ISRAEL JOURNAL OF MATHEMATICS 93 (1996), 359–372

A C^1 EXPANDING MAP OF THE CIRCLE WHICH IS NOT WEAK-MIXING

BY

ANTHONY N. QUAS

Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SB, England
e-mail: A.Quas@statslab.cam.ac.uk

ABSTRACT

In this paper, we construct an example of a C^1 expanding map of the circle which preserves Lebesgue measure such that the system is ergodic, but not weak-mixing. This contrasts with the case of $C^{1+\epsilon}$ maps, where any such map preserving Lebesgue measure has a Bernoulli natural extension and hence is weak-mixing.

1. Introduction

In this paper, we apply techniques of [7] to prove the following theorem.

THEOREM 1: *There is a C^1 expanding map of the circle preserving Lebesgue measure, such that Lebesgue measure is ergodic for the map, but not weak-mixing.*

This is in contrast with results for the $C^{1+\epsilon}$ case, where it is known that if such a map preserves Lebesgue measure, then the natural extension of the transformation is Bernoulli [8]. Previously, Bose [2] has established the existence of a piecewise monotone and continuous expansive map preserving Lebesgue measure which is weak-mixing but not ergodic. (He also found piecewise monotone and continuous maps which are weak- but not strong-mixing; and strong-mixing but not exact). These proofs were based on the construction of generalized baker's transformations (see [1] for details).

Received April 17, 1994 and in revised form July 11, 1994

We will make extensive use of g-measures in what follows. For a fuller description of g-measures, the reader is referred to [4], [6] and [7]. Here, we will construct a g-function on the symbol space

$$\Sigma_{10} \equiv \{0, \dots, 9\}^{\mathbb{Z}^+} = \{x_0 x_1 x_2 \cdots : x_i \in \{0, \dots, 9\}\}$$

with shift map σ (that is a continuous function g satisfying $0 < g(x) < 1$ for all x and $\sum_{y \in \sigma^{-1}(x)} g(y) = 1$ for all x). Given such a g, we consider sequences of random variables $(X_n)\colon \Omega \to \{0, \dots, 9\}$ satisfying

$$\mathbb{P}(X_n = i |\, X_{n-1} = a_1,\ X_{n-2} = a_2,\ \dots) = g(i, a_1, a_2, \dots), \tag{1}$$

for all n. There are then natural maps $\rho_n\colon \Omega \to \Sigma_{10}$ defined by $\rho_n(\omega) = X_{n-i}(\omega)$. These maps induce natural push-forward maps of probability distributions on Ω to probability measures on Σ_{10} defined by $\rho_n^*(\mathbb{P})(A) = \mathbb{P}(\rho_n^{-1}(A))$. A g-measure is a push-forward under ρ_0^* of any stationary distribution. Another way of characterizing g-measures on symbol spaces is that a g-measure is a measure ν satisfying

$$\lim_{n \to \infty} \frac{\nu([ix]^{n+1})}{\nu([x]^n)} = g(ix), \tag{2}$$

for all $x \in \Sigma_{10}$, where $[x]^n$ denotes the cylinder of those points of Σ_{10} which agree with x for the first n terms, and ix denotes the sequence in Σ_{10} which consists of the symbol i followed by the sequence x.

We will need to consider g-functions which have the property of **compatibility** introduced in [6], that is $g(000\dots) = g(999\dots)$ and $g(ai999\dots) = g(aj000\dots)$, for any $0 \le i < 9$, $j = i + 1$, and any finite word a. We will need the following result from [7].

Lemma 2: *Let g be a compatible g-function on Σ_r. Then if ν is a g-measure, there is a C^1 expanding map $T\colon S^1 \to S^1$ preserving Lebesgue measure λ, such that (σ, Σ_r, ν) is measure-theoretically isomorphic to (T, S^1, λ).*

Proof: Let π denote the map from Σ_r to S^1 given by $x \mapsto \sum_{i=0}^{\infty} x_i r^{-(i+1)}$ (mod 1). This is a semiconjugacy from (σ, Σ_r) to (S^1, T_r), where $T_r(x) = rx$ (mod 1). The semiconjugacy is one-to-one off a countable set. Let ν be as defined in the statement of the lemma and let μ be a measure defined on S^1 by $\mu(A) = \nu(\pi^{-1}(A))$. Next, let $h\colon S^1 \to S^1$ be given by $h(x) = \mu([0, x])$. This is an orientation-preserving homeomorphism of the circle (using the properties of g-measures that they are non-atomic and fully supported). Let $\tilde{\mu}$ be the push-forward of μ: $\tilde{\mu}(A) = \mu(h^{-1}(A))$. Then we see $\tilde{\mu}([0,x]) = \nu(h^{-1}[0,x]) = \nu([0,h^{-1}(x)]) = x$. It follows that $\tilde{\mu}$ is in fact Lebesgue measure λ. Letting $T = h \circ T_r \circ h^{-1}$, we see the systems (T, S^1, λ) and (σ, Σ_r, ν) are measure-theoretically isomorphic by the map $h \circ \pi$. Now take $x \in S^1$ and $y \le x \le z$ with y near, but not equal to z. Then we have

$$\begin{aligned}\frac{T(z)-T(y)}{z-y} &= \frac{\lambda(T([y,z]))}{\lambda([y,z])} = \frac{\mu(T_r[h^{-1}y,h^{-1}z])}{\mu([h^{-1}y,h^{-1}z])}\\ &= \frac{\nu(\pi^{-1}(T_r[h^{-1}y,h^{-1}z]))}{\nu(\pi^{-1}([h^{-1}y,h^{-1}z]))} = \frac{\nu(\sigma(\pi^{-1}([h^{-1}y,h^{-1}z])))}{\nu(\pi^{-1}([h^{-1}y,h^{-1}z]))}.\end{aligned}$$

This can be seen to converge to $1/g(\pi^{-1}h^{-1}(x))$ as y and z converge to x, using the compatibility of g if x is a preimage of 0. It follows that T is a C^1 expanding map preserving Lebesgue measure as claimed. ∎

It will then be sufficient to construct an example of a compatible g-function having a g-measure which is ergodic but not weak-mixing.

We start with some preliminary definitions. As in [7], we introduce a partial order on Σ_{10}. First define $3 \preceq i \preceq 6$ for any $0 \le i \le 9$. Then $x \preceq y$ if $x_i \preceq y_i$ for all $i \in \mathbb{Z}^+$. A function f: $\Sigma_{10} \to \mathbb{R}$ is called **monotonic** if $f(x) \le f(y)$ whenever $x \preceq y$. We will say that a function f: $\Sigma_{10} \to \mathbb{R}$ is **precompatible** if $f(090909\ldots) = f(909090\ldots)$ and $f(ai090909\ldots) = f(aj909090\ldots)$, where a is any finite word, i is any symbol with $0 \le i < 9$ and $j = i+1$. We write this second condition as $f(b, 090909\ldots) = f(b+1, 909090\ldots)$ for any finite word b not ending in a 9.

We will need to consider the involutions on Σ_{10} given by

$$F(x)_n = \begin{cases} 9 - x_n & \text{if } n \text{ is odd,} \\ x_n & \text{if } n \text{ is even;} \end{cases}$$
$$R(x)_n = 9 - x_n.$$

Write $\bar{x}$ for $R(x)$, $\hat{x}$ for $F(x)$ and $\tilde{x}$ for $R \circ F(x)$. We say that a function f is **symmetric** if $f(\bar{x}) = f(x)$ for all x.

Write π for the map $\Sigma_{10} \to I$, defined by $x \mapsto \sum_{i=0}^{\infty} x_i 10^{-(i+1)}$. We will identify Σ_{10} with I and often omit reference to π, when applying functions on I to arguments in Σ_{10}.

2. Construction of the Example

To construct the example, we will use the following lemma.

LEMMA 3: *There exists a precompatible, compatible, symmetric, monotonic g-function g with the property that if one considers random variables (X_n) evolving as*

$$\mathbb{P}(X_n = i |\, X_{n-1} = a_1,\ X_{n-2} = a_2,\ \ldots) = h(i, a_1, a_2, \ldots),$$

conditioned upon $X_i = 6$, for all $i < 0$, then there exists a $\beta > \frac{1}{2}$ such that $\mathbb{P}(X_n = 6) \geq \beta$ for all n.

We will write $\mathbb{P}_6$ for the probability distribution on (X_n) defined in this way. The construction shown here differs from the construction in [7] only in the initial stages. The reader should note that that paper in turn is based on [3].

Proof: Define $\delta(x) = \chi_6(x) - \chi_3(x)$, where $\chi_i(x)$ is 1 if $x_0 = i$ and 0 otherwise. Then let $\Delta_m(x) = \sum_{i=0}^{m-1} \delta(\sigma^i(x))$. To construct h, we will need to define a collection of functions $W^i_{m,n}\colon \Sigma_{10} \to (0,1)$ indexed by $0 \leq i \leq 9$ and $m > n > 0$. These will be based on a family of functions $V_{m,n}$ whose existence is asserted by the following lemma.

LEMMA 4: *There exists a family $V_{m,n}$ (where $m > n > 0$) of compatible, precompatible, monotonic Hölder continuous functions satisfying*

$$0 \leq V_{m,n}(x) \leq 1,$$
$$V_{m,n}(x) = \begin{cases} 1 & \text{if } \Delta_m(x) > n, \\ 0 & \text{if } \Delta_m(x) < n. \end{cases}$$

The construction of the $V_{m,n}$ is rather involved and is (in the author's opinion) a distraction from the main flow of the paper. It has therefore been relegated to an appendix to the paper. Once the $V_{m,n}$ have been defined, the $W_{m,n}$ are defined as follows:

$$W^6_{m,n}(x) = \tfrac{1}{10} + \tfrac{1}{2} V_{m,n}(x),$$
$$W^3_{m,n}(x) = W^6_{m,n}(\bar{x}),$$
$$W^i_{m,n}(x) = \tfrac{1}{10} - \tfrac{1}{16}(V_{m,n}(x) + V_{m,n}(\bar{x})) \quad \text{for } i \neq 3, 6.$$

Note that for each x, $\sum_{i=0}^{9} W^i_{m,n}(x) = 1$ and since we require $n > 0$, we have that for each x, only one of $V_{m,n}(x)$ and $V_{m,n}(\bar{x})$ is positive. This implies that

$W^i_{m,n}(x)$ is bounded below by $\frac{3}{80}$ for $i \neq 3, 6$. Let $q_j = \frac{1}{2}(\frac{2}{3})^j$, so $\sum_{j=1}^{\infty} q_j = 1$. We will choose n_j and m_j such that taking $g(ix) = \sum_{j=1}^{\infty} q_j W^i_{m_j,n_j}(x)$ will give a compatible continuous g with more than one g-measure. The choice of n_j and m_j will be made inductively, by considering certain Hölder continuous truncations of the final g-function. Suppose $n_1, \ldots, n_{k-1}$ and $m_1, \ldots, m_{k-1}$ are chosen. Then define vectors as follows:

$$u = (\tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{5}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{1}{10}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}),$$
$$v = (\tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{1}{10}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{5}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}),$$
$$z = (\tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{7}{20}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{7}{20}, \tfrac{3}{80}, \tfrac{3}{80}, \tfrac{3}{80}),$$

with indices running from 0 to 9. Now define

$$g^1_k(ix) = \sum_{j=1}^{k-1} q_j W^i_{m_j,n_j}(x) + \sum_{j=k}^{\infty} q_j z_i,$$
$$g^2_k(ix) = \sum_{j=1}^{k-1} q_j W^i_{m_j,n_j}(x) + q_k u_i + \sum_{j=k+1}^{\infty} q_j v_i,$$
$$g^3_k(ix) = \sum_{j=1}^{k-1} q_j W^i_{m_j,n_j}(x) + q_k W^i_{M,N}(x) + \sum_{j=k+1}^{\infty} q_j v_i,$$

where $M > N > 0$. These are all Hölder continuous g-functions and as such have unique g-measures (see [8]), which we call μ^e_k where $e = 1, 2, 3$. First note that g^1_k is symmetric: $g^1_k(1 - x) = 1 - g^1_k(x)$. This means the unique invariant measure must be preserved under the involution $x \mapsto 1 - x$. It follows that $\mu^1_k[6] = \mu^1_k[3]$. We will use the order-preserving properties of g to show that $\mu^3_k([6]) \geq \mu^2_k([6]) > \mu^1_k([6])$ and $\mu^3_k([3]) \leq \mu^2_k([3]) < \mu^1_k([3])$. Let $\alpha_k = \frac{1}{16}(\frac{2}{3})^k$.

Lemma 5: *We have $\mu^2_k([6]) \geq \mu^1_k([6]) + 2\alpha_k$ and $\mu^2_k([3]) \leq \mu^1_k([3]) - 2\alpha_k$. Further, suppose we are given $x \in \Sigma_{10}$. Then there is a coupling of the two processes (Y_n) and (Z_n) evolving under g^2_k and g^3_k respectively, conditioned on $Y_i = Z_i = x_{-i}$, for all $i \leq 0$ such that $Y_n \preceq Z_n$ with probability 1 for all n.*

Proof: The proof works by finding couplings of two processes evolving under different g-functions, which make it obvious that the required inequalities hold.

It is easy to check that $g^2_k(6x) - g^1_k(6x) = 2\alpha_k$ and $g^2_k(3x) - g^1_k(3x) = -2\alpha_k$, while $g^2_k(ix) = g^1_k(ix)$ for all $i \neq 3, 6$ and all x.

We use this to give an explicit coupling of two random processes (X_n) and (Y_n) evolving under g_k^1 and g_k^2 respectively as in (1). We write $P(ix, jy)$ for the probability that i is added to x and j is added to y. The transition probability will only be defined when $x \preceq y$, and it must therefore have $\mathbb{P}(ix \preceq jy) = 1$ in order that it can be applied repeatedly. Suppose $x \preceq y$. Then define

$$P(ix, jy) = \begin{cases} g_k^1(6x) & \text{if } i = j = 6, \\ \max\big(0, g_k^1(ix) - g_k^2(iy)\big) & \text{if } i \neq 3, 6 \text{ and } j = 6, \\ \min\big(g_k^1(ix), g_k^2(iy)\big) & \text{if } i = j \neq 3, 6, \\ \max\big(0, g_k^2(iy) - g_k^1(ix)\big) & \text{if } i = 3 \text{ and } j \neq 3, 6, \\ \min\big(g_k^2(6y) - g_k^1(6x), g_k^1(3x) - g_k^2(3y)\big) & \text{if } i = 3 \text{ and } j = 6, \\ g_k^2(3y) & \text{if } i = j = 3. \end{cases}$$

We note that all the transition probabilities are non-negative, and we must just check that the marginals of this coupling are as claimed. We compute one example as an illustration. We will show that under P, the probability that x goes to $3x$ is $g_k^1(3x)$ as required. By observation, we see that the probability that x goes to $3x$ is

$$\begin{aligned} & g_k^2(3y) + \min\big(g_k^2(6y) - g_k^1(6x), g_k^1(3x) - g_k^2(3y)\big) + 8\max(0, g_k^2(iy) - g_k^1(ix)) \\ &= g_k^1(3x) + \min\big(g_k^2(6y) - g_k^1(6x) - g_k^1(3x) + g_k^2(3y), 0\big) \\ &\qquad + 8\max\big(0, g_k^2(iy) - g_k^1(ix)\big) \\ &= g_k^1(3x) + \min\big(8g_k^1(ix) - 8g_k^2(ix), 0\big) + 8\max\big(0, g_k^2(iy) - g_k^1(ix)\big) \\ &= g_k^1(3x) \end{aligned}$$

as required, where i is any symbol distinct from 3 and 6. This shows that given that $x \preceq y$, we can choose i and j such that y evolves according to g_k^2 and x according to g_k^1 such that with probability 1, $ix \preceq jy$. Looking further at the coupling, we see that the probability that y is preceded by a 6 and x is not preceded by a 6 given that $x \preceq y$ is $g_k^2(6y) - g_k^1(6x)$, but $g_k^2(6y) - g_k^1(6y) = 2\alpha_k$ and $g_k^1(6y) \geq g_k^1(6x)$, so it follows that with (x, y) goes to $(ix, 6y)$ for some $i \neq 6$ with probability at least $2\alpha_k$. It follows that $\mu_k^2([6]) \geq \mu_k^1([6]) + 2\alpha_k$. A similar argument shows that $\mu_k^2([3]) \leq \mu_k^1([3]) - 2\alpha_k$.

To prove the remaining parts of the Lemma, it is necessary to consider a coupling of processes (Y_n) evolving under g_k^2 and (Z_n) evolving under g_k^3. This is done by a coupling exactly similar to the coupling above, with g_k^2 replacing g_k^1

and g_k^3 replacing g_k^2. The conclusion then is that given that $y \preceq z$, then y can be allowed to evolve under g_k^2 and z under g_k^3 in such a way that the ordering is preserved. For a more formal and general discussion of couplings, the reader is referred to Lindvall's book [5]. ∎

We now describe the inductive choice of m_k and n_k. In each case, n_k is given by $\lfloor \alpha_k m_k \rfloor$. Suppose we have chosen $m_1, m_2, \ldots, m_{k-1}$ and hence $n_1, n_2, \ldots, n_{k-1}$. Let $\eta(x) = \chi_{[6]}(x) - \chi_{[3]}(x)$ and $\Delta_j(x) = \sum_{i=0}^{m_j-1} \eta(\sigma^i(x))$. Let $G_j = \{x\colon \Delta_j(x) \geq n_j\}$ and $H_j = \{x\colon \Delta_j(x) \geq 3n_j\}$. Then note that $\sigma^{-n_j}(H_j) \subseteq G_j$. Note also that if $x \in H_j$ and $y \succeq x$, then $y \in H_j$. Assume that there are $t_1, t_2, \ldots t_{k-1}$ such that

$$\mathbb{P}\big((X_{t_j}, X_{t_j-1}, \ldots) \in H_j \,|\, X_{-i} = x_i,\ \forall i \geq 0\big) \geq 1 - 4^{-j}, \tag{3}$$

for all $j < k$ and $x \in \Sigma_{10}$, where the X_n evolve according to g_j^3.

Let $A_m = \{x\colon \Delta_m(x) > 3\alpha_k m\}$. We know $\int \eta(x)\, d\mu_k^2(x) \geq 4\alpha_k$ and we will use this to show $\mu_k^2(A_m) \to 1$ as $m \to \infty$.

Lemma 6: *We have $\mu_k^2(A_m) \to 1$ as $m \to \infty$.*

Proof: Suppose the claim does not hold. Since we have $\mu_k^2(A_m) \leq 1$ for all m, the only way the claim can fail is if there exists an $\epsilon > 0$ and a sequence M_i such that $\mu_k^2(A_{M_i}) < 1 - \epsilon$ for all i. In this case, we have

$$\mu_k^2\left(\bigcup_{i>j} A_i{}^c\right) > \epsilon \text{ for all } j, \qquad \text{so } \mu_k^2\left(\bigcap_j \bigcup_{i>j} A_i{}^c\right) \geq \epsilon.$$

Let S be $\bigcap_j \bigcup_{i>j} A_i{}^c$. If $x \in S$ then

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \eta(\sigma^i(x)) \leq 3\alpha_k.$$

We have however that μ_k^2 is ergodic, so for almost all x (with respect to μ_k^2), we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \eta(\sigma^i(x)) \geq 4\alpha_k.$$

This is a contradiction. ∎

Next, pick m_k such that $\mu_k^2(A_{m_k}) > 1 - 4^{-k}$ and $\alpha_k m_k > t_{k-1}$. Now $H_k = A_{m_k}$. Since g_k^2 is Hölder continuous, we can apply Walters' theorem [8] to get that

${\mathcal{L}_{g_k^2}}^n \chi_{H_k}(x)$ converges uniformly to $\mu_k^2(H_k)$, which is greater than $1-4^{-k}$. It follows that there exists a t_k such that ${\mathcal{L}_{g_k^2}}^{t_k} \chi_{H_k}(x) \geq 1-4^{-k}$, for all $x \in \Sigma_{10}$. This says that for all $x \in \Sigma_{10}$,

$$\mathbb{P}\big((X_{t_k}, X_{t_k-1}, \dots) \in H_k | \, X_{-i} = x_i, \; \forall i \geq 0\big) \geq 1 - 4^{-k},$$

where the X_n evolve according to g_k^2, but by the second statement of Lemma 5, it follows that the same equation holds when the evolution is according to g_k^3. This is precisely the statement of (3) when we take j to be k. This completes the inductive step.

To complete the inductive construction of the example, it remains only to specify an initial case for the induction. Taking $t_0 > 1$, applying the above induction step produces m_1, n_1 and t_1 which can be used as a starting point for the induction.

In the above section, the m_i and n_i were inductively constructed, so the g-function is now given by

$$g(ix) = \sum_{j=1}^{\infty} q_j W^i_{m_j, n_j}(x).$$

This is clearly compatible, precompatible, monotonic and continuous.

We consider the events E_k^t that $(X_t, X_{t-1}, \dots) \in H_k$. Write $\mathbb{P}_6$ for the probability distribution of the X_n conditioned on $X_i = 6$ for all $i < 0$ with subsequent evolution under g. Informally, E_k^t is the event that the process has a 'large majority of 6s over 3s at the m_k scale at time t'. We then consider letting the process evolve from an initial condition of all 6s (so $\mathbb{P}_6(E_k^0) = 1, \; \forall k$). We show inductively that the events E_k^t have a high probability by induction on t, using the result of (3), which says that if the process has a large majority of 6s on scales $m_{k+1}, m_{k+2}, \dots$ at time $t - t_k$, then with high probability, the process will have a majority of 6s on scale m_k at time t.

LEMMA 7: *We have*

$$\mathbb{P}_6(E_k^t) \geq 1 - \zeta_k, \tag{4}$$

for all $t \in \mathbb{Z}$ and $k \in \mathbb{N}$, where $\zeta_k = \frac{3}{2}4^{-k}$.

Proof: The proof is by induction on t. Note that the hypothesis is automatically true for all k if $t < 0$, so we need only prove the inductive step. Suppose

(4) holds for all $t < s$ then pick $k \in \mathbb{N}$. Let $S = \bigcap_{j>k} E_j^{s-t_k}$. Then by the induction hypothesis, $\mathbb{P}_6(S^c) \le \sum_{j>k} \zeta_j = \frac{1}{2}4^{-k}$. Now we decompose E_k^{sc} as $(E_k^{sc} \cap S) \cup (E_k^{sc} \cap S^c)$. We then have

$$\mathbb{P}_6(E_k^{sc}) \le \mathbb{P}_6(E_k^{sc} \cap S) + \mathbb{P}_6(S^c) \le \mathbb{P}_6(E_k^{sc}|S) + \tfrac{1}{2}4^{-k}.$$

But now suppose $\omega \in S$. Then let

$$x = (X_{s-t_k}, X_{s-t_k-1}, \dots) \quad \text{and} \quad z = (X_s, X_{s-1}, \dots).$$

Then $x \in \bigcap_{j>k} H_j$. It follows that if $y \in \sigma^{-t}(x)$, for some $t \le t_k$ then $y \in \bigcap_{j>k} G_j$. In particular, $g(y) = g_k^3(y)$, where the M and N in g_k^3 are taken to be m_k and n_k. It follows that the evolution of x for t_k steps takes place under g_k^3, but by (3), the probability that $z \in E_k^{sc}$ is no more than 4^{-k}. In particular, we have shown that $\mathbb{P}_6(E_k^{sc}) \le \zeta_k$ as required. This completes the proof of the inductive step and hence of the lemma. ∎

We apply this by calculating $\mathbb{P}_6(X_n = 6)$. Using the Lemma above, this is bounded below by $\sum_{j\ge1} q_j\big((1-\zeta_j)\frac{3}{5} + \zeta_j\frac{1}{10}\big)$. This turns out to be equal to $\frac{21}{40}$. Let $\mu_n = \rho_n^*(\mathbb{P}_6)$, as defined in §1. Then we have

$$\mu_{n+1}([ix]^{m+1}) = \int_{[x]^m} g(iy)\, d\mu_n(y).$$

Now let $\nu_n = \frac{1}{n}\sum_{j=0}^{n-1} \mu_j$. Then we see

$$\left| \nu_n([ix]^{m+1}) - \int_{[x]^m} g(iy)\, d\nu_n(y) \right| \le \frac{2}{n}.$$

Taking a weak*-convergent subsequence $\nu_{n_k} \to \nu$ of the ν_n, we find

$$\nu([ix]^{m+1}) = \int_{[x]^m} g(iy)\, d\nu(y).$$

As noted in §1, this implies that ν is a g-measure. However $\mu_n([6]) \ge \frac{21}{40}$, for all n, so it follows that $\nu([6]) \ge \frac{21}{40}$. This completes the proof of Lemma 3. ∎

3. Proof of Theorem 1

In this section, we use the results of §2 to prove Theorem 1, subject to the construction of $V_{m,n}$ in the appendix.

Proof of Theorem 1: Let g and $\mathbb{P}_6$ be as defined in the previous section. Take $\mu_n = \rho_n^*(\mathbb{P}_6)$ and form Cesàro sums $\nu_n = \frac{1}{n}\sum_{i=0}^{n-1}\mu_i$. Then we see (as in [7]) that if ν_{n_i} is a weak*-convergent subsequence, converging to a measure ν, then ν is a g-measure. We see also that $\nu([6])$, the measure of those members of Σ_{10} starting with a 6 is at least β. We may assume ν is ergodic, for otherwise, by ergodic decompostion, there is another g-measure with this property. If ν is not ergodic with respect to σ^2, then one can check that there exist sets A and B of measure $\frac{1}{2}$ such that $\sigma^{-1}(A) = B$ and $\sigma^{-1}(B) = A$. It then follows quickly that ν is ergodic but not weak-mixing and by Lemma 2 and the compatibility of g, Theorem 1 follows. It remains to consider the case where ν is ergodic with respect to σ^2. We note that the involution F defined above is not shift-commuting, but that F does commute with σ^2. Define a new measure μ by $\mu(A) = \frac{1}{2}\nu(\hat{A}) + \frac{1}{2}\nu(\tilde{A})$. This is shift-invariant. Now we have

$$\frac{\mu([ix]^{n+1})}{\mu([x]^n)} = \frac{\frac{1}{2}\nu([\widehat{ix}]^{n+1}) + \frac{1}{2}\nu([\widetilde{ix}]^{n+1})}{\frac{1}{2}\nu([\hat{x}]^n) + \frac{1}{2}\nu([\tilde{x}]^n)} = \frac{\nu([i\tilde{x}]^{n+1}) + \nu([\bar{i}\hat{x}]^{n+1})}{\nu([\tilde{x}]^n) + \nu([\hat{x}]^n)}.$$

Then using the symmetry of g, we see $g(i\tilde{x}) = g(\bar{i}\hat{x})$, so we get

$$\lim_{n\to\infty}\frac{\mu([ix]^{n+1})}{\mu([x]^n)} = g(i\tilde{x}) = g\circ F(ix).$$

It follows that μ is an h-measure, where $h = g\circ F$. Note that by the precompatibility of g, h is compatible. It remains to show that μ is ergodic but not weak-mixing. Suppose for a contradiction that $\sigma^{-1}(A) = A$ and $0 < \mu(A) < 1$. Then $\mu(A) = \frac{1}{2}\nu(\hat{A}) + \frac{1}{2}\nu(\tilde{A})$, but $\sigma^{-1}(\tilde{A}) = \hat{A}$ and $\sigma^{-1}(\hat{A}) = \tilde{A}$. It follows that $\nu(\hat{A}) = \nu(\tilde{A})$, so $0 < \nu(\hat{A}) < 1$. But this is a contradiction as $\sigma^{-2}(\hat{A}) = \hat{A}$ and ν is assumed to be ergodic with respect to σ^2, proving that μ is ergodic.

Next, note that μ is not ergodic with respect to σ^2 as $\mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$, where μ_1 and μ_2 are σ^2-invariant measures defined by $\mu_1(A) = \nu(\hat{A})$ and $\mu_2(A) = \nu(\tilde{A})$. These are not equal as $\mu_1([6]) > \frac{1}{2} > \mu_2([6])$. It follows that μ is not weak-mixing, thus completing the proof of Theorem 1 subject to the proof of Lemma 4 in the appendix. ∎

Appendix. Construction of $V_{m,n}$

Proof of Lemma 4: In this appendix, we give the construction of the function $V_{m,n}$, which was introduced in §2. First we define a contraction map $\mathcal{L}$ on the subspace X of $(C[0,1])^4$ with the metric induced by the uniform norm:

$$X = \{(f_1, f_2, f_3, f_4)\colon f_i\colon [0,1] \to [0,1];\ f_1(0) = f_3(0) = 0,\ f_1(1) = f_3(1) = 1,\\ f_2(0) = f_4(0) = 1,\ f_2(1) = f_4(1) = 0\}.$$

We will identify I with Σ_{10} so σ^2 will denote the map $x \mapsto 100x \bmod 1$. The map $\mathcal{L}$ is defined by $\mathcal{L}(f_1, f_2, f_3, f_4) = (g_1, g_2, g_3, g_4)$, where

$$g_1(x) = \begin{cases} 0 & 0 \le x < .04 \\ \frac{1}{2} f_1(\sigma^2(x)) & .04 \le x < .05 \\ \frac{1}{2} + \frac{1}{2} f_1(\sigma^2(x)) & .05 \le x < .06 \\ 1 & .06 \le x < .09 \\ \frac{1}{2} + \frac{1}{2} f_4(\sigma^2(x)) & .09 \le x < .10 \\ \frac{1}{2} - \frac{1}{2} f_4(1 - \sigma^2(x)) & .10 \le x < .11 \\ 0 & .11 \le x < .15 \\ \frac{1}{2} f_1(\sigma^2(x)) & .15 \le x < .16 \\ \frac{1}{2} & .16 \le x < .17 \\ \frac{1}{2} f_2(\sigma^2(x)) & .17 \le x < .18 \\ 0 & .18 \le x < .40 \\ \frac{1}{2} - \frac{1}{2} f_3(1 - \sigma^2(x)) & .40 \le x < .41 \\ \frac{1}{2} f_2(\sigma^2(x)) & .41 \le x < .42 \\ 0 & .42 \le x < .45 \\ \frac{1}{2} f_1(\sigma^2(x)) & .45 \le x < .46 \\ \frac{1}{2} & .46 \le x < .47 \\ \frac{1}{2} f_2(\sigma^2(x)) & .47 \le x < .48 \\ 0 & .48 \le x < .49 \\ \frac{1}{2} f_3(\sigma^2(x)) & .49 \le x < .50 \\ 1 - g_1(1 - x) & .50 \le x \le 1, \end{cases}$$

$$g_2(x) = \begin{cases} 1 - \frac{1}{2}f_4(1-\sigma^2(x)) & 0 \le x \le .01 \\ \frac{1}{2}f_2(\sigma^2(x)) & .01 \le x \le .02 \\ 0 & .02 \le x \le .04 \\ \frac{1}{2}f_1(\sigma^2(x)) & .04 \le x < .05 \\ \frac{1}{2} + \frac{1}{2}f_1(\sigma^2(x)) & .05 \le x < .06 \\ 1 & .06 \le x < .07 \\ \frac{1}{2} + \frac{1}{2}f_2(\sigma^2(x)) & .07 \le x < .08 \\ \frac{1}{2}f_2(\sigma^2(x)) & .08 \le x < .09 \\ 0 & .09 \le x < .15 \\ \frac{1}{2}f_1(\sigma^2(x)) & .15 \le x < .16 \\ \frac{1}{2} & .16 \le x < .19 \\ \frac{1}{2}f_4(\sigma^2(x)) & .19 \le x < .20 \\ g_1(x) & .20 \le x \le .80 \\ 1 - g_2(1-x) & .80 \le x \le 1, \end{cases}$$

$$g_3(x) = \begin{cases} g_1(x) & 0 \le x \le .07 \\ g_2(x) & .07 \le x \le .15 \\ 0 & .15 \le x \le .2 \\ g_1(x) & .2 \le x \le 1, \end{cases}$$

$$g_4(x) = \begin{cases} g_2(x) & 0 \le x \le .07 \\ g_1(x) & .07 \le x \le .15 \\ g_2(x) & .15 \le x \le 1. \end{cases}$$

It is then straightforward to check that $\mathcal{L}$ is indeed a contraction map from X to X, and it follows that there is a unique fixed point, $e = (e_1, e_2, e_3, e_4)$. Using the fact that these form a fixed point of $\mathcal{L}$, it is straightforward to check that if x and y agree for $2n$ digits, then the difference between $e_i(x)$ and $e_i(y)$ is at most 2^{-n}. It follows that the functions e_i are Hölder continuous when considered as functions $\Sigma_{10} \to [0,1]$. Since the functions are continuous as maps $[0,1] \to [0,1]$, it follows that considered as functions $\Sigma_{10} \to [0,1]$, they are compatible.

Next, suppose that $x \prec y$ and x and y differ in either the zeroth or first place. Then it is easy to see that $e_i(x) \le e_i(y)$ for each i just by examining the condition that e is a fixed point of $\mathcal{L}$. Then one checks that $x \prec y$ implies $e_i(x) \le e_i(y)$ for each i by induction on the first place in which they differ. It follows that the functions e_i are monotonic.

We also need to check the precompatibility of the functions e_i. First note the following table of values of the functions e_i. For later use, we include also two additional functions e_5 and e_6 defined by $e_5(x) = 1 - e_3(1-x)$ and $e_6(x) = 1 - e_4(1-x)$.

(5)

	0	.0909...	.9090...	0
e_1	0	1	0	1
e_2	1	0	1	0
e_3	0	0	0	1
e_4	1	1	1	0
e_5	0	1	1	1
e_6	1	0	0	0

It is then a routine matter to check that $e_i(a0909\ldots) = e_i(a+1, 9090\ldots)$ for each $i \le 4$, where a is any word of length 1 or 2 whose last digit is not a 9. Then by induction on the length of the word, as before, we see that the e_i are precompatible for each $i \le 4$.

We have therefore checked that the e_i $(1 \le i \le 4)$ are monotonic, compatible, precompatible, Hölder continuous and take values as shown in (5). One can check that e_5 and e_6 also have these properties. Further the functions e_i are all equal on the range $0.2 \le x \le 0.8$. This implies that forming f_{ij} defined by

$$f_{ij}(x) = \begin{cases} e_i(x) & x \le .5 \\ e_j(x) & x \ge .5 \end{cases}$$

for $3 \le i, j \le 6$ gives 16 functions, each of which is monotonic, compatible, precompatible and Hölder continuous. Looking at (5), we see that these functions take all combinations of values of 0 and 1 on the set $\{0, .0909\ldots, .9090\ldots, 1\}$. We label the functions according to their values on each of these four points as $d_{i_1 i_2 i_3 i_4}$ so for example d_{0110} takes values 0,1,1 and 0 at $0, .0909\ldots, .9090\ldots$ and 1 respectively, so $d_{0110} = f_{54}$.

To define $V_{m,n}$, we also need to define two further maps defined on words of $S_m = \{0, \ldots, 9\}^m$. We have already made implicit use of the equivalence relation $\sim$ generated by $a0909\ldots \sim a+1, 9090\ldots$, for any word a not ending with a 9 when discussing precompatibility. Given a word $a \in S_m$, define $\phi(a)$ by the requirement that $a0909\ldots \sim \phi(a)9090\ldots$ and $\psi(a)$ by the requirement that $a9090\ldots \sim \psi(a)0909\ldots$. We are now in a position to specify $V_{m,n}$. This is defined cylinder by cylinder. If $a \in S_m$, write $[a]$ for those elements of Σ_{10} whose first m digits are given by a. Define κ: $S_m \to \{0,1\}$ by $\kappa(b) = 1$ if

$\Delta_m(b) > n$ and $\kappa(b) = 0$ otherwise. By $a+1$, we mean the word obtained by adding 1 (with carry if necessary). The word $a-1$ is defined similarly, so for example, $99999 + 1 = 00000$ and $88900 - 1 = 88899$. Then given $a \in S_m$, define $N(a) = \kappa(a-1), \kappa(\phi(a)), \kappa(\psi(a)), \kappa(a+1)$. Note that $|\Delta_m(a) - \Delta_m(a+1)|$, $|\Delta_m(a) - \Delta_m(a-1)|$, $|\Delta_m(a) - \Delta_m(\phi(a))|$ and $|\Delta_m(a) - \Delta_m(\psi(a))|$ are all bounded above by 1. Given this, we set

$$V_{m,n}\big|_{[a]}(x) = \begin{cases} 1 & \Delta_m(a) > n, \\ 0 & \Delta_m(a) < n, \\ d_{N(a)}(\sigma^m(x)) & \Delta_m(a) = n. \end{cases}$$

The function $V_{m,n}$ defined in this way is then seen to be Hölder continuous, monotonic, compatible and precompatible. Further, it satisfies $0 \le V_{m,n} \le 1$, $V_{m,n}(x) = 1$ when $\Delta_m(x) > n$ and $V_{m,n}(x) = 0$ when $\Delta_m(x) < n$ as required. This completes the construction and hence the proof of Lemma 4 and Theorem 1. ∎

ACKNOWLEDGEMENT: The author would like to extend his thanks to Chris Bose who suggested the problem. It would be interesting to see if the other examples in [2] could also be made to be C^1, rather than just continuous.

References

[1] C. J. Bose, *Generalized baker's transformations*, Ergodic Theory and Dynamical Systems **9** (1989), 1–17.

[2] C. J. Bose, *Mixing examples in the class of piecewise monotone and continuous maps of the unit interval*, Israel Journal of Mathematics **83** (1993), 129–152.

[3] M. Bramson and S. A. Kalikow, *Nonuniqueness in g-functions*, Israel Journal of Mathematics **84** (1993), 153–160.

[4] M. Keane, *Strongly mixing g-measures*, Inventiones mathematicae **16** (1972), 309–324.

[5] T. Lindvall, *Lectures on the Coupling Method*, Wiley, New York, 1992.

[6] A. N. Quas, *Some problems in ergodic theory*, University of Warwick Thesis, 1993.

[7] A. N. Quas, *Non-ergodicity for C^1 expanding maps and g-measures*, Ergodic Theory and Dynamical Systems (to appear).

[8] P. Walters, *Ruelle's operator theorem and g-measures*, Transactions of the American Mathematical Society **214** (1975), 375–387.