MIXING EXAMPLES IN THE CLASS OF PIECEWISE MONOTONE AND CONTINUOUS MAPS OF THE UNIT INTERVAL

by

CHRISTOPHER BOSE

DMS–546–IR

July 1990

MIXING EXAMPLES IN THE CLASS OF PIECEWISE MONOTONE AND CONTINUOUS MAPS OF THE UNIT INTERVAL

by

CHRISTOPHER BOSE[1]

Department of Mathematics and Statistics
University of Victoria

[1]Supported by NSERC grant OGP0046586 90.

Section 1

Let $I = [0,1]$ equipped with Lebesgue measure λ on the Lebesgue subsets $\mathcal{B}$. We shall be concerned with measurable point mappings $g : I \longrightarrow I$ which satisfy:

1.1 There exists a finite or countably infinite collection of closed intervals $I_n \subseteq I$, $I_n = [a_n, b_n]$, such that $\lambda(I_n \cap I_m) = 0$ iff $m \neq n$ and $\bigcup_n I_n = I$, and such that

1.2 Restricted to each I_n, $g|_{I_n}$ is continuous and monotone (non–increasing or non–decreasing.)

We shall call such mappings *piecewise monotone and continuous* (p.m.c.). To study the dynamics of such a map one looks for an invariant measure μ for g; to avoid trivialities one specifies $\mu << \lambda$. This is an old idea going back at least to Renyi [R], Khintchine [Kh] and Doeblin [D], and the numerous articles cited there concerning number theoretic transformations.

The problem becomes tractable when we impose additional conditions on g. Here is a much studied situation. Let g be p.m.c. and satisfy

1.3 There exists a $\lambda > 1$ such that, for all n

$$\operatorname*{ess\,inf}_{x \in I_n} \{|g'(x)|\} \geq \lambda$$

1.4 $g|_{I_n}$. is twice continuously differentiable.

In case there are finitely many intervals I_n Lasota and Yorke [L&Y] proved the existence of absolutely continuous invariant measures for g satisfying 1.1) – 1.4). Under the same conditions Bowen [Bow] went on to show that if g is weak-mixing with respect to such an invariant μ, then the natural extension of g (with respect to μ) is automatically a Bernoulli shift.

Earlier, Adler [Ad] had found unique invariant measures for p.m.c. g satisfying 1.3) and 1.4) with infinitely many intervals I_n but with the additional conditions

1.5 $$\exists M, \ \forall n \quad \sup_{x,y,z \in I_n} \left| \frac{g''(x)}{g'(y)g'(z)} \right| \leq M$$

(trivial for finitely many intervals)

1.6 $$g(I_n) = I.$$

1.5) is known as Renyi's condition. Bowen and Series [Bow & S] observed that 1.6) may be weakened to the *Markov Condition*

1.7 If $$\Omega = \bigcup_n \left\{ \lim_{x \to a_n^+} g(x), \ \lim_{x \to b_n^-} g(x) \right\} \quad \text{then} \quad \Omega \subseteq \bigcup_n \{a_n, b_n\}$$

and, for all ℓ, k

$$I_\ell \subseteq \bigcup_{n \geq 0} g^n(I_k).$$

Under conditions 1.1) – 1.5) and with 1.7) their result ensures a *unique* invariant measure $\mu << \lambda$ and it turns out that the natural extension of g with respect to μ will be a Bernoulli shift.

A number of authors have weakened Renyi's condition 1.5) obtaining analogues of the Bowen & Series theorem for not necessarily C^2 maps. See, for example, [W], [Ke], and [Bo2]. A common feature amongst these results is that conditions sufficient to ensure the existence of a unique absolutely continuous invariant measure imply, with respect to this measure, the natural extension of g is a Bernoulli shift.

The purpose of this article is to show that, at least in the class of p.m.c. maps with unique absolutely continuous invariant measures, this type of behavior (*i.e.*, weak–mixing $\Rightarrow$ Bernoulli) is not inevitable. In particular, we construct p.m.c. maps g with Lebesgue measure the unique absolutely continuous invariant measure and which are, in order, ergodic but not weak–mixing, weak–mixing but not mixing and mixing but not exact (natural extension not a K–automorphism).

The central idea is contained in Section 3 where we describe a general method for extracting a p.m.c. interval map as a factor of the direct product of an abstract dynamical system (S,ν) and a Bernoulli shift. We require the system (S,ν) to possess the "$\bar{d} > \delta$ property": a lowerbounded $\bar{d}$ separation between (almost every) pair of names with respect to some generating partition. This property seems new to the literature and appears to be of interest beyond its use in our constructions. The bridge between the above dynamical system and the class of p.m.c. interval maps is provided by the generalized baker's transformation.

In Section 4 we construct our first two examples. Section 5 is reserved for the mixing, but not exact case, where it turns out any mixing rank 1 automorphism will suffice for the transformation S. In Section 2 we will establish notation, conventions and sketch the few facts about rank 1 transformations we shall be using.

The results in this article have been greatly clarified through the helpful comments of D. Rudolph and A. del Junco. A good deal of background material appears in King [Ki] and Kalikow [Ka]. We are please to acknowledge these contributions.

Section 2

In this section we establish our notation and a few preliminary lemmas, wherever possible adhering to what is standard. The knowledgeable reader, particularly one familiar with the rank–1 block construction, may prefer to begin directly with Section 3.

By a *dynamical system* $(X, \mathcal{B}, \mu, T)$ we shall mean a Lebesgue probability space $(X, \mathcal{B}, \mu)$ equipped with a point mapping $T : X \longrightarrow X$ which is measurable and measure–preserving: $T^{-1}A \in \mathcal{B}$ and $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$.

A *partition* $P = \{P_i\}_{i \in II}$ will be a finite or countable collection of disjoint subsets $P_i \in \mathcal{B}$ satisfying $\mu(\cup P_i) = 1$. Partition elements P_i satisfying $\mu(P_i) > 0$ will be called *atoms* of P. Given $P = \{P_i\}_{i \in II}$, and $Q = \{Q_j\}_{j \in II'}$ of the same space X we define the *join* $P \vee Q = \{P_i \cap Q_j\}_{\substack{i \in II \\ j \in II'}}$ or *common refinement* of P and Q. We say that P *refines* Q (and write $Q < P$) if each atom of Q is (up to null sets) a union of atoms of P. We say that P ϵ *refines* Q (and write $Q \overset{\epsilon}{<} P$) if there exists a $\tilde{Q} < P$ having the same number of atoms as Q and $\sum_i \mu(Q_i \vartriangle \tilde{Q}_i) < \epsilon$. In the presence of T, P gives rise to countably many partitions $T^kP = \{T^kP_i\}_{i \in II}$ ($k \in \mathbb{Z}$ if T invertible, $k \in \mathbb{Z}^-$ if T is not.) We say that P *generates* $\mathcal{B}$ under T if the smallest σ–algebra containing all the T^kP is $\mathcal{B}$. It is an important observation for us that, since X is a Lebesgue space, this statement is equivalent to requiring that the collection $\{T^kP\}$ separates (μ–almost all) points of X.

By a *process* we shall mean a pair consisting of a dynamical system

$(X, \mathcal{B}, \mu, T)$ with T invertible and a partition P of X. We write $(X, \mathcal{B}, \mu, T, P)$ or when there is little danger of confusion about the underlaying measure space, simply (T,P).

Given a process (T,P) and a point $x \in X$ we may consider the *P-T name* of x, a member of $II^{\mathbb{Z}}$, denoted $x_{-\infty}^{\infty}$ and defined by

$$(x_{-\infty}^{\infty})_j = i_0 \Leftrightarrow T^j x \in P_{i_0}$$

$$(\Leftrightarrow x \in T^{-j} P_{i_0}).$$

Evidently, μ a.e. $x \in X$ has a well defined P–T name. If $x_{-\infty}^{\infty} \in II^{\mathbb{Z}}$ and $m \leq n$ we denote by $x_m^n = x_m x_{m+1} \cdots x_n$. x_k means x_k^k.

If (T,P) is a process, $m \leq 0 \leq n$ integers and $i_m i_{m+1} \cdots i_n \in II^{n-m+1}$ a finite string of symbols we form the *elementary cylinder*

$$\left\{x \in X \mid X_m^n = i_m\, i_{m+1} \cdots i_n\right\} = T^{-m} P_{i_m} \cap T^{-m-1} P_{i_{m+1}} \cdots \cap T^{-n} P_{i_n}$$

measurable w.r.t. $\bigvee_m^n T^{-i} P$. If T is ergodic, we may apply the Birkhoff Ergodic Theorem obtaining a subset $X' \subseteq X$, $\mu(X') = 1$ so that if $x' \in X'$ and C is an elementary cylinder, $C = \left\{x \mid x_m^n = i_m\, i_{m+1} \cdots i_n\right\}$,

$$\mu(C) = \lim_{N \to \infty} \frac{1}{N - (n-m+1)} \left\{i \in [0\ N - (n-m+1)) \mid x_i^{i+(n-m)} = i_m\, i_{m+1} \cdots i_n\right\}.$$

We call such x' *generic* for (T,P).

Let $\mathcal{A}$ be a finite set of distinct symbols and let $X = \mathcal{A}^{\mathbb{Z}}$. Define $S : X \longrightarrow X$ by $(Sx)_i = x_{i+1}$, the "shift to the left". Any S invariant probability measure μ on the product (of discrete) σ-algebra gives rise to a *shift dynamical system*. Let $P_a = \{x \in \mathcal{A}^{\mathbb{Z}} \mid x_0 = a\}$. Then $P = \{P_a\}_{a \in \mathcal{A}}$ is called the *time-zero partition* and P is trivially a generator under S.

Shift systems are fundamental in the following sense. Let $(X, \mathcal{B}, \mu, T)$ be any dynamical system with T invertible and suppose $P = \{P_i\}_{i \in II}$ generates under T. The P–T name of $x \in X$ defines a 1–1 mapping from (μ a.e.) $x \in X$ to $x_{-\infty}^{\infty} \in II^{\mathbb{Z}}$ which carries T to the shift S, and μ to a shift invariant measure on $II^{\mathbb{Z}}$. The original dynamical system is therefore measurably isomorphic to this shift system. The partition P is sent to the time-zero partition of $II^{\mathbb{Z}}$.

Let $A \in \mathcal{A}^N$ and $B \in \mathcal{A}^M$ be two finite strings, say $A = a_0 a_1 \cdots a_{N-1}$ and $B = b_0 b_1 \cdots b_{M-1}$. We define the *concatenation* of A and B, an element of $\mathcal{A}^{N+M}$ by

$$A \otimes B = a_0 a_1 \cdots a_{N-1}\, b_0 b_1 \cdots b_{M-1}.$$

If $N = M$ then we measure the *$\bar{d}$-distance* between A and B

$$\bar{d}(A,B) = \frac{1}{N} \quad \{i \in [0,N) \mid a_i \neq b_i\}.$$

We denote by $|A|$, $|B|$ (or sometimes $a = |A|$, $b = |B|$) the lengths of A and B.

As usual, a *Rohlin stack* of height b is a partition $R = \{A_1, A_2, \cdots, A_b, E\}$ with $TA_i = A_{i+1}$, $1 \le i \le b-1$. The set E is called the *error set*. A fundamental and useful observation in ergodic theory is that every aperiodic T on a Lebesgue space X admits arbitrarily long Rohlin stacks with arbitrarily small E. Restricting this notion further gives a geometric (and possibly the most intuitive) definition of a rank 1 automorphism.

Definition 2.1. *We say* $T : X \to X$ *is* rank 1 *if there exists a sequence of Rohlin stacks* $R_1, R_2, \cdots, R_n, \cdots$ *with* $\lim_{n \to \infty} \mu(E_n) = 0$ *and so that if* Q *is any finite partition, and* $\epsilon > 0$, *there exists* k *so* $Q \overset{\epsilon}{<} R_k$.

Remarks 2.2.

(i) If (X,μ) is non-atomic the condition $\mu(E_n) \to 0$ is redundant.

(ii) Suppose one finds a T-invariant set A with $0 < \mu(A) \le \frac{1}{2}$. Obtain a stack $R_n = \{A_1^{(n)}, \cdots, A_{b_n}^{(n)}, E_n\}$ whose levels approximate A very well; in particular there exists a level $A_i^{(n)} \in R_n$ so $\mu(A \cap A_i^{(n)}) > \frac{4}{5} \mu(A_i^{(n)})$. But then this same inequality must be true for all $A_j^{(n)}$, $1 \le j \le b_n$ and we conclude that A has large measure. Thus any rank 1 T is automatically *ergodic*.

(iii) A similar argument assures us that if T has a non-null set of points of period $\ell > 0$ then T itself must be of period ℓ. Thus, if T is rank 1 then it is either periodic, or aperiodic.

(iv) Let $P = \{P_a\}_{a \in \mathcal{A}}$ be a generator for T and let $S : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ be the related (isomorphic) shift automorphism. Let $\epsilon > 0$ be given. Then there is a single string $\beta = a_1 a_2 \cdots a_{b_n}$ of symbols from $\mathcal{A}$ so that for almost every $x \in \mathcal{A}^{\mathbb{Z}}$ and for all large L we may make a "phantom" copy of x_0^L, denoted $\tilde{x}_0^L$, a concatenation of copies of β interspersed with spacer symbols from $\mathcal{A}$, the spacers occupying at most a fraction ϵ of the indices, so that

$$\bar{d}(x_0^L, \tilde{x}_0^L) < \epsilon.$$

Thus, P–names are, within small $\bar{d}$ error, copies of β and a small proportion of spacer symbols.

It turns out that with a careful choice of P and the stack sequence a more rigid structure for P–names is ensured. This well–known construction is elegantly described in King [Ki]; we will simply sketch here what is possible. We will make essential use of these properties in what follows.

First, one may simultaneously construct a new sequence of *refining* stacks

$$\tilde{R}_1 < \tilde{R}_2 < \cdots < \tilde{R}_n < \cdots \qquad (\mu(\tilde{E}_n) \longrightarrow 0)$$

which generate the σ algebra, and a two set partition $P = \{P_0, P_1\}$ so that each stack level is contained in an atom of P (*i.e.*: $P < \tilde{R}_n$). The implications of this for P–names follows.

2.3. There exists a sequence of strings of zeros and ones (n–blocks) so that for all n, an (n+1)–block is a concatenation of n–blocks interspersed with spacers (for convenience always the symbol "1") and, $\epsilon > 0$ given, for almost every $x \in \{0,1\}^{\mathbb{Z}}$, for all large L, x_0^L is a concatenation of n–blocks interspersed with spacers occupying a proportion of at most ϵ of the indices in x_0^L.

In case T is aperiodic we may also ensure that the base $A_1^{(n)}$ of the stack $\tilde{R}_n = \{\tilde{A}_1^{(n)}, \tilde{A}_2^{(n)}, \cdots, \tilde{A}_{b_n}^{(n)}, \tilde{E}_n\}$ is measurable with respect to

$\bigvee_{0}^{b_n-1} T^{-i}P$; the implication for n-blocks being:

2.4. The position of n-blocks inside the (n+1)-blocks is uniquely determined.

Since $\mu(\tilde{E}_n) \downarrow 0$ we also see that

2.5. For almost every $x \in \{0,1\}^{\mathbb{Z}}$, for all large enough n, the time-zero coordinate of x, x_0 is interior to the appearance of an n-block in $x_{-\infty}^{\infty}$. We shall denote this "time-zero" n-block by $B_n(x)$.

One way to obtain **2.4)** is to ensure that the P-T name of $x \in A_1^{(n)}$ begins with a unique string:

2.6. We may assume an n-block begins with "1" followed by a string of n "0"'s followed by a "1", and that an n-block ends with a "1". All spacers are assigned the label "1".

In practice, a rank 1 example is usually constructed by giving a recursive definition of the n-blocks using a fixed symbol for spacers and obtaining the shift invariant measure μ on cylinders A as the limiting frequency of the appearance of A in an n-block. This point of view is nicely described by Kalikow in [Ka]. One thereby obtains **2.3)** and **2.5)** but not necessarily the conditions **2.4)** and **2.6)** for the time-zero partition.

We conclude this section with the observation that, in shift setting for aperiodic rank 1 T, without loss of generality, most appearances of an n-block in $x_{-\infty}^{\infty}$ will be followed by a spacer string whose length is a "small" fraction of the length of an n–block. To this end we define, if x_0 is in a time–zero n–block,

$$\ell_n(x) = \#\text{ of spacer symbols between } B_n(x) \text{ and the next n-block to the right of } B_n(x) \text{ in } x_{-\infty}^{\infty}.$$

Set $\ell_n(x) = 0$ if x_0 is not in an n–block. Observe that

$$\frac{1}{\mu(\tilde{A}_1^{(n)})} \int_{\tilde{A}_1^{(n)}} \ell_n(x)\, d\mu(x) = \mu(\tilde{E}_n).$$

Let $\delta_0 > 0$ be fixed and set

$$H_n = \{x \in \tilde{A}_1^{(n)} \mid \ell_n(x) > \delta_0 |B_n(x)|\}.$$

Since

$$\mu(\tilde{E}_n) \geq \delta_0 |B_n(x)|\ \mu(H_n) = \delta_0 |B_n(x)|\ \mu(\tilde{A}_1^{(n)})\ \frac{\mu(H_n)}{\mu(\tilde{A}_1^{(n)})}$$

we obtain

$$\frac{\mu(\tilde{E}_n)}{\delta_0 |B_n(x)|\ \mu(\tilde{A}_1^{(n)})} \geq \frac{\mu(H_n)}{\mu(\tilde{A}_1^{(n)})} \geq 0.$$

Since $|B_n(x)|\mu(\tilde{A}_1^{(n)}) = 1 - \mu(\tilde{E}_n) \longrightarrow 1$, the left hand side converges to zero. For large enough n, only a small fraction of x will have its time–zero n–block followed by a spacer string longer than $\delta_0|B_n(x)|$. We may now apply the Borel–Cantelli Lemma to obtain

Lemma 2.7. *Let* $\delta_0 > 0$ *be given. Then we may choose a subsequence of the* n*-blocks* B_{n_k} *so that, for almost every* $x \in \{0,1\}^{\mathbb{Z}}$ *and for all sufficiently large* k, *the time-zero* n_k*-block for* x, $B_{n_k}(x)$ *is followed by a spacer string of length less than* $\delta_0 |B_{n_k}(x)|$.

Section 3

We describe a general construction which embeds an abstractly constructed dynamical system as a factor of the natural extension of a piecewise monotone and continuous interval map. The main idea here appears to be new and it is hoped that it will find application beyond its use in this article.

The following condition will be imposed on our abstract dynamical system.

DEFINTION 3.1. *Let* $\delta > 0$. *We say that the process* $(X, \mathcal{B}, \mu, T, P)$ *satisfies* <u>$\bar{d} > \delta$ *property*</u> *if there is an* $X_0 \subseteq X$, $\mu(X_0) = 1$, *so that if* $x, y \in X_0$, $x \neq y$ *then either*

$$\limsup_{n \to \infty} \bar{d}(x_0^n, y_0^n) > \delta$$

or

$$\limsup_{n \to \infty} \bar{d}(x_{-n}^0, y_{-n}^0) > \delta$$

where $x_{-\infty}^{\infty}$ *and* $y_{-\infty}^{\infty}$ *are the* P–T *names of* x *and* y *respectively.*

Evidently, the above is a condition on a process. We say that the dynamical system $(X, \mathcal{B}, \mu, T)$ satisfies the $\bar{d} > \delta$ property if there exists a generating partition P so that the process $(X, \mathcal{B}, \mu, T, P)$ satisfies $\bar{d} > \delta$. The main tool we shall use in the constructions to follow is

THEOREM 3.2. *Let* $(Y, \mathcal{F}, \nu, S)$ *be a dynamical system satisfying the* $\bar{d} > \delta$ *property with respect to a generating partition containing* ℓ *atoms. Then* *there exists a p.m.c. (with respect to* ℓ *subintervals) map* $g : [0,1] \to [0,1]$ *which satisfies* 1.1), 1.2), 1.3) *and* 1.6), *which is Lebesgue–measure–preserving and whose natural extension is measurably isomorphic to a direct product of the transformation* S *and a Bernoulli shift.*

Let $R = \{R_0, R_1, \cdots, R_{\ell-1}\}$ be the generating partition of Y. Let $(\Omega, \mathcal{G}, p, \sigma)$ be a Bernoulli shift with independent generator $Q = \{Q_0, Q_1, \cdots, Q_{\ell-1}\}$ satisfying

$$p(Q_0) = 1 - \frac{\delta}{3}, \quad p(Q_1) = p(Q_2) = \cdots = p(Q_{\ell-1}) = \frac{\delta}{3(\ell-1)} .$$

Form the product dynamical system

$$(X, \mathcal{B}, \mu, T) = (Y \times \Omega, \mathcal{F} \times \mathcal{G}, \nu \times p, S \times \sigma)$$

and define a measurable partition of X as follows: $P = \{P_0, P_1, \cdots, P_{\ell-1}\}$ with

$$P_j = \bigcup_{s+t=j \pmod{\ell}} R_s \times Q_t, \qquad j = 0,1,\cdots,\ell-1.$$

There is a simple formula to construct the P–T name of $x = (y,\omega)$ given the R–S name of y and the Q–σ name of ω:

$$x_i = y_i + \omega_i \pmod{\ell}; \qquad i \in \mathbb{Z}.$$

We will need the following:

LEMMA 3.3. *The partition* P *is a generator for* T.

Proof. Let $Y_0 \subseteq Y$ be the set of full measure given by the $\bar{d} > \delta$ property and let $\Omega_0 \subseteq \Omega$ be the set of full measure whose points are separated by $\bigvee_{-\infty}^{\infty} \sigma^{-i}Q$ and are generic for Q. We will show that the P–T names of $x = (y,\omega)$ and $x' = (y',\omega')$ with $x \neq x'$ in $Y_0 \times \Omega_0$ are distinct.

Suppose not. Observe first that $y \neq y'$ since $\{\sigma^{-i}Q\}_{i\in\mathbb{Z}}$ separates the points of Ω_0 and for each s there is exactly one t so that $s + t = j$.

Next, obtain a sequence of integers $n_k \longrightarrow \infty$ (or $-n_k \longrightarrow -\infty$, the argument will be the same) so that $\bar{d}(y_0^{n_k}, y'^{n_k}_0) > \delta$ for all k. Conclude that $\bar{d}(\omega_0^{n_k}, \omega'^{n_k}_0) > \delta$ since under our supposition, ω_i and ω'_i will differ on every index i such that $y_i \neq y'_i$.

But this gives a contradiction, since for sufficiently large k, the number of indices $i \in [0,n_k)$ for which ω_i and ω'_i will both be zero exceeds $\left[1 - \left[\frac{2\delta}{3} + \frac{\delta}{3}\right]\right]n_k = (1-\delta)n_k$ (the extra $\frac{\delta}{3}$ to take care of the difference between frequency of zero and $\mu(Q_0)$) whenever ω and ω' are generic for Q. ■

Now, for $j = 0,1,\cdots,\ell-1$ and $x \in X$ define

$$\varphi_j(x) = E\left[\chi_{P_j} \mid \bigvee_{i=1}^{\infty} T^{-i}P\right](x) = \lim_N E\left[\chi_{P_j} \mid \bigvee_{i=1}^{N} T^{-i}P\right](x) = \lim_N \varphi_j^{(N)}(x).$$

LEMMA 3.4. *For all* j, $x \in X$, $\varphi_j(x) \in \left[\frac{\delta}{3(\ell-1)}, 1 - \frac{\delta}{3}\right]$.

Proof. It is enough to show this bound for $\varphi_j^{(N)}(x)$. Let A be an atom of $\bigvee_{i=1}^{N} T^{-i}P$. Observe that for each atom $\alpha_k \in \bigvee_{i=1}^{N} S^{-i}R$ there is exactly one atom β_k such that $\mu((\alpha_k \times \beta_k) \cap A) > 0$. Moreover, using this we may write A as a disjoint union

$$A = \bigcup_{k=1}^{\ell^N} (\alpha_k \times \beta_k).$$

Expanding,

$$E\left[\chi_{P_j} \mid \bigvee_{i=1}^{N} T^{-i}P\right](x) = \sum_A \chi_A(x)\ \mu(P_j|A)$$

$$= \sum_A \chi_A(x)\left[\sum_k \mu(\alpha_k \times \beta_k|A)\ \mu(P_j|\alpha_k \times \beta_k)\right]$$

we see that it is enough to obtain the advertised bound for each term $\mu(P_j \mid \alpha_k \times \beta_k)$. But

$$\mu(P_j|\alpha_k \times \beta_k) = \sum_i \mu\left[P_j \cap (R_i \times \Omega) \mid \alpha_k \times \beta_k\right]$$

$$= \sum_i \mu\left[Q_{s(i)} \times R_i \mid \alpha_k \times \beta_k\right]$$

where $s(i)$ is that $s \in \{0,1,\cdots,\ell-1\}$ satisfying $s + i = j \pmod{\ell}$.

Finally using the independence we may rewrite the above sum as

$$\sum_i \frac{\nu(R_i \cap \alpha_k)\ p(Q_{s(i)} \cap \beta_k)}{\nu(\alpha_k)\ p(\beta_k)}$$

$$= \sum_i \frac{\nu(R_i \cap a_k)}{\nu(a_k)} p(Q_{s(i)}) \in \left[\frac{\delta}{3(\ell-1)}, 1 - \frac{\delta}{3}\right]. \quad \blacksquare$$

In applying these lemmas to a proof of Theorem 3.1, we assume the reader is familiar with the generalized baker's transformation (g.b.t.) representation of an abstract measure-preserving automorphism as a Lebesgue-measure- preserving automorphism of the unit square $\mathscr{S} = [0,1] \times [0,1]$. Those wishing for a more complete discussion are referred to [Bo1] where this construction is described and all of the properties we shall need are discussed in detail.

We summarize the results of applying this representation to our automorphism T with respect to the partition P:

3.5. Let $\varphi_j : [0,1] \to [0,1]$, $j = 0,1,\cdots,\ell-1$ be a collection of measurable functions satisfying

$$\Sigma\, \varphi_j = 1.$$

Let $\underline{f} = \langle \varphi_0, \varphi_1, \cdots, \varphi_{\ell-1} \rangle$ and define $P_{\underline{f}} = \{P_{\underline{f}}^{(0)}, P_{\underline{f}}^{(1)}, \cdots, P_{\underline{f}}^{(\ell-1)}\}$, the "natural" partition of $\mathscr{S}$ induced by $\underline{f}$, where

$$P_{\underline{f}}^{(j)} = \left\{(x,y) \in \mathscr{S} \mid \sum_{i<j} \varphi_i(x) \le y \le \sum_{i \le j} \varphi_i(x)\right\}.$$

The φ_i, in the terminology of [Bo1] are the "cutting functions". The action of $T_{\underline{f}}$ is best described in terms of $T_{\underline{f}}^{-1}$, which maps an atom $P_{\underline{f}}^{(j)}$ onto the vertical column $C_j = \left[\sum_{i<j} \int_0^1 \varphi_i(x)dx, \sum_{i \le j} \int_0^1 \varphi_i(x)dx\right] \times [0,1]$ in a Lebesgue-measure-preserving way and so as to send vertical fibres inside $P_{\underline{f}}^{(j)}$

onto full vertical fibres in C_j. It is easy to write an explicit formula for such a $T_{\underline{f}}$ but we shall not need this. The functions φ_j may be chosen in such a way that the process $(T_{\underline{f}},P_{\underline{f}})$ is isomorphic to the process (T,P).

3.6. The σ-algebra $(P_{\underline{f}})_{-\infty}^{-1} = \bigvee_{-\infty}^{-1} T_{\underline{f}}^i P_{\underline{f}}$ is measurable with respect to $\mathscr{V}$, the σ-algebra of vertical fibres on $\mathscr{S}$. Let $\pi_2 : \mathscr{S} \to [0,1]$ be first coordinate projection. The cutting functions can be shown to be measurable with respect to $\pi_1(P_{\underline{f}})_{-\infty}^{-1}$ and hence

$$\varphi_j(x) = E\left[\chi_{P_{\underline{f}}(j)} \mid (P_{\underline{f}})_{-\infty}^{-1}\right](\pi^{-1}x).$$

3.7. In view of Lemma 3.4 we see that the φ_j satisfy

$$\frac{\delta}{3(\ell-1)} \leq \varphi_j \leq 1 - \frac{\delta}{3} .$$

In particular, $P_{\underline{f}}$ is a generator for $T_{\underline{f}}$ and the isomorphism of processes in 3.5) is in fact measurable isomorphism of the two dynamical systems (X, T, μ) and $(\mathscr{S}, T_{\underline{f}}, \lambda)$.

3.8. We may specify that, restricted to each column C_j, $T_{\underline{f}}$ preserves order in each coordinate. Let I_j be the base of the column C_j, *i.e.* $I_j \times [0,1] = C_j$. Let $g_j : I_j \to [0,1]$ be $g_j(x) = \pi_1(T_{\underline{f}}(\pi^{-1}x))$ and define $g : [0,1] \to [0,1]$ by $g|_{I_j} = g_j$. g is the Lebesgue-measure-preserving factor of $T_{\underline{f}}$ on vertical fibres of $\mathscr{S}$ and is our advertized interval map. We collect the following facts about g.

3.9. The intervals I_j satisfy 1.1), $j = 0,1,\cdots,\ell-1$. The lengths of the I_j satisfy

$$\frac{\delta}{3(\ell-1)} \leq |I_j| \leq 1 - \frac{\delta}{3} .$$

3.10. In view of 3.5) and 3.8) above g satisfies 1.2), indeed each g_j is continuous and monotone increasing on I_j.

3.11. For each j, $g(I_j) = [0,1]$, in particular g is a Markov Map (property 1.7)) of the unit interval.

3.12. For each j, $g_j'(x) = \dfrac{1}{\varphi_j(g(x))} \geq \dfrac{1}{1-\frac{\delta}{3}} > 1$ for almost every $x \in I_j$. Thus g is expanding (property 1.3).

3.13. Since g is Lebesgue–measure–preserving we have $\sum_{y \in g^{-1}(x)} \frac{1}{g'(y)} = 1$, and if g is ergodic, Lebesgue measure is the unique (amongst all measures absolutely continuous with respect to Lebesgue measure) invariant measure for g.

Section 4

We obtain our examples by specifying the process (S,Q) in the construction from Section 3. In this section we will discuss two preliminary cases before proceeding, in Section 5, with the deeper mixing examples.

Example 4.1. *An interval map satisfying* 1.1) – 1.3), *and* 1.6) *which is ergodic but not weak-mixing with respect to its unique absolutely continuous invariant measure.*

Let $Y = \{0,1\}$, $\mathcal{B}$ = discrete σ–algebra and ν = equidistributed probability measure on the two points. Let S be the two point flip: $S(0) = 1$, $S(1) = 0$. Let $Q = \{\{0\},\{1\}\}$. Evidently (S,Q) satisfies the $\bar{d} > \lambda$ property for every $\lambda < 1$. Fix $\lambda = \frac{3}{4}$.

The product automorphism T is ergodic but not weak–mixing, possessing a rotation factor. Conclude that the interval map g from Section 3 is ergodic, being a factor of T, but cannot be weak–mixing, since it's natural extension, T, is not.

Example 4.2. *An interval map satisfying* 1.1) – 1.3) *and* 1.6) *which is weak–mixing but not mixing with respect to its unique absolutely continuous invariant measure.*

The automorphism S is chosen to be Chacon's automorphism, first defined and studied in [C1] using a cutting–and–stacking description. It will be more convenient for us to use the rank 1 shift description of S. The symbol set is $\{0,1\}$. Recall from Section 2 that it suffices to specify the n-blocks:

$$B_0 = 0$$

$$B_{n+1} = B_n \otimes B_n \otimes 1 \otimes B_n.$$

It is a routine check that the rank 1 conditions 2.3), 2.4) and 2.5) are satisfied. Let Q be the time zero partition of $Y \subsetneq \{0,1\}^{\mathbb{Z}}$ and recall the convention $b_n = |B_n|$.

We say that B_n *appears at* i_0 in the Q-name of $x \in Y$ if $x_{i_0}^{i_0+b_n-1} = B_n$. Given two n-blocks $B_n(x)$ and $B_n(y)$ appearing at $i_0(x)$ in $x_{-\infty}^{\infty}$ and $i_0(y)$ in $y_{-\infty}^{\infty}$ respectively we say these two n-blocks *line up* if $i_0(x) = i_0(y)$ and we say they *overlap* if

$$i_0(y) - b_n + 1 \leq i_0(x) \leq i_0(y) + b_n - 1.$$

Finally, we say that these two n-blocks *overlap substantially* (for the purposes of this example only) if they overlap by more than $\frac{b_n}{3}$ indices, that is,

$$i_0(y) - b_n + 1 + \frac{b_n}{3} < i_0(x) < i_0(y) + b_n - 1 - \frac{b_n}{3}.$$

We begin by showing that two n-blocks which do not line up, but overlap substantially must generate an *a priori* lowerbounded amount of $\bar{d}$ disagreement over this overlap. We will be arguing in a "coordinate free" manner where possible, using two copies of B_n, $\overline{B_n}$ and $\underline{B_n}$, with $\overline{B_n}$

laying above $\underline{B_n}$. We enumerate the three copies of B_{n-1} inside B_n, starting from the left as $B_{n-1}^{(1)}$, $B_{n-1}^{(2)}$ and $B_{n-1}^{(3)}$. When we wish to describe which copy of B_n these subblocks occupy we shall write $\overline{B_{n-1}^{(i)}}$ or $\underline{B_{n-1}^{(i)}}$ as appropriate. Finally, given the two copies $\overline{B_n}$ and $\underline{B_n}$ we will agree that $\bar{d}(B_n, S^j B_n)$ refers to the distance measured over the overlap of $\overline{B_n}$ and $\underline{B_n}$ shifted j indices to the right relative to $\overline{B_n}$.

Lemma 4.3. *Let* $B = B_n$ $(n \geq 1)$ *and let* $1 \leq j < \frac{2}{3} b_n$. *Then*

$$\bar{d}(B, S^j B) > \frac{1}{7} .$$

This will certainly be true if we can prove the stronger

Lemma 4.4. *Let* $B = B_n$ $(n \geq 1)$. *Then*

(a) $\bar{d}(B, S^j B) > \frac{1}{7}$ *if* $1 \leq j < \frac{2}{3} |B|$.

(b) *If* $\bar{d}(B, S^j B) \leq \frac{1}{7}$ *then* $\bar{d}(B, S^{j\pm 1} B) > \frac{1}{2}$ *where we allow* $j = 0,1,\cdots,b_n-1$.

(c) *If* $j = b_k$ *or* $j = b_k + 1$ *for some* $k < n$ *then* $\bar{d}(B, S^j B) > \frac{1}{4}$.

Proof. These three statements are easily verified for the case $n = 1$ by direct calculation. We assume $n \geq 2$ and the result true for $1 \leq k < n$. Note that $b_n \geq 13$ and $b_{n-1} \geq 4$.

Case 1. $j = 1$. There are 3 substantial overlaps of (n–1)–blocks each yielding $\bar{d} > \frac{1}{2}$ by application of b). Of the 3 remaining indices, 2 record a disagreement (the two spacer "1"'s) yielding error rate $\frac{2}{3}$ over this collection of indices. Combining we have $\bar{d} > \frac{1}{2}$.

Case 2. $1 < j < \frac{2}{3} b_{n-1}$. There are 3 substantial (n–1)–block overlaps each of which gives $\bar{d} > \frac{1}{7}$ by a). There are two (possibly) unsubstantial (n–1)–block overlaps and the two indices corresponding to spacer "1"'s remaining. If each (n–1)–block (possibly) unsubstantial overlap yields $\bar{d} > \frac{1}{7}$ then we proceed as follows:

If one of the spacer "1"'s aligns with a zero we easily calculate $\bar{d} > \frac{1}{7}$ on the entire overlap. If $\bar{d} = 0$ measured over the two spacer indices we proceed as follows. Suppose k is the length of (all 3) substantial (n–1)–block overlaps, ℓ and $(\ell-1)$ respectively, the length of the (possibly) unsubstantial ones. The number of indices of disagreement on, say, an interval of length ℓ is *strictly* greater than $\frac{1}{7}\ell$ and so is (an integer) $\geq \frac{1}{7}(\ell+1)$. Counting errors this way we obtain $\geq \frac{1}{7}(3k+3+2\ell+1)$ over the entire overlap of length $3k + 2\ell + 1$. Conclude, again $\bar{d} > \frac{1}{7}$ on the n–block overlap. Let us agree to call the above argument ***absorbing*** $\bar{d} = 0$ ***over the spacers*** and observe that we need one more overlap interval than spacer index to make it work.

Next, if one of (possibly) the unsubstantial overlaps yields $\bar{d} \leq \frac{1}{7}$, since the other overlap appears shifted by one, it sees $\bar{d} > \frac{1}{2}$. This yields at least $\frac{1}{2}\ell$ disagreements over the two overlaps of total length $2\ell - 1$ (lengths ℓ and $(\ell-1)$, respectively) giving $\bar{d} \geq \frac{\ell}{4\ell - 2} > \frac{1}{4}$. Again, we have enough intervals to absorb $\bar{d} = 0$ for the two spacer "1"'s.

Case 3. $\frac{2}{3} b_{n-1} \leq j < b_{n-1}$. The $\overline{B^{(1)}_{n-1}}/\underline{B^{(1)}_{n-1}}$ overlap is not substantial. Choose $0 \leq k < n-1$ to be smallest so that the last B_k block in $\overline{B^{(1)}_{n-1}}$ covers the above mentioned overlap. There are now two possibilities. If the last B_k in $\overline{B^{(1)}_{n-1}}$ and the first B_k in $\underline{B^{(1)}_{n-1}}$ do not line up then they overlap substantially and we obtain $\bar{d} > \frac{1}{7}$ on all three unsubstantial overlaps. Absorb $\bar{d} = 0$ over the spacers if necessary. Otherwise these two copies of B_k line up and the substantial B_{n-1} overlaps are shifts by b_k and $b_k + 1$ respectively, $k < n-1$. Applying c) we obtain $\bar{d} > \frac{1}{4}$ on these two substantial overlaps. The spacer "1" in $\overline{B_n}$ aligns with "0" in $\underline{B_n}$. The substantial overlaps plus spacer indices amount to at least $\frac{4}{7}$ of the total overlap length. Conclude, again, $\bar{d} > \frac{1}{7}$.

Case 4. $j = b_{n-1}$ or $j = b_{n-1} + 1$. One easily obtains $\bar{d} > \frac{1}{4}$ as one of the B_{n-1} blocks lines up and the other is shifted by one. In both cases any spacer "1" is aligned with "0".

Case 5. $b_{n-1} + 1 < j < b_{n-1} + \frac{2}{3} b_{n-1}$. Two of the (n–1)–block overlaps are substantial yielding $\bar{d} > \frac{1}{7}$ there. Again choose k smallest so the first B_k in $\overline{B^{(3)}_{n-1}}$ covers the $\overline{B^{(3)}_{n-1}}/\underline{B^{(1)}_{n-1}}$ overlap. If this B_k block does not align with the last B_k block in $\underline{B^{(1)}_{n-1}}$ then the overlap is substantial (for k) and we obtain $\bar{d} > \frac{1}{7}$. One absorbs $\bar{d} = 0$ over the single spacer with the three overlap intervals. Otherwise these B_k blocks line up and the two (n–1)–block shifts are $b_k + 1$ and b_k respectively. The spacer "1" aligns with "1", nevertheless we calculate, using c) from the hypothesis, with

ℓ the length of the $\overline{B_{n-1}^{(2)}}/\underline{B_{n-1}^{(1)}}$ overlap

$$\bar{d} \geq \frac{\frac{1}{4}(2\ell+3)}{2\ell + 2 + b_k} \geq \frac{\frac{1}{4}(2\ell+3)}{2\ell + 2 + \frac{1}{2}\ell}$$

since $\ell \geq 2\ b_k$. The latter is clearly $> \frac{1}{7}$.

Case 6. $b_{n-1} + \frac{2}{3}\ b_{n-1} < j < 2\ b_{n-1}$. The $\overline{B_{n-1}^{(3)}}/\underline{B_{n-1}^{(1)}}$ overlap is substantial yielding $\bar{d} > \frac{1}{7}$ there. If both remaining two overlaps yield $\bar{d} > \frac{1}{7}$ one may absorb $\bar{d} = 0$ over the spacer. Otherwise, if one unsubstantial overlap yields $\bar{d} \leq \frac{1}{7}$, the other yields $\bar{d} > \frac{1}{2}$ being a shift by one index (apply b)). Counting disagreements we get at least $\frac{1}{2}\ell$ (ℓ the length of the longer overlap) over the 2ℓ indices plus the spacer. One therefore gets $\bar{d} > \frac{1}{4}$ there, and hence $\bar{d} > \frac{1}{7}$ over the whole n-block overlap.

Case 7. $j = 2\ b_{n-1}$. Obtain $\bar{d} > \frac{1}{2}$ easily.

This finishes the proof of a) for $B = B_n$. If $\bar{d}(B, S^j B) \leq \frac{1}{7}$ then either $j = 0$ or $j \geq 2\ b_{n-1} + 1$. The first value is handled by case 1 and for the second value, note the overlap is entirely contained in the $\overline{B_{n-1}^{(3)}}/\underline{B_{n-1}^{(1)}}$ overlap and the induction hypothesis applies directly. b) has been shown. To obtain c) observe that if $k = n - 1$ this was handled by Case 4. Otherwise the overlap consists of three (n–1)–blocks shifted by b_k $(k < n-1)$ yielding $\bar{d} > \frac{1}{4}$, one aligned B_k block $(\bar{d} = 0)$ and one B_k block shifted against a B_k block by one unit. One of the spacer "1"'s aligns with "0". Evidently

$\bar{d} > \frac{1}{4}$. This verifies c) and completes the proof. ■

Let $x \in Y \subseteq \{0,1\}^{\mathbb{Z}}$ and let $m,n \in \mathbb{Z}$. We say that an n-block in $x_{-\infty}^{\infty}$ covers x_n^m if there exist $n_0 \leq n \leq m \leq m_0$ so that $x_{n_0}^{m_0} = B_n$. It remains to lift the $\bar{d}$ separation of n-blocks to full names.

Lemma 4.5. *The process* (S,Q) *satisfies* $\bar{d} > \frac{1}{42}$.

Proof. Let $x \neq y \in Y$ be fixed. We may as well assume $x_0 \neq y_0$ for we can always shift the picture: If I_n is a sequence of arbitrarily long intervals of integers starting at zero so $\lim_{n \to \infty} |I_n| = +\infty$ and

$$\liminf_n \bar{d}\left[S^{n_0} x \,|_{I_n}\; S^{n_0} y|_{I_n}\right] > \delta$$

then

$$\liminf_n \bar{d}\left[x|_{I_n}, y|_{I_n}\right] > \delta.$$

Obtain, for all sufficiently large n, time-zero n-blocks $B_n(x)$ and $B_n(y)$ covering x_0 and y_0 respectively. Evidently these n-blocks do not line up.

Case 1. There are ∞-many n so that $B_n(x)$ and $B_n(y)$ are shifted by one index with respect to each other. Enumerate such n by $\{n_k\}_{k=1}^{\infty}$. By dropping to a further subsequence we obtain intervals of integers

$$I_{n_k'} = [\alpha_{n_k'}, \beta_{n_k'}] \cap \mathbb{Z}$$

satisfying either

(i) for all k $I_{n'_k} \subseteq \mathbb{Z}^+ \cup \{0\}$ and corresponds to the indices of overlap between $B^{(3)}_{n'_k-1}(x)$ and $B^{(3)}_{n'_k-1}(y)$.

(ii) for all k $I_{n'_k} \subseteq \mathbb{Z}^- \cup \{0\}$ and corresponds to the indices of the $B^{(1)}_{n'_k-1}(x)$ and $B^{(1)}_{n'_k-1}(y)$ overlap

In both cases $|I_{n'_k}| = b_{n'_k-1} \longrightarrow +\infty$. Also in both cases

$$\max\{|n| \mid n \in I_{n'_k}\} \leq b_{n'_k} - 1 = 3b_{n_k-1}.$$

In the first instance set $N_k = \beta_{n'_k}$ and observe $\bar{d}(x_0^{N_k}, y_0^{N_k}) > \frac{1}{3} \cdot \frac{1}{7} = \frac{1}{21}$.
In the second instance, set $N_k = \alpha_{n'_k}$ obtaining

$$\bar{d}(x^0_{N_k}, y^0_{N_k}) > \frac{1}{3} \cdot \frac{1}{7} = \frac{1}{21}$$

using Lemma 4.3.

Case 2. For all sufficiently large n, $B_n(x)$ and $B_n(y)$ appear shifted by more than one index relative to each other. Let $B^+_n(x)$ and $B^+_n(y)$ be the first n–blocks appearing to the right of the previously defined $B_n(x)$ and $B_n(y)$. Evidently $B^+_n(x)$ and $B^+_n(y)$ do not line up. We define intervals of indices $I_n = [\alpha_n, \beta_n] \cap \mathbb{Z}^+$ as follows: If $B^+_n(x)$ and $B^+_n(y)$ overlap substantially then I_n is defined by this overlap. If not, then I_n is defined by the overlap of either $B_n(x)$ and $B^+_n(y)$ or of $B^+_n(x)$ and $B_n(y)$, one of which must be a substantial n–block overlap. We thereby obtain intervals I_n satisfying:

(i) $I_n \subseteq \mathbb{Z}^+$.

(ii) For each n, I_n is the set of indices of a substantial n–block overlap.

(iii) $|I_n| > \frac{1}{3} b_n$ and $\max\{k \in I_n\} \leq 2 b_n$.

Setting $N_n = \beta_n$ conclude $N_n \longrightarrow \infty$ and

$$\bar{d}(x_0^{N_n}, y_0^{N_n}) > \frac{1}{6} \cdot \frac{1}{7} = \frac{1}{42}$$

where we have used Lemma 4.3 to obtain $\bar{d} > \frac{1}{7}$ on the interval I_n. ∎

Remark 4.5.

(i) The $\bar{d} > \delta$ property for Chacon's Automorphism has not been observed before in the literature.

(ii) It is well known (and easy to prove, see [C1]) that S is weak–mixing but not mixing. The same will be true for T, the product of S and a Bernoulli shift.

The interval map g constructed in Section 3 will have Lebesgue measure as its unique (abs. cont.) invariant measure, with respect to which g is weak–mixing but not mixing, since the latter property lifts to natural extensions.

Section 5

In this section we discuss our final example in which S is a mixing rank 1 automorphism. Ornstein [Or] was the first to show that such transformations exist using a "random spacer" construction. With some effort, using the ideas of King and Weiss [Ki&We], one may modify the Ornstein construction obtaining a mixing rank 1 with $\bar{d} > \frac{1}{4}$. It turns out, however, that one need not be so industrious.

Theorem 5.1. *Let* S *be a* rank 1 *transformation which is mixing. Then there exists* $\delta > 0$ *so that* S *satisfies* $\bar{d} > \delta$.

Remark 5.2. *(i)* Having chosen S as above the results of Section 3 are invoked to produce an automorphism T (product of S and a Bernoulli shift) which is mixing but not a K–automorphism, posessing a non–trivial zero–entropy factor. It is in fact only the zero–entropy property of S we need — any mixing, zero–entropy automorphism with $\bar{d} > \delta$ would serve for S.
(ii) The interval map g associated to T in Section 3 has Lebesgue measure as its unique (amongst those abs. cont. to Lebesgue measure) invariant measure, and is mixing but not exact (in the terminology of endomorphisms) since its natural extension is not a K–automorphism.

In what follows S will be a rank 1 mixing automorphism viewed as a μ–invariant shift on sequences of symbols from A with time–zero partition Q. As discussed in Section 2 there is no loss of generality in assuming $A = \{0,1\}$, and the rank 1 properties 2.3) through 2.6) there. We begin with two observations about mixing (rank 1 not needed here) automorphisms S.

These have already been noted and used by Kalikow [Ka], (stated and proved there as Lemmas 4.1) and 4.2)); we simply restate them in our notation. The first is a type of uniform Blum–Hanson Theorem.

Lemma 5.3. *Let* A *be a measurable set and* $\epsilon > 0$. *Then there exists a* $K = K(A,\epsilon)$ *so that if* $k \geq K$ *and* $n_1 < n_2 < \cdots < n_k$ *are integers then*

$$\mu\left\{x \mid \left|\frac{1}{k}\,\#\left\{A \cap \{S^{n_1}x, S^{n_2}x, \cdots, S^{n_k}x\}\right\} - \mu(A)\right| > \epsilon\right\} < \epsilon.$$

Lemma 5.4. *Let* A *and* B *be measurable subsets and* $\epsilon > 0$. *Then there exists an* $M = M(A,B,\epsilon)$ *so that if both* $m \geq M$ *and* $n \geq M$ *one has*

$$\int_{X=A^{\mathbb{Z}}} \left|\frac{1}{n}\,\#\left\{T^{-m}A \cap B \cap \{x, Sx, \cdots, S^{n-1}x\}\right\} - \mu(A)\mu(B)\right| \, d\mu(x) < \epsilon.$$

We also note that, since S is mixing, the set of periodic points (periodic names in $X = A^{\mathbb{Z}}$) is of measure zero — we will henceforth assume all strings to be chosen disjoint from this set. Recall that B_N denotes an N–block; its length is $b_N = |B_N|$.

Our main observation is the following:

Proposition 5.5. *Let* $\epsilon_0 > 0$ *be fixed. Then there exists a* $\delta > 0$ *and* M *so that for all sufficiently large* N, *if* α *and* β *are two substrings of* B_N *appearing at* i_0 *and* $i_0 + m$ *respectively, with* $m \geq M$ *and* $q = |\alpha| = |\beta| \geq \epsilon_0\, b_N$ *then*

$$\bar{d}(\alpha,\beta) > \delta_0.$$

Proof. Let $A = Q_0 = \{x|x_0 = 0\}$, $B = Q_1 = \{x|x_0 = 1\}$ and let $\gamma = \mu(A)\mu(B) > 0$. Choose $\epsilon_1 > 0$ so that $\delta_0 = \gamma - (6+\epsilon_0)\sqrt{\epsilon_1} \geq \frac{\gamma}{2} > 0$. Apply Lemma 5.4 to the sets A, B and with $\epsilon = \epsilon_0^2\epsilon_1$, obtaining M such that if $m,n \geq M$

$$\int_X \left|\frac{1}{n}\#\left\{T^{-m}A \cap B \cap \{x,Sx,\cdots,S^{n-1}x\}\right\} - \gamma\right| d\mu(x) < \epsilon_0^2\epsilon_1.$$

Fix any $m \geq M$ and set $n = M$. Let

$$B_{m,n} = B_m = \left\{x\left|\left|\frac{1}{M}\#\left\{T^{-m}A \cap B \cap \{x,Sx,\cdots,S^{M-1}x\}\right\} - \gamma\right| > \epsilon_0\sqrt{\epsilon_1}\right.\right\}$$

and observe that $\mu(B_m^c) \geq 1 - \epsilon_0\sqrt{\epsilon_1}$. Rewriting, if $x \in B_m^c$ then

5.6 $$\frac{1}{M}\#\left\{i|0 \leq i \leq M-1,\ x_i = 1,\ x_{i+m} = 0\right\} > \gamma - \epsilon_0\sqrt{\epsilon_1}.$$

Now apply Lemma 5.3) to the set B_m^c with $\epsilon = \epsilon_0\sqrt{\epsilon_1}$, obtaining K. Let N be so large that $b_N = |B_N| >> MK$ and pick $k = k(N)$ so $|B_N| = (k-1)M + r$; $0 \leq r < M$. Note that $k > K$, and consider the set of integers $\mathcal{J} = \{\ell M|\ \ell = 0,\pm1,\cdots,\pm k\}$. It follows from our application of Lemma 5.3) that nearly a $\mu(B_m^c)$ proportion of the points $\{S^jx\}_{j\in\mathcal{J}}$ will lie in B_m^c for a large set of x; more precisely:

5.7 $$\mu\left\{x\left|\ \left|\frac{1}{|\mathcal{J}|}\#\left\{B_m^c \cap \{S^jx\ |\ j \in \mathcal{J}\}\right\} - \mu(B_m^c)\right| > \epsilon_0\sqrt{\epsilon_1}\right.\right\} < \epsilon_0\sqrt{\epsilon_1}.$$

Thus we may, for the rest of the argument, fix $x \in X$ so that both $x_0 \in B_N$ and

5.8) $$\frac{1}{|\mathscr{J}|}\#\left\{B_m^c \cap \{S_x^j\}_{j\in\mathscr{J}}\right\} > 1 - 2\epsilon_0\sqrt{\epsilon_1}.$$

Let us agree to say an index $j \in \mathscr{J}$ is *good* if $S^j x \in B_m$; otherwise j is *bad*. Let $B_N(x)$ be the time-zero N-block in $x_{-\infty}^{\infty}$ and observe that this appears inside the string x_{-kM}^{kM}, yet $B_N(x)$ covers at least $\frac{|\mathscr{J}| - 2}{2}$ of the indices in $\mathscr{J}$. Let a appear in $B_N(x)$ at x_r^{r+q-1} and let us first assume, for simplicity, that $r = r_1 M$, $r + q - 1 = r_2 M - 1$ so that a covers exactly $r_2 - r_1$ consecutive complete strings of length M starting at indices from $\mathscr{J}$. An easy calculation shows, since $|a| \geq \epsilon_0\, b_N$ that at least $\epsilon_0\left[\frac{|\mathscr{J}| - 2}{2} - 1\right]$ indices from $\mathscr{J}$ are covered by a. We now show that a proportion of at least $1 - 6\sqrt{\epsilon_1}$ of these indices covered by a are good. For, if not then there are at least $3\sqrt{\epsilon_1}\,\epsilon_0(|\mathscr{J}|-4)$ indices which are bad, hence at most

5.9 $$|\mathscr{J}|\left[1 - 3\sqrt{\epsilon_1}\,\epsilon_0 + \frac{12\sqrt{\epsilon_1}\,\epsilon_0}{|\mathscr{J}|}\right]$$

indices in $\mathscr{J}$ can be good. We could easily have chosen N (knowing just ϵ_0, ϵ_1 and M) so that $\frac{12\sqrt{\epsilon_1}\,\epsilon_0}{|\mathscr{J}|} < \sqrt{\epsilon_1}\,\epsilon_0$ and this contradicts 5.8).

It is now easy to see, since a is a union of strings of length M, most of which begin with a good index, that

$$\frac{1}{|a|}\#\left\{i \,\middle|\, x_i \in a,\ x_i = 1,\ x_{i+m} = 0\right\}$$

$$> \left[1 - 6\sqrt{\epsilon_1}\right]\left[\gamma - \epsilon_0\sqrt{\epsilon_1}\right]$$

$$> \gamma - (6+\epsilon_0)\sqrt{\epsilon_1} = \delta_0.$$

Conclude $\bar{d}(\alpha,\beta) > \delta_0$.

We leave it to the reader to check that in case α is not exactly a concatenation of full M-strings beginning at indices in $\mathcal{J}$ that a slight modification of the above argument for "end effects" (necessitating possibly an increase in N) will give the same result. ∎

Remark 5.9. *(i)* We have shown that, provided we restrict our attention to large enough N there is an *a priori* lowerbound on the frequency of seeing "1" in α and "0" in β when α and β are two "substantial" substrings inside B_N. Reversing the roles of "1" and "0" in the above will give a disjoint set of disagreements, with a corresponding increase in the lowerbound on $\bar{d}$, but this is not necessary for the arguments which follow.
(ii) The previous lemma could have been shown for an arbitrary generating partition with little increased complication. In the following proposition we make much greater use of the special partition described in Section 2.

It remains to handle substrings α and β shifted by fewer than M indices. In the following, we refer to two copies of B_N, one above the other with the lower copy shifted j indices to the right and we use $\bar{d}(B_N,S^jB_N)$ to denote the $\bar{d}$ distance measured over the overlap of these two strings.

Proposition 5.10. *Let* M *and* $\epsilon_0 > 0$ *be fixed. Then there is a* $\delta_0 > 0$ *so that for all sufficiently large* N, *if* α *and* β *are two substrings of* B_N *with* $|\alpha| = |\beta| \geq \epsilon_0|B_N|$ *and* α *appears at* i_0 *in* B_N *while* β *appears at* $i_0 + j$, $1 \leq j \leq M$, *then*

$$\bar{d}(\alpha,\beta) > \delta_0.$$

Proof. An n–block begins with a single symbol "1" followed by n occurrences of the symbol "0" so it is immediate that for fixed $n > M$

$$\inf_{1 \le j \le M} \bar{d}(B_n, S^j B_n) > 0.$$

Fix $N_0 > M$ large enough so that for all $N > N_0$, at most $\frac{1}{10}\epsilon_0$ of the indices in an N–block are not covered by N_0–blocks inside B_N, and so that $\frac{M}{|B_{N_0}|} < \frac{1}{5}$. Let α be fixed in B_N and let β be the string below α in $S^j B_N$. Observe that at most $\frac{1}{5}|\alpha|$ of the indices covered by α can correspond to a spacer (between N_0–blocks in B_N) in either α or β. The remaining indices correspond to N_0–block indices in both α and β. At most a further $\frac{1}{5}$ of these, (and hence at most another $\frac{1}{5}|\alpha|$ of the indices covered by α) can arise from the rightmost part of an N_0–block in β overlapping the leftmost part of an N_0–block in α. The remaining $|\alpha| - \frac{2}{5}|\alpha|$ indices correspond to N_0–blocks shifted by j. Removing the two (possibly) partial N_0–block overlaps at the ends of α we calculate

$$\bar{d}(\alpha,\beta) \ge \left[\frac{3}{5} - \frac{2|B_{N_0}|}{|\alpha|}\right] \min_{1 \le j \le M} \bar{d}(B_{N_0}, S^j B_{N_0}).$$

We can specify N large enough so $\frac{3}{5} - \frac{2|B_{N_0}|}{|\alpha|} > \frac{1}{2}$ and use

$$\delta_0 = \frac{1}{2} \min_{1 \le j \le M} \bar{d}(B_{N_0}, S^j B_{N_0}). \quad ■$$

We combine the two previous results as

PROPOSITION 5.11. *Let* $\epsilon_0 > 0$ *be fixed. Then there exists a* $\delta_0 > 0$ *and* $N_0 < \infty$ *so that if* $N \geq N_0$ *and* α *and* β *are two substrings of* B_N *appearing at* i_0 *and* $i_0 + j$ *respectively with* $j \geq 1$ *and* $|\alpha| = |\beta| \geq \epsilon_0 |B_N|$, *then*

$$\bar{d}(\alpha,\beta) > \delta_0.$$

It remains to lift this $\bar{d}$ separation property of n–blocks to full names. The issues are essentially as in Example 4.2) — a modification of the argument there gives

PROPOSITION 5.12. S *satisfies* $\bar{d} > \frac{\delta_0}{20}$.

Proof. Again, as in Lemma 4.5, fix $x \neq y \in A^{\mathbb{Z}}$ where we may assume without loss of generality that $x_0 \neq y_0$. Obtain $\delta_0 > 0$ and $N_0 < \infty$ from Proposition 5.11 using $\epsilon_0 = \frac{1}{10}$. Let us also assume that x and y have been chosen from the set of full measure given by property 2.5) and Lemma 2.7 with $\delta_0 = \frac{1}{10}$ so that, for all large enough N

5.13) x_0 and y_0 are inside an N–block $B_N(x)$ and $B_N(y)$ respectively, (the time–zero N–blocks).

5.14) The length of the string of spacer symbols in $x_{-\infty}^{\infty}$ (resp. $y_{-\infty}^{\infty}$) immediately to the right of $B_N(x)$ (resp. $B_N(y)$) does not exceed $\frac{1}{10}|B_N|$.

<u>Case I</u>. For infinitely many N, $B_N(x)$ and $B_N(y)$ are shifted relative to each other by j indices with $\frac{1}{10}|B_N| < j < \frac{8}{10}|B_N|$.

For such N, let $B_N^+(x)$ (resp. $B_N^+(y)$) denote the first N-block to the right of $B_N(x)$ (resp. $B_N(y)$). Evidently $B_N^+(x)$ and $B_N^+(y)$ do not line up and overlap by more than $\frac{1}{10}|B_N|$ units. By Proposition 5.11), the $\bar{d}$ distance measured on this overlap is at least δ_0 and it appears within an interval of $2|B_N|$ indices from the origin so the $\bar{d}$ distance measured over the interval from the origin to the rightmost index of the overlap interval is at least $\frac{1}{20}\,\delta_0$.

<u>Case II</u>. For infinitely many N the shift between $B_N(x)$ and $B_N(y)$ is at least $\frac{8}{10}|B_N|$. Then we can replace either $B_N(x)$ with $B_N^+(x)$ or $B_N(y)$ with $B_N^+(y)$ to obtain an interval in $\mathbb{Z}^+$ over which $\bar{d} > \delta_0$. Complete the argument as in Case I.

Otherwise, we have

<u>Case III</u>. For all large N, $B_N(x)$ and $B_N(y)$ overlap by at least $\frac{2}{10}|B_N|$ and hence, for infinitely many N an interval of at least $\frac{1}{10}|B_N|$ in the time-zero N-block overlap lies in $\mathbb{Z}^+ \cup \{0\}$ (or $\mathbb{Z}^- \cup \{0\}$) and we may apply Proposition 5.11) directly obtaining again, $\bar{d} > \delta_0$ over these intervals.

In all cases we find

$$\limsup_n \bar{d}(x_0^n, y_0^n) > \frac{1}{20}\,\delta_0$$

or

$$\lim_n \sup \bar{d}(x^0_{-n}, y^0_{-n}) > \frac{1}{20}\,\delta_0$$

and the result has been shown. ■

References

[Ad] R. Adler, Continued Fractions and Bernoulli Trials, *Ergodic Theory, A Seminar, Editors, J. Moser, E. Phillips and S. Varadhan, Courant Inst. Math. Sci.*, (1975), Chapter 16, 111–120.

[Bo1] C. Bose, Generalized baker's transformations, *Ergod. Th. and Dynam. Sys.*, **9**(1989), 1–17.

[Bo2] C. Bose, Invariant measures and equilibrium states for piecewise $C^{1+\alpha}$ endomorphisms of the unit interval, *Trans. Amer. Math. Soc.*, **315**(1989), 105–125.

[Bow] R. Bowen, Bernoulli maps of the interval, *Israel J. Math.*, **28**(1977), 161–168.

[Bow&S] R. Bowen and C. Series, Markov maps associated with Fuchsian groups, *Inst. Hautes Etudes Sci. Publ. Math.*, **50**(1979), 401–418.

[C1] R.V. Chacon, A geometric construction of measure preserving transformations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press*, 1965, Vol. 2, Part 2, 335–360.

[D] W. Doeblin, Remarques sur la théorie métrique des fractions continues, *Compositio Math.*, **7**(1940), 353–371.

[Ka] S. Kalikow, Twofold mixing implies threefold mixing for rank one transformations, *Ergod. Th. and Dynam. Sys.*, 4(1984), 237–259.

[Ke] G. Keller, Generalized bounded variation and applications to piecewise monotone transformations, *Z. Wahrsch. Verw. Gebiete*, **69**(1985), 461–478.

[Kh] A. Khintchine, Metrishe kettenbruchprobleme, *Compositio Mathematics*, **1**(1935), 359–382.

[Ki] J. King, Joining rank and the structure of finite rank mixing transformations, *J. d'Analyse Math.*, **51**(1988), 182–227.

[Ki&We] J. King and B. Weiss, A group rotation factor of a non-rigid rank-1 map. *Dynamical systems; Proceedings of the Special Year held at the University of Maryland, College Park, 1986-87, Springer LNM*, **1342**(1988), 425–439.

[L&Y] A Lasota and J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations, *Trans. Amer. Math. Soc.*, **186**(1973), 481–488.

[Or] D. Ornstein, On the root problem in ergodic theory. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, (1970), 347–356.

[R] A. Renyi, Representations of real numbers and their ergodic properties, *Acta. Math. Akad. Sci. Hungar.*, **8**(1957), 477–493.

[W] P. Walters, Invariant measures and equilibrium states for some mappings which expand distances, *Trans. Amer. Math. Soc.*, **236**(1978), 121–153.