0$, then by (4) and (5),

$$|J_j(a)|(1-\frac{\varepsilon}{8s_j 2^j})(1-\frac{\varepsilon}{8C_1}) \leq \frac{|Z_{e_1\dots e_n j}|}{|Z_{e_1\dots e_n}|} \leq |J_j(a)|(1+\frac{\varepsilon}{8s_j 2^j})(1+\frac{\varepsilon}{8C_1}).$$

If $s_j \geq t_0$, we don't know whether $J_j(a)$ persists in A. An extra set of arguments is necessary.

Let $(a_1, a_2) = Z_{e_1\dots e_n} \subset A$ be any cylinder. By Lemma 4, there exists m such that $c_m(a_1) = c_1(a_1)$ or $c_2(a_1)$. Hence $c_2(a_1) \in T_{a_1}^{-m+\gamma}(c)$ for $\gamma \in \{1, 2\}$. Let $x(a)$ be the continuation of this preimage in (a_1, a_2). Let

$$W_{e_1\dots e_n} = \{a \in Z_{e_1\dots e_n} \mid \Phi_n(a) < x(a)\}.$$

As $|x(a_2) - c_2(a_2)| \approx |Z_{e_1\dots e_n}|$, it follows that $|W_{e_1\dots e_n}| \approx |Z_{e_1\dots e_n}|^2$. Next take $W_n = \bigcup_{Z_{e_1\dots e_n} \subset A} W_{e_1\dots e_n}$. As $|Z_{e_1\dots e_n}| \leq a^{-n}$, it follows that $W_n \leq \mathcal{O}(a^{-n})|A|$, as asserted.

¿From now on we concentrate on parameters $a \in Z_{e_1\dots e_n} \setminus W_n$. Assume $\Phi_n(a) \in J_i(a)$, where $s_i \geq t_0$. We will try to reconstruct what happens to $J_i(a)$ as a moves down to a_1. Because $J_i(a) \geq x(a)$ we can indeed trace back J_i and remain in the core $[c_2(a), c_1(a)]$. As we remarked in section 4, $\frac{d}{da}|J_i(a)| < 0$. If $J_i(a)$ already existed at a_1, then $|J_i(a_1)| \geq |J_i(a)|$. If $J_i(a)$ is created between a_1 and a, then it was created from countably many merging branch-domains with larger transfer-times. Each of these domains may have been created in another merging process and so on. But in any case, way arrive at

$$|\bigcup_{s_i \geq t} J_i(a)| \leq |\bigcup_{s_i \geq t} J_i(a_1)| \leq C_2 \sum_{s \geq t} r^s.$$

Using the small distortion of Φ_n, we obtain

$$\sum_{\substack{s_j \geq t \\ Z_{e_1\dots e_n j} \not\subset W_{e_1\dots e_n}}} t|Z_{e_1\dots e_n j}| \leq C_2 |Z_{e_1\dots e_n} \setminus W_{e_1\dots e_n}| \sum_{s \geq t} (1 + \frac{\varepsilon}{8C_1}) s r^s$$

Combining all this, we get

$$\begin{aligned}
&\mathbb{E}(s(\Phi_n(a)) \mid Z_{e_1\dots e_n} \setminus W_n) \\
&\le \sum_{t<t_0} t \sum_{s_j=t} \frac{|Z_{e_1\dots e_n j}|}{|Z_{e_1\dots e_n} \setminus W_n|} + \sum_{t\ge t_0} \sum_{\substack{s_j \ge t \\ Z_{e_1\dots e_n j} \not\subset W_{e_1\dots e_n}}} s_j \frac{|Z_{e_1\dots e_n j}|}{|Z_{e_1\dots e_n} \setminus W_n|} \\
&\le \sum_{s_j<t_0} s_j |J_j(a_0)| \frac{|Z_{e_1\dots e_n}|}{|Z_{e_1\dots e_n} \setminus W_{e_1\dots e_n}|} (1+\frac{\varepsilon}{8 s_j 2^j})(1+\frac{\varepsilon}{8C_1}) \\
&\qquad + \sum_{t\ge t_0} \sum_{s\ge t} s(1+\frac{\varepsilon}{8C_1}) C_2 r^s \\
&\le \sum_{s_j<t_0} s_j |J_j(a_0)| (1+\mathcal{O}(1)|Z_{e_1\dots e_n}|) + \frac{\varepsilon}{2} \le M(a_0)+\varepsilon.
\end{aligned}$$

A similar proof shows that also $\mathbb{E}(s(\Phi_n(a)) \mid Z_{e_1\dots e_n} \setminus W_n) \ge M(a_0) - \varepsilon$. For the variance one obtains:

$$\begin{aligned}
Var(s(\Phi_n(a) \mid Z_{e_1\dots e_n} \setminus W_n) &\le \mathbb{E}(s(\Phi_n(a))^2 \mid Z_{e_1\dots e_n} \setminus W_n) \\
&\le \mathcal{O}(1) \sum_t \sum_{s\ge t} s^2 C_2 r^s < \infty.
\end{aligned}$$

□

Proposition 3. *For a.e.* $a \in [\sqrt{2}, 2]$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(\Phi_i(a)) = \sum_i s_i(a)|J_i(a)|.$$

In other words, condition iii) is fulfilled for $x = c_3(a)$ *for a.e.* $a \in [\sqrt{2}, 2]$.

Proof. Take a_0 as in the previous lemma. Apply Lemma 8 with $X_m = s(\Phi_m(a_0))$ on $A \setminus \bigcup_{n\ge N} W_n$. Then the conditions of Lemma 8 are satisfied. For every $\varepsilon > 0$

$$(6)\quad \limsup_n |\frac{1}{n} \sum_{i=0}^{n-1} s(\Phi_i(a)) - M(a_0)| \le \varepsilon \quad \text{for a.e. } a \in A \setminus \bigcup_{n\ge N} W_n.$$

Now $\frac{|\bigcup_{n\ge N} W_n|}{|A|} \le \mathcal{O}(1) \sum_{n\ge N} a^{-n} = \mathcal{O}(a^{-N}) \to 0$ as $N \to \infty$. Because (6) is true for every N, we indeed obtain

$$\limsup_n |\frac{1}{n} \sum_{i=0}^{n-1} s(\Phi_i(a)) - M(a_0)| \le \varepsilon \quad \text{for a.e. } a \in A.$$

Let us show that $M : [\sqrt{2}, 2] \to \mathbb{R}$ is continuous in a_0. Let $\eta > 0$ be arbitrary. Find a neighbourhood $A \ni a_0$ such that for each $a \in A$ the following properties hold:

- The integer $N > 0$ (by Lemma 7) such that

$$\sum_{s_j(a)>N} s_j(a)|J_i(a)| \leq \frac{\eta}{3},$$

- No interval J_j with $s_j \leq N$ are created as a varies in A.
- For each j such that $s_j(a) \leq N$,

$$||J_j(a)| - |J_j(a_0)|| \leq \frac{\eta}{2^j}.$$

Then it follows that $|M(a) - M(a_0)| < \eta$ for all $a \in A$, proving continuity.

Hence we can apply Lemma 9, with $g_n(a) = \frac{1}{n}\sum_{i=0}^{n-1} s(\Phi_i(a))$. The proposition follows. $\square$

References

[BC] M. Benedicks, L. Carleson, *On iterations of* $1 - ax^2$ *on* $(-1, 1)$., Ann. Math. (1985), 1-25.

[BGMY] L. Block, J. Guckenheimer, M. Misiurewicz, L.-S. Young, *Periodic points and topological entropy of one dimensional maps*, in LNM 819 (1980), 18-34.

[BM] K. Brucks, M. Misiurewicz, *Trajectory of the turning point is dense for almost all tent maps*, Erg. Th. and Dyn. Sys. **16** (1996), 1173-1183.

[B] H. Bruin, *Induced maps, Markov extension and invariant measures in one-dimensional dynamics*, Commun. Math. Phys. **168** (1995), 571-580.

[B2] H. Bruin, *Combinatorics of the kneading map,*, Int. Jour. of Bifur. and Chaos **5** (1995), 1339-1349.

[DGP] B. Derrida, A. Gervois, Y. Pomeau, *Iteration of endomorphisms on the real axis and representations of numbers*, Ann. Inst. Henri Poincaré. Physique Théorique **XXIX 3** (1978), 305-356.

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

[K] G. Keller, *Lifting measures to Markov extensions*, Monath Math. **108** (1989), 183-200.

[MS] W. de Melo, S. van Strien, *One-dimensional dynamics*, Ergebnisse der Mathematik und ihrer Grensgebiete (1993), Springer Verlag, Berlin and New York.

[M] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Publ. Math. I.H.E.S. **53** (1981), 17-51.

[P] K. R. Parthasaraty, *Probability Measures on Metric Spaces.*, Acad. Press, New York, London (1967).

[Sa] D. Sands, *Topological conditions for positive Lyapunov exponent in unimodal maps*, Ph.D. Thesis, Cambridge (1994).

[Sc] J. Schmeling, *Symbolic dynamics for* β*-shifts and self-normal numbers*, Ergod. Th. and Dyn. Sys. **17** (1997), 675-694.

[T] H. Thunberg, *Absolutely continuous invariant measures and superstable periodic orbits: weak*-convergence of natural measures*, in Ph.D. Thesis, Stockholm (1996).