

Generalized baker's transformations

CHRISTOPHER J. BOSE

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210, USA

(Received 10 November 1987 and revised 21 March 1988)

Abstract. A class of automorphisms of the unit square called generalized baker's transformations (gbt) is defined in such a way that every stationary stochastic process may be represented as the movement of a simple partition of the square under a gbt. This extends the classical example of the representation of independent processes by the well-known baker's transformation.

Every ergodic, positive-entropy automorphism is measurably isomorphic to some gbt (again generalizing the classical result about Bernoulli shifts), and we show that a large class of gbt's satisfying certain continuity restrictions are actually measurably isomorphic to Bernoulli shifts.

1. Introduction

Recall the classical baker's transformation on the unit square as described, for instance, by Halmos [H]. The purpose of this article is to give a generalization of this old example which will allow us to represent any countable-state, stationary, stochastic process as the movement of a partition on the square, whose atoms are regions bounded by the graphs of 'cutting functions', under a Lebesgue-measure-preserving point transformation. This latter mapping we shall call a generalized baker's transformation (gbt).

Amongst the features of the classical example which are preserved by our generalization we shall see that our gbt's map vertical fibres into vertical fibres on the square and, restricted to the σ -algebra of vertical fibres, a gbt is an expanding endomorphism of the unit interval. When the process we are representing by a gbt happens to be the action of an ergodic, positive-entropy transformation on a generating partition, the gbt is actually measurably isomorphic to the original point transformation (Corollary 4). Thus our representation is still very general even at the level of measurable isomorphism.

This last feature leads us naturally to investigate specific classes of gbt's. In the main part of this article we study gbt's with Hölder-continuous 'cutting-functions', a large class whose members we show to be measurably isomorphic to Bernoulli shifts.

The techniques and results of this last section are reminiscent of similar facts from the theory of expanding endomorphisms and g -measures. (See [L], [Bo], [A], [Be] and the publications cited therein, where the central problem is the existence

of a ‘reasonable’ invariant measure.) Since our transformations are measure-preserving our setting is more restrictive and subsequently we obtain better results than would follow from these known facts. Nonetheless we choose to deduce our main theorem (Theorem 3.4) from a result of Walters [W].

In [B2] we have used the ideas and results presented here to investigate expanding endomorphisms and to prove a $C^{1+\epsilon}$ version of Adler’s Folklore Theorem.

2.

Let λ be Lebesgue measure on the Borel subsets \mathcal{B}_1 of $[0, 1]$. Let $\mu = \lambda \times \lambda$ be two dimensional Lebesgue measure on the Borel subsets \mathcal{B}_2 of the unit square $S = \{(x, y) | 0 \leq x, y \leq 1\}$, and let $f_i \geq 0$ $i = 1, 2$ be measurable functions on $[0, 1]$ satisfying $f_1 + f_2 = 1$ (λ - a.e.). Define two mappings of $[0, 1]$ into $[0, 1]$ by the formulae

$$\begin{aligned} \varphi_1^*(x) &= \int_0^x f_1(t) dt, \\ \varphi_2^*(x) &= \varphi_1^*(1) + \int_0^x f_2(t) dt. \end{aligned}$$

Observe that the φ_i^* are increasing and map the interval continuously onto $[0, \varphi_1^*(1)]$ and $[\varphi_1^*(1), 1]$ respectively. Define, for $x \in [0, \varphi_1^*(1)]$

$$\varphi_1(x) = \inf \{t | \varphi_1^*(t) \geq x\},$$

and for $x \in [\varphi_1^*(1), 1]$,

$$\varphi_2(x) = \inf \{t | \varphi_2^*(t) \geq x\}.$$

Let $f = (f_1, f_2)$, $P_1 = \{(x, y) | 0 \leq y \leq f_1(x)\}$, $P_2 = S - P_1$ and $P_f = \{P_1, P_2\}$. Let

$$T_f(x, y) = \begin{cases} (\varphi_1(x), f_1(\varphi_1(x)) \cdot y) & \text{if } x \in [0, \varphi_1^*(1)], \\ (\varphi_2(x), 1 - f_2(\varphi_2(x)) \cdot (1 - y)) & \text{if } x \in [\varphi_1^*(1), 1]. \end{cases}$$

If $0 \leq a < b < \varphi_1^*(1)$, T_f moves the subset $[a, b] \times [0, 1]$ to $([\varphi_1(a), \varphi_1(b)] \times [0, 1]) \cap P_1$ and if $\varphi_1^*(1) \leq a < b \leq 1$, T_f moves the subset $[a, b] \times [0, 1]$ to $([\varphi_2(a), \varphi_2(b)] \times [0, 1]) \cap P_2$. Of course here, and in all that follows equalities and inequalities will be taken modulo a set of (the relevant) measure zero. In particular if $g = (g_1, g_2)$ with $g_i = f_i$ (λ - a.e.), then $T_f = T_g$ (μ - a.e.). If the φ_i^* are invertible (as would be the case if, for example, $\lambda\{x | \text{either } f_1(x) = 0 \text{ or } f_2(x) = 0\} = 0$) then $\varphi_i^* = \varphi_i^{-1}$, $i = 1, 2$ and we may write a formula for T_f^{-1} :

$$T_f^{-1}(x, y) = \begin{cases} \left(\varphi_1^*(x), \frac{y}{f_1(x)} \right) & \text{if } (x, y) \in P_1, \\ \left(\varphi_2^*(x), 1 - \frac{1-y}{f_2(x)} \right) & \text{if } (x, y) \in P_2. \end{cases}$$

In any case, as $\varphi_i^* \varphi_i = id$ we have $T_f^{-1} T_f = id$, and $T_f T_f^{-1}(x, y) = (x, y)$ provided $\varphi_i \varphi_i^*(x) = x$ for that i satisfying $(x, y) \in P_i$. If $\varphi_i \varphi_i^*(x_0) < x_0$ then $f_i \equiv 0$ on the interval $[\varphi_i \varphi_i^*(x_0), x_0]$ and so $P_i \cap ([\varphi_i \varphi_i^*(x_0), x_0] \times [0, 1]) = \emptyset$. We conclude T_f is invertible off a set of measure zero and that a formula for T_f^{-1} may be taken to be the above even when φ_i^* fails to be invertible.

Suppose for the moment $f_i \in C^1[0, 1]$. Differentiating our formula for T_f we obtain $|DT_f| = 1$. It follows then that T_f is measure-preserving. In fact, this last statement is true for any measurable f_i as the following elementary calculation shows.

LEMMA 2.1. T_f is Lebesgue-measure-preserving.

Proof. It will be convenient, and sufficient to check the fact on subsets of the form $\tilde{A} = A \times [0, c)$ where $A \in \mathcal{B}_1$ and $0 < c \leq 1$. Decompose \tilde{A} by setting $A_1 = \{x \in A \mid f_1(x) < c\}$, $A_2 = A - A_1$, $\tilde{A}_1 = A_1 \times [0, c)$, $\tilde{A}_2 = A_2 \times [0, c)$. Note that $\varphi_1^{*'} = f_1$, and hence the Fubini theorem followed by a change of variable gives

$$(i) \mu(T_f^{-1}\tilde{A}_2) = \int_{\varphi_1^{*'} A_2} \frac{c}{f_1(\varphi_1(x))} dx = \int_{A_2} c dx = \mu(\tilde{A}_2).$$

By a similar computation

$$(ii) \mu(T_f^{-1}(A_1 \times [c, 1])) = \mu(A_1 \times [c, 1]),$$

and as observed before

$$(iii) \mu(T_f^{-1}(\tilde{A}_1 \cap P_1)) = \mu(\tilde{A}_1 \cap P_1),$$

$$(iv) \mu(T_f^{-1}((A_1 \times [0, 1]) \cap P_2)) = \mu((A_1 \times [0, 1]) \cap P_2).$$

Putting all the pieces together one concludes

$$\mu(T_f^{-1}\tilde{A}) = \mu(\tilde{A}). \quad \square$$

If $f_1 = f_2 \equiv \frac{1}{2}$, T_f is the classical baker's transformation mentioned in the Introduction. Given $f = (f_1, f_2)$ the automorphism T_f will be called the *generalized baker's transformation associated with the pair $f = (f_1, f_2)$* (abbreviated gbt, with tacit dependence on f) and f_1 will be called the *cutting function* for T_f .

Moser [M] has studied a class of automorphisms on S similar to the collection defined above but with cutting functions restricted to those satisfying a Lipschitz condition. His transformations are not necessarily measure-preserving. Alexander and Yorke [A, Y] have studied another generalization of the classical baker's transformation which is neither one-to-one nor measure-preserving unless equal to the classical baker's transformation.

We conclude this section with a brief exposition of the conventions and notations we shall be using.

By a *measure space* we shall mean a triple (X, \mathcal{F}, μ) where X is some set, \mathcal{F} a σ -algebra of subsets of X and μ is a countably additive and *normalized* measure on \mathcal{F} . All our measure spaces will be *Lebesgue spaces*, that is, measurably isomorphic to an interval plus an (at most) countable collection of point masses. Let I be a finite or countably infinite and linearly ordered set. A *partition P of (X, \mathcal{F}, μ) indexed by I* is a collection of disjoint members of \mathcal{F} , say $P = \{P_s\}_{s \in I}$, each $P_s \in \mathcal{F}$ so that $X = \bigcup_s P_s$. Each P_s with $\mu(P_s) > 0$ is called an *atom* of the partition P . If $P = \{P_s\}_{s \in I}$ and $Q = \{Q_t\}_{t \in J}$ are two partitions of (X, \mathcal{F}, μ) then $P \vee Q$, the *join* of P and Q is the partition $P \vee Q = \{P_s \cap Q_t\}_{s \in I, t \in J}$ where $I \times J$ is given the lexicographic order. The apparently unnecessary assumption of a linear order on our partition will facilitate our constructions in the later sections and may always be assumed. $|P|$ will denote the number of atoms in P .

Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub- σ -algebra and denote by $E_{\mathcal{G}}$ the conditional expectation, $\mu(\cdot | \mathcal{G})$ the conditional measure, and $\text{dist}(P | \mathcal{G})$ the conditional distribution of a partition P with respect to \mathcal{G} . For two σ -algebras \mathcal{G} and \mathcal{H} we write $\mathcal{G} < \mathcal{H}$ if for each $G \in \mathcal{G}$ there is an $H \in \mathcal{H}$ so that $\mu((H \cap G^c) \cup (H^c \cap G)) = 0$. A pair (T, P) is called a *process* if T is an automorphism of a measure space (X, \mathcal{F}, μ) and P is a partition. Let (S, Q) be another process and suppose P and Q have the same index set I . Then we say that (T, P) and (S, Q) are *isomorphic processes* if for each $n, m \in \mathbb{N}$,

$$\text{dist} \left(\bigvee_{-m}^n T^i P \right) = \text{dist} \left(\bigvee_{-m}^n S^i Q \right).$$

Equip I with the discrete σ -algebra. Each choice of $m \leq n \in \mathbb{Z}$ and string of symbols $s \in I^{n-m+1}$, $s = s_m s_{m+1} \cdots s_n$, can be identified with a cylinder subset $\hat{s} \in I^{\mathbb{Z}}$, namely $\hat{s} = \{x \in I^{\mathbb{Z}} | x_i = s_i, m \leq i \leq n\}$. Define a map $\Phi_P : X \rightarrow I^{\mathbb{Z}}$ by $\Phi_P(x) = \hat{x}$ where $\hat{x}_i = s_i$ if and only if $T^i x \in P_{s_i}$. $\Phi_P(x)$ is called the *P-name of the point* $x \in X$. Φ_P carries the measure μ to a measure $\hat{\mu}$ on the σ -algebra generated by cylinder subsets of $I^{\mathbb{Z}}$. Let $\hat{\mathcal{F}}$ be the completion of this σ -algebra with respect to $\hat{\mu}$. Define the *shift* S on $I^{\mathbb{Z}}$ by the rule $(Sx)_i = x_{i+1}$ and check that $S \circ \Phi_P = \Phi_P \circ T$; S is evidently an invertible, measure-preserving transformation and if we set $Q = \Phi_P P$, the *time-zero partition*, (T, P) and (S, Q) are isomorphic. We say that P is a *generator for* T if we may find a subset $X' \subseteq X$ with $\mu(X') = 1$ so that $\Phi_P|_{X'}$ is one-to-one; in this event T and S are *measurably isomorphic transformations*.

We denote the *information* and *conditional information functions* of the partition P , respectively, by $I_{\mu}(P) = -\sum \chi_{P_i} \log \mu(P_i)$ and $I_{\mu}(P | \mathcal{G}) = -\sum \chi_{P_i} \log \mu(P_i | \mathcal{G})$. The *entropy* and *conditional entropy* of P will then become

$$h_{\mu}(P) = \mu(I_{\mu}(P)) \quad \text{and} \\ h_{\mu}(P | \mathcal{G}) = \mu(I_{\mu}(P | \mathcal{G}))$$

and these may be infinite. If $h_{\mu}(P) < \infty$ then the *entropy of* T with respect to P is

$$h_{\mu}(T, P) = h_{\mu} \left(P \bigg| \bigvee_{i=1}^{\infty} T^{-i} P \right)$$

whereafter the entropy of T is

$$h_{\mu}(T) = \sup_P \{h_{\mu}(T, P) | h_{\mu}(P) < \infty\}.$$

If $\mathcal{H} < \mathcal{F}$ is another sub- σ -algebra then $I_{\mu}(\mathcal{H} | \mathcal{G})$ will be the increasing limit of $I_{\mu}(P_i | \mathcal{G})$ where $P_i < P_{i+1}$ are finite partitions and $\bigvee P_i = \mathcal{H}$; this limit may be seen to be independent of the choice of the sequence P_i .

3.

If $B \in \mathcal{B}_1$ and $x \in [0, 1]$ then we call $B \times [0, 1]$ the *full column over* B , and $\{x\} \times [0, 1]$ the *vertical fibre over* x . Let C be the σ -algebra of full columns in S . For each $n \in \mathbb{N}$ let $C_n = \bigvee_{i=1}^n T_f^{-1} P_f$ and let C_{∞} be the smallest σ -algebra containing the $\bigcup_n C_n$. Clearly $C_n < C$ for each $n \geq 1$ and hence $C_{\infty} < C$. We begin with a simple, nonetheless useful observation.

LEMMA 3.1. *If A is an atom of C_∞ , with $A = A_1 \times [0, 1]$ then either $f_1|_{A_1} = 0$ or $f_1|_{A_1} = 1$.*

Proof. We will show that $\int_{A_1} f_1(t) dt$ is either zero or $\lambda(A_1)$. Suppose not. Then $T_f^{-1}A$ is the disjoint union of two full columns $T_f^{-1}(A \cap P_1) = \varphi_1^*(A_1) \times [0, 1]$ and $T_f^{-1}(A \cap P_2) = \varphi_2^*(A_1) \times [0, 1]$ with measures strictly between zero and $\mu(A)$. Since A is an atom, we conclude that A and the two full columns of $T_f^{-1}A$ are pairwise disjoint. Hence $\mu(T_f^{-1}A \cup A) = 2\mu(A)$. Now consider the set $T_f^{-2}A = \bigcup_{i=1,2} T_f^{-1}(T_f^{-1}(A \cap P_i))$. The columns of $T_f^{-1}(T_f^{-1}(A \cap P_i))$ must be disjoint from A and both columns of $T_f^{-1}A$, hence $\mu(T_f^{-2}A \cup T_f^{-1}A \cup A) = 3\mu(A)$. Continuing in this way we obtain a contradiction. \square

As our first application of this fact we show

LEMMA 3.2. *For $i = 1, 2$ and $(x, y) \in S$ we have*

$$\mu(P_i | C_\infty)(x, y) = f_i(x).$$

Proof. Setting $\bar{f}_i(x, y) = f_i(x)$ it will be enough to show $\mu(P_i | C_\infty) = \bar{f}_i$. Since $\int_A \bar{f}_i d\lambda = \int_A \chi_{P_i} d\lambda$ for any $A \in C_\infty$ we need only check that the \bar{f}_i are C_∞ measurable. $C_\infty = \bigvee_{n=1}^\infty C_n$ and each C_n is a partition of S into full columns over intervals in $[0, 1]$ so on the non-atomic part of $([0, 1], C_\infty, \mu)$, C_∞ separates points. The previous lemma implies that the \bar{f}_i are constant on any atoms of C_∞ (in fact either zero or one) and hence the result follows. \square

COROLLARY 3.3. *The entropy of T_f with respect to the two set partition P_f is calculated as*

$$h_\mu(T_f, P_f) = \int_0^1 - \sum_{i=1}^2 f_i(t) \log f_i(t) dt,$$

(where $0 \log 0 = 0$.) *In particular $h_\mu(T_f, P_f) = 0$ iff $f_1 = \chi_A$ for a measurable subset $A \in \mathcal{B}$.*

Lemma 3.2 also says that the pair $f = (f_1, f_2)$ is a 'graph' of the distribution of P_f on the past C_∞ . It should not be hard then to turn this around and find, given a stationary distribution, appropriate $f = (f_1, f_2)$ to represent the process.

THEOREM 3.4. *Let (T, P) be a process with $|P| = 2$. Then there exist two measurable functions f_1 and f_2 with $f_1 + f_2 = 1$, so that if $f = (f_1, f_2)$ we have (T_f, P_f) and (T, P) isomorphic processes.*

Before giving a proof we establish some notation. If $s \in \{0, 1\}^n$ and $s' \in \{0, 1\}^m$, say $s = s_1 s_2 \cdots s_n$, $s' = s'_1 s'_2 \cdots s'_m$, then denote by ss' the concatenation of s and s'

$$(ss')_i = \begin{cases} s_i & \text{if } 1 \leq i \leq n \\ s'_{i-n} & \text{if } n < i \leq n + m. \end{cases}$$

Let \check{s} denote the *truncation* of s ; $\check{s} \in \{0, 1\}^{n-1}$ with $\check{s}_i = s_i$ for $1 \leq i \leq n-1$ that is, \check{s} is obtained from s by dropping the rightmost entry of s .

Proof of Theorem 3.2. It will be enough to show that for our advertised $f = (f_1, f_2)$ and each $n \geq 1$ that $\text{dist}(C_n) = \text{dist}(\bigvee_{i=1}^n T^{-i}P)$. Also, by the remarks in the first section we may assume that T is the shift on $\{0, 1\}^{\mathbb{Z}}$ with invariant measure $\hat{\mu}$ on

the σ -algebra generated by the cylinder subsets and that P is the time-zero partition of $\{0, 1\}^Z$. On $[0, 1]$ we construct a sequence of partitions into intervals, beginning with $I_0^1 = [0, \hat{\mu}(0)]$, $I_1^1 = [0, 1] - I_0^1$ and $I^1 = \{I_s^1\}_{s \in \{0,1\}}$. Given the partition $I^n = \{I_s^n\}_{s \in \{0,1\}^n}$, with each $I_s^n = [a_s, b_s]$ and $\lambda(I_s^n) = \hat{\mu}(s)$ define $I_{s0}^{n+1} = [a_s, a_s + \hat{\mu}(s0)]$, $I_{s1}^{n+1} = [b_s - \hat{\mu}(s1), b_s]$. Clearly the I_s^{n+1} satisfy $\lambda(I_s^{n+1}) = \hat{\mu}(s)$ for each $s \in \{0, 1\}^{n+1}$. Evidently the partition $I^{n+1} = \{I_s^{n+1}\}_{s \in \{0,1\}^{n+1}}$ is a collection of closed intervals indexed by $s \in \{0, 1\}^{n+1}$, $\lambda(I_s^{n+1}) = \hat{\mu}(s)$ and moreover the I_s^{n+1} are placed on the interval so that the lexicographic order on $s \in \{0, 1\}^{n+1}$ is consistent with the usual order on $[0, 1]$. Almost every $x \in [0, 1]$ lies in a unique sequence of nested intervals of the form

$$I_n(x) = I_{s_1}^1 \cap I_{s_1 s_2}^2 \cap \dots \cap I_{s_1 s_2 \dots s_n}^n,$$

and in this case we define

$$\Phi(x) = \{x \in \{0, 1\}^Z \mid x_i = s_i, i = 1, 2, \dots\}.$$

Φ is, by construction, measure-preserving and we may let $f_1(x) = \hat{\mu}(P_0 \mid \bigvee_1^\infty T^{-i}P)(\Phi(x))$, $f_2(x) = 1 - f_1(x)$ and $f = (f_1, f_2)$. It remains only to show that $\bigvee_1^\infty T_f^{-1}P_f = I^n \times [0, 1]$ when viewed as two partitions of the square. This will follow if both $\varphi_1^* I_s^n = I_{0s}^{n+1}$ and $\varphi_2^* I_s^n = I_{1s}^{n+1}$ for each $s \in \{0, 1\}^n$. Write $I_s^{n-1} = I_{s0}^n \cup I_{s1}^n$ (with I_{s0}^n immediately to the left of I_{s1}^n) and assume $\varphi_1^* I_s^{n-1} = I_{0s}^n$. Then $I_{0s}^n = \varphi_1^* I_{s0}^n \cup \varphi_1^* I_{s1}^n$. Since the φ_i^* are continuous and increasing, the rightmost side of this equality is a union of two intervals $\varphi_1^* I_{s0}^n$ immediately to the left of $\varphi_1^* I_{s1}^n$ and having measures $\int_{I_{s0}^n} f_1(x) dx = \hat{\mu}(0s0)$ and $\int_{I_{s1}^n} f_1(x) dx = \hat{\mu}(0s1)$ respectively.

But $I_{0s}^n = I_{0s0}^{n+1} \cup I_{0s1}^{n+1}$ is also a decomposition of I_{0s}^n into two intervals, I_{0s0}^{n+1} immediately to the left of I_{0s1}^{n+1} with measures $\hat{\mu}(0s0)$ and $\hat{\mu}(0s1)$ respectively. We conclude that $\varphi_1^* I_{s0}^n = I_{0s0}^{n+1}$ and $\varphi_1^* I_{s1}^n = I_{0s1}^{n+1}$ as desired. The same argument applied to φ_2^* instead of φ_1^* completes our proof. \square

It will be worthwhile at this point to mention that our setting of two-state stationary stochastic processes has been chosen for notational convenience only. For, suppose we have a countable, linearly ordered index set I and a collection of measurable functions $f_i : [0, 1] \rightarrow [0, 1]$ satisfying $\sum_I f_i = 1$,

$$\left(\text{here, if } \{a_i\}_{i \in I} \text{ are positive numbers then } \sum_I a_i = \sup_{\substack{F \in I \\ |F| < \infty}} \left\{ \sum_{i \in F} a_i \right\} \right).$$

Define, for each $i_0 \in I$ and $x \in [0, 1]$,

$$\varphi_{i_0}^*(x) = \sum_{i < i_0} \int_0^1 f_i(t) dt + \int_0^x f_{i_0}(t) dt.$$

The $\varphi_{i_0}^*$ are continuous, increasing functions and if $i < i_0$ then $\varphi_i(1) \leq \varphi_{i_0}(0)$. Also, $[0, 1] = \bigcup_i \varphi_i^*[0, 1]$. For $x \in \varphi_{i_0}^*[0, 1] = [\varphi_{i_0}^*(0), \varphi_{i_0}^*(1)]$ let $\varphi_{i_0}(x) = \inf \{t \mid \varphi_{i_0}(t) \geq x\}$ and let f denote the (countably infinite) vector $\{f_i\}_{i \in I}$. If $(x, y) \in S$ and $x \in \varphi_{i_0}^*[0, 1]$ set

$$T_f(x, y) = \left(\varphi_{i_0}(x), \sum_{i < i_0} f_i(\varphi_{i_0}(x)) + f_{i_0}(\varphi_{i_0} x) y \right).$$

Let $P_{i_0} = \{(x, y) \in S \mid \sum_{i < i_0} f_i(x) \leq y \leq \sum_{i \leq i_0} f_i(x)\}$ and $P_f = \{P_i\}_{i \in I}$. Then T_f is an invertible, Lebesgue-measure-preserving transformation on the square S . If $C_\infty = \bigvee_1^\infty T_f^{-1} P_f$ then for each $i_0 \in I$, and (almost every) $(x, y) \in S$ we have $\mu(P_{i_0} \mid C_\infty)(x, y) = f_{i_0}(x)$. If

$$h_\mu(P_f) = - \sum_{i \in I} \int_0^1 f_i(x) dx \log \left(\int_0^1 f_i(x) dx \right) < \infty$$

then the entropy formula in Corollary 3.3 is still correct for countable I . On the other hand, by considering a sequence of finite partitions $P_f^{(n)} = \{P_i\}_{i \in F_n} \cup \{\bigcup_{I - F_n} P_j\}$ associated to a fixed increasing sequence of finite subsets $F_n \uparrow I$ one easily shows that the right hand side of the equality in Corollary 3.3 computes the entropy of the factor of T_f on the σ -algebra generated by P_f . In case $h_\mu(P_f) = \infty$ this could be infinite, while in the other direction it is easy to construct a zero-entropy gbt with $h_\mu(P_f) = \infty$. Given a countable state process (T, P) (here, the index set I of $P = \{P_i\}_{i \in I}$ comes equipped with a linear ordering by our convention) one may follow the construction in Theorem 1.4 to obtain functions $\{f_i\}_{i \in I}$ so that (T_f, P_f) and (T, P) are isomorphic.

4.

We would like to investigate ergodic properties of the transformation T_f , and as is often the case it will be convenient to do so by studying a suitable process constructed from T_f . The previous section provides evidence that the process (T_f, P_f) is a natural choice for our attention, and so we shall investigate conditions under which P_f will be a generator for T_f . In this section we shall frame our results in the context of a single cutting function (i.e., a two-state process.) The reader may check that our arguments work equally well for countable-state processes with the obvious modifications.

Consider the factor of T_f obtained by restricting the action of T_f to the invariant σ -algebra of vertical fibres. This is (measurably isomorphic in an obvious way to) the mapping

$$S_f(x) = \begin{cases} \varphi_1(x) & \text{if } x \in [0, \varphi_1^*(1)] \\ \varphi_2(x) & \text{if } x \in [\varphi_1^*(1), 1], \end{cases}$$

a 2-1 Lebesgue-measure-preserving endomorphism of the unit interval. We should mention at this point that T_f restricted to the σ -algebra generated by P_f will be the natural extension of S_f (see [R] for a definition of this). We will make essential use of this fact in § 5.

Using the two branches of S_f^{-1} , ${}_1S_f^{-1} = \varphi_1^*$ and ${}_2S_f^{-1} = \varphi_2^*$ we define the Frobenius-Perron operator F_f on $L^\infty[0, 1]$ associated with S_f :

$$\begin{aligned} F_f g(x) &= ({}_1S_f^{-1})'(x)g({}_1S_f^{-1}(x)) + ({}_2S_f^{-1})'(x)g({}_2S_f^{-1}(x)) \\ &= f_1(x)g(\varphi_1^*(x)) + f_2(x)g(\varphi_2^*(x)). \end{aligned}$$

We will need two easily checked properties of the operator F_f :

- (i) $\int F_f g d\lambda = \int g d\lambda$.

(ii) If $g \geq 0$ and $F_f g = g$ then g is the density of an S_f -invariant measure absolutely continuous with respect to λ .

We will say that S_f is *ergodic* if whenever $A \in \mathcal{B}_1$ with $S_f^{-1}A = A$, then either $\lambda(A) = 0$ or $\lambda(A) = 1$. We now present two sufficient conditions for P_f to be a generator.

LEMMA 4.1. *Suppose $f = (f_1, f_2)$ satisfies both*

- (i) $\lambda(\bigcup_{i=1,2} \{x \mid f_i(x) = 0\}) < 1$, and
- (ii) *The endomorphism S_f is ergodic.*

Then P_f is a generator for T_f .

LEMMA 4.2. *Suppose $f = (f_1, f_2)$ satisfies*

- (i) *There exists a $\delta > 0$ so that for $i = 1, 2, f_i \leq 1 - \delta$.*

Then P_f is a generator for T_f .

COROLLARY 4.3. *Let $f = (f_1, f_2)$ with f_1 not the characteristic function of a set in $[0, 1]$. Then T_f is ergodic if and only if S_f is ergodic.*

Proof. f satisfies condition (i) of Lemma 4.1.

COROLLARY 4.4. *Let T be an ergodic automorphism with $0 < h(T) < \log 2$. Then T is measurably isomorphic to a gbt T_f with a single cutting function.*

Proof. By a theorem of W. Krieger [K] we may find a generating partition P for the automorphism T with $|P| = 2$. By Theorem 3.4 we may find functions $f = (f_1, f_2)$ so that (T, P) and (T_f, P_f) are isomorphic processes. Of necessity $0 < h(T) = h(T, P) = h(T_f, P_f)$ and so by Corollary 3.3 $\lambda(\bigcup_{i=1,2} \{x \mid f_i(x) = 0\}) < 1$. Furthermore $\lim_{n \rightarrow \infty} \sup \{\mu(A) \mid A \in \bigvee_1^n T^i P\} = 0$ (one may see this, for example, from the Shannon-McMillan-Breiman Theorem) so we may conclude that the σ -algebra C_∞ contains no atoms and hence $C_\infty = C$ the σ -algebra of vertical fibres. We conclude that S_f is (isomorphic to) a factor of T and so is ergodic. Lemma 4.1 now implies the result. □

Proof of Lemma 4.2. Without loss of generality we may assume $f_i(x) < 1 - \delta$ for $i = 1, 2$ and all $x \in [0, 1]$. It is clear from Lemma 1.1 that C_∞ is the σ -algebra of vertical fibres on S so it only remains to show that two points on the same fibre, say (x, y_1) and (x, y_2) with $y_1 < y_2$ eventually lie in different atoms of $\bigvee_{i=0}^n T_f^i P_f$. Suppose not. Then for each $i \geq 1, T_f^{-i}(x, y_1)$ and $T_f^{-i}(x, y_2)$ lie on the same vertical fibre and in the same atom of P_f . A simple induction however shows that

$$|T_f^{-i}(x, y_1) - T_f^{-i}(x, y_2)| \geq \left(\frac{1}{1 - \delta}\right)^i (y_2 - y_1),$$

a contradiction for sufficiently large i . □

Example 4.5. The hypotheses of Lemma 4.2 are not necessary, for, let T_x denote the gbt with cutting function $f_1(x) = x$. Let $(x, y_1) \neq (x, y_2)$ and

$$\delta_0 = \min \left\{ \frac{|y_2 - y_1|}{2}, \frac{1}{10} \right\}.$$

If $T_x^{-n}(x, y_1)$ and $T_y^{-n}(x, y_2)$ lie in the same atom of P_x for each $n = 0, 1, 2 \dots$ (and hence on a common vertical fibre for each $n \geq 0$), then no two consecutive iterates of these points may lie on a fibre over the subset $[0, \delta_0] \cup [1 - \delta_0, 1]$. Conclude that

$$|T^{-n}(x, y_1) - T^{-n}(x, y_2)| \geq \left(\frac{1}{1 - \delta_0}\right)^{n/2} |y_1 - y_2|$$

and hence that P_x is a generator.

Proof of Lemma 2.1. We show first that C_∞ separates the vertical fibres on S . For, if not, then we may get $x_1 < x_2$, the vertical fibres over which always lie in the same atom of $T_f^{-k}P_f$ for each $k \geq 1$. Denote by I_f the partition of $[0, 1]$ induced by $T_f^{-1}P_f$. Then $[x_1, x_2]$ lies in an atom of $\bigvee_{n=0}^\infty S_f^{-n}I_f$ and for each $p > 0$, $S_f^p[x_1, x_2]$ is a single interval and lies in an atom of $S_f^{-k}I_f$. Conclude that for each $p > 0$ the column over $S_f^p[x_1, x_2]$ is an atom of C_∞ and hence, by Lemma 3.1 either $f_1|_{S_f^p[x_1, x_2]} = 0$ or $f_2|_{S_f^p[x_1, x_2]} = 0$. Set $J = [x_1, x_2]$ and $A = J \cup S_f J \cup S_f^2 J \cup \dots$ and observe that $S_f^{-1}A = A$. ($S_f A \subseteq A$ so $S_f^{-1}(S_f A) \subseteq S_f^{-1}A$ and obviously $A \subseteq S_f^{-1}(S_f A)$. Hence $A \subseteq S_f^{-1}A$.) Thus $\bigcup_{n=1}^\infty S_f^n J = [0, 1]$ and this contradicts condition (i).

Given a subset $B \in \mathcal{B}_2$ and $x \in [0, 1]$ denote by B_x the intersection of B with the vertical fibre over x . Let ν denote the product measure on $S \times S$ relative to the factor S_f ; its values on cylinder subsets are easily described, for if $A, B \in \mathcal{B}_2$ then $\nu(A \times B) = \int_0^1 \mu(A_x \times B_x) dx$. For each $n \geq 0$ let $D_n = \{((x, y_1), (x, y_2)) \mid (x, y_1) \text{ and } (x, y_2) \text{ lie in the same atom of } \bigvee_0^n T_f^i P_f\} \subseteq S \times S$. The D_n are measurable and $D_{n+1} \subseteq D_n$. If

$$d_n(x) = \mu\{(y_1, y_2) \mid ((x, y_1), (x, y_2)) \in D_n\}$$

then $\nu(D_n) = \int d_n(x) d\lambda(x)$. Since $\bigvee_0^n T_f^i P_f$ induces an interval partition on each vertical fibre in the square, in order to show $\bigvee_0^\infty T_f^i P_f$ separates points on these vertical fibres it will be enough to show $\nu(D_n) \downarrow 0$ or equivalently, since $d_n \downarrow$, that $d_n \downarrow 0$ (λ - a.e.).

Observe that the functions d_n satisfy the following recursion relation: for λ - a.e. $x \in [0, 1]$

$$d_n(x) = f_1^2(x)d_{n-1}(\varphi_1^*(x)) - f_2^2(x)d_{n-1}(\varphi_2^*(x)).$$

Setting $d(x) = \lim_n d_n(x)$ and taking limits in the above expression we see for λ - a.e. $x \in [0, 1]$

$$d(x) = f_1^2(x)d(\varphi_1^*(x)) + f_2^2(x)d(\varphi_2^*(x)),$$

from which we may conclude (since $d \geq 0$)

$$\sqrt{d(x)} \leq f_1(x)\sqrt{d(\varphi_1^*(x))} + f_2(x)\sqrt{d(\varphi_2^*(x))} = F_f \sqrt{d(x)}.$$

But $\int \sqrt{d} d\lambda = \int F_f \sqrt{d} d\lambda$ so the previous inequality must be (λ - a.e.) equality and we conclude that \sqrt{d} is the density of an absolutely continuous S_f invariant measure; consequently $\sqrt{d} = \text{const.}$ (λ - a.e.). Hence d is a constant function, but this is in contradiction with the equality

$$d(x) = f_1^2(x)d(\varphi_1^*(x)) + f_2^2(x)d(\varphi_2^*(x))$$

at any x where $0 < f_1(x), f_2(x) < 1$ unless $d = 0$. The lemma now follows. □

Example 4.6. Lemma 4.2 does not follow from Lemma 4.1. We construct a non-ergodic gbt with cutting function f bounded away from zero and one.

Let Y denote the measure space of two points a and b equipped with equidistributed probability measure p_2 . Let B denote the Bernoulli $(\frac{1}{3}, \frac{2}{3})$ shift on $\Omega = \{0, 1\}^{\mathbb{Z}}$ with $\nu(0) = \frac{1}{3}, \nu(1) = \frac{2}{3}$ and time-zero partition R . Denote the atoms of R by '0' and '1'. Let $X = Y \times \Omega, T = \text{identity} \times B$ and let P be the partition of X into two atoms

$$P_0 = (\{a\} \times '0') \cup (\{b\} \times '1'); P_1 = X - P_0.$$

Then P is clearly a generator for T as no two points $(a, x), (b, y)$ can have the same P -name by the ergodic theorem applied to B . Each atom $A \in \bigvee_1^n T^{-i}P$ may be written as $A = A_a \cup A_b$ where $A_a = \{(a, y) \in A\}, A_b = \{(b, y) \in A\}$. If $m = p_2 \times \nu$ then

$$m(P_0|A) = \frac{m(A_a)}{m(A)} m(P_0|A_a) + \frac{m(A_b)}{m(A)} m(P_0|A_b).$$

But $m(P_0|A_a) = \frac{1}{3}$ and $m(P_0|A_b) = \frac{2}{3}$ so $\frac{1}{3} \leq m(P_0|A) \leq \frac{2}{3}$. It follows that the functions $f = (f_1, f_2)$ constructed in Theorem 3.4 satisfy $\frac{1}{3} \leq f_i \leq \frac{2}{3}$ and so P_f is a generator for T_f . Since T_f is not ergodic, neither can S_f be ergodic.

5.

We shall now investigate connections between topological restrictions on the functions f_i and ergodic properties of the associated generalized baker's transformation T_f . So far, we have made the trivial (and well known) observation that if the f_i are constant functions then P_f is an independent generator for T_f and hence T_f is measurably isomorphic to a Bernoulli shift with entropy determined by Corollary 3.3 (which may be infinite.) In Example 4.6 we saw that even when there exists a $\delta < 1$ with $f_i \leq \delta$ for all $i \in I, T_f$ may fail to be ergodic. In this example the cutting function was discontinuous. In what follows we shall show, roughly speaking, that if the f_i are Hölder-continuous functions and are uniformly bounded below 1, T_f is measurably isomorphic to a Bernoulli shift with the appropriate entropy. This is achieved by verifying the weak-Bernoulli condition on the generator P_f , from which the result follows by the Ornstein theory, [F, O], [O]. We shall give our arguments in the context of a countable number of cutting functions $\{f_i\}_{i \in I}$, a result we shall need for later applications.

Recall the definitions of S_f and $I_f = \{I_i\}_{i \in I}$ from § 4, and extend these in the obvious way to the setting of a countable number of cutting functions.

Definition 5.1. We will say the family $\{f_i\}_{i \in I}$ is *piecewise Hölder-continuous with respect to I_f* if we may find numbers $0 < \alpha \leq 1$ and $M < \infty$, and a partition of $[0, 1]$ into finitely many disjoint open intervals $J_n = (a_n, b_n)$, each J_n a union (within a set of measure zero) of atoms from I_f and so that if $x, y \in J_n$

$$\left| \frac{f_i(x)}{f_i(y)} - 1 \right| \leq M |x - y|^\alpha.$$

Here we set

$$\frac{0}{0} = 1$$

by convention.

Definition 5.2. We say that P_f is a *Markov mixing partition* for T_f if there is an N so that $\mu(P_k | T_f^n P_l) > 0$ for all $k, l \in I$ and $n \geq N$.

Remark 5.3. Definition 5.2 implies that for all sufficiently large n , and all atoms $I_i \in I_f$ and atoms $A \in T_f^n P_f$ (the later a union of horizontal strips bounded by measurable functions) there is an $x \in I_i$ so that the fibre over x has a non-trivial interval of intersection with A . If the $\{f_i\}_{i \in I}$ also satisfy Definition 5.1 then the fibres over every $x \in \text{interior } A = \overset{\circ}{A}$ will have this property.

Recall that C denotes the σ -algebra of vertical fibres on the square. Write S_0 for the square S minus the collection of vertical fibres over endpoints of the J_n . If $(x, y) \in S, (x, y) \in P_i$ then set $\varphi(x, y) = -\log f_i(x); \varphi : S \rightarrow \mathbb{R} \cup \{+\infty\}$. Let $M(S_0)$ denote the probability measures on S_0 and $M_{T_f}(S_0)$ the T_f -invariant members of $M(S_0)$.

Let us now state the main result of this section.

THEOREM 5.4. Let $\{f_i\}_{i \in I}$ satisfy the conditions:

- (i) $\sup_i \{\sup_{x \in [0,1]} f_i(x)\} = \delta < 1$.
- (ii) $\{f_i\}_{i \in I}$ is a piecewise Hölder-continuous collection with respect to I_f .
- (iii) P_f is a Markov mixing partition for T_f .

Then the partition P_f is a weak-Bernoulli generator for the transformation T_f and hence T_f is measurably isomorphic to a (generalized) Bernoulli shift. Moreover, Lebesgue measure μ is the unique member of $M_{T_f}(S_0)$ which satisfies the variational inequality:

$$\mu(I_\mu(T_f C | C) - \varphi) \geq m(I_m(T_f C | C) - \varphi)$$

for all $m \in M_{T_f}(S_0)$, the left-hand side of this inequality being zero.

Remark 5.5. In case $f_i > 0$ condition (iii) is trivially satisfied. In this simpler situation a direct verification of the weak-Bernoullicity of P_f is possible and in fact a lower bound on the rate of convergence to independence may be given. The details are worked out in [B1]. On the other hand, in order to include the well known case of finite state mixing Markov shifts, whose cutting functions will be piecewise constant and which were shown to have weak-Bernoulli generators in [F, O], the more general hypotheses of our theorem are needed. In all that follows we shall assume that the $\{f_i\}_{i \in I}$ satisfy the conditions of this theorem.

It will be convenient to work with the endomorphism S_f and the partition I_f , drawing our conclusions from Theorem 15 of Walters [W]. We must first, however, change slightly the space upon which S_f acts. Note first that without loss of generality we may assume that the support of each f_i is exactly one interval J_m ; in terms of the endomorphism this implies S_f sends each I_k onto exactly one interval J_m . Denote by $|x_1 - x_2|$ the usual distance between two points $x_1, x_2 \in \mathbb{R}^2$. We now follow the notation of [W].

Let $\bar{X} = \bigcup_n \bar{J}_n \times \{5n\}$, a compact subset of \mathbb{R}^2 . Let $X = \bigcup_n \tilde{J}_n \times \{5n\}$, so that $X \subseteq \bar{X}$ is an open, dense subset, and let $\tilde{J}_n = J_n - \{\text{all endpoints of the intervals } I_j\}$. Then $X_0 = \bigcup_n \tilde{J}_n \times \{5n\}$ is an open, dense subset of \bar{X} and $X_0 \subseteq X$. If $(x, 5m) \in \tilde{J}_m \times \{5m\}$

and $S_f x \in J_n$ then set $T(x, 5m) = (S_f x, 5n)$, thereby defining a continuous map $T: X_0 \xrightarrow{\text{onto}} X$. To keep our notation reasonably compact we shall often write x for $(x, 5n) \in \bar{X}$ when $x \in I_n$, and so, for example both $T(x)$ and $S_f(x)$ make sense. Continuing in this spirit we shall denote by λ , the one-dimensional Lebesgue measure on \bar{X} and by I_f the partition of \bar{X} naturally induced by the partition I_f on $[0, 1]$.

Restricted to the interior of any atom of I_f , T is a diffeomorphism onto a subset $\dot{J}_m \times \{5m\}$. If $x \in I_i$ then

$$T'(x) = S'_f(x) = \frac{1}{f_i(\varphi_i(x))},$$

the right-hand side being defined since φ_i takes I_i onto a set where f_i is strictly positive by hypothesis (ii).

T is measure-preserving on \bar{X} , moreover, by the above, if $\lambda(E) = 0$ then $\lambda(TE) = 0$; in the terminology of [W] λ is positively nonsingular for T . If $x \in \bar{X}$, $\varepsilon > 0$, let $B_\varepsilon(x) = \{x' \in \bar{X} \mid |x - x'| < \varepsilon\}$.

LEMMA 5.6. (1) *There exists an $\varepsilon_0 > 0$ such that for every $x \in X$, $T^{-1}(B_{2\varepsilon_0}(x) \cap X)$ can be written uniquely as a disjoint union of a finite or countable number of open subsets $A_1(x), A_2(x) \dots$ of X_0 with, for each i , $T|_{A_i(x)}$ being a homeomorphism of $A_i(x)$ onto $B_{2\varepsilon_0}(x) \cap X$ and so that if $y, y' \in A_i(x)$, $|Ty - Ty'| \geq |y - y'|$.*

(2) *For all $\varepsilon > 0$ there exists an $M > 0$ such that for all $x \in X$, $T^{-M}x$ is an ε -dense subset of X .*

(3) $d\lambda \circ T^{-1} / d\lambda = 1$.

(4) $\sup_{n \geq 1} \{\sup_{y \in T^{-n}x} (d\lambda \circ T^n / d\lambda)(y') / (d\lambda \circ T^n / d\lambda)(y)\}$ is bounded if $|x - x'| < \varepsilon_0$, where y' is chosen to be that $y' \in T^{-n}x'$ which lies in the same atom of $\bigvee_{j=0}^n T^{-j}I_f$ as does y . Moreover the above quantity converges to 1 as $|x - x'| \rightarrow 0$.

(5) *The partition I_f restricted to X_0 with atoms $\{\dot{I}_i\}_{i \in I}$ satisfies.*

(a) $T|_{\dot{I}_i}$ is one-to-one.

(b) $\lambda(\partial \dot{I}_i) = 0$.

(c) For each $i \in I$, $T\dot{I}_i$ is λ -a.e. a union of \dot{I}_k .

(d) For all choices of the sequence $\{i_n\}$ in $I^{\mathbb{N}}$ the set $\bigcap_n T^{-n}(\dot{I}_{i_n})$ contains at most one point.

(e) Each \dot{I}_i is a subset of some component of $T^{-1}B_{\varepsilon_0}(x_i)$ for some $x_i \in X$; ε_0 from (1).

Proof. (1) Let $\varepsilon_0 = 1$. The $A_i(x)$ then become the interiors of the I_i for any $x \in X$.

(2) Consider the interval partition of $[0, 1]$, $\bigvee_0^n S_f^{-i}I_f$. Hypothesis (i) of the theorem implies the maximal length of an atom in this partition is δ^{n+1} and so the same is true for the partition $\bigvee_0^n T^{-i}I_f$ on X . Thus it is enough to show that there exists an N so that for all $x \in X$, $\{T^{-n}x\} \cap I_i \neq \emptyset$ for all $n \geq N$, and $i \in I$. But this follows immediately from the Markov mixing hypothesis (iii): the N from Definition 5.2 will do the job.

(3) T is measure-preserving.

(4) It will be enough to show there is a function $M(x, x')$ defined for $(x, x') \in \bigcup_m J_m \times J_m$ with $M(x, x') \leq L|x - x'|^\alpha$ for some constants L and $0 < \alpha \leq 1$, and so

that for all $n \geq 1$ if y and y' are in the same atom of $\bigvee_{j=0}^{n-1} S_f^{-j} I_f$ with $S_f^n y = x$ and $S_f^n y' = x'$ then

$$(S_f^n)'y' / (S_f^n)'y \leq 1 + M(x, x').$$

Let M , α , and δ be from hypothesis (i) and (ii) of the theorem. Choose $\tilde{M} > M$ so

$$M \left(\frac{1}{\tilde{M}} + 1 \right) \leq \tilde{M} \delta^\alpha.$$

We shall find $M(x, x')$ by induction. First, we calculate

$$S_f'(y') / S_f'(y) = \frac{f_i(\varphi_i y')}{f_i(\varphi_i y)} = \frac{f_i(x')}{f_i(x)}$$

if y and $y' \in I_i$. But this latter quantity is bounded, in view of hypothesis (ii) by

$$1 + M|x - x'|^\alpha \leq 1 + (1 + \tilde{M})|x - x'|^\alpha.$$

Take $M_1 = 1 + \tilde{M}$. For $n > 1$ we apply the chain rule and obtain, if $S_f y \in I_i$

$$\begin{aligned} (S_f^n)'y' / (S_f^n)'y &= \frac{(S_f^{n-1})'(S_f y')}{(S_f^{n-1})'(S_f y)} \frac{S_f'(y')}{S_f'(y)} \\ &\leq (1 + M_{n-1}|x' - x|^\alpha)(1 + M|S_f y' - S_f y|^\alpha) \\ &\leq (1 + M_{n-1}|x' - x|^\alpha)(1 + M\delta^{\alpha(n-1)}|x' - x|^\alpha), \end{aligned}$$

this last inequality in view of the fact that $|S_f y' - S_f y| \leq \delta^{\alpha(n-1)}|x' - x|$ by hypothesis (i).

Expanding this last expression and factoring out $|x' - x|^\alpha$ one obtains, assuming $M_{n-1} \geq \tilde{M}$

$$\begin{aligned} &\leq 1 + |x' - x|^\alpha (M_{n-1} + M\delta^{\alpha(n-1)} + M_{n-1}M\delta^{\alpha(n-1)}) \\ &\leq 1 + M_{n-1}|x' - x|^\alpha \left(1 + \delta^{\alpha(n-1)}M \left(\frac{1}{\tilde{M}} + 1 \right) \right) \\ &\leq 1 + M_{n-1}|x' - x|^\alpha (1 + \tilde{M}\delta^{\alpha n}). \end{aligned}$$

Now if $M_{n-1} = \prod_{j=0}^{n-1} (1 + \tilde{M}\delta^{\alpha j})$ we have, indeed, $M_{n-1} > \tilde{M}$ and one finally obtains a bound of

$$1 + \prod_{j=0}^n (1 + \tilde{m}\delta^{\alpha j})|x' - x|^\alpha.$$

Set $M_n = \prod_{j=0}^n (1 + \tilde{M}\delta^{\alpha j})$ and $M^* = \sup_n M_n$. It follows that we may take the function $M(x, x')$ to be $M^*|x - x'|^\alpha$.

(5) All properties here are immediate consequences of the definitions except perhaps (d) which follows from the observation already made in (2) about the maximal length of an atom from $\bigvee_{j=0}^n S_f^{-j} I_f$. \square

Proof of Theorem 5.4. We have, in properties (1)–(5) of the previous lemma, verified the hypothesis of Theorem 15 in [W] for the mapping T defined on the open, dense subset X_0 of the compact, metric space \bar{X} , and with respect to the (non-singular and positively non-singular) measure λ on \bar{X} and the partition I_f . Since λ is also in $M_T(\bar{X})$ we conclude that the natural extension of T with respect to λ (and hence the natural extension of S_f with respect to λ) is a Bernoulli shift (a generalized Bernoulli shift in case $h(T_f)$ is infinite.) By Lemma 4.2, or more precisely by the

countably infinite cutting function version of Lemma 4.2, P_f is a generator for T_f and so, as observed before T_f is the natural extension of S_f .

Also, by Theorem 15 λ is the unique element of $M_T(X)$ satisfying

$$\lambda(I_\lambda(\mathcal{B}|T^{-1}\mathcal{B}) - \log T') \geq m(I_m(\mathcal{B}|T^{-1}\mathcal{B}) - \log T')$$

for all $m \in M_T(X)$. Here \mathcal{B} denotes the σ -algebra of Borel subsets of X and is generated by the $\{T^{-n}I_f\}_{n=0}^\infty$.

Setting $I_0 = [0, 1] - \{\text{endpoints of the intervals } J_n\}$ and writing this last inequality in terms of S_f we see

$$\lambda(I_\lambda(\mathcal{B}_1|S_f^{-1}\mathcal{B}_1) - \log S'_f) \geq m(I_m(\mathcal{B}_1|S_f^{-1}\mathcal{B}_1) - \log S'_f)$$

for all $m \in M_{S_f}(I_0)$. Furthermore, since $I_\lambda(\mathcal{B}_1|S_f^{-1}\mathcal{B}_1)(x) = -\log f_i(\varphi_i(x))$ when $x \in I_i$, and

$$S'_f(x) = \frac{1}{f_i(\varphi_i(x))},$$

the left-hand side of the above inequality is zero.

Finally, since P_f is a generator for T_f we may lift the above inequality to the square

$$\mu(I_\mu(C|T_f^{-1}C) - \varphi \circ T_f) \geq m(I_m(C|T_f^{-1}C) - \varphi \circ T_f)$$

or equivalently

$$\mu(I_\mu(T_f C|C) - \varphi) \geq m(I_m(T_f C|C) - \varphi)$$

for all $m \in M_{T_f}(S_0)$, the left-hand side of this inequality being zero. □

We conclude by mentioning that one may construct various generalized baker's transformations with single cutting functions bounded away from zero and one, and with T_f being ergodic but not weak mixing or weak mixing but not a K -automorphism. The constructions are similar to that of Example 4.6. We intend to publish the details in a future article.

6.

We say that $g : [0, 1] \rightarrow [0, 1]$ is a *piecewise increasing and C^1 mapping with finite image* if

- (1) There is a countable partition of $[0, 1]$ into subintervals; $\mathcal{I} = \{I_n\}_{n \in \mathbb{N}}$, $I_n = [a_n, b_n]$, $\bigcup_{n \in \mathbb{N}} I_n = [0, 1]$ and $\lambda(I_k \cap I_l) = 0$ if $k \neq l$, so that
- (2) for every $n \in \mathbb{N}$, $g|_{(a_n, b_n)}$ is strictly increasing, continuously differentiable and extends to a continuously differentiable function on an open interval $\tilde{I}_n \supseteq I_n$.
- (3) $\Omega = \bigcup_{n \in \mathbb{N}} \{\lim_{x \rightarrow a_n^+} g(x), \lim_{x \rightarrow b_n^-} g(x)\}$ is a finite set contained in $\bigcup_{n \in \mathbb{N}} \{a_n, b_n\}$.

In particular, each interval gI_k is λ -a.e. a union of atoms from \mathcal{I} .

If $x \in g(a_n, b_n)$ set $h_n(x) = (g|_{(a_n, b_n)})^{-1}(x)$ and extend this definition of h_n to all of $[0, 1]$ so as to make h_n continuous on $[0, 1]$ with $h'_n = 0$ on the interiors of the (at most) two intervals making up the complement of $g(a_n, b_n)$.

g will be *Lebesgue-measure-preserving* (i.e. $\lambda(g^{-1}B) = \lambda(B)$ for all $B \in \mathcal{B}_1$) if the following is verified for λ -a.e. $x \in [0, 1]$

$$F_g 1(x) = \sum_{n \in \mathbb{N}} h'_n(x) = 1.$$

Give N the linear order it inherits from the position of the intervals J_n on the interval $[0, 1]$, that is $n_1 < n_2$ iff (a_{n_1}, b_{n_1}) lies to the left of (a_{n_2}, b_{n_2}) . Theorem 5.4 then implies the following result.

COROLLARY 6.1. *Let $g : [0, 1] \rightarrow [0, 1]$ be a piecewise increasing and C^1 mapping which is Lebesgue-measure-preserving. Assume g also satisfies*

- (i) *There exists a $D > 0$ so that if $d \geq D$ and $k, l \in N, \lambda(I_k \cap g^d I_l) > 0$.*
- (ii) $\inf_n \{ \inf_{x \in (a_n, b_n)} g'(x) \} = C > 1$.
- (iii) *There exists $M < \infty$ and $0 < \alpha \leq 1$ so that if x and y lie in the interior of the same interval I_n one has*

$$\left| \frac{g'(x)}{g'(y)} - 1 \right| \leq M |g(x) - g(y)|^\alpha.$$

Then, the partition \mathcal{I} is a weak-Bernoulli generator for the endomorphism g and the natural extension of g is the gbt associated with the cutting functions $\{h'_n\}_{n \in N}$. This extension is measurably isomorphic to a Bernoulli shift with entropy

$$\sum_n \int_{a_n}^{b_n} \log g'(x) dx$$

when this sum converges (a generalized Bernoulli shift when the sum diverges). Moreover, if we set $I_0 = (0, 1) - \Omega$, λ is the unique member of $M_g(I_0)$ satisfying

$$\lambda(I_\lambda(\mathcal{B}_1 | g^{-1} \mathcal{B}_1) - \log g') \geq m(I_m(\mathcal{B}_1 | g^{-1} \mathcal{B}_1) - \log g')$$

for all $m \in M_g(I_0)$.

Proof. With $f_n = h'_n$ (defined λ -a.e.) we have the collection $f = \{f_n\}_{n \in N}$ a piecewise Hölder-continuous family with respect to the finite interval partition formed by the points $\Omega \cup \{0, 1\}$, for, if $(x, y) \cap \Omega = \phi$

$$\left| \frac{f_n(x)}{f_n(y)} - 1 \right| = \left| \frac{h'_n(x)}{h'_n(y)} - 1 \right| = \left| \frac{g'(h_n(y))}{g'(h_n(x))} - 1 \right| \leq M |x - y|^\alpha$$

by hypothesis (iii). Also, $\sup_n \{ \sup_{x \in [0, 1]} f_n(x) \} = C^{-1} < 1$. One easily sees that the partition $T_f^{-1} P_f$ is $\mathcal{I} \times [0, 1] = \{I_n \times [0, 1]\}_{n \in N}$ and that hypothesis (i) gives the Markov mixing condition on $T_f^{-1} P_f$, and hence on P_f . If $x \in (a_n, b_n)$ one computes, using the formula of § 3

$$\begin{aligned} S_f(x) &= \inf_{t \geq 0} \left\{ \int_0^t h'_n(x) d\lambda(s) \geq x - a_n \right\} \\ &= \inf_{t \geq 0} \{ h_n(t) - a_n \geq x - a_n \} \\ &= \inf_{t \geq 0} \{ h_n(t) \geq x \} \\ &= g(x), \end{aligned}$$

this last inequality since h_n is increasing and continuous. By the countable cutting function version of Lemma 4.2, P_f is a generator for T_f and so T_f is the natural extension of S_f and hence of g . The conclusions now follow from Theorem 5.4 and

the entropy formula in § 3:

$$\begin{aligned} h(g) = h(T_f) &= -\sum_n \left(\int_0^1 h'_n(x) \log h'_n(x) d\lambda(x) \right) \\ &= -\sum_n \left(\int_{g(a_n, b_n)} h'_n(x) \log h'_n(x) d\lambda(x) \right) \\ &= \sum_n \left(\int_{g(a_n, b_n)} h'_n(x) \log g'(h_n(x)) d\lambda(x) \right) \\ &= \sum_n \left(\int_{a_n}^{b_n} \log g'(u) d\lambda(u) \right). \end{aligned}$$

Finally, the variational inequality stated is precisely the one which appears for S_f in the proof of Theorem 5.4. \square

Remark 6.2. Corollary 6.1 is an addition to a family of results, the prototype for which is the 'Adler's Folklore Theorem' [A]. The main point of interest here is the relative weakness of the differentiability condition (iii); the corresponding condition in Adler's result being

$$\sup_n \left\{ \sup_{x, y \in I_n} \frac{|g''(x)|}{|g'(y)|^2} \right\} < +\infty$$

which has come to be known as Renyi's condition. On the other hand, the mappings treated by the Folklore Theorem need not be Lebesgue-measure-preserving. Indeed part of the point in that result is to give an invariant measure equivalent to Lebesgue measure which subsequently proves to have the strong ergodic properties we have shown in our Corollary.

In [B2] we have shown the existence of invariant measures equivalent to Lebesgue measure for mappings satisfying conditions (i)–(iii) of the Corollary but which are not a-priori measure-preserving and so present a true generalization of Adler's Theorem.

Acknowledgement. This work constitutes part of the authors doctoral thesis, written at the University of Toronto under the direction of M. A. Akcoglu, supported in part by the Natural Sciences and Engineering Research Council of Canada and the government of Ontario, Canada.

REFERENCES

- [A] R. Adler. Continued fractions and Bernoulli trials. *Ergodic Theory, A Seminar, Courant Institute of Mathematical Sciences*. J. Moser, E. Phillips & S. Varadhan, eds. (1975), Ch. 16, pp 111–120.
- [A, Y] J. C. Alexander & J. A. Yorke. Fat baker's transformations. *Ergod. Th. & Dynam. Sys.* 4 (1984), 1–23.
- [Be] H. C. P. Berbee. Chains with infinite connections: Uniqueness and Markov representation. *Centrum Voor Wiskunde en Informatica*. Report MS-R8509 (November 1985).
- [B1] C. Bose. Generalized baker's transformations. *Thesis, University of Toronto*. (1986).
- [B2] C. Bose. Invariant measures and equilibrium states for piecewise $C^{1+\delta}$ mappings of the interval. *Trans. Amer. Math. Soc.* (1989) to appear.

- [Bo] R. Bowen. Bernoulli maps of the interval. *Israel J. Math.* **28** (1977), 161–168.
- [F, O] N. A. Friedman & D. S. Ornstein. On isomorphism of weak-Bernoulli transformations. *Adv. Math.* **5** (1971), 365–394.
- [H] P. Halmos. *Lectures on Ergodic Theory*. Chelsea Publishing Company: New York, 1956.
- [K] W. Krieger. On entropy and generators of measure preserving transformations. *Trans. Amer. Math. Soc.* **149** (1970), 453–464.
- [L] F. Ledrappier. Principe variationnel et systèmes dynamiques symboliques. *Z. Wahrscheinlichkeitstheorie verw. Gebiete.* **30** (1974), 185–202.
- [M] J. Moser. Stable and random motions in dynamical systems. *Hermann Weyl Lectures, Annals of Mathematical Studies Number 77*. Princeton University Press: Princeton, 1973.
- [O] D. S. Ornstein. Two Bernoulli shifts with infinite entropy are isomorphic. *Adv. Math.* **5** (1970), 339–348.
- [R] V. A. Rohlin. Exact endomorphisms of a Lebesgue space. *Transl. Amer. Math. Soc. Series II.* **39** (1964), 1–36.
- [W] P. Walters. Invariant measures and equilibrium states for some mappings which expand distances. *Trans. Amer. Math. Soc.* **236** (1978), 121–153.