

**Polynomial Decay of Correlations in the Generalized  
Baker's Map**

**Chris Bose and Rua Murray**

**DMS-862-IR**

**October 2007**

# Polynomial decay of correlations in the generalized baker's map\*

Chris Bose and Rua Murray

October 15, 2007

The decay of correlations of the 2-dimensional generalized baker's maps can be studied via their one-dimensional projections (onto the expanding direction). A suitable tower can be used to reveal the decay rates.

## 1 A family of maps

The generalized baker's construction has been used to define Lebesgue-measure-preserving maps of the unit square  $S = [0, 1]^2$ . Specifically, a two-dimensional map  $g$  is determined by a 'cut function'  $\phi$  whose graph  $y = \phi(x)$  partitions the square  $S$  into lower and upper pieces. The cut function is assumed to be measurable and to satisfy  $0 \leq \phi(x) \leq 1$ ; these are the only constraints in the construction. The two-dimensional dynamics are defined by mapping the vertical lines  $\{x = 0\}, \{x = 1\}$ , into themselves, and sending vertical fibres into vertical fibres so that areas are preserved, the rectangle  $[0, a] \times [0, 1]$  maps to the lower portion of the square and  $[a, 1] \times [0, 1]$  maps to the upper portion, where  $a = \int_0^1 \phi$ ; the map has a discontinuity along  $\{x = a\}$ .

---

\*Preliminary Version complete to Aug 23, 2007

Explicitly,

$$(x, y) \mapsto (x_+, y_+) = g(x, y)$$

where

$$y_+ = \begin{cases} \phi(x_+)y & \text{if } x \leq a, \\ y + \phi(x_+)(1-y) & \text{if } x \geq a, \end{cases} \quad (1)$$

$$\text{and } \begin{cases} x & = \int_0^{x_+} \phi(t) dt & \text{if } x \leq a, \\ 1-x & = \int_{x_+}^1 (1-\phi(t)) dt & \text{if } x \geq a. \end{cases} \quad (2)$$

Remark: The equations (1)-(2) for  $g$  in terms of  $\phi$  are implicit and may not uniquely define the values  $x_+$  and  $y_+$  in cases where  $\phi$  takes the value 0 or 1 on sets of positive measure – this technical point is easily resolved, but in any event, it will not arise for the examples we have in mind.

By construction,  $g$  preserves the two-dimensional Lebesgue measure on the square  $S$ . The sub-sigma-algebra of vertical fibres on  $S$  is invariant for  $g$  and the associated (non-invertible) factor is naturally identified with the one-dimensional map

$$f : x \mapsto x_+,$$

a two branched, piecewise monotone and increasing map on  $[0, 1]$ .

In fact, defining  $\pi : [0, 1]^2 \rightarrow [0, 1]$  by  $\pi(x, y) = x$ , we have

$$\pi \circ g = f \circ \pi$$

so that if  $\mu$  denotes Lebesgue measure on  $[0, 1]^2$ , then<sup>1</sup>  $\pi_*\mu = m$  is the Lebesgue measure on  $[0, 1]$  and it satisfies

$$f_* m = f_* \pi_* \mu = \pi_* g_* \mu = \pi_* \mu = m.$$

From the definition of  $y_+$ ,  $y_+ \leq \phi(x_+)$  according to whether  $x \leq a$ ; thus, the position on vertical fibres records the inverse history of possible  $f$ -orbits.

We can make this more precise as follows. The cut function  $\phi$  determines a natural partition  $P$  of the square  $S$  into the region below the graph of  $\phi$  and

---

<sup>1</sup>We adopt the standard notation  $T_*\nu = \nu \circ T^{-1}$  for a map  $T$  and measure  $\nu$ .

the region above the graph (which we shall denote by  $P_0$  and  $P_1$  respectively). For many examples (including the maps discussed in this paper) it will happen that the partition  $P$  is a generator<sup>2</sup>. In this case,  $g : [0, 1]^2 \rightarrow [0, 1]^2$  is the natural extension (or inverse limit) of the one-dimensional dynamical system  $f : [0, 1] \rightarrow [0, 1]$ . (Otherwise, the natural extension will appear as a proper factor of  $g$ .)

We will study a parameterized family of baker's transformations  $g_\alpha$ ,  $0 < \alpha < \infty$ , with cut functions

$$\phi_\alpha(x) = \begin{cases} 1 - 2^{\alpha-1}x^\alpha & \text{if } x \leq \frac{1}{2}, \\ 2^{\alpha-1}(1-x)^\alpha & \text{if } x \geq \frac{1}{2}, \end{cases} \quad (3)$$

Within this family, the values  $x_+$ ,  $y_+$  are uniquely defined by equations (1)-(2) and the factor  $x_+ = f_\alpha(x)$  on  $[0, 1]$  is measure-preserving with two surjective branches separated by  $a = \frac{1}{2}$ .

The simplest case, and the most geometrically appealing is when  $\alpha = 1$ , so that  $\phi(x) = 1 - x$ . Then the integrals defining  $x_+$  in (2) and the map  $x_+ = f(x)$  are computed explicitly as

$$f(x) = \begin{cases} 1 - \sqrt{1 - 2x} & \text{if } x < \frac{1}{2}, \\ \sqrt{2x - 1} & \text{if } x > \frac{1}{2}. \end{cases}$$

We emphasized again that  $f$  is a measure-preserving circle endomorphism on  $[0, 1)$  with a discontinuity in  $f'$  at the single point  $a = \frac{1}{2}$ , and a neutral fixed point at  $x = 0$ , but in this case, with quadratic order of contact. Thus the example does not fit into the usual picture for maps with indifferent fixed points (eg. [3], [8]). In fact, the branches of  $f$  do not have bounded distortion, since  $f'(x) \rightarrow \infty$  as  $x \rightarrow \frac{1}{2}$ . Observe, however, that the slow escape of mass in the neighbourhood of the neutral point is perfectly balanced by a very small rate of arrival in these intervals (for example,  $f^{-1}([0, \epsilon]) \setminus [0, \epsilon) = [\frac{1}{2}, \frac{1}{2} + O(\epsilon^2))$ ). This is the mechanism which allows for a finite invariant measure even though the map has only a weakly repelling fixed point. The same observations apply to all members of our parameterized family of maps; our goal is to study the ergodic properties and rates of mixing for these simple examples.

---

<sup>2</sup>That is,  $\bigvee_{n=-\infty}^{\infty} g^{-n}P = \mathcal{B}$ , the Borel sigma-algebra on the square.

## Young towers

In order to proceed, we briefly outline the general machinery developed by L. S. Young [7, 8] for analysis of non-uniformly hyperbolic dynamics using an abstract tower construction. A set  $\Delta_0$  is assumed to be given, along with a function  $R : \Delta_0 \rightarrow \mathbb{Z}^+$  (the return time function) from which one defines a tower

$$\Delta := \{(z, l) : z \in \Delta_0, l \in \mathbb{Z}, 0 \leq l < R(x)\}.$$

It is helpful to consider  $\Delta$  as a subset of  $\Delta_0 \times \mathbb{Z}^+$ . We denote

$$\Delta_l = \Delta \cap \Delta_0 \times \{l\},$$

the  $l$ -th level of the tower. We also assume there is a countable partition  $\{\Delta_{0,i}\}$  of  $\Delta_0$  such that  $R$  is constant on each atom of the partition.

A map  $F : \Delta \rightarrow \Delta$  is provided on the tower such that  $F(z, l) = (z, l + 1)$  if  $l < R(x) - 1$  and  $F(z, R(x) - 1) \in \Delta_0$ . In this way  $F^R : \Delta_0 \rightarrow \Delta_0$  is the first return map to  $\Delta_0$ .  $F$  carries the partition of  $\Delta_0$  into a partition  $\eta$  of the tower:  $\Delta_{l,i} = \{(z, l) \in \Delta : z \in \Delta_{0,i}\}$  and one assumes that the partition generates, in the sense that  $\bigvee_{j=0}^{\infty} F^{-j}\eta$  separates the points of  $\Delta$ .

It is assumed there is a measure (the reference measure)  $\mu$  on  $\Delta$  with respect to which  $\eta$  is measurable (and hence  $R$  is a measurable function) and such that  $F_*(\mu|_{\Delta_{l,i}}) = \mu|_{\Delta_{l-1,i}}$  for  $0 < l < r$  whenever  $R|_{\Delta_{0,i}} = r$ . Finally, it is assumed that, for each  $i$ ,  $F^R : \Delta_{0,i} \rightarrow \Delta_0$  is bijective, and both  $F^R_{\Delta_{0,i}}$  and its inverse are nonsingular with respect to  $\mu$ . The Jacobian of the return map will be denoted by  $JF^R$  and on each  $\Delta_{0,i}$ ,  $JF^R > 0$ , again by assumption.

A measure of separation  $s$  is defined for pairs of points in the base of the tower. If <sup>3</sup>  $x, y \in \Delta_0$  we set  $s(x, y)$  to be the minimum  $n$  such that  $(F^R)^n(x), (F^R)^n(y)$  lie in different atoms  $\Delta_{0,i}$ . Clearly  $s < \infty$  since  $\bigvee_{j=0}^{\infty} F^{-j}\eta$  separates points.

---

<sup>3</sup>From this point on we will generally simplify notation and write  $x$  instead of  $(z, l)$  for points in the tower.

The separation  $s$  is next used to give a notion of regularity for the map  $F$ : Assume there exists  $0 < \beta < 1$  and  $C$  such that

$$\left| \frac{JF^R(x)}{JF^R(y)} - 1 \right| \leq C\beta^{s(F^R(x), F^R(y))}, \quad \forall i, \forall x, y \in \Delta_{0,i}. \quad (4)$$

Finally, the measure of separation is extended to the tower  $\Delta$  as follows. If  $x, y \in \Delta$  lie on different levels of the tower, then  $s(x, y) = 0$ . Otherwise,  $s(x, y)$  is set equal to  $s(x', y')$  for the unique first pre-images  $x', y' \in \Delta_0$  under iteration by  $F^{-1}$ .

With  $s$  extended to  $\Delta$  one distinguishes two classes of regular functions. For  $0 < \beta < 1$  as in (4),

$$C_\beta(\Delta) = \{\phi : \Delta \rightarrow \mathbb{R} : \exists c_\phi \text{ s.t. } \forall x, y \in \Delta, |\phi(x) - \phi(y)| \leq c_\phi \beta^{s(x,y)}\}$$

and

$$C_\beta^+(\Delta) = \{\phi : \Delta \rightarrow [0, \infty) : \exists c_f \text{ s.t. for each } l, i \text{ either } \phi \equiv 0 \text{ on } \Delta_{l,i} \text{ or } \phi > 0 \text{ on } \Delta_{l,i} \text{ and } \left| \frac{\phi(x)}{\phi(y)} - 1 \right| \leq c_\phi \beta^{s(x,y)} \forall x, y \in \Delta_{l,i}\}.$$

We now quote the abstract results from [8] which we will be using. For economy of notation throughout the paper we will adopt the standard conventions for asymptotics of sequences:  $x_n = O(y_n)$  means there exists a constant  $C < \infty$  such that for all large  $n$ ,  $x_n \leq Cy_n$  and  $x_n \approx y_n$  if both  $x_n = O(y_n)$  and  $y_n = O(x_n)$ .

**Theorem 1** ([8], part of Theorems 1-3) *Assume the setting and notation above. Assume also that  $\int_{\Delta_0} R d\mu < \infty$  and that  $\gcd\{R_i\} = 1$  where  $R_i := R|_{\Delta_{0,i}}$ . Then,*

- $F$  admits an absolutely continuous (w.r.t.  $\mu$ ) invariant probability measure  $\nu$  on  $\Delta$  with  $\frac{d\nu}{d\mu} > 0$ . Moreover, the system  $(F, \nu)$  is exact.

Extend the definition of  $R$  to  $\Delta$  in the natural way:  $\hat{R}(x) =$  the smallest  $n > 0$  such that  $F^{\hat{R}}(x) \in \Delta_0$ . Assume that there is a constant  $\gamma > 0$  such that  $\mu\{\hat{R} > n\} = O(n^{-\gamma})$ . Then the following upper bounds on the speed of convergence for measures and decay of correlations for functions are obtained:

- For any measure  $\lambda$  with  $\frac{d\lambda}{dm} \in C_\beta^+(\Delta)$  we have

$$|F_*^n \lambda - \nu| = O(n^{-\gamma}).$$

- For each  $\varphi \in L^\infty$  and  $\psi \in C^\beta(\Delta)$  we have

$$\left| \int_\Delta (\varphi \circ F^n) \psi d\nu - \int_\Delta \varphi d\nu \int_\Delta \psi d\nu \right| = O(n^{-\gamma}).$$

Remarks:

- Observe that  $\mu\{\hat{R} > n\} = \sum_{l>n} \mu(\Delta_l)$  so the asymptotics above are exactly the decay rate of the mass in the top of the tower.
- The decay of correlation statement follows immediately from the speed of convergence to equilibrium for measures, once it is shown that each function in  $C_\beta$  can be decomposed into a difference of functions in  $C_\beta^+$ . Details are found in [8].

## A tower for $f_\alpha$

For the rest of this article we will assume that a value  $\alpha \in (0, \infty)$  and the map  $f_\alpha$  have been chosen. Unless necessary for clarity, we will suppress the subscript and simply write  $f$ . We now show how the abstract tower construction can be realized for our map.

Note that  $f$  admits a period-2 orbit  $\{x_0, y_0\}$  since  $f^2$  is a four-branched, piecewise continuous and onto map. We may assume that<sup>4</sup>  $x_0 < 1/2$  and

---

<sup>4</sup>Let  $x_0$  be the fixed point for  $f^2$  on the second branch.

$y_0 > 1/2$ . By symmetry,  $y_0 = 1 - x_0$ . For example, when  $\alpha = 1$  one easily computes  $x_0 = \sqrt{2} - 1$  and  $y_0 = 2 - \sqrt{2}$ . Let  $\Delta_0 = [x_0, \frac{1}{2}) \cup (\frac{1}{2}, y_0]$ . Let  $\{x_n\}$  be defined under the left branch of  $f$  (recursively) by  $f(x_n) = x_{n-1}$ . Put  $J_n = [x_{n+1}, x_n)$ , with symmetrically chosen intervals  $J'_n$  in  $[y_0, 1]$ . Finally, put  $I_{n+1} = f^{-1}(J_n) \setminus J_{n+1}$  (and similarly for  $\{I'_n\}$ ). Observe that  $I_n \subseteq (\frac{1}{2}, y_0]$  while  $I'_n \subseteq [x_0, \frac{1}{2})$ . Let  $R$  denote the return time function to  $\Delta_0$ . Under  $f$ ,

$$I_k \rightarrow J_{k-1} \rightarrow J_{k-2} \rightarrow \cdots \rightarrow J_0 \rightarrow \Delta_0,$$

and similarly for the  $I'_n$  and  $J'_n$  intervals. Note that each application is injective and onto. Thus,  $R(x) = k + 1$  when  $x \in I_k^{(l)}$ ; moreover,  $f^R$  maps bijectively to  $\Delta_0$  from each  $I_k^{(l)}$ . The base of the tower is taken to be  $\Delta_0$ , which is partitioned by two infinite sets of intervals  $\Delta_{0,i} = I_i \times \{0\}$  and  $\Delta'_{0,i} = I'_i \times \{0\}$ . Then,  $R|_{\Delta_{0,i}^{(l)}} = i + 1$  ( $i \geq 1$ ) and the tower is

$$\Delta = \bigcup_{i=1}^{\infty} \bigcup_{l=0}^i (\Delta_{l,i} \cup \Delta'_{l,i}),$$

where  $\Delta_{l,i}^{(l)} := \Delta_{0,i}^{(l)} \times \{l\}$ , embedding the tower in  $\Delta_0 \times \mathbb{Z}^+$

The tower map is

$$F(x, l) = \begin{cases} (x, l + 1) & \text{if } l < R(x) - 1, \\ (f^R(x), 0) & \text{if } l = R(x) - 1 \text{ and } R = R(x). \end{cases}$$

We next establish asymptotics on the  $x_n$  and intervals  $I_n$  and  $J_n$ .

**Lemma 1** (i)  $x_n \approx (\frac{1}{n})^{1/\alpha}$ ;

(ii)  $m(J_k) \approx (\frac{1}{k})^{1+1/\alpha}$ ;

(iii)  $m(I_k) \approx (\frac{1}{k})^{2+1/\alpha}$ ;

(iv) if  $\rho > 0$  then  $x_k - x_{k+n} \approx x_k \frac{n}{k}$  when  $n \leq \rho k$ .

The same estimates hold for the  $'$  intervals.

*Proof:* See Appendix.

In the previous section, a separation function  $s$  was defined on the tower  $\Delta$  with respect to the partition  $\eta$ . We emphasize that in our construction,  $\eta$  is given by the intervals  $\Delta_{l,i}^{(l)}$  and  $\Delta_{l,i} \neq \Delta'_{l,i}$  are taken to be different atoms in  $\eta$ .

**Lemma 2** *There exists a constant  $\beta < 1$  such that if  $x, y \in \Delta_0$  and  $s(x, y) = n$  then  $|x - y| \leq \beta^n$*

*Proof:* Set  $\beta(= \beta(\alpha)) := [f'(y_0)]^{-1} = [f'(x_0)]^{-1}$  and observe that on the set  $\Delta_0$ ,  $f' \geq \beta$ . Therefore, if  $x, y$  lie in a common atom  $\Delta_{0,i}^{(l)} \subseteq (f^R)^{-1}[x_0, y_0]$  with  $x = (f^R)^{-1}(x')$ ,  $y = (f^R)^{-1}(y')$  then  $|x - y| \leq \beta$ . The result follows by induction on  $i \leq n$ .  $\square$

Remark. For the case  $\alpha = 1$ ,  $[f'(y_0)]^{-1} = (3 - 2\sqrt{2})^{1/2} \sim 0.414$ .

**Lemma 3 (Uniform distortion)** *Let  $y, z \in \Delta_0$  and suppose that  $s(y, z) \geq 1$ . Then there is a constant  $D > 1$  (depending on  $\alpha$  but not  $y, z$ ) such that*

$$\left| \frac{f^{R'}(y)}{f^{R'}(z)} - 1 \right| \leq \frac{D(D-1)}{m(\Delta_0)} |f^R(y) - f^R(z)|.$$

*Proof:* See Appendix.

Remarks:

- The ambient measure  $\mu$  in the abstract tower construction is taken to be Lebesgue measure  $m_{[0,1]}$  on  $\Delta_0$  which is carried by the action of  $F$  to Lebesgue measure  $m_\Delta$  on the tower. Note, however, that since  $m_{[0,1]}$  is

invariant for  $f$ ,  $m_{\Delta_0}$  is  $f^R$ -invariant on  $\Delta_0$ . Since  $F^R(x) = f^R(x) \forall x \in \Delta_0$ ,  $m_{\Delta}$  is  $F$ -invariant on the tower. Therefore  $F^R$  and its inverse satisfy assumed nonsingularity as maps between  $\Delta_{0,i}^{(r)}$  and  $\Delta_0$ .

- Combining the two Lemmas above and the previous remark gives us the basic regularity estimate on the tower map  $F$  in equation (4). Simply observe that  $|f^R(y) - f^R(z)| \leq \beta^{s(f^R(y), f^R(z))}$  and that  $F^R = f^R$ .

## Mixing rates I – upper bounds on the tower

We begin by reformulating the results in the abstract mixing theorem for our version of the Young tower.

**Theorem 2** *Fix  $0 < \alpha < \infty$  and any  $\beta \geq \beta(\alpha)$  as in Lemma 2. Denote by  $m$  the Lebesgue measure  $m_{\Delta}$  on the tower. Then*

1.  *$m$  is the unique absolutely continuous  $F$ -invariant probability measure on  $\Delta$ . Moreover, the system  $(F, m)$  is exact, hence ergodic and mixing.*
2. *For each absolutely continuous measure  $\lambda$  such that  $\frac{d\lambda}{dm} \in C_{\beta}^+$  we have*

$$|F_*^n \lambda - m| = O(n^{-\frac{1}{\alpha}})$$

3. *For every  $\varphi \in L^{\infty}(\Delta)$  and  $\psi \in C_{\beta}(\Delta)$  we have*

$$\left| \int \varphi \circ F^n \psi dm - \int \varphi dm \int \psi dm \right| = O(n^{-\frac{1}{\alpha}})$$

Remark: Note that in this theorem, Lebesgue measure  $m = m_{\Delta}$  on the tower is both the original ambient (or reference) measure and the ACIM. Therefore, the estimates on decay of correlation in (2) and (3) can be interpreted in two ways, first as relaxation rates for iterates for the Frobenius-Perron operator in (2) but also as mixing rates for the measure-preserving system in (3).

*Proof:* (1) The regularity condition in equation (4) is satisfied for every  $\beta \geq \beta(\alpha)$ . Next, using Lemma 1 we can estimate

$$\int_{\Delta_0} R(x) dm(x) = \sum_{k=1}^{\infty} (k+1)m(I_k \cup I'_k) \approx \sum_{k=1}^{\infty} (k+1) \left(\frac{1}{k}\right)^{2+\frac{1}{\alpha}} \approx \sum_{k=1}^{\infty} \left(\frac{1}{k}\right)^{1+\frac{1}{\alpha}} < \infty.$$

(In fact, since  $F$  preserves  $m$  Kac's Theorem applies and the value of the sum is exactly 1!) Finally, we note that the values taken by the return time function are  $R = 2, 3, \dots$  so the g.c.d condition in Theorem 1 also holds. Applying the theorem to our tower yields an invariant measure  $\nu$  on  $\Delta$  equivalent (i.e. mutually absolutely continuous) to  $m$ . Of course, we hope to identify  $m$  with  $\nu$ , and there is a simple way to do this. The use of Kac's Theorem above showed that  $m(\Delta) = 1$ , so  $F$  is equipped with two equivalent invariant probability measures, and one of them ( $\nu$ ) is ergodic. Now suppose  $m(F^{-1}A\Delta A) = 0$ . Then the same holds for  $\nu$  and hence either  $A$  or  $A^c$  are  $\nu$ -null sets. Therefore one of  $A$  or  $A^c$  is also  $m$ -null, proving that  $m$  is ergodic. Since two invariant and ergodic probability measures must be either identical or singular, we conclude that  $m = \nu$ .

(2)-(3) We make another simple calculation using Lemma 1 to estimate the mass in the top of the tower.

$$m(\hat{R} > n) = \sum_{l>n} m(\Delta_l) = 2 \sum_{l>n} (l-n)m(I_l) \approx \left(\frac{1}{n}\right)^{\frac{1}{\alpha}}$$

Theorem 1 now gives the rates of mixing and correlation decay as claimed.  $\square$

## Obtaining $f$ as a factor of $F$

The tower  $(F, \Delta)$  provides a representation for the dynamics of  $f$  oriented around the *induced transformation*  $f^R$  of first returns to  $\Delta_0$ . In order to interpret the mixing results of Theorem 2 in terms of the original map  $f$  we

first extract  $f$  as a factor of  $F$  in the following simple way<sup>5</sup>.

For  $(x, l) \in \Delta$  define

$$\Phi(x, l) = f^l(x).$$

We collect a few simple facts about  $\Phi$ .

- $\Phi|_{\Delta_0} \equiv \text{id}_{[x_0, y_0]}$
- $\Phi$  maps  $\Delta_l$  injectively onto  $(0, x_0) \cup (y_0, 1)$
- $\Phi^{-1}(J_k) = \bigcup_{l=1}^{\infty} I_{l+k} \times \{l\}$  (with a similar equality for  $\cdot'$ )
- There exists a  $D'$  such that for all  $i, l$ , if  $A \subseteq I_i \times \{l\} \subset \Delta_{l,i}$  then

$$D'^{-1} \leq \frac{m(A)}{m(I_i)} \frac{m(J_{i-l})}{m(\Phi(A))} \leq D'$$

(with a similar inequality for  $\cdot'$ ).

- The semi-conjuacy property:

$$\begin{aligned} \Phi \circ F(x, l) &= \begin{cases} \Phi(f^{l+1}(x), 0) & \text{if } x \in \Delta_{0,l}, \\ \Phi(x, l+1) & \text{if } x \in \Delta_{0,k}, k > l \end{cases} \\ &= f^{l+1}(x) = f(f^l(x)) = f \circ \Phi(x, l). \end{aligned}$$

- That  $\Phi_* m_\Delta = m_{[0,1]}$ . This computation can be done by bare hands, or one can use the  $F$ -invariance of  $m_\Delta$  as follows: Set  $m_0 = \Phi_* m_\Delta$ , and invariant measure for the factor  $f$ . From Theorem 2 we know that  $m_0$  is in fact ergodic for  $f$  and from the distortion relation derived above, it is equivalent to  $m_{[0,1]}$ . Then the argument in the proof of the first part of Theorem 2 again applies to show that  $m_0 = m_{[0,1]}$ .

Now suppose  $\phi$  is a function on  $[0, 1]$ . We denote by  $\hat{\phi}$  it's natural lift to  $\Delta$ :  $\hat{\phi}(x, l) = \phi(\Phi(x, l))$

---

<sup>5</sup>There are a number of natural ways to identify  $f$  as a factor, we simply choose one which is convenient for our purpose.

**Lemma 4** Let  $\beta = \beta(\alpha)$ . If  $\phi$  is a Hölder continuous function<sup>6</sup> on  $[0, 1]$ ,  $\hat{\phi} \in C_{\beta_0}(\Delta)$ , where  $\beta_0 = (\beta)^\gamma$

*Proof:* We need to check the regularity condition on  $\hat{\phi}$ . First, if  $(x, l), (y, k)$  are not on the same level of the tower, then  $s((x, l), (y, k)) = 0$  and we estimate (for any choice of  $\beta$ )

$$|\hat{\phi}(x, l) - \hat{\phi}(y, k)| \leq 2\|\phi\|_\infty \beta^0$$

In fact, the same inequality holds also whenever  $s((x, l), (y, l)) = 0$  on the same level of the tower in which case  $c_\phi = 2\|\phi\|_\infty$  will do the job. Now suppose  $s((x, l), (y, l)) = n > 0$ . Then, with  $C$  and  $\gamma > 0$  from the Hölder condition on  $\phi$  and applying Lemma 2 we obtain

$$\begin{aligned} |\hat{\phi}(x, l) - \hat{\phi}(y, l)| &= |\phi(f^l(x)) - \phi(f^l(y))| \\ &\leq C|f^l(x) - f^l(y)|^\gamma \leq C|(F^R(x)) - (F^R(y))|^\gamma . \\ &\leq C\beta^{(n-1)\gamma} = C\beta^{-\gamma}(\beta^\gamma)^n \end{aligned}$$

Therefore it suffices to take  $c_\phi = \max\{C\beta^{-\gamma}, 2\|\phi\|_\infty\}$  and  $\beta_0 = \beta^\gamma$  in the definition of  $C_{\beta_0}(\Delta)$ .  $\square$

**Theorem 3** Let  $m = m_{[0,1]}$  denote Lebesgue measure on  $[0, 1]$ .

1. The system  $(f, m)$  is exact and hence  $g$  acting on  $S$  is a  $K$ -automorphism.
2. If  $d\lambda = \phi dm$  is any absolutely continuous probability measure with  $\phi$  Hölder continuous, then

$$|f_*^n \lambda - m| = O(n^{-\frac{1}{\alpha}}).$$

3. If  $\varphi \in L^\infty[0, 1]$  and  $\psi : [0, 1] \rightarrow \mathbb{R}$  is Hölder continuous, then

$$\left| \int_0^1 \varphi \circ f^n \psi dm - \int_0^1 \varphi dm \int_0^1 \psi dm \right| = O(n^{-\frac{1}{\alpha}})$$

---

<sup>6</sup>Meaning,  $|\phi(x) - \phi(y)| \leq C|x - y|^\gamma$ , for some  $C$ ,  $\gamma > 0$  and all  $x, y$

*Proof:* Denote again by  $m_\Delta$  Lebesgue measure on the tower. Since  $(f, m)$  is a factor of the exact system  $(F, m_\Delta)$ , it is also exact, and hence its natural extension  $g$  on  $S$  is a  $K$ -automorphism. Next we may assume  $\gamma \leq 1$  in the Hölder condition, so  $\beta^\gamma \geq \beta$ . Finally, observe the elementary identity

$$\int_{[0,1]} q(x) dm(x) = \int_{[0,1]} q(x) d\Phi_* m_\Delta = \int_\Delta \hat{q} dm_\Delta$$

Now an application of Lemma 4, combined with the decay of correlations result in Theorem 2, using the value of  $\beta^\gamma \geq \beta(\alpha)$  yields the result  $\square$

## 2 Mixing rates II – slow decay for most densities

The upper bounds on speed of convergence to equilibrium and correlation decay obtained in the previous section are in fact sharp in most situations. Lower bounds on the decay rate are effectively determined by the behaviour of initial densities in the neighbourhoods of the indifferent fixed points at 0 and 1. We say a measure  $\lambda$  is separated from  $m$  at  $x$  if either

$$\limsup_{\epsilon \rightarrow 0} \frac{\lambda(x-\epsilon, x+\epsilon)}{m(x-\epsilon, x+\epsilon)} < 1 \text{ or } \liminf_{\epsilon \rightarrow 0} \frac{\lambda(x-\epsilon, x+\epsilon)}{m(x-\epsilon, x+\epsilon)} > 1.$$

**Theorem 4** (i) If  $\lambda \ll m$ ,  $\frac{d\lambda}{dm} \in L^\infty$ , and  $\lambda$  is separated from  $m$  at 0 or 1 then for  $n \in \mathbb{N}$ ,  $|f_*^n \lambda - m| \geq c n^{-1/\alpha}$  ( $c$  is a constant depending on  $\lambda$  and  $\alpha$ ).

(ii) Let  $\psi(x) = \text{sgn}(\frac{1}{2} - x)$ . Let  $h : [1, \infty) \rightarrow [0, \infty)$  be a decreasing, differentiable function such that  $\lim_{x \rightarrow \infty} h(x) = 0$ .

(a) There is  $\varphi_h \in L^1$  such that

$$\int_0^1 \psi \circ f^n \varphi_h dm - \int_0^1 \psi dm \int_0^1 \varphi_h dm = |f_*^n(\varphi_h m) - m| \approx h(n).$$

Moreover,

(b) if  $x^{1+1/\alpha} h'(x)$  is bounded then  $\varphi_h$  can be chosen to be an  $L^\infty$  density,

such that  $\hat{\varphi}_h \in C_\beta^+$  when  $\beta \geq \beta_\alpha$ ;  
(c) if  $\lim_{x \rightarrow \infty} x^{1+1/\alpha} h'(x) = 0$  then also

$$\lim_{x \rightarrow 0} \varphi_h(x) = \lim_{x \rightarrow 1} \varphi_h(x) = 1.$$

Theorem 4 shows that the speed of measure convergence given in the previous section is sharp when the initial densities do not assume the correct value at the endpoints. Using  $h(x) = x^{-1/\alpha}$ , part (ii)(b) shows that the implied rates of decay of correlation are essentially optimal. Finally, part (ii)(c) shows that faster decay of correlations can occur for integration against densities which take the value 1 at the ends of the interval. For example,  $\frac{x^{-1/\alpha}}{\log x}$ ,  $x^{-1/\alpha-\gamma}$ , or  $\exp(-\gamma x)$  ( $\gamma > 0$ ) are valid choices for  $h(x)$ .

**Lemma 5** *If  $\limsup_{x \rightarrow 0} \frac{\lambda[0,x]}{x} < 1$  then  $|f_*^n \lambda - m| \geq \frac{c}{n^{1/\alpha}}$  where the constant  $c$  depends on  $\lambda$  and  $\alpha$ .*

*Proof:* Let  $\epsilon, \delta > 0$  be such that  $\lambda[0, u] < (1 - \delta)u$  for all  $u \in (0, \epsilon)$ . Write  $f^{-n}[0, u] = [0, v] \cup A_n$  where  $f^n(v) = u$ . Then,  $f_*^n \lambda[0, u] \leq \lambda[0, v] + \left| \frac{d\lambda}{dm} \right|_\infty m(A_n)$ . Since  $m$  is  $f$  invariant,  $u = m[0, u] = m \circ f^{-n}[0, u] = v + m(A_n)$ . Now let  $u = x_k$ , where  $k$  is large enough that  $x_k < \epsilon$ . Then,  $v = x_{k+n}$  and

$$f_*^n \lambda[0, u] \leq (1 - \delta)v + \left| \frac{d\lambda}{dm} \right|_\infty (u - v) \leq (1 - \delta)x_k + \left| \frac{d\lambda}{dm} \right|_\infty c_2 x_k \frac{n}{k}$$

(where the finite  $c_2$  is chosen corresponding to  $\rho = 1$  in Lemma 1 (iv)). Now, choose  $N \in \mathbb{N}$  such that  $\frac{\left| \frac{d\lambda}{dm} \right|_\infty c_2}{N} < \frac{\delta}{2}$  and  $x_N < \epsilon$ . Using  $u = x_k = x_{nN}$ ,

$$f_*^n \lambda[0, x_{nN}] \leq (1 - \delta)x_{nN} + \frac{\delta}{2}x_{nN}.$$

Consequently,  $|f_*^n \lambda - m| \geq |f_*^n \lambda[0, x_{nN}] - m[0, x_{nN}]| \geq \frac{\delta}{2}x_{nN} \geq \frac{\delta}{2}c_1 \left(\frac{1}{nN}\right)^{1/\alpha}$ , by Lemma 1 (i).  $\square$

**Lemma 6** *If  $\liminf_{x \rightarrow 0} \frac{\lambda[0,x]}{x} > 1$  then  $|f_*^n \lambda - m| \geq \frac{c}{n^{1/\alpha}}$  where the constant  $c$  depends on  $\lambda$  and  $\alpha$ .*

*Proof:* Let  $\epsilon, \delta > 0$  be such that  $\lambda[0, u] > (1 + \delta)u$  for all  $u \in (0, \epsilon)$ . Let  $c_1$  be such that  $x_k - x_{n+k} \leq c_2 x_k \frac{n}{k}$  when  $n \leq k$ . Now choose  $N \in \mathbb{N}$  such that  $(1 - \frac{c_2}{N}) \geq (1 + \delta)^{-1/2}$  and  $x_N < \epsilon$ . Then,

$$\frac{x_{nN+n}}{x_{nN}} = 1 - \frac{x_{nN} - x_{nN+n}}{x_{nN}} \geq 1 - \frac{c_2}{N} \geq (1 + \delta)^{-1/2}.$$

Now, for any  $n \geq 1$ ,

$$f_*^n \lambda[0, x_{nN}] = \lambda[0, x_{nN+n}] \geq (1 + \delta) x_{nN+n} \geq (1 + \delta) (1 + \delta)^{-1/2} x_{nN}.$$

In particular,  $|f_*^n \lambda - m| \geq |f_*^n \lambda[0, x_{nN}] - m[0, x_{nN}]| \geq (\sqrt{1 + \delta} - 1) x_{nN} \geq (\sqrt{1 + \delta} - 1) c_1 (\frac{1}{nN})^{1/\alpha}$ , by Lemma 1 (i).  $\square$

*Proof of Theorem 4 (i):* Symmetrical versions of the previous two lemmas hold near 1. Together they prove part (i) of the theorem.  $\square$

**Lemma 7** *For  $k \in \mathbb{N}$  there is a function  $\eta_k$ , supported on  $J_k$  such that*

$$f_*^{k+1}(\eta_k m) = m_{\Delta_0},$$

*$\|\eta_k\|_{L^1} = m(\Delta_0)$  and  $\|\eta_k\|_{L^\infty} \leq C' k^{1+1/\alpha}$ . The constant  $C'$  is independent of  $k$ , but may depend on  $\alpha$ .*

*Proof:* Put  $\eta_k = (f^{k+1})' 1_{J_k}$ . Since  $f^{k+1}$  maps  $J_k$  injectively onto  $\Delta_0$ , the Frobenius–Perron operator maps  $\eta_k$  onto  $\mathbf{1}|_{\Delta_0}$ . The proof of Lemma 3 reveals that for any  $u, v \in J_k$

$$\left| \frac{(f^{k+1})'(u)}{(f^{k+1})'(v)} \right| \leq D$$

(where  $D$  is the constant in Lemma 3). Thus, restricted to  $J_k$ ,

$$|(f^{k+1})'| \leq D \frac{m(\Delta_0)}{m(J_k)} \leq D \frac{m(\Delta_0)}{c k^{-1-1/\alpha}}$$

for some  $c$ .  $\square$

*Proof of Theorem 4 (ii)(a):* Let  $\eta_k$ , supported on  $J_k$  be given by Lemma 7. A similar construction applies on  $J'_k$  so  $\eta_k - \eta'_k$  is supported on  $J_k \cup J'_k$ , positive on  $J_k$  and negative on  $J'_k$  such that

$$|f_*^n[(\eta_k - \eta'_k)m]| = \begin{cases} 2m(\Delta_0) & \text{if } n \leq k, \\ 0 & \text{if } n > k. \end{cases}$$

Define  $\eta = \sum_{k=1}^{\infty} \{h(k) - h(k+1)\}(\eta_k - \eta'_k)$ . Then

$$\int_0^1 \eta \, dm = 0 \text{ and } \int_0^1 \psi \eta \, dm = |\eta m| = 2m(\Delta_0)h(1).$$

Moreover,  $f_*^n(\eta m)|_{(1/2,1)} \leq 0 \leq f_*^n(\eta m)|_{(0,1/2)}$ . Now put  $\varphi_h = 1 + \eta$ . Clearly,  $\varphi_h \in L^1$ . Next, since  $\int_0^1 \varphi_h \, dm = 1$  and  $m = f_*^n m$ ,

$$\begin{aligned} \int_0^1 \psi \circ f^n \varphi_h \, dm - \int_0^1 \psi \, dm \int_0^1 \varphi_h \, dm &= \int_0^1 \psi \, d(f_*^n \{(\varphi_h - 1)m\}) \\ &= \int_0^1 \psi \, d(f_*^n \{\eta m\}) \\ &= |f_*^n \{\eta m\}| \\ &= 2m(\Delta_0) \sum_{k=n}^{\infty} \{h(k) - h(k+1)\} \\ &\approx h(n). \end{aligned}$$

*Proof of (ii)(b):* Combining the constructions and Lemma 7,

$$\sup |\eta| \leq C' \sup_{k \geq 1} |k^{1+1/\alpha} \{h(k) - h(k+1)\}|.$$

But  $|h(k) - h(k+1)| = h'(k + \theta_k)$  for some  $\theta_k \in [0, 1]$  so that

$$k^{1+1/\alpha} |h(k) - h(k+1)| \leq (k + \theta_k)^{1+1/\alpha} |h'(k + \theta_k)|$$

is bounded. Thus,  $\eta$  is bounded, so one can put  $\varphi_h = 1 + \frac{\eta}{2|\eta|_\infty}$ , to obtain an  $L^\infty$  density which is bounded below by  $\frac{1}{2}$ . Finally, let  $\hat{\varphi}_h = \varphi_h \circ \Phi$  (where  $\Phi$  is the semi-conjugacy mapping  $F$  onto  $f$ ) and let  $(x, l), (y, l) \in \Delta_{l, k+l}$  ( $k \geq 1$ ). ( $\hat{\varphi}_h = 0$  on  $\Delta_0$  and  $\Delta_{l,l}^{(l)}$  and a similar argument works on  $\Delta'_{l, k+l}$ .) Then

$$\begin{aligned} \left| \frac{\hat{\varphi}_h(x, l)}{\hat{\varphi}_h(y, l)} - 1 \right| &= \frac{h(k) - h(k+1)}{2|\eta|_\infty \varphi_h(y, l) (f^{k+1})'(f^l(y))} \left| \frac{(f^{k+1})'(f^l(y))}{(f^{k+1})'(f^l(x))} - 1 \right| \\ &\leq \frac{h(1)}{|\eta|_\infty} \left| \frac{(f^{k+1})'(f^l(y))}{(f^{k+1})'(f^l(x))} - 1 \right| \leq C |f^R(y) - f^R(z)| \end{aligned}$$

by standard distortion estimates (cf. proof of Lemma 3). Now  $\hat{\varphi}_h \in C_\beta^+$  by Lemma 2.  $\square$

*Proof of part (ii)(c):* Under the additional hypothesis,

$$\lim_{x \rightarrow 0} |\eta(x)| = \lim_{k \rightarrow \infty} |\eta \mathbf{1}_{J_k}|_\infty \leq C' \lim_{k \rightarrow \infty} |k^{1+1/\alpha} (h(k) - h(k+1))| = 0.$$

A similar argument works for the limit at 1.  $\square$

### 3 Entropy

We have already noted without proof that the upper/lower partition  $P$  induced by  $\phi_\alpha$  is a generator for the map  $g_\alpha$ . In order to verify this, it is sufficient to observe that  $P$  separates (almost all) points in  $S$ . We now quote a simple theorem which applies in this situation.

**Lemma 8** ([1] 4.1) *Suppose the cutting function  $\phi$  defining a generalized baker's transformation  $g$  satisfies  $m\{\phi(x) = 0 \text{ or } \phi(x) = 1\} < 1$  and the factor  $f$  on vertical fibres is ergodic. Then the upper/lower partition  $P$  is a generator.*

A simple but useful consequence of this observation is that  $P_{-1}^{-\infty} = \bigvee_{i=1}^{\infty} g_\alpha^{-i} P$  is exactly to the sub-sigma-algebra of vertical fibres on  $S$ .

Using this, Kolmogorov's theorem and Rohklin's formula for entropy, we obtain the following intuitive formula:

$$h(g_\alpha) = h(f) = -2 \int_0^1 \phi_\alpha(x) \log \phi_\alpha(x) dx.$$

Next, we look at this from a slightly different point of view. At each point  $(x, y) \in S$  the stable manifold is simply the vertical fibre at  $x$  while the unstable manifold is the graph of a (continuous) function on  $[0, 1]$ . By Theorem 3, our maps  $f$  are ergodic (in fact, exact) and so the Lyapunov exponents at  $(x, y)$  are easily computed:

$$\begin{aligned} \lambda_+(x, y) &= - \int_0^{1/2} \log \phi_\alpha(f(x)) dx - \int_{1/2}^1 \log(1 - \phi_\alpha(f(x))) dx \\ &= - \int_0^1 \phi_\alpha(x) \log \phi_\alpha(x) dx - \int_0^1 (1 - \phi_\alpha(x)) \log(1 - \phi_\alpha(x)) dx \\ &= -2 \int_0^1 \phi_\alpha(x) \ln \phi_\alpha(x) dx. \end{aligned}$$

while  $\lambda_- = -\lambda_+$ , both independent of the point  $(x, y)$ . Thus Pesin's formula [6] for the entropy

$$h_\mu(g_\alpha) = \int_S \lambda_+(x, y) d\mu(x, y)$$

has a transparent link with the purely measure-theoretic calculation obtained in the first paragraph.

## 4 The Bernoulli Property

We show that for every  $\alpha \in (0, \infty)$ , the baker's map  $g_\alpha$  is measurably isomorphic to a Bernoulli shift (exactly which shift is determined by the entropy  $h(g_\alpha)$  computed by the formula derived in Section 3.) In the case of  $\alpha = 1$  the Bernoulli property has previously been observed by M. Rahe [4]. Therefore we include in this section proofs for all values of  $\alpha$  simultaneously. Our argument follows [4] in form.

Isomorphism to a Bernoulli shift follows if we can show that the partition  $P$  of  $S$  into upper and lower regions relative to the cut function  $\phi_\alpha$  is weak-Bernoulli [5]. This is, in turn, implied by the following result of del Junco and Rahe [2] on the rate of convergence of conditional entropy.

**Theorem 5** *Suppose  $(X, T, \mu)$  is an invertible, mixing, measure-preserving system and  $P$  is a generating partition for  $T$ . Suppose that*

$$\sum_n \int_X \nu \circ \mu(P|P_{-1}^{-n})(x) - \nu \circ \mu(P|P_{-1}^{-\infty})(x) d\mu(x) < \infty$$

where  $\nu(t) = -t \log t$ ,  $t \geq 0$  and  $\mu(\cdot|\cdot)$  denotes conditional probability. Then  $P$  is a weak-Bernoulli partition for  $T$ .

Application of this result is particularly transparent for our examples since  $P_{-1}^{-n}$  and  $P_{-1}^{-\infty}$  are partitions into vertical columns and vertical fibres on  $S$  respectively. We adopt the convention that the atom  $P_0$  of  $P$  is the area under the cut function and  $P_1$  the area above. Then  $\mu(P_0|P_{-1}^{-\infty})(x, y) = \phi_\alpha(x)$  and, for  $(x, y) \in A \times [0, 1] \in P_{-1}^{-n}$ ,  $\mu(P_0|P_{-1}^{-n})(x, y) = 1/\mu(A) \int_A \phi_\alpha(s) dm(s) := \phi_\alpha^n(x)$ , the average value of  $\phi_\alpha$  over the base of the column. Corresponding (and symmetric) statements hold for  $\mu(P_1|\cdot)$  by replacing  $\phi_\alpha$  with  $1 - \phi_\alpha(x)$ .

Note that in view of Theorem 3,  $g_\alpha$  is a  $K$ -automorphism, hence mixing.

Summability of the conditional entropy differences is a consequence of the following estimate.

**Lemma 9** *For each  $\alpha \in (0, \infty)$ , there is a constant  $C = C(\alpha) < \infty$  such that for all sufficiently large (again, depending on  $\alpha$ )  $n$  we have*

$$0 \leq \int_0^1 \nu \circ \phi_\alpha^n(x) - \nu \circ \phi_\alpha(x) dx \leq \frac{C \log n}{n^{1+\frac{1}{\alpha}}}$$

*Proof:* Fix  $\alpha$  and write  $f$  for  $f_\alpha$  and  $\phi$  for  $\phi_\alpha$ . For  $n \geq 3$  fixed, consider the atoms of  $P_{-1}^{-n}$ ; full vertical columns over intervals in  $[0, 1]$ . Label these intervals  $A_1^n, A_2^n, \dots, A_{2^n}^n$  from left to right and the endpoints of these intervals as  $0 = a_0 < a_1 < a_2 < \dots < a_{2^n} = 1$ . For example,  $A_1^n = [0, a_1]$ ,  $A_2^n = [a_1, a_2]$  ... and so-on. We begin with estimates on the position of the  $a_j$ . Let  $l(x), r(x)$  denote the left and right branches of  $f$  respectively and observe that  $a_1 = l^{-n+1}(\frac{1}{2})$ . Since  $l^{-1}(\frac{1}{2}) \in (x_1, x_0)$  we have  $a_1 \in (x_{n-1}, x_{n-2})$ . Now  $a_2 = l^{-n+2}(\frac{1}{2})$  since  $f^{n-1}(A_2^n) = [\frac{1}{2}, 1]$ , so we can estimate  $a_2 \in (x_{n-2}, x_{n-3})$ . In particular,  $a_1 \approx (\frac{1}{n})^{1/\alpha}$  while  $a_2 \approx (\frac{1}{n-1})^{1/\alpha}$ . Therefore,  $m(A_2^n) \leq c_1 (\frac{1}{n})^{1+1/\alpha}$  for some constant  $c_1$  (independent of  $n$ ). A final preliminary observation is that, for  $j = 3, 4, \dots, 2^{n-1}$ ,  $m(A_j^n) \leq m(A_2^n)$ . This is easily proved by induction on  $n \geq 2$  using the fact that on the left branch of  $f$ ,  $f'$  is increasing. By symmetry, we have

- $1 - a_{2^{n-1}} = a_1 \approx (\frac{1}{n})^{1/\alpha}$
- $m(A_{2^{n-1}}^n) = m(A_2^n) \approx (\frac{1}{n})^{1+1/\alpha}$
- For  $j = 2, 3, 4, \dots, 2^{n-1} - 1$ ,  $m(A_{2^n-j}^n) \leq m(A_{2^{n-1}}^n)$

We now estimate the integral. There exists a constant  $c_2$  such that For all  $n$  and  $x \in A_j^n$ ,  $j = 2, \dots, 2^n - 1$

$$|\phi^n(x) - \phi(x)| \leq \|\phi'\|_\infty m(A_j^n) \leq c_2 \|\phi'\|_\infty (\frac{1}{n})^{1+1/\alpha}$$

Using this pointwise estimate in the integral for the conditional entropy difference, along with the mean value theorem on  $\nu$  yields

$$\int_{a_1}^{a_{2^{n-1}}} |\nu \circ \phi^n(x) - \nu \circ \phi(x)| dx \leq \frac{c_3 \log n}{n^{1+1/\alpha}}$$

for an appropriate choice of constant  $c_3$  independent of  $n$ .

On the interval  $A_1^n$  we also make a simple estimate. Since both  $\phi(x)$  and  $\phi^n(x)$  lie in  $[1 - 2^{\alpha-1}a_1^\alpha, 1]$ , using a linear estimate on  $\nu$  near 1 implies there exists a constant  $c_4$  such that for all  $0 \leq x \leq a_1$  we have

$$|\nu \circ \phi^n(x) - \nu \circ \phi(x)| \leq c_4/n$$

Integrating gives, for appropriately chosen constant  $c_5$

$$\left| \int_0^{a_1} \nu \circ \phi^n(x) - \nu \circ \phi(x) dx \right| \leq c_5 \left( \frac{1}{n} \right)^{1+1/\alpha}$$

which is clearly smaller than our claimed upper bound.

Finally, on the interval  $A_{2^n}$  we use Jensen's inequality.

$$\frac{1}{m(A_{2^n})} \int_{a_{2^{n-1}}}^1 \nu \circ \phi(x) dx \leq \nu \left( \frac{1}{m(A_{2^n})} \int_{a_{2^{n-1}}}^1 \phi(x) dx \right)$$

But

$$\frac{1}{m(A_{2^n})} \int_{a_{2^{n-1}}}^1 \phi(x) dx = \frac{1}{m(A_1)} \int_0^{a_1} 2^{\alpha-1} x^\alpha dx \tag{5}$$

$$= \frac{2^{\alpha-1}}{a_1(\alpha+1)} a_1^{\alpha+1} \tag{6}$$

$$\leq c_6/n \tag{7}$$

for an appropriately chosen constant  $c_6$ . Then we have  $\nu(\phi^n(x)) = \frac{c_6}{n} (\log n - \log c_6)$  at each  $x \in A_{2^n}$ . Using this estimate, we see that

$$\int_{a_{2^{n-1}}}^1 \nu \circ \phi^n(x) - \nu \circ \phi(x) dx \tag{8}$$

$$= a_1 \frac{1}{m(A_{2^n})} \int_{a_{2^{n-1}}}^1 \nu \circ \phi^n(x) - \nu \circ \phi(x) dx \tag{9}$$

$$\leq a_1 \nu(\phi^n(x))|_{A_{2^n}} \tag{10}$$

$$\leq \frac{c_7 \log n}{n^{1+1/\alpha}} \tag{11}$$

Note that the integral in the second line above is positive (by Jensen's) and hence so is the first integral. Therefore we have provided an upper estimate on the absolute value of the integral over  $A_{2^n}$ . Putting all three estimates together gives the upper bound claimed in the lemma.  $\square$

By symmetry we obtain similar estimates on the functions  $1 - \phi_\alpha$  and  $(1 - \phi_\alpha)^n$ , leading to the required summability for terms involving  $P_1$  instead of  $P_0$ . This completes the derivation of the summability in Theorem 5.

## Appendix: precise distortion and decay estimates

For each  $\alpha$ ,  $\phi_\alpha$  and  $f_\alpha$  are defined via (3) and (2). It is useful to observe

$$f_\alpha'(x) = \begin{cases} \frac{1}{\phi_\alpha(f_\alpha(x))} & x < \frac{1}{2}, \\ \frac{1}{1 - \phi_\alpha(f_\alpha(x))} & x > \frac{1}{2}. \end{cases}$$

One can also calculate for  $x < \frac{1}{2}$  and the left branch of  $f_\alpha$ ,

$$x - f_\alpha^{-1}(x) = \int_0^x (1 - \phi_\alpha(t)) dt = \frac{2^{\alpha-1}}{\alpha+1} x^{1+\alpha}. \quad (12)$$

Continue with the notation  $x_0$  the left most period-2 point and  $x_k = f_\alpha^{-1}(x_{k-1}) \cap [0, x_k)$ .

### Proof of Lemma 1 on asymptotics of the $x_n$

(i) First, for any  $y \geq 2^\alpha$ ,  $z \geq 0$ , the mean value theorem and (12) give

$$\begin{aligned} \frac{[\frac{1}{y}]^{1/\alpha} - [\frac{1}{y+z}]^{1/\alpha}}{[\frac{1}{y}]^{1/\alpha} - f_\alpha^{-1}([\frac{1}{y}]^{1/\alpha})} &= \frac{\alpha+1}{\alpha 2^{\alpha-1}} \left[ \frac{1}{y+\theta z} \right]^{1/\alpha-1} \left( \frac{1}{y} - \frac{1}{y+z} \right) y^{1+1/\alpha} \\ &= \frac{\alpha+1}{\alpha 2^{\alpha-1}} \left[ \frac{y}{y+\theta z} \right]^{1/\alpha} \left[ \frac{y+\theta z}{y+z} \right] z \end{aligned} \quad (13)$$

(where  $\theta \in [0, 1]$ ). The upper and lower bounds are obtained by distinct applications of (13). For the upper bound note that if  $z = \frac{\alpha 2^{\alpha-1}}{\alpha+1}$  then the RHS is bounded above by 1, so that

$$\left[\frac{1}{y}\right]^{1/\alpha} - \left[\frac{1}{y+z}\right]^{1/\alpha} \leq \left[\frac{1}{y}\right]^{1/\alpha} - f_\alpha^{-1}\left(\left[\frac{1}{y}\right]^{1/\alpha}\right).$$

In particular,  $f_\alpha^{-1}\left(\left[\frac{1}{y}\right]^{1/\alpha}\right) \leq \left[\frac{1}{y+z}\right]^{1/\alpha}$ , so that by using  $y = x_0^{-\alpha}$  and induction,

$$x_n = f_\alpha^{-n}(x_0) = f_\alpha^{-n}\left(\left[\frac{1}{y}\right]^{1/\alpha}\right) \leq \left[\frac{1}{y+nz}\right]^{1/\alpha} \leq \frac{1}{z^{1/\alpha}} \left[\frac{1}{n}\right]^{1/\alpha}.$$

On the other hand, if  $y \geq z$  then the RHS of (13) is bounded below by  $\frac{\alpha+1}{\alpha 2^{\alpha-1}} \frac{1}{2^{1+1/\alpha}} z$ . Pick  $z$  such that this quantity is equal to 1. Finally, choose  $y = \max\{z, x_0^{-\alpha}\}$  so that

$$x_n = f_\alpha^{-n}(x_0) \geq f_\alpha^{-n}\left(\left[\frac{1}{y}\right]^{1/\alpha}\right) \geq \left[\frac{1}{y+nz}\right]^{1/\alpha} \geq \frac{1}{(2y)^{1/\alpha}} \left[\frac{1}{n}\right]^{1/\alpha}.$$

(ii) Since  $J_k = [x_{k+1}, x_k]$ , we have  $m(J_k) = x_k - x_{k+1} \approx x_k^{1+\alpha} \approx \left[\frac{1}{k}\right]^{1+1/\alpha}$  by (12) and part (i) of the lemma.

(iii) Since  $f_\alpha : I_k \rightarrow J_k$  bijectively, there is an  $x \in I_k$  such that

$$m(I_k) = \frac{m(J_k)}{f_\alpha'(x)} = m(J_k) (1 - \phi_\alpha(f_\alpha(x))) \approx \left[\frac{1}{k}\right]^{1+1/\alpha} \left(\left[\frac{1}{k}\right]^{1/\alpha}\right)^\alpha$$

since  $f_\alpha(x) \in J_k \subset [0, \frac{1}{2}]$ .

(iv) When  $n \leq \rho k$ ,  $\left[\frac{1}{k+n}\right] \approx \left[\frac{1}{k}\right]$  so the estimate follows from parts (i) and (ii) and the fact that  $x_k - x_{k+n} = \sum_{k \leq i < k+n} m(J_i)$ .  $\square$

### Proof of Lemma 3 on uniform distortion

We assume that  $y, z \in I_i \subset \Delta_{0,i}$  since the case where  $y, z \in I_i'$  is similar. For each  $1 \leq k < i+1 = R$  let  $y_k = f^{R-k}(y)$  and  $z_k = f^{R-k}(z)$ . Thus  $y_k, z_k \in J_{k-1}$ . Now,

$$[\log(f')]'|_{J_k} = \frac{f''}{f'} \Big|_{J_k} = \left[ \frac{-\phi_{\alpha'}}{\phi_{\alpha^2}} \right] \circ f|_{J_k} \approx \left( \left[ \frac{1}{k+1} \right]^{1/\alpha} \right)^{\alpha-1}$$

(since  $\phi_\alpha|_{f(J_k)}$  is bounded below and  $\phi_\alpha' \circ f(x) \approx [f(x)]^{\alpha-1} \approx x^{\alpha-1}$  when  $x \in J_k$ ). Thus,

$$\begin{aligned} \left| \log \frac{f'(y_k)}{f'(z_k)} \right| &\leq c \left[ \frac{1}{k} \right]^{1-1/\alpha} |y_k - z_k| = c \left[ \frac{1}{k} \right]^{1-1/\alpha} m(J_{k-1}) \frac{|y_k - z_k|}{m(J_{k-1})} \\ &\leq c' \left[ \frac{1}{k} \right]^2 \frac{|y_k - z_k|}{m(J_{k-1})} \leq c' \left[ \frac{1}{k} \right]^2 \end{aligned} \quad (14)$$

since  $|y_k - z_k| \leq m(J_{k-1}) \approx m(J_k)$ , where the latter is estimated by Lemma 1 (ii). A similar calculation shows that

$$\left| \log \frac{f'(y)}{f'(z)} \right| \leq \frac{c''}{i} \frac{|y-z|}{m(I_i)} \leq \frac{c''}{i} \quad (15)$$

for some  $c''$  independent of  $y, z, i$  (but possibly depending on  $\alpha$ ). Now, since  $(f^R)'(y) = f'(y) f'(y_{R-1}) \cdots f'(y_1)$  (and similarly for  $z$ ),

$$\left| \log \frac{(f^R)'(y)}{(f^R)'(z)} \right| = \left| \log \frac{f'(y)}{f'(z)} \right| + \sum_{k < R} \left| \log \frac{f'(y_k)}{f'(z_k)} \right| < c'' + c' \sum_{k=1}^{\infty} \frac{1}{k^2} \stackrel{\text{def}}{=} C. \quad (16)$$

Now put  $D = e^C$ . Since this inequality holds uniformly for any choice of  $y, z \in I_i$  and the map  $f^R : I_i \rightarrow \Delta_0$  is bijective, we have

$$\frac{|y-z|}{m(I_i)} \leq D \frac{|f^R(y) - f^R(z)|}{m(\Delta_0)}.$$

Similarly,  $\frac{(f^k)'(y_k)}{(f^k)'(z_k)} \leq D$  and since  $f^k(y_k) = f^R(y)$  and  $f^k(z_k) = f^R(z)$ ,

$$\frac{|y_k - z_k|}{m(J_{k-1})} \leq D \frac{|f^R(y) - f^R(z)|}{m(f^k(J_{k-1}))} = D \frac{|f^R(y) - f^R(z)|}{m(\Delta_0)}.$$

The last two displayed expressions can now be used to refine (15) and (14), yielding

$$\left| \log \frac{f'(y)}{f'(z)} \right| \leq \frac{c''}{i} D \frac{|f^R(y) - f^R(z)|}{m(\Delta_0)} \quad \text{and} \quad \left| \log \frac{f'(y_k)}{f'(z_k)} \right| \leq c' \left[ \frac{1}{k} \right]^2 D \frac{|f^R(y) - f^R(z)|}{m(\Delta_0)}$$

from which:

$$\left| \log \frac{(f^R)'(y)}{(f^R)'(z)} \right| \leq C D \frac{|f^R(y) - f^R(z)|}{m(\Delta_0)}.$$

Finally, if  $|\log x| < C$  then  $|\log x| > \frac{C}{e^C - 1} |x - 1|$ . In view of (16),

$$\left| \frac{(f^R)'(y)}{(f^R)'(z)} - 1 \right| \leq \frac{D-1}{C} \left| \log \frac{(f^R)'(y)}{(f^R)'(z)} \right| \leq \frac{D(D-1)}{m(\Delta_0)} |f^R(y) - f^R(z)|.$$

This completes the proof.

## References

- [1] C. Bose, Generalized baker's transformations. *Erg. Th. Dyn. Syst.* **9**, (1989), 1–17.
- [2] A. del Junco and M. Rahe, Finitary codings and weak Bernoulli partitions. *Proc. Amer. Math. Soc.* **75** #2 (1979), 259–264.
- [3] G. Pianigiani, First Return Map and Invariant Measures. *Isr. J. Math.* **35** (1980), 32-48.
- [4] M. Rahe, On a class of generalized baker's transformations. *Can. J. Math* **45** (1993), 638–649.
- [5] N. Friedman and D. Ornstein, On isomorphism of weak Bernoulli transformations. *Adv. Math.* **5** (1971), 365-394.
- [6] R. Mane, Ergodic Theory and Differentiable Dynamics. Springer Verlag, Berlin, 1989.
- [7] L.S. Young. Statistical Properties of Dynamical Systems with Some Hyperbolicity. *Ann. Math.* **147** (3) (1998), 585-650.
- [8] L.S. Young, Recurrence times and rates of mixing. *Isr. J. Math.* **110** (1999), 153–188.