Intermittency and number expansions for random interval maps
Zeegers, B.P.

PART I

STATISTICAL PROPERTIES OF CRITICALLY INTERMITTENT SYSTEMS

CHAPTER 2

Absolutely continuous invariant measures for critically intermittent systems

This chapter is based on: [HKR+22].

Abstract

Critical intermittency stands for a type of intermittent dynamics caused by an interplay of a superstable fixed point and a repelling fixed point. We consider a large class of random interval maps that exhibit critical intermittency and demonstrate the existence of a phase transition when varying probabilities, where the absolutely continuous invariant measure changes between finite and infinite. We discuss further properties of this measure and show that its density is not in L^q for any $q > 1$. This provides a theory of critical intermittency alongside the theory for the well-studied Manneville-Pomeau maps where the intermittency is caused by a neutral fixed point.

§2.1 Introduction

The concept of critical intermittency has been illustrated in Subsection 1.1.1 by random i.i.d. applications of the two logistic maps $L_2(x) = 2x(1-x)$ and $L_4(x) = 4x(1-x)$ on the unit interval: Orbits converge superexponentially fast to $\frac{1}{2}$ under applications of L_2, and as soon as L_4 is applied then diverge exponentially fast from the repelling fixed point, behaving chaotically again for some time once escaped. See Figures 1.1(b) and 1.2(b). The alternation between chaotic periods and being in a seemingly steady state is related to the probability of choosing the maps as well as the critical order $\ell_2 = 2$ of L_2, which determines the speed of convergence to $\frac{1}{2}$. In [AGH18] it is shown that this random system admits a σ-finite acs measure which is infinite if the probability p_2 of choosing L_2 satisfies $p_2 > \frac{1}{2}$. We will see in this chapter that this acs measure in fact is finite if and only if $p_2 \cdot \ell_2 < 1$. More generally, we will prove such a phase transition for the acs measure for a large family of random interval maps with critical intermittency.

In this chapter we consider critically intermittent systems on $[0,1]$ that are defined by random i.i.d. applications of so-called *bad* maps that share a globally superattracting fixed point $c \in (0,1)$ and *good* maps that map c into the common invariant and repelling set $\{0,1\}$. To be precise, the families of maps we consider are defined as follows.

Throughout the text we fix a point $c \in (0,1)$ that will represent the single critical point of our maps, both good and bad.

A map $T_g : [0,1] \to [0,1]$ is in the class of *good maps*, denoted by $\mathfrak{G}$, if

(G1) $T_g|_{[0,c)}$ and $T_g|_{(c,1]}$ are C^3 diffeomorphisms onto $[0,1)$ or $(0,1]$ and $T_g(c) \in \{T_g(c-), T_g(c+)\}$;

(G2) T_g has non-positive Schwarzian derivative on $[0,c)$ and $(c,1]$;

(G3) to T_g we can associate three constants $r_g \geq 1$, $0 < K_g < 1$ and $M_g > r_g$ such that, for each $x \in [0,1]$,

$$K_g|x-c|^{r_g-1} \leq |DT_g(x)| \leq M_g|x-c|^{r_g-1}; \tag{2.1}$$

(G4) we have $|DT_g(0)|, |DT_g(1)| > 1$.

These conditions imply in particular that $T_g(\{0,c,1\}) \subseteq \{0,1\}$, that at least one of the maps $T_g|_{[0,c]}$ or $T_g|_{[c,1]}$ is continuous, and that both branches of T_g are strictly monotone. Note also that the conditions $K_g < 1$ and $M_g > r_g$ are superfluous, since we can always choose a smaller constant K_g and larger constant M_g to satisfy (2.1), but we need these specific bounds in our estimates later. The critical point c is mapped to either 0 or 1 under each of the good maps and both 0 and 1 are (eventually) fixed points or periodic points (with period 2) by (G1) that are repelling by (G4). Furthermore, a consequence of the Minimum Principle is that $|DT_g|$ has locally no strict minima in the intervals $(0,c)$ and $(c,1)$. In particular, there cannot be any attracting fixed points for T_g in $(0,c)$ and $(c,1)$. Examples of good maps include the doubling map and surjective unimodal maps, see Figures 2.1(a)-(d).

The choice of conditions (G1)-(G4) is based on two factors: firstly, these conditions incorporate the most important properties of the 'good' logistic map $L_4(x) = 4x(1-x)$, which is the primary motivating example for this chapter, and secondly, some of the techniques used in this chapter are motivated by the work of Nowicki and Van Strien [NvS91] where the following result has been proven. Let λ denote the Lebesgue measure on $[0,1]$.

Theorem 2.1.1 (Main Theorem in [NvS91]). *Suppose that $T : [0,1] \to [0,1]$ is unimodal, C^3, has negative Schwarzian derivative and that the critical point of T is of order $r \geq 1$. Moreover assume that the growth rate of $|DT^n(c_1)|$, $c_1 = T(c)$, is so fast that*

$$\sum_{n=0}^{\infty} |DT^n(c_1)|^{-1/r} < \infty. \tag{2.2}$$

Then T has a unique probability acim μ which is ergodic and of positive entropy. Furthermore, there exists a positive constant K such that

$$\mu(B) \leq K \cdot \lambda(B)^{1/r}, \tag{2.3}$$

for any Borel set $B \subseteq [0,1]$. Finally, the density $\frac{d\mu}{d\lambda}$ of the measure μ with respect to λ is an $L^{\tau-}$-function where $\tau = r/(r-1)$ and $L^{\tau-} = \bigcap_{1 \leq t < \tau} L^t$ and $L^t = \big\{ f \in L^1([0,1],\lambda) : \int_0^1 |f|^t d\lambda < \infty \big\}$.

Formally this result is not immediately applicable to the good maps we introduced. The difference, however, is not principal and the conclusion remains exactly the same, the main reason being that the conditions (G1) and (G4) imply the growth rate (2.2), and hence any good map admits a unique probability acim.

A map $T_b : [0,1] \to [0,1]$ is in the class of *bad maps*, denoted by $\mathfrak{B}$, if

(B1) $T_b|_{[0,c)}$ and $T_b|_{(c,1]}$ are C^3 diffeomorphisms onto $[0,c)$ or $(c,1]$ and $T_b(c) = c$;

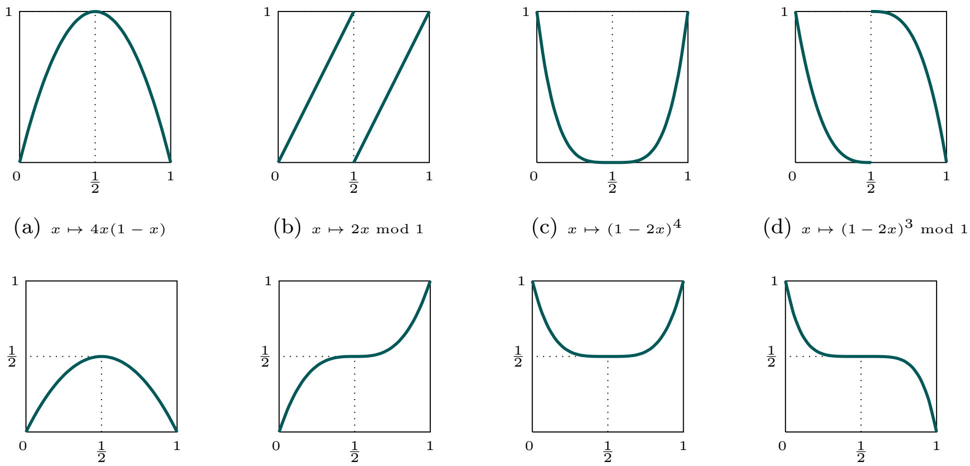
(B2) T_b has non-positive Schwarzian derivative on $[0,c)$ and $(c,1]$;

(B3) to T_b we can associate three constants $\ell_b > 1$, $0 < K_b < 1$ and $M_b > \ell_b$ such that, for each $x \in [0,1]$,

$$K_b |x-c|^{\ell_b - 1} \leq |DT_b(x)| \leq M_b |x-c|^{\ell_b - 1}; \tag{2.4}$$

(B4) we have $|DT_b(0)|, |DT_b(1)| > 1$.

In particular (B1) implies that $T_b(\{0,1\}) \subseteq \{0,1\}$, that T_b is continuous, and that T_b is strictly monotone on the intervals $[0,c]$ and $[c,1]$. In contrast to (G3), note that in (B3) we have assumed that ℓ_b is not equal to one. This means that $DT_b(c) = 0$, so c is a superstable fixed point for each bad map. Furthermore, c has $(0,1)$ as its basin of attraction, since it again follows from the Minimum Principle that $|DT_b|$ has locally no strict minima in the intervals $(0,c)$ and $(c,1)$. Combining this with the Poincaré Recurrence Theorem, an immediate consequence is that the only finite T_b-invariant



(a) $x \mapsto 4x(1-x)$ (b) $x \mapsto 2x \bmod 1$ (c) $x \mapsto (1-2x)^4$ (d) $x \mapsto (1-2x)^3 \bmod 1$

(e) $x \mapsto 2x(1-x)$ (f) $x \mapsto \frac{1}{2} + 4(x-\frac{1}{2})^3$ (g) $x \mapsto \frac{1}{2} + 8(x-\frac{1}{2})^4$ (h) $x \mapsto \frac{1}{2} - 16(x-\frac{1}{2})^5$

Figure 2.1: Eight maps with critical point $c = \frac{1}{2}$. The upper four graphs (a)-(d) *show good maps, while in* (e)-(h) *we see the graphs of four bad maps.*

measures are linear combinations of Dirac measures at $0, c$, and 1. For examples, see Figures 2.1(e)-(h).

The random systems we consider in this chapter are the following. Let $T_1, \ldots, T_N \in \mathfrak{G} \cup \mathfrak{B}$ be a finite collection of good and bad maps. Write $\Sigma_G = \{1 \le j \le N : T_j \in \mathfrak{G}\}$ and $\Sigma_B = \{1 \le j \le N : T_j \in \mathfrak{B}\}$ for the index sets of the good and bad maps respectively and assume that $\Sigma_G, \Sigma_B \neq \emptyset$. Write $\Sigma = \{1, \ldots, N\} = \Sigma_G \cup \Sigma_B$. Let F be the skew product associated to $\{T_j\}_{j \in \Sigma}$, i.e.

$$F : \Sigma^{\mathbb{N}} \times [0,1] \to \Sigma^{\mathbb{N}} \times [0,1],\ (\omega, x) \mapsto (\tau\omega, T_{\omega_1}(x)), \tag{2.5}$$

where τ denotes the left shift on sequences in $\Sigma^{\mathbb{N}}$. Let $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ be a probability vector representing the probabilities with which we choose the maps T_j, $j \in \Sigma$, and let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $\Sigma^{\mathbb{N}}$. Our main results are the following.

Theorem 2.1.2. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ a strictly positive probability vector.*

- (a) *There exists a unique (up to scalar multiplication) σ-finite acs measure μ for $(\mathcal{T}, \boldsymbol{p})$. Moreover, F is ergodic w.r.t. $m_{\boldsymbol{p}} \times \mu$.*
- (b) *The density $\frac{d\mu}{d\lambda}$ is bounded away from zero, is locally Lipschitz on $(0, c)$ and $(c, 1)$ and is not in L^q for any $q > 1$.*

Theorem 2.1.3. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ a strictly positive probability vector. Let μ be the unique acs measure from Theorem 2.1.2. Set $\theta = \sum_{b \in \Sigma_B} p_b \ell_b$. Then μ is finite if and only if $\theta < 1$. In this case, there exists a* *constant $C > 0$ such that*

$$\mu(B) \leq C \cdot \sum_{k=0}^{\infty} \theta^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}} \tag{2.6}$$

for any Borel set $B \subseteq [0,1]$, where $r_{\max} = \max\{r_g : g \in \Sigma_G\}$ and $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$.

As we shall see in (2.28) the bound in (2.6) can be improved by not bounding mixtures $\ell_{\boldsymbol{b}} r_g = \prod_{i=1}^k \ell_{b_i} r_g$ by their maximal value $\ell_{\max}^k r_{\max}$, but this improvement does not change the qualitative behaviour of the bound.

Theorem 2.1.3 shows that the system undergoes a phase transition where the acs measure changes between finite and infinite with threshold $\theta = 1$. Interestingly, this situation is significantly different than for the random LSV maps discussed in Example 1.4.3 where the existence of an acs probability measure only depends on whether there is a positive probability to choose an LSV map with parameter < 1.

It is also worth mentioning that it follows from Theorem 2.1.3 that not only does no finite acs measure exist if $\theta \geq 1$, but also that no physical measure[1] for F exists in this case. Indeed, this follows by applying Aaronson's Ergodic Theorem [A97, Theorem 2.4.2] to the infinite measure $m_{\boldsymbol{p}} \times \mu$.

Since μ is an acs measure, the density $\frac{d\mu}{d\lambda}$ is a fixed point of the associated Perron-Frobenius operator being of the form as in (1.20). Moreover, Theorem 2.1.2 tells that this density $\frac{d\mu}{d\lambda}$ is bounded away from zero. Using these two statements it is easy to see that $\frac{d\mu}{d\lambda}$ blows up to infinity at the points zero and one and also at least on one side of c. See Figure 2.2 for an example. Furthermore, Theorem 2.1.3 says that $\frac{d\mu}{d\lambda}$ is integrable if and only if θ is small enough, namely $\theta < 1$. This intuitively makes sense since for a smaller value of θ the attraction of orbits to c is weaker on average and consequently orbits typically spend less time near zero and one once a good map is applied.

The inequality (2.6) is the counterpart of the Nowicki-Van Strien inequality (2.3), and naturally gives a substantially worse bound due to the presence of bad maps. It is not immediately clear how much worse (2.6) is in comparison to (2.3). However, the following holds.

Corollary 2.1.4. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ a strictly positive probability vector. Suppose $\theta = \sum_{b\in\Sigma_B} p_b \ell_b < 1$. Then there exist $K > 0$ and $\varkappa > 0$ such that for any Borel set $B \subseteq [0,1]$ with $\lambda(B) \in (0,1)$ one has*

$$\mu(B) \leq K \frac{1}{\log^{\varkappa}(1/\lambda(B))}.$$

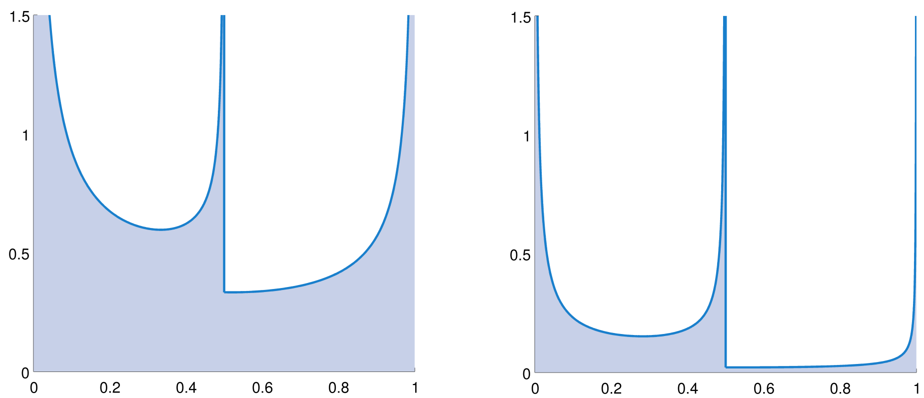
[1] A probability measure μ is a *physical measure* for a transformation $T : X \to X$ if it is T-invariant and there exists a set $U \subseteq X$ of positive Lebesgue measure (for our setting, this means $m_{\boldsymbol{p}} \times \lambda(U) > 0$) such that for each continuous function $f : X \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int_X f \, d\mu, \qquad \text{for each } x \in U.$$

Moreover, the acs measure from Theorem 2.1.2 depends continuously on the probability vector $\boldsymbol{p} \in \mathbb{R}^N$ as the next result shows. Here we write $\mu_{\boldsymbol{p}}$ for the acs probability measure that corresponds to the probability vector $\boldsymbol{p}$.

Corollary 2.1.5. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above. For each $n \geq 0$, let $\boldsymbol{p}_n = (p_{n,j})_{j \in \Sigma}$ be a strictly positive probability vector such that $\sup_n \sum_{b \in \Sigma_B} p_{n,b} \ell_b < 1$ and assume that $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ in $\mathbb{R}^N_+$. Then the sequence $\mu_{\boldsymbol{p}_n}$ converges weakly to $\mu_{\boldsymbol{p}}$.*

As discussed in Section 1.4 the problem of finding acs probability measures for random interval maps that are expanding on average is well studied and these results often rely on bounded variation techniques from Lasota and Yorke [LY73]. In [P84, Section 4] these techniques are extended to a class of i.i.d. random interval maps on the unit interval that are not expanding on average and are composed of a uniformly expanding map T and a contracting map S such that S contracts no faster than that T expands. An example is $T(x) = 2x \bmod 1$ and $S(x) = \frac{x}{2}$. Under additional conditions on S and T and assuming that the probability of choosing T is bigger than S, Pelikan shows that such a random map admits an acs probability measure. The proof of this result relies heavily on the explicit expression that Pelikan has for a conjugacy map between this random map and another random map that is expanding on average and consists of two Lasota-Yorke type maps. This allows Pelikan to find the order of the density of the acs measure near the pole at 0. As we see in Theorem 2.1.2 and prove in Subsection 2.2.3, for our systems the density of the acs measure is not in L^q for any $q > 1$. This suggests that a similar conjugacy for our systems, if it exists, might not have such a nice explicit expression. Therefore, we resort to techniques similar to the ones used by Nowicki and Van Strien in their proof of Theorem 2.1.1 rather than the techniques introduced by Lasota and Yorke.



(a) (b)

Figure 2.2: Approximation of $\frac{d\mu}{d\lambda}$ in case $\Sigma_G = \{1\}$, $\Sigma_B = \{2\}$, $T_1(x) = L_4(x) = 4x(1-x)$ and $T_2(x) = L_2(x) = 2x(1-x)$ for two different values of p_1. Both pictures depict $P^{20}_{\mathcal{T},\boldsymbol{p}}(1)$ with Perron-Frobenius operator $P_{\mathcal{T},\boldsymbol{p}}$, where in (a) *we have taken $p_1 = \frac{3}{4}$ and in* (b) *$p_1 = \frac{1}{4}$.*

To be more precise, for the existence result from Theorem 2.1.2 we use an inducing scheme. This approach is inspired by [AGH18], but the choice of the inducing domain needed some care. With the help of Kac's Lemma we then obtain that the acs measure is infinite in case $\theta \geq 1$. To prove that this measure is finite for $\theta < 1$ we use an approach similar to the one employed in [NvS91] by estimating the sizes of preimages of neighborhoods around points in the postcritical orbits. For this we apply the Minimum Principle and Koebe Principle to iterates of the maps T_j, which is possible due to (G2) and (B2). The main difficulty to obtain these estimates is that it may take an arbitrarily long time before the superattracting fixed point is mapped onto the repelling orbit by one of the good maps, which decreases the regularity of the density of the acs measure. Furthermore, we will use the following key lemma.

Recall the constants ℓ_b, K_b and M_b from (2.4) in condition (B3) and set $\ell_{\min} = \min\{\ell_b : b \in \Sigma_B\}$ and $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$. We prove the next lemma in Subsection 2.2.3.

Lemma 2.1.6. *For all $n \in \mathbb{N}$, $\omega \in \Sigma_B^{\mathbb{N}}$ and $x \in [0,1]$,*

$$\left(\tilde{K}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}} \leq |T_\omega^n(x) - c| \leq \left(\tilde{M}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}},$$

with $\tilde{K} = \left(\frac{\min\{K_b\,:\,b\in\Sigma_B\}}{\ell_{\max}}\right)^{\frac{1}{\ell_{\min}-1}} \in (0,1)$ *and* $\tilde{M} = \left(\frac{\max\{M_b\,:\,b\in\Sigma_B\}}{\ell_{\min}}\right)^{\frac{1}{\ell_{\min}-1}} > 1$.

It follows that under iterations of bad maps the distance $|T_\omega^n(x) - c|$ is eventually decreasing superexponentially fast in n. In Section 2.3 we will use the upper bound on $|T_\omega^n(x) - c|$ that we obtained in Lemma 2.1.6 to prove that μ in Theorem 2.1.3 is infinite if $\theta \geq 1$. The lower bound from Lemma 2.1.6 will be used to show that μ is finite if $\theta < 1$.

In (B3) we have assumed that for any bad map T_b the corresponding value ℓ_b is not equal to one. Note that a bad map T_b for which we allow $\ell_b = 1$ satisfies $|DT_b(c)| > 0$, so in this case c is an attracting fixed point for T_b but not superattracting. It should not come as a surprise that results similar to Theorem 2.1.2 and Theorem 2.1.3 also hold in case some or all of the bad maps T_b have $\ell_b = 1$. The proofs presented for these theorems, however, do not immediately carry over. This has mainly to do with the constants $\tilde{K}$ and $\tilde{M}$ from Lemma 2.1.6, which are not well defined in case $\ell_{\min} = 1$. In the last section we explain how the results are affected in case some or all maps T_b satisfy $\ell_b = 1$ and what the necessary changes in the proofs are.

The remainder of this chapter is organised as follows. Section 2.2 is devoted to the proof of Theorem 2.1.2 and in Section 2.3 we prove Theorem 2.1.3. In Section 2.4 we prove Corollaries 2.1.4 and 2.1.5 and explain what the analogues of Theorem 2.1.2 and 2.1.3 are in case $\ell_b = 1$ for one or more $b \in \Sigma_B$ and how the proofs of Theorem 2.1.2 and 2.1.3 need to be modified to get these results. We end this chapter with some final remarks.

§2.2 Existence of a σ-finite acs measure

From now on we fix an integer $N \geq 2$ and consider a finite collection $T_1, \ldots, T_N \in \mathfrak{G} \cup \mathfrak{B}$ of good and bad maps in the classes $\mathfrak{G}$ and $\mathfrak{B}$. As in Section 2.1 we write $\Sigma_G = \{1 \leq j \leq N : T_j \in \mathfrak{G}\}$ and $\Sigma_B = \{1 \leq j \leq N : T_j \in \mathfrak{B}\}$ for the corresponding index sets and assume that $\Sigma_G, \Sigma_B \neq \emptyset$. We write $\Sigma = \{1, 2, \ldots, N\}$. In this section we prove Theorem 2.1.2, i.e. we establish the existence of an ergodic acs measure and several of its properties using an inducing scheme for the random system F. We fix the index $g \in \Sigma_G$ of one good map T_g and start by constructing an inducing domain that depends on this g. Throughout this section and the next ones we use the notations for words and compositions of the maps T_j introduced in Section 1.4.

§2.2.1 The induced system and first return time partition

The first lemma is needed to specify the set on which we induce. For each $k \in \mathbb{N}$ let x_k and x'_k in $(0, c)$ denote the critical points of T_g^k closest to 0 and c, respectively. Furthermore, let y_k and y'_k in $(c, 1)$ denote the critical points of T_g^k closest to 1 and c, respectively.

Lemma 2.2.1. *We have $x_k \downarrow 0$, $x'_k \uparrow c$, $y'_k \downarrow c$, $y_k \uparrow 1$ as $k \to \infty$.*

Proof. Let a and b denote the critical points of T_g^2 in $(0, c)$ and $(c, 1)$, respectively. Then at least one of the branches $T_g^2|_{(0,a)}$ and $T_g^2|_{(b,1)}$ is increasing. Suppose that $T_g^2|_{(0,a)}$ is increasing. It then follows from the Minimum Principle that $T_g^2(x) \geq \min\{\frac{x}{a}, DT_g^2(0) \cdot x\}$ for each $x \in [0, a]$. To see this, suppose there is an $x \in (0, a)$ with $T_g^2(x) < \min\{\frac{x}{a}, DT_g^2(0) \cdot x\}$. Then there must be a $y \in (0, x)$ with $DT_g^2(y) < \min\{DT_g^2(0), \frac{1}{a}\}$ and a $z \in [x, a)$ with $DT_g^2(z) > \frac{1}{a}$. On the other hand, by the Minimum Principle, $DT_g^2(y) \geq \min\{DT_g^2(0), DT_g^2(z)\}$, a contradiction. Combining this with $DT_g^2(0) > 1$ and defining $L : (0, 1) \to (0, a)$ by $L = (T_g^2|_{(0,a)})^{-1}$, we see that $L^k(a) \downarrow 0$ as $k \to \infty$. Furthermore, define $R : (0, 1) \to (b, 1)$ by $R = (T_g^2|_{(b,1)})^{-1}$. If $T_g^2|_{(b,1)}$ is increasing, we see that similarly $R^k(b) \uparrow 1$ as $k \to \infty$. On the other hand, if $T_g^2|_{(b,1)}$ is decreasing, we have $RL^k(a) \uparrow 1$ as $k \to \infty$. Finally, if $T_g^2|_{(0,a)}$ is decreasing, then $T_g^2|_{(b,1)}$ must be increasing, which yields $LR^k(b) \downarrow 0$ as $k \to \infty$. We conclude that $x_k \downarrow 0$ and $y_k \uparrow 1$ as $k \to \infty$. It follows from (G1) that c is a limit point of both of the sets $\bigcup_{k \in \mathbb{N}} (T_g|_{(0,c)})^{-1}(\{x_k, y_k\})$ and $\bigcup_{k \in \mathbb{N}} (T_g|_{(c,1)})^{-1}(\{x_k, y_k\})$. So $x'_k \uparrow c$, $y'_k \downarrow c$ as $k \to \infty$. $\square$

By the previous lemma and (G1), for $k \in \mathbb{N}$ large enough it holds that

$$\begin{aligned} &T_g(x'_k) \leq x'_k \text{ or } T_g(x'_k) \geq y'_k, \text{ and} \\ &T_g(y'_k) \leq x'_k \text{ or } T_g(y'_k) \geq y'_k, \end{aligned} \tag{2.7}$$

and, using also (G4), (B1) and (B4), for every $j \in \Sigma$,

$$\begin{aligned} &T_j\big([0, x_k] \cup [y_k, 1]\big) \subseteq [0, x'_k) \cup (y'_k, 1] \text{ and} \\ &|DT_j(x)| > d > 1 \;\text{ for all } x \in [0, x_k) \cup (y_k, 1] \;\text{ and some constant } d. \end{aligned} \tag{2.8}$$

Fix a $\kappa \in \mathbb{N}$ for which (2.7) and (2.8) hold. We introduce some notation. Let $t \in \Sigma$ be such that $t \neq g$, and define

$$\begin{aligned} C &= [\underbrace{g \cdots g}_{\kappa \text{ times}}\, t] = [g^\kappa t], \\ J_0 &= (x_\kappa, x'_\kappa), \quad J_1 = (y'_\kappa, y_\kappa), \quad J = J_0 \cup J_1, \\ Y &= C \times J. \end{aligned}$$

Lemma 2.2.2. *The set Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \lambda$.*

Proof. For $m_{\boldsymbol{p}}$-almost all $\omega \in \Sigma^{\mathbb{N}}$ we have $\tau^n \omega \in [g]$ for infinitely many $n \in \mathbb{N}$. For any such n and each $x \in (0,c) \cup (c,1)$ either $T_\omega^n(x) \in J$ or $T_\omega^n(x) \notin J$. If $T_\omega^n(x) \in (0, x_\kappa] \cup [y_\kappa, 1)$, then it follows from (2.8) that there is an $m \geq 1$ such that $T_\omega^{n+m}(x) \in J$. If $T_\omega^n(x) \in [x'_\kappa, c) \cup (c, y'_\kappa]$ it follows from (2.7) that $T_\omega^{n+1}(x) = T_g \circ T_\omega^n(x) \in (0, x'_\kappa] \cup [y'_\kappa, 1)$, which means that we are in the first case if $T_\omega^{n+1}(x) \notin J$. Hence, there exists a measurable set $A \subseteq \Sigma^{\mathbb{N}} \times [0,1]$ with $m_{\boldsymbol{p}} \times \lambda(A) = 1$ such that for each $(\omega, x) \in A$ we have $T_\omega^n(x) \in J$ for infinitely many $n \in \mathbb{N}$.

We define

$$\mathcal{E} = A \setminus \bigcup_{n=0}^{\infty} F^{-n} Y$$

and for each $x \in [0,1]$ we define

$$\mathcal{E}_x = \Big\{ \omega \in \Sigma^{\mathbb{N}} : (\omega, x) \in A \setminus \bigcup_{n=0}^{\infty} F^{-n} Y \Big\},$$

which is the x-section of $\mathcal{E}$. It follows from Fubini's Theorem that $\mathcal{E}_x$ is measurable for λ-almost all $x \in [0,1]$ and that

$$m_{\boldsymbol{p}} \times \lambda(\mathcal{E}) = \int_{[0,1]} m_{\boldsymbol{p}}(\mathcal{E}_x) d\lambda(x).$$

Combining this with $m_{\boldsymbol{p}} \times \lambda(A) = 1$, it remains to show that $m_{\boldsymbol{p}}(\mathcal{E}_x) = 0$ holds for λ-almost all $x \in [0,1]$ for which $\mathcal{E}_x$ is measurable.

Let $x \in [0,1]$ for which $\mathcal{E}_x$ is measurable. According to the Lebesgue Differentiation Theorem (see e.g. [T04]) we have that $m_{\boldsymbol{p}}$-almost all $\omega \in \Sigma^{\mathbb{N}}$ is a Lebesgue point of the function $1_{\mathcal{E}_x}$. Consider such an ω and suppose that $\omega \in \mathcal{E}_x$. Then $(\omega, x) \in A$, so there exists an increasing sequence $(n_j)_{j \in \mathbb{N}}$ in $\mathbb{N}$ that satisfies $T_\omega^{n_j}(x) \in J$ for each $j \in \mathbb{N}$. Recall that τ denotes the left shift on sequences. If $\omega' \in \tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]$, then $T_{\omega'}^{n_j}(x) = T_\omega^{n_j}(x) \in J$ and so $F^{n_j}(\omega', x) \in Y$, which gives $\omega' \notin \mathcal{E}_x$. So $\mathcal{E}_x$ and $\tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]$ are disjoint for each $j \in \mathbb{N}$, which together with ω being a Lebesgue point of $1_{\mathcal{E}_x}$ yields that

$$\begin{aligned} 1 \;&\geq\; \frac{m_{\boldsymbol{p}}\big((\mathcal{E}_x \cup \tau^{-n_j} C) \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} \\ &= \frac{m_{\boldsymbol{p}}\big(\mathcal{E}_x \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} + \frac{m_{\boldsymbol{p}}\big(\tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} \xrightarrow{j \to \infty} 1_{\mathcal{E}_x}(\omega) + m_{\boldsymbol{p}}(C). \end{aligned}$$

Since $m_{\boldsymbol{p}}(C) > 0$, we find that $\omega \in \mathcal{E}_x$ gives a contradiction. We conclude that $m_{\boldsymbol{p}}(\mathcal{E}_x) = 0$. □

Since F is non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$, it follows from Lemma 2.2.2 that in particular

$$m_{\boldsymbol{p}} \times \lambda\Big(Y \setminus \bigcup_{n=1}^{\infty} F^{-n}Y\Big) \le m_{\boldsymbol{p}} \times \lambda\Big(F^{-1}\Big(X \setminus \bigcup_{n=0}^{\infty} F^{-n}Y\Big)\Big) = 0.$$

Hence, the first return time map φ_Y of the form as in (1.6) and the induced transformation F_Y are well defined on the full measure subset of points in Y that return to Y infinitely often under iterations of F, which we call Y again. The set of points in Y that return to Y after n iterations of F can be described as

$$Y \cap F^{-n}(Y) = \bigcup_{\omega \in C \cap \tau^{-n}C} [\omega_1 \cdots \omega_n] \times (T_\omega^n|_J)^{-1}(J) \mod m_{\boldsymbol{p}} \times \lambda, \tag{2.9}$$

which is empty for $n \le \kappa$. Note that in (2.9) in fact $[\omega_1 \cdots \omega_n] = [g^\kappa t \omega_{\kappa+2} \cdots \omega_n g^\kappa t]$ and that by construction each map $T_\omega^n|_J$ in (2.9) consists of branches that all have range $(0, c)$ or $(c, 1)$ or $(0, 1)$, since any branch of $T_\omega^\kappa|_J$ maps onto $(0, 1)$. Therefore, $Y \cap F^{-n}(Y)$ can be written as a finite union of products $A = [\boldsymbol{u} g^\kappa t] \times I$ of cylinders $[\boldsymbol{u} g^\kappa t] \subseteq C$ with $|\boldsymbol{u}| = n$ and open intervals $I \subseteq J$, each of which is mapped under F^n onto $C \times J_0$ or $C \times J_1$. Call the collection of these sets P_n and let $\alpha = \bigcup_{n>\kappa} P_n$. Let $m_{\boldsymbol{p},C}$ and λ_J denote the normalised restrictions of $m_{\boldsymbol{p}}$ to C and λ to J respectively.

Lemma 2.2.3.

(1) The collection α forms a countable first return time partition *of Y, i.e. $m_{\boldsymbol{p},C} \times \lambda_J(\bigcup_{A \in \alpha} A) = 1$, any two different sets $A, A' \in \alpha$ are disjoint and on any $A \in \alpha$ the first return time map φ_Y is constant.*

(2) Let π denote the canonical projection of $\Sigma^{\mathbb{N}} \times [0, 1]$ onto the second coordinate. Any $x \in J$ is contained in a set $\pi(A)$ for some set $A \in \alpha$.

Proof. The fact that $m_{\boldsymbol{p},C} \times \lambda_J(\bigcup_{A \in \alpha} A) = 1$ follows from Lemma 2.2.2. Furthermore, it is clear from the construction that the first return time map φ_Y is constant on any element of α once we know that any two distinct elements of α are disjoint. To show the latter, note that for $A, A' \in P_n$ this is clear. Suppose there are $1 \le m < n$, $A = [\boldsymbol{u} g^\kappa t] \times I \in P_n$ and $A' = [\boldsymbol{v} g^\kappa t] \times I' \in P_m$ such that $A \cap A' \neq \emptyset$. Since $t \neq g$ we get $n \ge m + \kappa + 1$ and $[\boldsymbol{u} g^\kappa t] = [g^\kappa t v_{\kappa+2} \cdots v_m g^\kappa t u_{m+\kappa+2} \cdots u_n g^\kappa t]$. Moreover, $I \cap \partial I' \neq \emptyset$ or $I = I'$. In both cases, note that $F^{m+\kappa+1}([\boldsymbol{v} g^\kappa t] \times \partial I') \subseteq \Sigma^{\mathbb{N}} \times \{0, 1\}$, so by (G1) and (B1) also $F^n([\boldsymbol{v} g^\kappa t] \times \partial I') \subseteq \Sigma^{\mathbb{N}} \times \{0, 1\}$, contradicting that $F^n(A) \subseteq Y$. This proves (1).

For (2) note that, since α is a partition of Y, for each $x \in J$ it holds that there is an $A = [\boldsymbol{u} g^\kappa t] \times I \in \alpha$ with $x \in I$ or $x \in \partial I$. In the first case there is nothing to prove, so assume that $x \in \partial I$. Then $T_{\boldsymbol{u}}(x) \in \partial J_i$ for some $i \in \{0, 1\}$. From the first part of the proof of Lemma 2.2.2 it then follows that there is an $n > |\boldsymbol{u}|$ and an $\omega \in C$ such that $T_\omega^n(x) \in J$. If we write I' for the interval in $(T_\omega^n)^{-1}(J)$ containing x, then this means that there exists a set $A' = [\boldsymbol{v} g^\kappa t] \times I' \in \alpha$ with $x \in \pi(A')$. □

The second part of Lemma 2.2.3 shows that even though the partition elements of α are disjoint, their projections on the second coordinate are not. The same is true for the first coordinate as the same string $\boldsymbol{u}$ can lead points in J to J_0 and J_1.

§2.2.2 Properties of the induced transformation

It follows from (2.9) and Lemma 2.2.3 that for each $A \in \alpha$ we have either $F_Y(A) = C \times J_0$ or $F_Y(A) = C \times J_1$. For any $[\boldsymbol{u}g^\kappa t] \times I \in \alpha$, the transformation $T_{\boldsymbol{u}}|_I$ is invertible from I to one of the sets J_0 or J_1. Define the operator $\mathcal{P}_{\boldsymbol{u},I} : L^1(J, \lambda_J) \to L^1(J, \lambda_J)$ by

$$\mathcal{P}_{\boldsymbol{u},I} h(x) = \begin{cases} \dfrac{h(T_{\boldsymbol{u}}|_I^{-1}(x))}{\left|DT_{\boldsymbol{u}}|_I(T_{\boldsymbol{u}}|_I^{-1}(x))\right|}, & \text{if } T_{\boldsymbol{u}}|_I^{-1}\{x\} \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

The *random Perron-Frobenius-type operator* $\mathcal{P}_Y : L^1(J, \lambda_J) \to L^1(J, \lambda_J)$ is given by

$$\mathcal{P}_Y = \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}]) \mathcal{P}_{\boldsymbol{u},I}. \tag{2.10}$$

Note that $\mathcal{P}_Y$ is not exactly of the same form as the usual Perron-Frobenius operator in (1.20). Nonetheless, we have the following result.

Lemma 2.2.4. *If $\varphi \in L^1(J, \lambda_J)$ is a fixed point of $\mathcal{P}_Y$, then the measure $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi d\lambda_J$ is invariant for F_Y.*

Proof. Suppose $\varphi \in L^1(J, \lambda_J)$ is a fixed point of $\mathcal{P}_Y$. For each cylinder $K \subseteq C$ and each Borel set $E \subseteq J$ we have

$$\begin{aligned} m_{\boldsymbol{p},C} \times \nu\big(F_Y^{-1}(K \times E)\big) &= \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}g^\kappa t] \cap \tau^{-|\boldsymbol{u}|} K) \nu(I \cap T_{\boldsymbol{u}}^{-1} E) \\ &= m_{\boldsymbol{p},C}(K) \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}]) \int_E \mathcal{P}_{\boldsymbol{u},I} \varphi d\lambda_J \\ &= m_{\boldsymbol{p},C}(K) \int_E \mathcal{P}_Y \varphi d\lambda_J \\ &= m_{\boldsymbol{p},C} \times \nu(K \times E). \end{aligned}$$

This gives the result. □

In Lemma 2.2.5 below we show that a fixed point of $\mathcal{P}_Y$ exists. For $m \in \mathbb{N}$, set $\alpha_m = \bigvee_{j=0}^{m-1} F_Y^{-j} \alpha$. Atoms of this partition are the *m-cylinders of F_Y*. Introducing for each $Z = \bigcap_{j=0}^{m-1} F_Y^{-j}([\boldsymbol{u}_j g^\kappa t] \times I_j)$ in α_m the notation

$$C_Z = \bigcap_{j=0}^{m-1} \tau^{-\sum_{i=0}^{j-1} |\boldsymbol{u}_i|} [\boldsymbol{u}_j g^\kappa t] \quad \text{and} \quad J_Z = \bigcap_{j=0}^{m-1} T^{-1}_{\boldsymbol{u}_0 \boldsymbol{u}_1 \cdots \boldsymbol{u}_{j-1}}(I_j), \tag{2.11}$$

we obtain $Z = C_Z \times J_Z$. Writing $\tau_Z = \tau^{\sum_{i=0}^{m-1} |\boldsymbol{u}_i|}|_{C_Z}$ and $T_Z = T_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{m-1}}|_{J_Z}$ we have $F_Y^m|_Z = \tau_Z \times T_Z$. Each T_Z has non-positive Schwarzian derivative, so we can apply the Koebe Principle. The image $T_Z(J_Z)$ either equals J_0 or J_1. Choose a $\bar\rho > 0$ such that $J_0' := [x_\kappa - \bar\rho, x_\kappa' + \bar\rho] \subseteq (0, c)$ and $J_1' := [y_\kappa' - \bar\rho, y_\kappa + \bar\rho] \subseteq (c, 1)$. There is a canonical way to extend the domain of each T_Z to an interval J_Z' containing J_Z, such that $T_Z(J_Z')$ equals either J_0' or J_1' and $\boldsymbol{S}(T_Z) \leq 0$ on J_Z'. Then by the Koebe Principle, i.e. (1.14) and (1.15), there exist constants $K^{(\bar\rho)} > 1$ and $M^{(\bar\rho)} > 0$ such that for all $m \in \mathbb{N}$, $Z \in \alpha_m$ and $x, y \in J_Z$,

$$\frac{1}{K^{(\bar\rho)}} \leq \frac{DT_Z(x)}{DT_Z(y)} \leq K^{(\bar\rho)}, \tag{2.12}$$

$$\Big|\frac{DT_Z(x)}{DT_Z(y)} - 1\Big| \leq \frac{M^{(\bar\rho)}}{\min\{\lambda(J_0), \lambda(J_1)\}} \cdot |T_Z(x) - T_Z(y)|. \tag{2.13}$$

Note that for the random Perron-Frobenius-type operator from (2.10) we have for each $m \geq 1$ that

$$\mathcal{P}_Y^m = \frac{1}{m_{\boldsymbol{p}}(C)} \sum_{Z \in \alpha_m} m_{\boldsymbol{p},C}(C_Z) \mathcal{P}_{T_Z}, \tag{2.14}$$

where $\mathcal{P}_{T_Z}$ is of the form as in (1.10).

Lemma 2.2.5 (cf. Lemmas V.2.1 and V.2.2 of [dMvS93]). *$\mathcal{P}_Y$ admits a fixed point $\varphi \in L^1(J, \lambda_J)$ that is bounded, Lipschitz and bounded away from zero.*

Proof. For each $m \in \mathbb{N}$ and $x \in J$,

$$\mathcal{P}_Y^m 1(x) = \frac{1}{m_{\boldsymbol{p}}(C)} \sum_{\substack{Z \in \alpha_m: \\ x \in T_Z(J_Z)}} \frac{m_{\boldsymbol{p},C}(C_Z)}{|DT_Z(T_Z^{-1}x)|}.$$

Using the Mean Value Theorem, for all $m \in \mathbb{N}$ and $Z \in \alpha_m$ there exists a $\xi \in J_Z$ such that

$$\frac{\lambda\big(T_Z(J_Z)\big)}{\lambda(J_Z)} = |DT_Z(\xi)|. \tag{2.15}$$

Set $K_1 = \frac{\max\{K^{(\bar\rho)}, M^{(\bar\rho)}\}}{m_{\boldsymbol{p}}(C) \cdot \min\{\lambda(J_0), \lambda(J_1)\}}$, where $\bar\rho$ is as in (2.12) and (2.13). Since $DT_Z(\xi)$ and $DT_Z(y)$ have the same sign for any $y \in J_Z$, (2.15) together with (2.12) implies

$$\mathcal{P}_Y^m 1(x) \leq \sum_{Z \in \alpha_m} \frac{m_{\boldsymbol{p},C}(C_Z)}{m_{\boldsymbol{p}}(C)} \cdot K^{(\bar\rho)} \frac{\lambda(J_Z)}{\lambda(T_Z(J_Z))} \leq K_1 \sum_{Z \in \alpha_m} m_{\boldsymbol{p},C} \times \lambda_J(C_Z \times J_Z) = K_1. \tag{2.16}$$

Moreover, if for $A = [\boldsymbol{u} g^\kappa t] \times I \in \alpha$ we take $x, y \in I$, then for any $Z \in \alpha_m$ it holds that $x \in T_Z(J_Z)$ if and only if $y \in T_Z(J_Z)$. For such Z, let $x_Z, y_Z \in J_Z$ be such that

$T_Z(x_Z) = x$ and $T_Z(y_Z) = y$. Then by (2.13)

$$\begin{aligned} |\mathcal{P}_Y^m 1(x) - \mathcal{P}_Y^m 1(y)| &\le \sum_{\substack{Z\in\alpha_m:\\ x\in T_Z(J_Z)}} \frac{m_{\boldsymbol{p},C}(C_Z)}{m_{\boldsymbol{p}}(C)} \left| \frac{1}{|DT_Z(x_Z)|} - \frac{1}{|DT_Z(y_Z)|} \right| \\ &\le \sum_{\substack{Z\in\alpha_m:\\ x\in T_Z(J_Z)}} m_{\boldsymbol{p},C}(C_Z) \frac{1}{|DT_Z(x_Z)|} K_1 |T_Z(x_Z) - T_Z(y_Z)| \\ &= K_1 \mathcal{P}_Y^m 1(x)|x-y|. \end{aligned} \tag{2.17}$$

Together (2.16) and (2.17) imply that the sequence $\big(\frac{1}{m}\sum_{j=0}^{m-1} \mathcal{P}_Y^j 1\big)_m$ is uniformly bounded and equicontinuous on I for each $A = [\boldsymbol{u}g^\kappa t] \times I$. By Lemma 2.2.3(2) it follows that the same holds on J. Hence, by the Arzela-Ascoli Theorem there exists a subsequence

$$\Big(\frac{1}{m_k}\sum_{j=0}^{m_k-1} \mathcal{P}_Y^j 1\Big)_{m_k}$$

converging uniformly to a function $\varphi : J \to [0,\infty)$ satisfying $\varphi \le K_1$ and for each $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ and $x, y \in I$,

$$|\varphi(x) - \varphi(y)| \le K_1 \varphi(x)|x-y|. \tag{2.18}$$

Hence, φ is bounded and by Lemma 2.2.3(2) it is clear that φ is Lipschitz (with Lipschitz constant bounded by K_1^2). It is readily checked that φ is a fixed point of $\mathcal{P}_Y$, so that $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi\, d\lambda$ is an invariant probability measure for F_Y.

Because of Lemma 2.2.3(2) and because φ is continuous, in order to obtain that φ is bounded away from zero on J it suffices to verify for each $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ that φ on I is bounded away from zero. Suppose that there is $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ for which $\inf_{x\in I} \varphi(x) = 0$. Then from (2.18) it follows that $\varphi(y) = 0$ for all $y \in I$, hence $\nu(I) = 0$. Either $I \subseteq J_0$ or $I \subseteq J_1$. If $I \subseteq J_0$, then for any set $A' = [\boldsymbol{v}g^\kappa t] \times I' \in \alpha$ with $T_{\boldsymbol{v}}(I') = J_0$ it holds that

$$m_{\boldsymbol{p},C} \times \lambda_J(A' \cap F_Y^{-1}A) > 0$$

and, by the F_Y-invariance of $m_{\boldsymbol{p},C} \times \nu$,

$$m_{\boldsymbol{p},C} \times \nu(A' \cap F_Y^{-1}A) \le m_{\boldsymbol{p},C} \times \nu(F_Y^{-1}A) = m_{\boldsymbol{p},C} \times \nu(A) = 0,$$

which together give $\inf_{x\in I'} \varphi(x) = 0$ and therefore, like before, $\nu(I') = 0$. There are sets $A' = [\boldsymbol{v}g^\kappa t] \times I'$ with $I' \subseteq J_1$ and $T_{\boldsymbol{v}}(I') = J_0$, so we can repeat the argument to show that also for any set $A'' = [\boldsymbol{v}g^\kappa t] \times I'' \in \alpha$ with $T_{\boldsymbol{v}}(I'') = J_1$ we have $\nu(I'') = 0$. So $m_{\boldsymbol{p},C} \times \nu(A) = 0$ for all $A \in \alpha$. If $I \subseteq J_1$ we come to the same conclusion. This gives a contradiction, so φ is bounded from below on each interval I. □

It follows from Lemma 2.2.4 that $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi d\lambda_J$ is a finite F_Y-invariant measure. To show that $m_{\boldsymbol{p},C} \times \lambda_J$ is F_Y-ergodic we need the following result, which states that the sets $\pi(A)$ for $A \in \alpha_m$ shrink uniformly to λ-null sets as $m \to \infty$.

Lemma 2.2.6. $\lim_{m\to\infty} \sup\{\lambda_J(J_Z) : Z \in \alpha_m\} = 0.$

Proof. Set $\delta = \sup\{\lambda_J(J_Z) : Z \in \alpha\} < 1$. Let $m \geq 2$ and $Z = \bigcap_{j=0}^{m-1} F_Y^{-j}([\boldsymbol{u}_j g^\kappa t] \times I_j) = C_Z \times J_Z \in \alpha_m$ as in (2.11). Set

$$\tilde{J}_Z = \bigcap_{j=0}^{m-2} T^{-1}_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{j-1}}(I_j),$$

so that $J_Z = \tilde{J}_Z \cap T^{-1}_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(I_{m-1})$. Let J_i, $i \in \{0,1\}$, be such that $T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z) = J_i$. It holds that $T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(J_Z) = I_{m-1}$, so $\lambda(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(J_Z)) \leq \delta$ and thus

$$\lambda(T_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z \setminus J_Z) \geq \lambda(J_i) - \delta.$$

Since $\tilde{J}_Z \setminus J_Z$ consists of at most two intervals, with (2.12) and (1.12) applied to this setting this gives

$$1 - \frac{\lambda_J(J_Z)}{\lambda_J(\tilde{J}_Z)} = \frac{\lambda_J(\tilde{J}_Z \setminus J_Z)}{\lambda_J(\tilde{J}_Z)} \geq \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z \setminus J_Z)}{\lambda_J(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z))} \geq \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(J_i) - \delta}{\lambda_J(J_i)}.$$

Set $K_1 := \max\left\{1 - \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(J_i)-\delta}{\lambda_J(J_i)} : i = 0,1\right\} \in (0,1)$. Then by repeating the same steps, we obtain

$$\lambda_J(J_Z) \leq K_1 \lambda_J(\tilde{J}_Z) \leq \cdots \leq K_1^{m-1} \lambda_J(I_0) < K_1^{m-1},$$

which proves the lemma. □

Lemma 2.2.7. *The measure $m_{\boldsymbol{p},C} \times \lambda_J$ is F_Y-ergodic.*

Proof. Suppose $E \subseteq Y$ with $m_{\boldsymbol{p},C} \times \lambda_J(E) > 0$ satisfies $F_Y^{-1}E = E \bmod m_{\boldsymbol{p},C} \times \lambda_J$. We show that $m_{\boldsymbol{p},C} \times \lambda_J(E) = 1$. The Borel measure ρ on Y given by

$$\rho(V) = \int_V 1_E(\omega, x)\varphi(x) dm_{\boldsymbol{p},C}(\omega) d\lambda_J(x)$$

for Borel sets V is F_Y-invariant. According to Proposition 1.2.12 and Lemma 1.4.1 this yields a stationary measure $\tilde{\mu}$ on $[0,1]$ that is absolutely continuous w.r.t. λ and satisfies $(m_{\boldsymbol{p}} \times \tilde{\mu})|_Y = \rho$. Let $L := \{x \in J : \frac{d\tilde{\mu}}{d\lambda}(x) > 0\}$ denote the support of $\frac{d\tilde{\mu}}{d\lambda}|_J$. Since

$$m_{\boldsymbol{p}} \times \lambda(Y) \cdot \frac{d\tilde{\mu}}{d\lambda}(x) = \frac{d\rho}{dm_{\boldsymbol{p},C} \times \lambda_J}(\omega, x) = 1_E(\omega, x)\varphi(x), \qquad m_{\boldsymbol{p},C} \times \lambda_J\text{-a.e.}$$

we obtain, using that φ is bounded away from zero,

$$E = C \times L \mod m_{\boldsymbol{p},C} \times \lambda_J.$$

To obtain the result, it remains to show that $\lambda_J(J \setminus L) = 0$.

We have $C \times L = \bigcup_{Z \in \alpha_m} C_Z \times (J_Z \cap L)$ and $F_Y^{-m}(C \times L) = \bigcup_{Z \in \alpha_m} C_Z \times T_Z^{-1} L$. From the non-singularity of F_Y w.r.t. $m_{\boldsymbol{p},C} \times \lambda_J$ it follows that for each $m \in \mathbb{N}$,

$$C \times L = E = F_Y^{-m} E = F_Y^{-m}(C \times L) \quad \mod m_{\boldsymbol{p},C} \times \lambda_J,$$

which yields

$$J_Z \cap L = T_Z^{-1} L \quad \mod \lambda_J, \quad \text{for each } Z \in \alpha_m. \tag{2.19}$$

Let $\varepsilon > 0$. Since $\lambda_J(L) > 0$, it follows from Lemma 2.2.6 and the Lebesgue Density Theorem that there are $i \in \{0,1\}$, $m' \in \mathbb{N}$ and $Z' \in \alpha_{m'}$ such that

$$T_{Z'}(J_{Z'}) = J_i \quad \text{and} \quad \lambda_J(J_{Z'} \cap L) \geq (1-\varepsilon)\lambda_J(J_{Z'}).$$

By (2.19), $T_{Z'}^{-1}(J_i \backslash L) = J_{Z'} \backslash L \mod \lambda_J$. The Mean Value Theorem gives the existence of a $\xi \in J_{Z'}$ such that

$$\frac{\lambda_J(T_{Z'}(J_{Z'}))}{\lambda_J(J_{Z'})} = |DT_{Z'}(\xi)|,$$

and from (2.12) it follows that

$$\lambda_J(T_{Z'}(J_{Z'} \backslash L)) = \int_{J_{Z'} \backslash L} |DT_{Z'}| d\lambda \leq K^{(\bar{\rho})} |DT_{Z'}(\xi)| \lambda_J(J_{Z'} \backslash L).$$

Hence,

$$\frac{\lambda_J(J_i \backslash L)}{\lambda_J(J_i)} = \frac{\lambda_J(T_{Z'}(J_{Z'} \backslash L))}{\lambda_J(T_{Z'}(J_{Z'}))} \leq K^{(\bar{\rho})} \frac{\lambda_J(J_{Z'} \backslash L)}{\lambda_J(J_{Z'})} \leq K^{(\bar{\rho})} \varepsilon. \tag{2.20}$$

So, for each $\varepsilon > 0$ we can find $i = i(\varepsilon)$, $Z' = Z'(\varepsilon)$ and $m' = m'(\varepsilon)$ for which (2.20) holds. If for each $\varepsilon_0 > 0$ and each $i_0 \in \{0,1\}$ there exists an $\varepsilon \in (0, \varepsilon_0)$ such that $i(\varepsilon) = i_0$, we obtain from (2.20) that $\lambda_J(J \backslash L) = 0$. Otherwise, there exists $\varepsilon_0 > 0$ and $i_0 \in \{0,1\}$ such that $i(\varepsilon) = i_0$ for all $\varepsilon \in (0, \varepsilon_0)$. Without loss of generality, suppose that $i_0 = 0$. Then (2.20) gives $\lambda_J(J_0 \backslash L) = 0$. By the equivalence of ν and λ_J and the fact that every good map has full branches it follows that

$$m_{\boldsymbol{p},C} \times \nu\big((C \times J_0) \cap F_Y^{-1}(C \times J_1)\big) > 0.$$

Together with the Poincaré Recurrence Theorem this gives that

$$A = \{(\omega, x) \in C \times J_0 : F_Y^m(\omega, x) \in C \times J_1 \text{ for infinitely many } m \in \mathbb{N}\}$$

satisfies $m_{\boldsymbol{p},C} \times \nu(A) > 0$, and therefore $m_{\boldsymbol{p},C} \times \lambda_J(A) > 0$. Together with $\lambda_J(J_0 \backslash L) = 0$ it follows from the Lebesgue Density Theorem that there exists a Lebesgue point $x \in \pi(A) \cap L$ of $1_{\pi(A) \cap L}$. Since $x \in \pi(A)$, for infinitely many $m \in \mathbb{N}$ there exists $Z_m \in \alpha_m$ such that $x \in J_{Z_m}$ and $T_{Z_m}(J_{Z_m}) = J_1$. This again together with Lemma 2.2.6 yields that for each $\varepsilon > 0$ there exist $m \in \mathbb{N}$ and $Z \in \alpha_m$ such that

$$T_Z(J_Z) = J_1 \quad \text{and} \quad \lambda_J(J_Z \cap L) \geq (1-\varepsilon)\lambda_J(J_Z).$$

Similar as before, this gives $\lambda_J(J_1 \backslash L) = 0$, so $\lambda_J(J \backslash L) = 0$. □

§2.2.3 Proof of Theorem 2.1.2

We will first give the proof of Lemma 2.1.6.

Proof of Lemma 2.1.6. It follows from (B3) that for any $j \in \Sigma_B$ and $x \in [0,1]$,

$$|T_j(x) - c| = |T_j(x) - T_j(c)| = \left| \int_c^x DT_j(y)dy \right| \geq \frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} |x-c|^{\ell_j}.$$

By induction we get that for each $n \in \mathbb{N}$ and $\omega \in \Sigma_B^{\mathbb{N}}$,

$$|T_\omega^n(x) - c| \geq \left(\frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} \right)^{1+\sum_{i=0}^{n-2} \ell_{\omega_n} \cdots \ell_{\omega_{n-i}}} \cdot |x-c|^{\ell_{\omega_1} \cdots \ell_{\omega_n}}.$$

From (B3) we see that $\frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} < 1$. The lower bound now follows by observing that

$$\Big(1 + \sum_{i=0}^{n-2} \ell_{\omega_n} \cdots \ell_{\omega_{n-i}}\Big) / (\ell_{\omega_1} \cdots \ell_{\omega_n}) \leq \sum_{i=1}^{n} \frac{1}{\ell_{\min}^i} < \frac{1}{\ell_{\min} - 1}.$$

The result for the upper bound follows similarly, by noticing that in this case from (B3) it follows that $\frac{\max\{M_b \,:\, b \in \Sigma_B\}}{\ell_{\min}} > 1$. □

Together with the results from the previous two subsections we now collected all the ingredients necessary to prove Theorem 2.1.2.

Proof of Theorem 2.1.2. (1) We have constructed a finite F_Y-invariant measure $m_{\boldsymbol{p},C} \times \nu$ which is absolutely continuous with respect to $m_{\boldsymbol{p},C} \times \lambda_J$. Since F is non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$ we can therefore by Proposition 1.2.12 extend $m_{\boldsymbol{p},C} \times \nu$ to a σ-finite F-invariant measure $m_{\boldsymbol{p}} \times \mu$ that is absolutely continuous with respect to $m_{\boldsymbol{p}} \times \lambda$. According to Theorem 1.2.10 what is left to show is that F is conservative and ergodic w.r.t. $m_{\boldsymbol{p}} \times \lambda$.

Since $\mu \ll \lambda$ it follows from Lemma 2.2.2 that Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \mu$. Maharam's Recurrence Theorem then gives that F is conservative with respect to $m_{\boldsymbol{p}} \times \mu$. Furthermore, in the proof of part (2) below we will see that the density of $\frac{d\mu}{d\lambda}$ is bounded away from zero. Hence, $\lambda \ll \mu$ and therefore F is also conservative with respect to $m_{\boldsymbol{p}} \times \lambda$. Furthermore, combining the ergodicity of F_Y with respect to $m_{\boldsymbol{p},C} \times \lambda_J$ and the fact that Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \lambda$ gives using Proposition 1.2.11 that F is ergodic with respect to $m_{\boldsymbol{p}} \times \lambda$.

(2) For the density $\psi := \frac{d\mu}{d\lambda}$ it holds that $\psi|_J = \varphi$. Since we can take κ in the definition of J as large as we want, ψ is locally Lipschitz on $(0,c)$ and $(c,1)$. Moreover, it is a fixed point of the Perron-Frobenius operator $\mathcal{P}_{\mathcal{T},\boldsymbol{p}}$ being of the form as in (1.20) and thus for all $x \in (0,1)$,

$$\psi(x) = \mathcal{P}_{\mathcal{T},\boldsymbol{p}}^{\kappa} \psi(x) \geq p_g^{\kappa} \sum_{y \in J \cap T_g^{-\kappa}\{x\}} \frac{\varphi(y)}{|DT_g^{\kappa}(y)|}.$$

From Lemma 2.2.5 we conclude that ψ is bounded from below by some constant $C > 0$. It remains to show that ψ is not in L^q for any $q > 1$. To see this, fix a $b \in \Sigma_B$. Since ψ is bounded from below by $C > 0$, we have for all $k \in \mathbb{N}_0$ and $x \in [0,1]$ that

$$\psi(x) = \mathcal{P}^{k+1}_{\mathcal{T},\boldsymbol{p}}\psi(x) \geq C \cdot p_g p_b^k \sum_{y \in (T_g T_b^k)^{-1}\{x\}} \frac{1}{|D(T_g T_b^k)(y)|}. \tag{2.21}$$

Let $\ell_b, M_b, r_g, M_g, K_g$ be as in (B3) and (G3). From (B3), (G3) and Lemma 2.1.6 we get

$$\begin{aligned}
|D(T_g T_b^k)(y)| &= |DT_g(T_b^k(y))| \prod_{i=1}^{k} |DT_b(T_b^{k-i}(y))| \\
&\leq M_g |T_b^k(y) - c|^{r_g - 1} \prod_{i=0}^{k-1} (M_b |T_b^i(y) - c|^{\ell_b - 1}) \\
&\leq M_g M_b^k (\tilde{M}|y - c|)^{\ell_b^k (r_g - 1)} \prod_{i=0}^{k-1} (\tilde{M}|y - c|)^{\ell_b^i(\ell_b - 1)} \\
&= K_1 |y - c|^{\ell_b^k r_g - 1},
\end{aligned} \tag{2.22}$$

for the positive constant $K_1 = M_g M_b^k \tilde{M}^{\ell_b^k r_g - 1}$. On the other hand, from (G3) we obtain for any $y \in (T_g T_b^k)^{-1}\{x\}$ as in the proof of Lemma 2.1.6 that

$$|x - T_g(c)| = |T_g T_b^k(y) - T_g(c)| \geq \frac{K_g}{r_g} |T_b^k(y) - c|^{r_g}$$

and then Lemma 2.1.6 yields

$$|x - T_g(c)| \geq K_2 |y - c|^{\ell_b^k r_g} \tag{2.23}$$

for the positive constant $K_2 = \frac{K_g}{r_g} \tilde{K}^{\ell_b r_g}$. Now for any $q > 1$ we can choose $k \in \mathbb{N}_0$ large enough so that $\tau := (1 - \ell_b^{-k} r_g^{-1}) q \geq 1$. Combining (2.21), (2.22) and (2.23) we obtain

$$\begin{aligned}
\psi^q(x) &\geq \Big(\frac{C p_g p_b^k}{K_1}\Big)^q \Big(\sum_{y \in (T_g T_b^k)^{-1}\{x\}} |y - c|^{1 - \ell_b^k r_g} \Big)^q \\
&\geq K_3 |x - T_g(c)|^{-\tau}
\end{aligned}$$

for a positive constant K_3. This gives the result. □

Remark 2.2.8. The result from Theorem 2.1.2 still holds if we allow the critical order ℓ_b from (B3) to be equal to 1 for some b, as long as $\ell_{\max} > 1$. To see this, note that in the proof of Theorem 2.1.2 condition (B3) only plays a role in proving that $\frac{d\mu}{d\lambda} \notin L^q$ for any $q > 1$. Here we refer to Lemma 2.1.6 and the constants $\tilde{K}$ and $\tilde{M}$, which are not well defined if $\ell_{\min} = 1$. In (2.22) however, we use the estimates from

Lemma 2.1.6 only for one arbitrary fixed $b \in \Sigma_B$. By the same reasoning as in the proof of Lemma 2.1.6 it follows that

$$\left(\left(\frac{K_b}{\ell_b}\right)^{\frac{1}{\ell_b-1}}|x-c|\right)^{\ell_b^n} \le |T_b^n(x)-c| \le \left(\left(\frac{M_b}{\ell_b}\right)^{\frac{1}{\ell_b-1}}|x-c|\right)^{\ell_b^n}.$$

for any $b \in \Sigma_B$ with $\ell_b > 1$. Hence, if there exists at least one $b \in \Sigma_B$ with $\ell_b > 1$, then we can replace the bounds obtained from Lemma 2.1.6 in (2.22) and (2.23) by constants $K_1 = M_g M_b^k (\frac{K_b}{\ell_b})^{(\ell_b^k r_g - 1)/(\ell_b-1)}$ and $K_2 = \frac{K_g}{r_g}(\frac{M_b}{\ell_b})^{\ell_b r_g/(\ell_b-1)}$ and obtain the same result. In case $\ell_{\max} = 1$, then most parts from Theorem 2.1.2 still remain valid with the exception that in that case we can only say that $\frac{d\mu}{d\lambda} \notin L^q$ if $q \ge \frac{r_{\max}}{r_{\max}-1}$. This follows from the above reasoning by taking $k = 0$ in the definition of τ in the proof of Theorem 2.1.2 and by noting that $\tau = (1 - r_{\max}^{-1})q \ge 1$ if $q \ge \frac{r_{\max}}{r_{\max}-1}$.

§2.3 Estimates on the acs measure

In this section we prove Theorem 2.1.3. Recall the definition of θ from Theorem 2.1.3:

$$\theta = \sum_{b \in \Sigma_B} p_b \ell_b.$$

§2.3.1 The case $\theta \ge 1$

To prove one direction of Theorem 2.1.3, namely that the unique acs measure μ from Theorem 2.1.2 is infinite if $\theta \ge 1$, we introduce another induced transformation. Furthermore, we use that there exists a $\delta > 0$ such that $|DT_b(x)| < 1$ for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B$. This implies

$$|T_b(x) - c| < |x - c| \tag{2.24}$$

for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B$.

Proposition 2.3.1. *Suppose $\theta \ge 1$. Then the unique acs measure μ from Theorem 2.1.2 is infinite.*

Proof. Fix a $b \in \Sigma_B$. Recall the definitions of $\tilde{M}$ from Lemma 2.1.6 and set $\gamma = \min\{\delta, \frac{1}{2}\tilde{M}^{-1}\}$ with δ as given above. Let $a \in [c-\gamma, c)$. Then there exists a $\xi \in (a, c)$ such that $T_b(a) > \xi$ and $T_b^2(a) > \xi$. Take $[bb] \times (a, \xi)$ as the inducing domain and let

$$\kappa(\omega, x) = \inf\{k \in \mathbb{N} : F^k(\omega, x) \in [bb] \times (a, \xi)\}$$

be the first return time to $[bb] \times (a, \xi)$ under F. If $m_{\boldsymbol{p}} \times \mu([bb] \times (a, \xi)) = \infty$, then there is nothing left to prove. If not, then we compute $\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu$ and use Kac's Lemma, i.e. Lemma 1.2.13, to prove the result.

So, assume that $m_{\boldsymbol{p}} \times \mu([bb] \times (a, \xi)) < \infty$. The conditions that $T_b(a) > \xi$ and $T_b^2(a) > \xi$ together with the fact that any bad map has c as a fixed point and is strictly monotone on the intervals $(0, c)$ and $(c, 1)$, guarantee that for each $n \in \mathbb{N}$ and $\omega \in \Sigma_B^{\mathbb{N}} \cap [bb]$ we get

$$T_\omega^n((a,\xi)) \cap (a,\xi) = \emptyset. \tag{2.25}$$

For any $\omega \in [bb]$ and $x \in (a,\xi)$ it follows by (2.25) and (2.24) that $T_\omega^n(x)$ can only return to (a,ξ) after at least one application of a good map. Assume that $\omega \in [bb]$ is of the form

$$\omega = (b, b, \omega_3, \omega_4, \ldots, \omega_n, g, \omega_{n+2}, \ldots),$$

with $n \geq 2$, $\omega_i \in \Sigma_B$ for $3 \leq i \leq n$, $g \in \Sigma_g$, and $x \in (a,\xi)$. Then $\kappa(\omega, x) \geq n+1$. Lemma 2.1.6 yields that

$$|T_\omega^n(x) - c| \leq (\tilde{M}\gamma)^{\ell_{\omega_1}\cdots\ell_{\omega_n}} < 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n}}. \tag{2.26}$$

From (G3) and (2.26) we obtain that

$$|T_g T_\omega^n(x) - T_g(c)| = \left| \int_c^{T_\omega^n(x)} DT_g(y)\,dy \right| \leq \frac{M_g}{r_g} |T_\omega^n(x) - c|^{r_g} < \frac{M_g}{r_g} \cdot 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n} r_g}. \tag{2.27}$$

Set

$$\zeta = \sup\{|DT_j(x)| \,:\, j \in \Sigma,\, x \in [0,1]\}.$$

Then $\zeta > 1$ by (G4), (B4). Assume $\kappa(\omega, x) = m + n$ for some $m \geq 1$. Then $T_\omega^{m+n}(x) \in (a,\xi)$ so that by (G1),

$$|T_\omega^{m+n}(x) - T_g(c)| \geq \min\{a, 1-\xi\}.$$

Because of (2.27) this implies

$$\zeta^{m-1} \frac{M_g}{r_g} \cdot 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n} r_g} \geq \min\{a, 1-\xi\}.$$

Solving for m yields

$$m \geq K_1 + K_2 \ell_{\omega_1} \cdots \ell_{\omega_n}$$

for constants $K_1 = 1 + \log\left(\frac{\min\{a,1-\xi\} r_g}{M_g}\right)/\log\zeta \in \mathbb{R}$ and $K_2 = \log(2^{r_g})/\log\zeta > 0$. Note that K_1, K_2 are independent of ω, x, m and n.

We obtain that for any $g \in \Sigma_G$,

$$\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu \geq \sum_{n\in\mathbb{N}_{\geq 2}} \sum_{\omega_3,\ldots,\omega_n \in \Sigma_B} m_{\boldsymbol{p}}([bb\omega_3\cdots\omega_n g])\mu((a,\xi))\Big(n + K_1 + K_2 \ell_b^2 \prod_{i=3}^n \ell_{\omega_i}\Big).$$

Since

$$\sum_{n\in\mathbb{N}_{\geq 2}} \sum_{\omega_3,\ldots,\omega_n \in \Sigma_B} m_{\boldsymbol{p}}([\omega_3\cdots\omega_n]) \prod_{i=3}^n \ell_{\omega_i} = 1 + \sum_{n\in\mathbb{N}} \theta^n = \infty,$$

we get $\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu = \infty$ and from Lemma 1.2.13 we now conclude that μ is infinite. □

§2.3.2 The case $\theta < 1$

For the other direction of Theorem 2.1.3, assume $\theta < 1$. We first obtain a stationary probability measure $\tilde{\mu}$ for F as in (2.5) using a standard Krylov-Bogolyubov type argument. For this, let $\mathcal{M}$ denote the set of all finite Borel measures on $[0,1]$, and define the operator $\mathcal{P} : \mathcal{M} \to \mathcal{M}$ by

$$\mathcal{P}\nu = \sum_{j \in \Sigma} p_j \nu \circ T_j^{-1}, \qquad \nu \in \mathcal{M},$$

where $\nu \circ T_j^{-1}$ denotes the pushforward measure of ν under T_j. Then $\mathcal{P}$ is a *Markov-Feller* operator (see e.g. [LMS04]) with dual operator U on the space $BM([0,1])$ of all bounded Borel measurable functions given by[2] $Uf = \sum_{j\in\Sigma} p_j f \circ T_j$ for $f \in BM([0,1])$. As before, let λ denote the Lebesgue measure on $[0,1]$, and set $\lambda_n = \mathcal{P}^n \lambda$ for each $n \geq 0$. Furthermore, for each $n \in \mathbb{N}$ define the Cesáro mean $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \lambda_k$. Since the space of probability measures on $[0,1]$ equipped with the weak topology is sequentially compact, there exists a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ of $(\mu_n)_{n\in\mathbb{N}}$ that converges weakly to a probability measure $\tilde{\mu}$ on $[0,1]$. Using that a Markov-Feller operator is weakly continuous, it then follows from a standard argument that $\mathcal{P}\tilde{\mu} = \tilde{\mu}$, that is, $\tilde{\mu}$ is a stationary probability measure for F. The next theorem will lead to the estimate (2.6) from Theorem 2.1.3. For any $\boldsymbol{u} = u_1 \cdots u_k \in \Sigma^k$, $k \geq 0$, recall that we abbreviate $p_{\boldsymbol{u}} = \prod_{i=1}^k p_{u_i}$ and also let $\ell_{\boldsymbol{u}} = \prod_{i=1}^k \ell_{u_i}$ if $\boldsymbol{u} \in \Sigma_B^k$, where we use $p_{\boldsymbol{u}} = \ell_{\boldsymbol{u}} = 1$ in case $k = 0$.

Theorem 2.3.2. *There exists a constant $C > 0$ such that for all $n \in \mathbb{N}$ and all Borel sets $B \subseteq [0,1]$ we have*

$$\lambda_n(B) \leq C \cdot \sum_{g \in \Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot \lambda(B)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.$$

Before we prove this theorem, we first show how it gives Theorem 2.1.3.

Proof of Theorem 2.1.3. The first part of the statement follows from Proposition 2.3.1. For the second part, assume that $\theta < 1$ and that Theorem 2.3.2 holds. Let $B \subseteq [0,1]$ be a Borel set. Using the regularity of λ, for any $\delta > 0$ there exists an open set $G \subseteq [0,1]$ such that $B \subseteq G$ and $\lambda(G) \leq \lambda(B) + \delta$. Using that $(\mu_{n_k})_{k\in\mathbb{N}}$ converges weakly to $\tilde{\mu}$, we obtain from the Portmanteau Theorem together with Theorem 2.3.2 that

$$\begin{aligned}\tilde{\mu}(B) &\leq \tilde{\mu}(G) \leq \liminf_k \mu_{n_k}(G) \\ &\leq C \cdot \sum_{g \in \Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot (\lambda(B) + \delta)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.\end{aligned}$$

[2] By definition of a Markov-Feller operator, the space of bounded *continuous* functions is required to be invariant under the dual operator U. If there is a $g \in \Sigma_G$ for which T_g is discontinuous (namely at c), we then first identify $[0,1]$ with the unit circle S^1 so that T_g can be viewed as a continuous map on S^1. With the same identification any acs measure on S^1 then gives an acs measure on $[0,1]$.

Since $\theta < 1$, the sum is bounded and with the Dominated Convergence Theorem we can take the limit as $\delta \to 0$ to obtain

$$\tilde{\mu}(B) \le C \cdot \sum_{g \in \Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot \lambda(B)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}. \tag{2.28}$$

This proves that $\tilde{\mu}$ is absolutely continuous with respect to the Lebesgue measure on $[0,1]$. It follows that the probability measure $\tilde{\mu}$ is equal to the unique acs measure μ from Theorem 2.1.2. The estimate (2.6) follows directly from (2.28). □

It remains to give the proof of Theorem 2.3.2. We shall do this in a number of steps.

Proposition 2.3.3. *There exists a constant $K_1 > 0$ such that for all $n \in \mathbb{N}$, all $\boldsymbol{u} \in \Sigma^n$ and all Borel sets $A \subseteq [0,1]$ with $0 < 3\lambda(B) < \frac{1}{2}\min\{c, 1-c\}$ we have*

$$\lambda(T_{\boldsymbol{u}}^{-1}B) \le K_1\big(\lambda(T_{\boldsymbol{u}}^{-1}[0,3\eta)) + \lambda(T_{\boldsymbol{u}}^{-1}(c-3\eta, c+3\eta)) + \lambda(T_{\boldsymbol{u}}^{-1}(1-3\eta, 1])\big),$$

where $\eta = \lambda(B)$.

Proof. Let $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma^n$ and a Borel set $B \subseteq [0,1]$ with $0 < 3\lambda(B) < \frac{1}{2}\min\{c, 1-c\} < 1$ be given and write $\eta = \lambda(B)$. The map $T_{\boldsymbol{u}}$ has non-positive Schwarzian derivative on any of its intervals of monotonicity (see (1.13)) and the interior of the image of any such interval is $(0,c), (c,1)$ or $(0,1)$. Set $B_1 = (\eta, c-\eta)$ and $B_2 = (2\eta, c-2\eta)$. Let I be a connected component of $T_{\boldsymbol{u}}^{-1}B_1$, and set $f = T_{\boldsymbol{u}}|_I$ and $I^* = f^{-1}B_2$. The Minimum Principle yields

$$|Df(x)| \ge \min_{z \in \partial I^*} |Df(z)|, \qquad \text{for all } x \in I^*. \tag{2.29}$$

Suppose the minimal value is attained at $f^{-1}(2\eta)$ and set $B_3 = (2\eta, 3\eta)$ and $J = f^{-1}B_3$. By the condition on the size of B it follows from the Koebe Principle that

$$K^{(\eta)}|Df(f^{-1}(2\eta))| \ge |Df(x)|, \qquad \text{for all } x \in J. \tag{2.30}$$

Combining (2.29) and (2.30) gives

$$\begin{aligned}\lambda(f^{-1}(B \cap B_2)) &= \int_{B \cap B_2} \frac{1}{|Df(f^{-1}y)|}\, d\lambda(y) \le \lambda(B) \cdot \frac{1}{|Df(f^{-1}(2\eta))|} \\ &\le K^{(\eta)} \int_{B_3} \frac{1}{|Df(f^{-1}y)|}\, d\lambda(y) = K^{(\eta)}\lambda(f^{-1}(B_3)).\end{aligned}$$

We conclude that

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (2\eta, c-2\eta)\big)\big) \le K^{(\eta)}\lambda\big(T_{\boldsymbol{u}}^{-1}(2\eta, 3\eta)\big).$$

In case $\min_{z \in \partial I^*} |Df(z)| = f^{-1}(c-2\eta)$, a similar reasoning yields

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (2\eta, c-2\eta)\big)\big) \le K^{(\eta)}\lambda\big(T_{\boldsymbol{u}}^{-1}(c-3\eta, c-2\eta)\big).$$

Furthermore, a similar reasoning can be done for the interval $[c,1]$ to conclude that

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (c+2\eta, 1-2\eta)\big)\big) \le K^{(\eta)}\Big(\lambda\big(T_{\boldsymbol{u}}^{-1}(c+2\eta, c+3\eta)\big) + \lambda\big(T_{\boldsymbol{u}}^{-1}(1-3\eta, 1-2\eta)\big)\Big).$$

Hence, setting $K_1 = \max\{K^{(\eta)}, 1\}$ gives the desired result. □

Proposition 2.3.3 shows that to get the desired estimate from Theorem 2.3.2 it suffices to consider small intervals on the left and right of $[0,1]$ and around c, i.e. sets of the form

$$I_c(\varepsilon) := (c-\varepsilon, c+\varepsilon) \quad \text{and} \quad I_0(\varepsilon) := [0,\varepsilon) \cup (1-\varepsilon, 1]$$

for $\varepsilon > 0$. We first focus on estimating the measure of the intervals $I_c(\varepsilon)$.

Lemma 2.3.4. *There exists a constant $K_2 \geq 1$ such that for all $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma^{n-1} \times \Sigma_G$ and all $\varepsilon > 0$ we have*

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) \leq K_2 \varepsilon.$$

Proof. Let $n \in \mathbb{N}$ and $\boldsymbol{u} \in \Sigma^{n-1} \times \Sigma_G$. Let $\varepsilon > 0$. Suppose that $\varepsilon \geq \frac{1}{4}\min\{c, 1-c\}$. Then

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) \leq 1 \leq \frac{4\varepsilon}{\min\{c,1-c\}}. \tag{2.31}$$

Now suppose $\varepsilon < \frac{1}{4}\min\{c,1-c\}$. Again the map $T_{\boldsymbol{u}}$ has non-positive Schwarzian derivative on the interior of any of its intervals of monotonicity and since $u_n \in \Sigma_G$ the interior of the image of any such interval is $(0,1)$. Use $\mathcal{I}$ to denote the collection of connected components of $T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)$. Let $A \in \mathcal{I}$ and write $J = J_A$ and $I = I_A$ for the intervals that satisfy $A \subseteq J$, $A \subseteq I$ and

$$\begin{aligned} T_{\boldsymbol{u}}(J) &= \Big[c - \frac{1}{2}\min\{c,1-c\}, c + \frac{1}{2}\min\{c,1-c\}\Big], \\ T_{\boldsymbol{u}}(I) &= \Big[c - \frac{3}{4}\min\{c,1-c\}, c + \frac{3}{4}\min\{c,1-c\}\Big]. \end{aligned}$$

Also, write $f = T_{\boldsymbol{u}}|_I$. Since f has non-positive Schwarzian derivative, it follows from (1.12) applied to this setting

$$\frac{\lambda(A)}{\lambda(J)} \leq K^{(\frac{1}{4})} \frac{\lambda(f(A))}{\lambda(f(J))} = K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c,1-c\}}.$$

We conclude that

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) = \sum_{A \in \mathcal{I}} \lambda(A) \leq K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c,1-c\}} \sum_{A \in \mathcal{I}} \lambda(J_A) \leq K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c,1-c\}}. \tag{2.32}$$

Defining $K_2 = \frac{2\max\{2, K^{(\frac{1}{4})}\}}{\min\{c,1-c\}}$, the desired result now follows from (2.31) and (2.32). □

To find $\lambda_n\big(I_c(\varepsilon)\big)$, first note that from Lemma 2.1.6 it follows that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma_B^n$,

$$T_{\boldsymbol{u}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(\tilde{K}^{-1} \varepsilon^{\ell_{\boldsymbol{u}}^{-1}}\big). \tag{2.33}$$

By splitting Σ^n according to the final block of bad indices, we can then write using (2.33) and Lemma 2.3.4 that

$$\begin{aligned}\lambda_n\big(I_c(\varepsilon)\big) &= \sum_{k=0}^{n-1}\sum_{\boldsymbol{v}\in\Sigma^{n-k-1}}\sum_{g\in\Sigma_G}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{v}g\boldsymbol{b}}\lambda\big(T_{\boldsymbol{v}g\boldsymbol{b}}^{-1}I_c(\varepsilon)\big)+\sum_{\boldsymbol{b}\in\Sigma_B^n}p_{\boldsymbol{b}}\lambda\big(T_{\boldsymbol{b}}^{-1}I_c(\varepsilon)\big)\\ &\le \sum_{k=0}^{n-1}\sum_{\boldsymbol{v}\in\Sigma^{n-k-1}}\sum_{g\in\Sigma_G}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{v}g\boldsymbol{b}}\lambda\big(T_{\boldsymbol{v}g}^{-1}I_c(\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}})\big)+\sum_{\boldsymbol{b}\in\Sigma_B^n}p_{\boldsymbol{b}}\lambda\big(I_c(\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}})\big)\\ &\le \sum_{k=0}^{n-1}\sum_{g\in\Sigma_G}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_g p_{\boldsymbol{b}} K_2\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}}+\sum_{\boldsymbol{b}\in\Sigma_B^n}p_{\boldsymbol{b}}2\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}}.\end{aligned}$$

Taking $K_3=\max\big\{K_2, 2\big(\sum_{g\in\Sigma_G}p_g\big)^{-1}\big\}\cdot\tilde{K}^{-1}\ge 1$ then gives

$$\lambda_n\big(I_c(\varepsilon)\big)\le K_3\sum_{g\in\Sigma_G}\sum_{k=0}^{n}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_g p_{\boldsymbol{b}}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}}. \tag{2.34}$$

We now focus on $I_0(\varepsilon)=[0,\varepsilon)\cup(1-\varepsilon,1]$. Fix an $0<\varepsilon_0<\frac{1}{2}\min\{c,1-c\}$ and a $t>1$ that satisfy

$$|DT_j(x)|>t,\qquad \text{for all } x\in I_0(\varepsilon_0) \text{ and each } j\in\Sigma.$$

Such ε_0 and t exist because of (G4) and (B4). From (G3) it follows that for each $0<\varepsilon<\varepsilon_0$ and $g\in\Sigma_G$,

$$|T_g(x)-T_g(c)|=\left|\int_c^x DT_g(y)dy\right|\ge\frac{K_g}{r_g}\cdot|x-c|^{r_g}.$$

Set $K_4=\max\{(K_g^{-1}r_g)^{r_g^{-1}}:g\in\Sigma_G\}\ge 1$. Then (G1) implies that

$$T_g^{-1}I_0(\varepsilon)\subseteq I_0(\varepsilon t^{-1})\cup I_c(K_4\varepsilon^{r_g^{-1}}). \tag{2.35}$$

Furthermore, from (B1) it follows that for each $\varepsilon\in(0,\varepsilon_0)$ and $b\in\Sigma_B$,

$$T_b^{-1}I_0(\varepsilon)\subseteq I_0(\varepsilon t^{-1}). \tag{2.36}$$

Write each $\boldsymbol{u}\in\Sigma^n$ as

$$\boldsymbol{u}=\boldsymbol{b}_1\boldsymbol{g}_1\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}} \tag{2.37}$$

for some $\tilde{s}\in\{1,\dots,n\}$, where for each i we have $\boldsymbol{b}_i=b_{i,1}\cdots b_{i,k_i}\in\Sigma_B^{k_i}$ and $\boldsymbol{g}_i=g_{i,1}\cdots g_{i,m_i}\in\Sigma_G^{m_i}$ for some $k_1,m_{\tilde{s}}\in\mathbb{N}_0$ and $k_2,\dots,k_{\tilde{s}},m_1,\dots,m_{\tilde{s}-1}\in\mathbb{N}$. Define

$$s=\begin{cases}\tilde{s}, & \text{if } m_{\tilde{s}}\ge 1,\\ \tilde{s}-1, & \text{if } m_{\tilde{s}}=0.\end{cases}$$

Moreover, we introduce notation to indicate the length of the tails of the block $\boldsymbol{u}$:

$$\begin{aligned} d_i &= |\boldsymbol{b}_i \boldsymbol{g}_i \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}|, && i \in \{1, \ldots, \tilde{s}\}, \\ q_{i,j} &= |g_{i,j+1} \cdots g_{i,m_i} \boldsymbol{b}_{i+1} \boldsymbol{g}_{i+1} \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}|, && i \in \{1, \ldots, \tilde{s}\},\ j \in \{0, \ldots, m_i\}. \end{aligned}$$

If necessary to avoid confusion, we write $s(\boldsymbol{u})$, $k_i(\boldsymbol{u})$, etc., to emphasise the dependence on $\boldsymbol{u}$.

Lemma 2.3.5. *There exists a constant $K_5 > 0$ such that for each $0 < \varepsilon < \varepsilon_0$, $n \in \mathbb{N}$ and $\boldsymbol{u} = \boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}} \in \Sigma^n$,*

$$\begin{aligned} T_{\boldsymbol{u}}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-d_1}) &\cup \bigcup_{i=1}^{s} T^{-1}_{\boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{i-1} \boldsymbol{g}_{i-1}} I_c\Big(K_5 (\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}}\Big) \\ &\cup \bigcup_{i=1}^{s} \bigcup_{j=2}^{m_i} T^{-1}_{\boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{i-1} \boldsymbol{g}_{i-1} \boldsymbol{b}_i g_{i,1} \cdots g_{i,j-1}} I_c(K_5 (\varepsilon t^{-q_{i,j}})^{r_{g_{i,j}}^{-1}}). \end{aligned}$$

Proof. We prove the statement by an induction argument for $\tilde{s}$. Let $\boldsymbol{u}$ be a word with symbols in Σ, and write $\boldsymbol{u} = \boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}$ for its decomposition as in (2.37). First suppose that $\tilde{s} = 1$. If $m_1 = 0$, then the statement immediately follows from repeated application of (2.36). If $m_1 \geq 1$, then repeated application of (2.35) gives

$$\begin{aligned} T_{\boldsymbol{g}_1}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-q_{1,0}}) &\cup I_c\Big(K_4 (\varepsilon t^{-q_{1,1}})^{r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_4 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

By setting $K_5 = \tilde{K}^{-1} K_4$, applying (2.33) and (2.36) then yields

$$\begin{aligned} T_{\boldsymbol{b}_1 \boldsymbol{g}_1}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-d_1}) &\cup I_c\Big(K_5 (\varepsilon t^{-q_{1,1}})^{\ell_{\boldsymbol{b}_1}^{-1} r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{\boldsymbol{b}_1 g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_5 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

Note that this is true for the case that $k_1 = 0$ as well. This proves the statement if $\tilde{s} = 1$. Now suppose $\tilde{s}(\boldsymbol{u}) > 1$ and suppose that the statement holds for all words $\boldsymbol{v}$ with $\tilde{s}(\boldsymbol{v}) = \tilde{s}(\boldsymbol{u}) - 1$. In particular, the statement then holds for the word $\boldsymbol{b}_2 \boldsymbol{g}_2 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}$. Note that $m_1 \geq 1$. Again, by repeated application of (2.35) it follows that

$$\begin{aligned} T_{\boldsymbol{g}_1}^{-1} I_0(\varepsilon t^{-d_2}) \subseteq I_0(\varepsilon t^{-q_{1,0}}) &\cup I_c\Big(K_4 (\varepsilon t^{-q_{1,1}})^{r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_4 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

Furthermore, applying (2.33) and (2.36) then yields

$$\begin{aligned} T_{\boldsymbol{b}_1 \boldsymbol{g}_1}^{-1} I_0(\varepsilon t^{-d_2}) \subseteq I_0(\varepsilon t^{-d_1}) &\cup I_c\Big(K_5 (\varepsilon t^{-q_{1,1}})^{\ell_{\boldsymbol{b}_1}^{-1} r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{\boldsymbol{b}_1 g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_5 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

This together with the statement being true for the word $\boldsymbol{b}_2\boldsymbol{g}_2\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}$ yields the statement for $\boldsymbol{u}$. ☐

Combining Lemma 2.3.4 and Lemma 2.3.5 gives

$$\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon)) \le 2\varepsilon t^{-d_1} + \sum_{i=1}^{s} K_2K_5(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}} + \sum_{i=1}^{s}\sum_{j=2}^{m_i} K_2K_5(\varepsilon t^{-q_{i,j}})^{r_{g_{i,j}}^{-1}}.$$

Let $r_{\max} = \max\{r_g : g \in \Sigma_G\}$ and set $\alpha := t^{1/r_{\max}} > 1$. Then

$$\sum_{i=1}^{s}\sum_{j=2}^{m_i} \alpha^{-q_{i,j}} \le \sum_{\ell=0}^{\infty} \alpha^{-\ell} = \frac{1}{1-1/\alpha},$$

so that

$$\begin{aligned}\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon)) &\le 2\varepsilon^{1/r_{\max}} + K_2K_5\sum_{i=1}^{s}\sum_{j=2}^{m_i} \varepsilon^{1/r_{\max}}\alpha^{-q_{i,j}} + \sum_{i=1}^{s} K_2K_5(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}} \\ &\le \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{i=1}^{s}(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}}.\end{aligned} \tag{2.38}$$

Proposition 2.3.6. *There exists a constant $K_6 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0)$ and $n \in \mathbb{N}$,*

$$\lambda_n(I_0(\varepsilon)) \le K_6 \sum_{g\in\Sigma_G} p_g \sum_{k=0}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{b}}\ell_{\boldsymbol{b}} \cdot \varepsilon^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.$$

Proof. Let $n \in \mathbb{N}$. Then with (2.38) we obtain

$$\begin{aligned}\lambda_n(I_0(\varepsilon)) &= \sum_{\boldsymbol{u}\in\Sigma^n} p_{\boldsymbol{u}}\lambda\big(T_{\boldsymbol{u}}^{-1}(I_0(\varepsilon))\big) \\ &\le \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{\boldsymbol{u}\in\Sigma^n} p_{\boldsymbol{u}} \sum_{i=1}^{s(\boldsymbol{u})} \big(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\big)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}} \\ &= \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} \\ &\quad + K_2K_5\sum_{i=1}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\big(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\big)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}},\end{aligned} \tag{2.39}$$

where we defined $\tau = \lfloor \frac{n+1}{2} \rfloor$ which is the largest value $s(\boldsymbol{u})$ can take. Let us consider the term in (2.39). First of all, note that a word $\boldsymbol{u} \in \Sigma^n$ satisfies $s(\boldsymbol{u}) \ge 1$ if and only if $m_1(\boldsymbol{u}) \ge 1$. Therefore,

$$\{\boldsymbol{u} \in \Sigma^n : s(\boldsymbol{u}) \ge 1\} = \bigcup_{k=0}^{n-1} \Sigma_B^k \times \Sigma_G \times \Sigma^{n-k-1}.$$

Hence, defining the function χ on $\{0,\dots,n-1\}^2$ by

$$\chi(k,q) = \sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G} p_{\boldsymbol{b}}p_g\big(\varepsilon t^{-q}\big)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}, \qquad (k,q)\in\{0,\dots,n-1\}^2. \tag{2.40}$$

we can rewrite and bound the term with $i=1$ in (2.39) as follows:

$$\begin{aligned}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(1)p_{\boldsymbol{u}}\big(\varepsilon t^{-q_{1,1}(\boldsymbol{u})}\big)^{\ell_{\boldsymbol{b}_1(\boldsymbol{u})}^{-1}r_{g_{1,1}(\boldsymbol{u})}^{-1}} &= \sum_{k=0}^{n-1}\sum_{\boldsymbol{v}\in\Sigma^{n-k-1}} p_{\boldsymbol{v}}\chi(k,n-k-1)\\ &\le \varepsilon^{1/r_{\max}} + \sum_{k=1}^{n-1}\chi(k,n-k-1).\end{aligned} \tag{2.41}$$

Secondly, note that for each $i\in\{2,\dots,\tau\}$ a word $\boldsymbol{u}\in\Sigma^n$ satisfies $s(\boldsymbol{u})\ge i$ if and only if $m_{i-1}(\boldsymbol{u}),k_i(\boldsymbol{u}),m_i(\boldsymbol{u})\ge 1$. For each $k\in\{1,\dots,n-2\}$ and $q\in\{0,\dots,n-k-2\}$ and $i\in\{2,\dots,\tau\}$ we define

$$A_{i,k,q} = \{\boldsymbol{v}\in\Sigma^{n-k-q-1} : \tilde{s}(\boldsymbol{v}) = i-1, v_{n-k-q-1}\in\Sigma_G\}.$$

The set $A_{i,k,q}$ contains all words of length $n-k-q-1$ that can precede the word $\boldsymbol{b}_i\boldsymbol{g}_i\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}$ with $|\boldsymbol{b}_i| = k$ and $|g_{i,2}\cdots g_{i,m_i}\boldsymbol{b}_{i+1}\boldsymbol{g}_{i+1}\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}| = q$. So

$$\{\boldsymbol{u}\in\Sigma^n : s(\boldsymbol{u})\ge i\} = \bigcup_{k=1}^{n-2}\bigcup_{q=0}^{n-k-2} A_{i,k,q}\times\Sigma_B^k\times\Sigma_G\times\Sigma^q, \qquad i\in\{2,\dots,\tau\}.$$

Hence, using (2.40) we can rewrite and bound the sum in (2.39) that runs from $i=2$ to τ as follows:

$$\begin{aligned}&\sum_{i=2}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\big(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\big)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1}r_{g_{i,1}(\boldsymbol{u})}^{-1}}\\ &\quad= \sum_{i=2}^{\tau}\sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\sum_{\boldsymbol{v}_1\in A_{i,k,q}}\sum_{\boldsymbol{v}_2\in\Sigma^q} p_{\boldsymbol{v}_1}p_{\boldsymbol{v}_2}\chi(k,q)\\ &\quad= \sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\chi(k,q)\sum_{i=2}^{\tau}\sum_{\boldsymbol{v}_1\in A_{i,k,q}}\sum_{\boldsymbol{v}_2\in\Sigma^q} p_{\boldsymbol{v}_1}p_{\boldsymbol{v}_2}\\ &\quad\le \sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\chi(k,q).\end{aligned} \tag{2.42}$$

Here the last step follows from the fact that

$$\sum_{i=2}^{\tau}\sum_{\boldsymbol{v}_1\in A_{i,k,q}} p_{\boldsymbol{v}_1} \le \sum_{\boldsymbol{v}\in\Sigma^{n-k-q-2}}\sum_{g\in\Sigma_G} p_{\boldsymbol{v}}p_g \le 1.$$

Combining (2.41) and (2.42) gives

$$\sum_{i=1}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\big(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\big)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}} \leq \varepsilon^{1/r_{\max}} + \sum_{k=1}^{n-1}\sum_{q=0}^{n-k-1}\chi(k,q). \tag{2.43}$$

Furthermore, for each $\boldsymbol{b}\in\Sigma_B^k$ and $g\in\Sigma_G$ we have again by setting $r_{\max}=\max\{r_j : j\in\Sigma_G\}$ and $\alpha = t^{1/r_{\max}}$ that

$$\sum_{q=0}^{n-k-1}(t^{-q})^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}} \leq \sum_{q=0}^{n-k-1}\big(\alpha^{-\ell_{\boldsymbol{b}}^{-1}}\big)^q \leq \frac{1}{1-\alpha^{-\ell_{\boldsymbol{b}}^{-1}}} \leq \frac{\alpha\ell_{\boldsymbol{b}}^{-1}}{\alpha^{\ell_{\boldsymbol{b}}^{-1}}-1}\ell_{\boldsymbol{b}} \leq \frac{\alpha}{\log(\alpha)}\ell_{\boldsymbol{b}}, \tag{2.44}$$

where the last step follows from the fact that $f(x) = \frac{x}{\alpha^x - 1}$ is a decreasing function and $\lim_{x\downarrow 0} f(x) = \frac{1}{\log\alpha}$. Hence, combining (2.39), (2.43) and (2.44) gives

$$\begin{aligned}
&\lambda_n(I_0(\varepsilon))\\
&\leq \left(2+\frac{K_2K_5}{1-1/\alpha}+K_2K_5\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{k=1}^{n-1}\sum_{q=0}^{n-k-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G}p_{\boldsymbol{b}}p_g\big(\varepsilon t^{-q}\big)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}\\
&\leq \left(2+K_2K_5\frac{2\alpha-1}{\alpha-1}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{k=1}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G}p_{\boldsymbol{b}}p_g\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}\frac{\alpha\ell_{\boldsymbol{b}}}{\log(\alpha)}\\
&\leq K_6\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}},
\end{aligned}$$

where $K_6 = \frac{1}{\min\{p_g : g\in\Sigma_G\}}\big(2+K_2K_5\frac{2\alpha-1}{\alpha-1}\big)+\frac{K_2K_5\alpha}{\log\alpha}$. □

We are now ready to prove Theorem 2.3.2.

Proof of Theorem 2.3.2. Let $B\subseteq[0,1]$ be a Borel set. First suppose that $\lambda(B)\geq\frac{\varepsilon_0}{3}$. Then there exists a constant $C = C(\varepsilon_0) > 0$ such that

$$\lambda_n(B)\leq 1\leq C\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\cdot\lambda(B)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}.$$

Now suppose that $\lambda(B) < \frac{\varepsilon_0}{3}$ and set $\varepsilon = 3\lambda(B)$. It follows from Proposition 2.3.3 that for all $n\in\mathbb{N}$ and all $\boldsymbol{u}\in\Sigma^n$ we have

$$\lambda\big(T_{\boldsymbol{u}}^{-1}B\big)\leq K_1\big(\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon))+\lambda(T_{\boldsymbol{u}}^{-1}I_c(\varepsilon))\big).$$

Together with (2.34) and Proposition 2.3.6 this yields for all $n\in\mathbb{N}$ that

$$\lambda_n(B)\leq K_1\cdot(K_3+K_6)\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\cdot\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}.$$

This gives the result. □

§2.4 Further results and final remarks

§2.4.1 Proof of Corollaries 2.1.4 and 2.1.5

In this section we prove Corollaries 2.1.4 and 2.1.5.

Proof of Corollary 2.1.4. We use the bound (2.6) obtained in Theorem 2.1.3. For convenience, we set $\ell = \ell_{\max}$ and $x = \lambda(B)^{1/r_{\max}}$. The asymptotics are determined by the interplay between $\theta^k \searrow 0$ and $x^{1/\ell^k} \nearrow 1$. First suppose $\theta < x^{1/\ell}$. Then $\lambda(B) > \theta^{\ell r_{\max}}$, so for each $\varkappa > 0$ there exists $C > 0$ sufficiently large such that

$$\mu(B) \leq 1 \leq C \cdot \frac{1}{\log^{\varkappa}(1/\lambda(B))}.$$

Now suppose $\theta \geq x^{1/\ell}$. Note that $\theta^N \geq x^{1/\ell^N}$ if and only if

$$\log N + N \log \ell \leq \log\left(\frac{\log x}{\log \theta}\right).$$

Since $\log N \leq N$, this last inequality is satisfied if we take for example

$$N = \left\lfloor \frac{1}{1+\log \ell} \log\left(\frac{\log x}{\log \theta}\right)\right\rfloor = \left\lfloor \frac{1}{1+\log \ell} \log\left(\frac{\log(1/x)}{\log(1/\theta)}\right)\right\rfloor, \tag{2.45}$$

where $\lfloor y \rfloor$ denotes the largest integer not exceeding y. Taking N as in (2.45), note that it follows from $\theta \geq x^{1/\ell}$ that $N \geq 0$. Then $\theta^k \geq x^{1/\ell^k}$ for all $k \leq N$ as well, and hence

$$\begin{aligned}\sum_{k=0}^{\infty} \theta^k x^{1/\ell^k} &= \sum_{k=0}^{N} \theta^k x^{1/\ell^k} + \sum_{k=N+1}^{\infty} \theta^k x^{1/\ell^k} \leq \sum_{k=0}^{N} \theta^k \cdot x^{1/\ell^N} + \sum_{k=N+1}^{\infty} \theta^k \cdot 1 \\ &\leq \frac{1}{1-\theta} x^{1/\ell^N} + \frac{\theta^{N+1}}{1-\theta} \leq \frac{1}{1-\theta}(1+\theta)\theta^N.\end{aligned}$$

From (2.45) we see that $N \geq \frac{1}{1+\log \ell} \log\left(\frac{\log x}{\log \theta}\right) - 1$, thus

$$\begin{aligned}\theta^N &= \exp(N \log \theta) \leq \exp\left(\left(\frac{1}{1+\log \ell} \log\left(\frac{\log(1/x)}{\log(1/\theta)}\right) - 1\right) \log \theta\right) \\ &= \exp\left(\frac{\log \theta}{1+\log \ell} \log\log(1/x) + C(\ell,\theta)\right) \\ &= \overline{C}(\ell,\theta)\big(\log(1/x)\big)^{\frac{\log \theta}{1+\log \ell}} = \overline{C}(\ell,\theta)\left(\frac{r_{\max}}{\log(1/\lambda(B))}\right)^{\varkappa},\end{aligned}$$

where we set $\varkappa = \frac{\log(1/\theta)}{1+\log \ell} > 0$, and where $C(\ell,\theta) \in \mathbb{R}$ and $\overline{C}(\ell,\theta) > 0$ are constants that only depend on ℓ and θ. We conclude from the bound (2.6) that

$$\mu(B) \leq K \cdot \frac{1}{\log^{\varkappa}(1/\lambda(B))}$$

for some positive constant K. □

The proof of Corollary 2.1.5 consists of two steps. Firstly we show that any weak limit point of $\mu_{\boldsymbol{p}_n}$ is a stationary measure, i.e. satisfies (1.18), and secondly that any weak limit point of $\mu_{\boldsymbol{p}_n}$ is absolutely continuous with respect to the Lebesgue measure. The corollary then follows from the uniqueness of absolutely continuous stationary measures given by Theorem 2.1.2.

Proof of Corollary 2.1.5. For each $n \geq 0$, let $\boldsymbol{p}_n = (p_{n,j})_{j\in\Sigma}$ be a strictly positive probability vector such that $\sup_n \sum_{b\in\Sigma_B} p_{n,b}\ell_b < 1$ and assume that $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ in $\mathbb{R}^N_+$ for some strictly positive probability vector $\boldsymbol{p} = (p_j)_{j\in\Sigma}$. Let $\tilde{\mu}$ be a weak limit point of an arbitrary subsequence of $\mu_{\boldsymbol{p}_n}$. Again, note that such a $\tilde{\mu}$ exists because the space of probability measures on $[0,1]$ equipped with the weak topology is sequentially compact. After passing to a further subsequence we have for any continuous function $\varphi : [0,1] \to \mathbb{R}$ that

$$\lim_{n\to\infty} \int_{[0,1]} \varphi \, d\mu_{\boldsymbol{p}_n} = \int_{[0,1]} \varphi \, d\tilde{\mu}.$$

Moreover, by the stationarity of the measures $\mu_{\boldsymbol{p}_n}$ it follows that for each $n \geq 1$,

$$\int_{[0,1]} \varphi \, d\mu_{\boldsymbol{p}_n} = \sum_{j\in\Sigma} p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n}.$$

To prove that $\tilde{\mu}$ is stationary for $\boldsymbol{p}$, it is sufficient to show that for each $j \in \Sigma$,

$$\lim_{n\to\infty} p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n} = p_j \int_{[0,1]} \varphi \circ T_j \, d\tilde{\mu}. \tag{2.46}$$

If $j \in \Sigma_B$ this is obvious, since then $\varphi \circ T_j$ is continuous. For $j \in \Sigma_G$ the map $\varphi \circ T_j$ might have a discontinuity at c. In this case, we let φ_δ be the continuous function given by $\varphi_\delta(x) = \varphi \circ T_j(x)$ for $x \in I \setminus (c-\delta, c+\delta)$ and φ_δ is linear otherwise. Then we have

$$\lim_{n\to\infty} \left| p_{n,j} \int_{[0,1]} \varphi_\delta \, d\mu_{\boldsymbol{p}_n} - p_j \int_{[0,1]} \varphi_\delta \, d\tilde{\mu} \right| = 0,$$

by the weak convergence and since $p_{n,j} \to p_j$ as $n \to \infty$. Also, we have

$$\left| p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n} - p_{n,j} \int_{[0,1]} \varphi_\delta \, d\mu_{\boldsymbol{p}_n} \right| \leq C\mu_{\boldsymbol{p}_n}([c-\delta, c+\delta]) \to 0 \text{ as } \delta \to 0,$$

where the convergence is uniform in n because of (2.6). Similarly,

$$\left| p_j \int_{[0,1]} \varphi \circ T_j d\tilde{\mu} - p_j \int_{[0,1]} \varphi_\delta d\tilde{\mu} \right| \leq C\tilde{\mu}([c-\delta, c+\delta]) \to 0 \text{ as } \delta \to 0,$$

The last three relations imply (2.46).

To show that $\tilde{\mu}$ is absolutely continuous with respect to the Lebesgue measure λ we proceed as in the proof of Theorem 2.1.3. We set $\tilde{\theta} = \sup_n \sum_{b\in\Sigma_B} p_{n,b}\ell_b < 1$. Let $B \subseteq [0,1]$ be a Borel set. Every $\mu_{\boldsymbol{p}_n}$ satisfies the conclusion of Theorem 2.1.3, so

$$\mu_{\boldsymbol{p}_n}(B) \leq C_n \sum_{k=0}^{\infty} \tilde{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}},$$

where the constant C_n depends on $\left(\sum_{g\in\Sigma_G} p_{n,g}\right)^{-1}$ and $(\min\{p_{n,g} : g \in \Sigma_G\})^{-1}$ (and properties of the good and bad maps themselves that are not linked to the probabilities). Since $\boldsymbol{p}$ and each $\boldsymbol{p}_n$, $n \geq 0$, are strictly positive probability vector and $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$, both these quantities can be bounded from above and $\tilde{C} := \sup_n C_n < \infty$. From the weak convergence of $\mu_{\boldsymbol{p}_n}$ to $\tilde{\mu}$ we obtain as in (2.28) using the Portmanteau Theorem that

$$\tilde{\mu}(B) \leq \tilde{C} \sum_{k=0}^{\infty} \tilde{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}}.$$

Hence, $\tilde{\mu} \ll \lambda$. By Theorem 2.1.2 we know that $\mu_{\boldsymbol{p}}$ is the unique acs probability measure for $(\mathcal{T}, \boldsymbol{p})$. So, $\tilde{\mu} = \mu_{\boldsymbol{p}}$. □

§2.4.2 The non-superattracting case

With some modifications the results from Theorem 2.1.2 and Theorem 2.1.3 can be extended to the class $\mathfrak{B}^1 \supseteq \mathfrak{B}$ of bad maps whose critical order ℓ_b in (B3) is allowed to be equal to 1. We will list the modified statements and the necessary modifications to the proofs here. Note that for each $T \in \mathfrak{B}^1 \setminus \mathfrak{B}$, we have $DT(c) \neq 0$, and due to the minimal principle, $|DT(c)| < 1$. We consider $T_1, \ldots, T_N \in \mathfrak{G} \cup \mathfrak{B}^1$ with $\Sigma_B^1 = \{1 \leq j \leq N : T_j \in \mathfrak{B}^1\}$ and Σ_G, Σ_B as before and such that $\Sigma_G, \Sigma_B^1 \setminus \Sigma_B \neq \emptyset$. Furthermore, we write again $\Sigma = \{1, \ldots, N\} = \Sigma_G \cup \Sigma_B^1$. We again set $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$.

Theorem 2.4.1. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ a strictly positive probability vector.*

(a) *There exists a unique (up to scalar multiplication) σ-finite acs measure μ for $(\mathcal{T}, \boldsymbol{p})$. Moreover, F is ergodic w.r.t. $m_{\boldsymbol{p}} \times \mu$ and the density $\frac{d\mu}{d\lambda}$ is bounded away from zero and is locally Lipschitz on $(0, c)$ and $(c, 1)$.*

(b) *Suppose $\ell_{\max} > 1$.*

  (i) *The measure μ is finite if and only if $\theta = \sum_{b\in\Sigma_B^1} p_b \ell_b < 1$. In this case, for each $\hat{\theta} \in (\theta, 1)$ there exists a constant $C(\hat{\theta}) > 0$ such that*

$$\mu(B) \leq C(\hat{\theta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}}$$

  for any Borel set $B \subseteq [0, 1]$, where $r_{\max} = \max\{r_g : g \in \Sigma_G\}$.

  (ii) *The density $\frac{d\mu}{d\lambda}$ is not in L^q for any $q > 1$.*

(c) *Suppose $\ell_{\max} = 1$.*

  (i) *The measure μ is finite, and for each $\boldsymbol{\eta} = (\eta_b)_{b\in\Sigma_B^1}$ such that $\eta_b > 1$ for each $b \in \Sigma_B^1$ and $\hat{\theta}(\boldsymbol{\eta}) = \sum_{b\in\Sigma_B^1} p_b \eta_b < 1$ there exists a constant $C(\boldsymbol{\eta}) > 0$ such that*

$$\mu(B) \leq C(\boldsymbol{\eta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}(\boldsymbol{\eta})^k \lambda(B)^{\eta_{\max}^{-k} r_{\max}^{-1}}$$

for any Borel set $B \subseteq [0,1]$, *where* $\eta_{\max} = \max\{\eta_b : b \in \Sigma_B^1\}$. *If* $\sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, *so if the bad maps are expanding on average at the point* c, *then we can get the estimate*

$$\mu(B) \leq C \cdot \lambda(B)^{r_{\max}^{-1}} \tag{2.47}$$

for some constant $C > 0$ *and any Borel set* $B \subseteq [0,1]$.

(ii) If $r_{\max} > 1$, *then* $\frac{d\mu}{d\lambda} \notin L^q$ *for any* $q \geq \frac{r_{\max}}{r_{\max}-1}$. *If also* $\sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, *then* $\frac{d\mu}{d\lambda} \in L^q$ *for all* $1 \leq q < \frac{r_{\max}}{r_{\max}-1}$.

(iii) If $r_{\max} = 1$ *and* $\sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, *then* $\frac{d\mu}{d\lambda} \in L^\infty$.

The main issue we need to deal with in order to get Theorem 2.4.1 is adapting Lemma 2.1.6, i.e. finding suitable bounds for $|T_\omega^n(x) - c|$, since the constants $\tilde{K}$ and $\tilde{M}$ from Lemma 2.1.6 are not well defined in case $\ell_{\min} = 1$. This is done in the next two lemmas. For the upper bound of $|T_\omega^n(x) - c|$ we assume $\ell_{\max} > 1$ since we only need it for the proof of part (b)(i).

Lemma 2.4.2. *Let* $\{T_j : j \in \Sigma\}$ *be as above. Suppose* $\ell_{\max} > 1$. *There are constants* $\hat{M} > 1$ *and* $\delta > 0$ *such that for all* $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ *and* $x \in [c-\delta, c+\delta]$ *we have*

$$|T_\omega^n(x) - c| \leq \left(\hat{M}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}}.$$

Proof. Similar as in the proof of Lemma 2.1.6 it follows that there exists an $M > 1$ such that for any $b \in \Sigma_B$ and $x \in [0,1]$ we have

$$|T_b(x) - c| \leq M|x-c|^{\ell_b}. \tag{2.48}$$

Furthermore, there exists a $\delta > 0$ such that $|DT_b(x)| < 1$ for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B^1$. This implies

$$|T_b(x) - c| < |x-c| \tag{2.49}$$

for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B^1$. Note that $\Sigma_B \neq \emptyset$ because $\ell_{\max} > 1$. We set $\upsilon = \min\{\ell_b : b \in \Sigma_B\} > 1$ and $\hat{M} = M^{\frac{1}{\upsilon-1}}$. For each $n \in \mathbb{N}$ and $\omega \in (\Sigma_B^1)^{\mathbb{N}}$, write

$$m(n,\omega) = \#\{1 \leq \omega_i \leq n : \ell_{\omega_i} > 1\}.$$

The statement follows by showing that for all $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $x \in [c-\delta, c+\delta]$ we have

$$|T_\omega^n(x) - c| \leq \left(M^{(1-\upsilon^{-m(n,\omega)})/(\upsilon-1)}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}}. \tag{2.50}$$

We prove (2.50) by induction. From (2.48) and (2.49) it follows that (2.50) holds for $n=1$. Now suppose (2.50) holds for some $n \in \mathbb{N}$. Let $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $y \in [c-\delta, c+\delta]$. If $\ell_{\omega_{n+1}} = 1$, then the desired result follows by applying (2.49) with $b = \omega_{n+1}$ and $x = T_\omega^n(y)$. Suppose $\ell_{\omega_{n+1}} > 1$. Then, using (2.48),

$$\begin{aligned} |T_\omega^{n+1}(y) - c| &\leq M|T_\omega^n(y) - c|^{\ell_{\omega_{n+1}}} \\ &\leq \left(M^{(1-\upsilon^{-m(n,\omega)})/(\upsilon-1)+\upsilon^{-m(n+1,\omega)}}|y-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_{n+1}}}. \end{aligned}$$

Using that

$$\upsilon^{-m(n+1,\omega)} = \frac{\upsilon^{-m(n,\omega)} - \upsilon^{-m(n+1,\omega)}}{\upsilon - 1},$$

the desired result follows. □

Lemma 2.4.3. *Let $\{T_j : j \in \Sigma\}$ be as above. Let $\boldsymbol{\eta} = (\eta_b)_{b \in \Sigma_B^1}$ be a vector such that $\eta_b > 1$ for each $b \in \Sigma_B^1$. Set $\hat{\eta}_b = \max\{\eta_b, \ell_b\}$ for each $b \in \Sigma_B^1$. Then there exists a constant $\hat{K}(\boldsymbol{\eta}) \in (0,1)$ such that for all $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $x \in [0,1]$ we have*

$$\Big(\hat{K}(\boldsymbol{\eta})|x-c|\Big)^{\hat{\eta}_{\omega_1}\cdots\hat{\eta}_{\omega_n}} \leq |T_\omega^n(x) - c|.$$

Proof. Note from (B3) that for each $b \in \Sigma_B^1$ we have

$$K_b|x-c|^{\hat{\eta}_b - 1} \leq K_b|x-c|^{\ell_b - 1} \leq |DT_b(x)|.$$

By setting $\hat{\eta}_{\min} = \min\{\hat{\eta}_b : b \in \Sigma_B^1\}$, $\hat{\eta}_{\max} = \max\{\hat{\eta}_b : b \in \Sigma_B^1\}$ and $\hat{K}(\boldsymbol{\eta}) = \big(\frac{\min\{K_b : b \in \Sigma_B^1\}}{\hat{\eta}_{\max}}\big)^{\frac{1}{\hat{\eta}_{\min}-1}}$ the result now follows in the same way as in the proof of Lemma 2.1.6. □

Proof of Theorem 2.4.1. Firstly, note that (a), (b)(ii) and the first part of (c)(ii) immediately follow from Remark 2.2.8. Moreover, as in [dMvS93, Section 5.4] it can be shown that (2.47) implies that $\frac{d\mu}{d\lambda}$ is in L^q if $r_{\max} > 1$ and $1 \leq q < \frac{r_{\max}}{r_{\max}-1}$, giving the remainder of 3(ii). It is immediate that (2.47) implies that $\frac{d\mu}{d\lambda}$ is in L^∞ if $r_{\max} = 1$, so (3)(iii) holds. Hence, it remains to prove (b)(i) and (c)(i).

Suppose $\theta = \sum_{b \in \Sigma_B^1} p_b \ell_b \geq 1$, which means that $\ell_{\max} > 1$. The proof that in this case $\mu_{\boldsymbol{p}}$ is infinite follows by the same reasoning as in Subsection 2.3.1 by now taking $\gamma = \min\{\delta, \frac{1}{2}\hat{M}^{-1}\}$ with δ and $\hat{M}$ as in the proof of Lemma 2.4.2. Now suppose $\theta < 1$. Let $\boldsymbol{\eta} = (\eta_b)_{b \in \Sigma_B^1}$ be a vector such that $\eta_b > 1$ for each $b \in \Sigma_B^1$ and $\hat{\theta}(\boldsymbol{\eta}) = \sum_{b \in \Sigma_B^1} p_b \hat{\eta}_b < 1$ with again $\hat{\eta}_b = \max\{\eta_b, \ell_b\}$. Applying Lemma 2.4.3 yields that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{b} \in (\Sigma_B^1)^n$,

$$T_{\boldsymbol{b}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(\hat{K}(\boldsymbol{\eta})^{-1}\varepsilon^{\hat{\eta}_{\boldsymbol{b}}^{-1}}\big), \tag{2.51}$$

where we used the notation $\hat{\eta}_{\boldsymbol{b}} = \hat{\eta}_{b_1}\cdots\hat{\eta}_{b_n}$ for a word $\boldsymbol{b} = b_1 \cdots b_n$. Following the line of reasoning in Subsection 2.3.2 with (2.51) instead of (2.33), we obtain that there exists a constant $C(\boldsymbol{\eta}) > 0$ such that

$$\mu(B) \leq C(\boldsymbol{\eta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}(\boldsymbol{\eta})^k \lambda(B)^{\hat{\eta}_{\max}^{-k} r_{\max}^{-1}} \tag{2.52}$$

for any Borel set $B \subseteq [0,1]$. In case $\ell_{\max} > 1$ we can choose $\boldsymbol{\eta}$ to satisfy $\hat{\eta}_{\max} = \ell_{\max}$ and such that $\hat{\theta}(\boldsymbol{\eta}) - \theta > 0$ is arbitrarily small, which yields (b)(i). In case $\ell_{\max} = 1$, then $\hat{\eta}_{\max} = \eta_{\max}$, so this together with (2.52) yields the first part of (c)(i).

Finally, for the second part of (c)(i), suppose $\ell_{\max} = 1$ and $\Lambda = \sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$. Setting $K_{\boldsymbol{b}} = |DT_{\boldsymbol{b}}(c)|$ for each $\boldsymbol{b} \in (\Sigma_B^1)^n$ and $n \in \mathbb{N}$, note that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{b} \in (\Sigma_B^1)^n$,

$$T_{\boldsymbol{b}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(K_{\boldsymbol{b}}^{-1}\varepsilon\big). \tag{2.53}$$

By using (2.53) instead of (2.33), letting $\tilde{p}_{\boldsymbol{b}} = K_{\boldsymbol{b}}^{-1} p_{\boldsymbol{b}}$ play the role of $p_{\boldsymbol{b}}$ in the reasoning of Subsection 2.3.2 and noting that $\Lambda^k = \sum_{\boldsymbol{b} \in (\Sigma_B^1)^k} \tilde{p}_{\boldsymbol{b}}$, we arrive similarly as for Theorem 2.3.2 to the conclusion that there exists a constant $\tilde{C} > 0$ such that for all $n \in \mathbb{N}$ and all Borel sets $B \subseteq [0,1]$,

$$\lambda_n(B) \leq \tilde{C} \cdot \sum_{g \in \Sigma_G} p_g \Big(\sum_{k=0}^{\infty} \Lambda^k \Big) \lambda(B)^{r_g^{-1}}.$$

This proves the remaining part of (c)(i). □

§2.4.3 Final remarks

The results from Theorem 2.4.1 contain one possible extension of our main results to another set of conditions (G1)–(G4), (B1)–(B4). In this section we discuss some of the questions that our main results brought up in this respect, i.e. about whether or not some of the conditions (G1)–(G4), (B1)–(B4) can be relaxed, and questions about other possible future extensions.

A condition that plays a fundamental role in the proofs of Theorem 2.1.2 and Theorem 2.1.3 is the fact that the critical point is mapped to a point that is a common repelling fixed point for all maps T_j. We considered whether this condition can be relaxed, for instance by assuming that the branches of one of the good maps are not full. However, in this case the critical values of the random system are not just $0, c, 1$ but contain all the values of all possible postcritical orbits of c. This has several consequences:

- An invariant density (if it exists) clearly cannot be locally Lipschitz on $(0,c)$ and $(c,1)$.

- Proposition 2.3.3 and all subsequent arguments fail, since it is not sufficient to restrict to neighbourhoods around only 0, c and 1. One might try to solve this issue by requiring that (on average) the postcritical orbits 'gain enough expansion' as is done in for instance [NvS91] for deterministic maps (see Theorem 2.1.1). An analogous condition for random systems, however, would be much more difficult to verify since it would involve all possible random orbits of c.

- The argument using Kac's Lemma might fail, because in that case there exist words $\boldsymbol{u}$ with symbols in Σ and neighbourhoods U of c such that $T_{\boldsymbol{u}}(x)$ is bounded away from 0 and 1 uniformly in $x \in U$.

There are also some additional questions that our main results raise. It would be interesting for example to study further statistical properties of the random systems such as mixing properties and if possible mixing rates in case the acs measure is finite.

As is well known from e.g. [KN92, Y92, Y98, BLvS03] the good maps individually have exponential decay of correlations. But since trajectories in the random systems spend long periods of time near the points 0 or 1 or both, polynomial mixing rates are expected rather than exponential. This will be the topic of Chapter 3.

The dynamical behaviour of the system is governed by the interplay between the superexponential convergence to c and the exponential divergence from 0 and 1. In this chapter we fixed the exponential divergence away from 0 and 1 and the two regimes $\theta < 1$ and $\theta \geq 1$ in Theorem 2.1.3 only refer to the convergence to c: For smaller (bigger) θ orbits are less (more) attracted to c. It would be interesting to see under what other conditions on the rates of convergence to c and divergence from 0 and 1 the system admits an acs measure. Could one for example

- take exponential convergence to c and polynomial divergence from 0 and 1? We will investigate this question in Chapter 4.

- replace the conditions (G4) and (B4) stating that all good and bad maps are expanding at 0 and 1 by the condition that the random system is expanding on average at a sufficiently large neighbourhood of 0 and 1?

Finally, in Theorems 2.1.2 and 2.4.1 we have seen that the regularity of the density $\frac{d\mu}{d\lambda}$ depends on whether or not there is a bad map for which c is superstable: If $\ell_{\max} > 1$, then $\frac{d\mu}{d\lambda}$ is not in L^q for any $q > 1$. On the other hand, if $\ell_{\max} = 1$ and the bad maps are expanding on average at c, i.e. $\sum_{b\in\Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, then the density has the same regularity as in the setting of Theorem 2.1.1 by Nowicki and Van Strien. Indeed, in this case, if $r_{\max} > 1$, we have $\frac{d\mu}{d\lambda} \in L^q$ if and only if $1 \leq q < \frac{r_{\max}}{r_{\max}-1}$ and in the case that $r_{\max} = 1$ we have $\frac{d\mu}{d\lambda} \in L^q$ for all $q \in [1,\infty]$. In view of this, one could wonder for which $q > 1$ we have $\frac{d\mu}{d\lambda} \in L^q$ in the intermediate case that $\ell_{\max} = 1$ and $\sum_{b\in\Sigma_B^1} \frac{p_b}{|DT_b(c)|} \geq 1$, i.e. if c is not superattracting for any bad map and the bad maps are not expanding on average at c.