ISRAEL JOURNAL OF MATHEMATICS, Vol. 35, No. 4, 1980

A NON-BERNOULLI SKEW PRODUCT WHICH IS LOOSELY BERNOULLI

BY
ROBERT M. BURTON, JR.

ABSTRACT

Let $T: Y \to Y$ be the Bernoulli two shift with independent generator $Q = \{Q_0, Q_1\}$ and let $S: X \to X$ be a measure preserving bijection. If (S, X) is ergodic then the skew product on $X \times Y$ defined by

$$\hat{S}(x, y) = \begin{cases} (x, Ty) & \text{if } y \in Q_0, \\ (Sx, Ty) & \text{if } y \in Q_1, \end{cases}$$

is a K-automorphism. If $\hat{S}$ is also Bernoulli we say S is pre-Bernoulli. J. Feldman showed that if S is pre-Bernoulli then S must be loosely Bernoulli. We construct an example to show the converse is false, i.e. an S that is loosely Bernoulli but not pre-Bernoulli.

1. Introduction and notations

For our purposes a *dynamical system* is a pair (S, X) where X is a Lebesgue probability space and $S: X \to X$ is a measurable, measure-preserving bijection. A *process* (S, P, X) is a dynamical system together with a finite measurable partition, P on X (usually we will just write (S, P) if the space X is understood). If $x \in X$ then the P-name of x is the sequence $\{x(n)\}$ where $x(n) = P_j$ if $S^{-n}x \in P_j \in P$ for $n \in \mathbf{Z}$. The *future P-name* of x is the one sided sequence $\{x(n)\}_{n=0}^{\infty}$. If $M \geqq N$ we call $\{x(n)\}_{n=N}^{M}$ a *string* and denote it by X_N^M. A process (S, P, X) is called *Bernoulli* if (S, X) is a Bernoulli shift. (T, Q, Y) will always denote the Bernoulli two shift so $Q = \{Q_0, Q_1\}$ is an independent generating partition with $m(Q_0) = m(Q_1) = \frac{1}{2}$. (Note: we will always use m to denote the measure regardless of which space we are working on.)

Let (S, P, X) be an ergodic process with $P = \{P_2, \cdots, P_k\}$. We define a skew product as $X \times Y$ which we will consider for the remainder of this paper. Let

Received March 1, 1979

$$\hat{S}(x,y)=\begin{cases}(x,Ty) & \text{if } y\in Q_0,\\ (Sx,Ty) & \text{if } y\in Q_1.\end{cases} \tag{1}$$

We put a partition $\hat{P}=\{\hat{P}_0,\hat{P}_2,\cdots,\hat{P}_k\}$ on $X\times Y$ by setting $\hat{P}_0=X\times Q_0$ and $\hat{P}_i=P_i\times Q_1$ for $i=2,\cdots,k$. Let $\hat{P}_1=\hat{P}_2\cup\cdots\cup\hat{P}_k$ and set $\hat{Q}=\{\hat{P}_0,\hat{P}_1\}$. Note that the $(\hat{S},\hat{Q})$ name of $(x,y)\in X\times Y$ coincides with the (S,Q) name of y.

Thus the $\hat{P}$-name of (x,y) is formed by taking the Q-name of y and replacing Q_1 occurrences in order by the P-name of x.

Noιe that if P is a generator of the σ-algebra on X then $\hat{P}$ is a generator of the σ-algebra on $X\times Y$. This is clear because if P iterates separate points on X then $\hat{P}$ iterates separate points on $X\times Y$.

In this paper we are concerned with the relationship between $(\hat{S},\hat{P})$ and (S,P) for ergodic (S,P). It follows from a result of Meilijson [4] that $(\hat{S},\hat{P})$ is always a K process. In particular we are concerned with properties of processes (S,P) such that $(\hat{S},\hat{P})$ is Bernoulli. Such (S,P) we call *pre-Bernoulli.*

In 1976 J. Feldman [1] introduced the concept of a process being loosely Bernoulli. This may be defined as follows. Let (S,P,X) be a process. Two points $x,y\in X$ and said to be $\bar{f}$-equivalent if there are subsequences $\{n_i\}$, $\{m_i\}$, of N for $i=0,1,2,\cdots$ with density one so that $x(n_i)=y(m_i)$ for all i. In practice it is usually easier to verify that the condition holds for subsequences of density $1-\varepsilon$ with ε arbitrarily small.

(S,P) is said to be *loosely Bernoulli* if for almost all atoms $\mathcal{P}_1,\mathcal{P}_2\in\bigvee_{-1}^{-\infty}S^iP$ of the past there is a bijection $\phi:\mathcal{P}_1\to\mathcal{P}_2$ that is measurable, preserves the normalized fiber measure and for almost all $x\in P_1$, $\phi(x)$ and x are $\bar{f}$-equivalent.

If the entropy of (S,P) is zero then the definition of loosely Bernoulli reduces to the condition that for all $x,y\in X$ belonging to a set of full measure are $\bar{f}$-equivalent.

One of the reasons Feldman introduced the notion of loosely Bernoulli was to show not all ergodic processes are pre-Bernoulli. He proved that all pre-Bernoulli processes are loosely Bernoulli and exhibited a non-loosely Bernoulli process, [1]. Feldman asked whether all loosely Bernoulli processes are pre-Bernoulli. The main result of this paper shows this is untrue.

2. Rank one processes

A *Rochlin tower* of a dynamical system (S,X) is a set $G=\{A_i\}_{i=1}^n$ where the $A_i\subseteq X$ are measurable and where $SA_i=A_{i+1}$ for $1\leqq i<n$. A well known theorem says any ergodic dynamical system has Rochlin towers with $m(\bigcup_{i=1}^n A_i)$

arbitrarily close to 1. We say a sequence $G_n = \{A_{i,n}\}_{i=1}^{h(n)}$ of Rochlin towers is *increasing* if $n < m$ implies $h(n) < h(m)$, $\bigcup_{i=1}^{h(n)} A_{i,n} \subseteq \bigcup_{j=1}^{n(m)} A_{j,m}$ and if $x \in A_{1,n}$ then $x \in A_{j,m}$ for some $j, 1 \leqq j \leqq h(m) - h(n) + 1$. We say a process (S, P, X) is of *rank one* if (S, X) is an increasing limit of Rochlin towers $G_n = \{A_{1,n}\}_{i=1}^{h(n)}$ of height $h(n)$ for each $n \geqq 1$ so $SA_{i,n} = A_{i+1,n}$ if $i < H(n)$ and so that for each i, n there is a $P_j \in P$ so $A_{i,n} \subseteq P_j$. Thus each tower G_n has a unique P-name which we call an n-block of (S, P). The above definition tells us that almost every $x \in X$ has a P-name composed of nested n-blocks (i.e., each n-block is a substring of an $(n+1)$-block). So a rank one process is well-defined if we prescribe how to build $(n+1)$-blocks out of n-blocks and check to see that we have a finite measure space. (For exmaple, we could use a standard cutting and stacking construction — see [7].)

It is well known that rank one processes are ergodic [3]. It is also easy to see rank one processes have zero entropy and are loosely Bernoulli [8].

We want to extend the definition of rank one to the positive entropy case. There is a standard and natural way a rank one process may be used to construct a K-process of arbitrary entropy while preserving much of the block structure of the original process.

Let (S, P) be a rank one process and P_f a symbol not occurring in $P = \{P_1, \cdots, P_k\}$. Let $f(n)$ for $n \geqq 1$ be a non-negative integer valued function with $\limsup f(n) = +\infty$ and let μ_n be a probability distribution on $\{1, \cdots, f(n) - 1\}$ if $f(n) \geqq 1$, (S^*, P^*) will have a nested block structure similar to that of (S, P) but there will be more than just one n-block for large enough n.

We describe the distribution of n-blocks inductively. For $j \geqq 1$ let $\boldsymbol{b}_j$ be a string of j consecutive P_f symbols. Then if $\boldsymbol{a}$ is a 1-block of (S, P), 1-blocks of (S^*, P^*) are of the form $\boldsymbol{b}_j * \boldsymbol{a} * \boldsymbol{b}_{f(1)-j}$ where $*$ is the concatenation operator and j is a random integer $1 \leqq j \leqq f(1) - 1$ distributed as μ_1. Note that all 1-blocks have the same length. If $n > 1$ and $\boldsymbol{a}_1 * \boldsymbol{c} * \boldsymbol{a}_2 * \cdots * \boldsymbol{c} * \boldsymbol{a}_{L+1}$ is the name of an (X, P) n-block and $\boldsymbol{c}$ the name of $(n-1)$-subblocks then an (S^*, P^*) n-block has form

$$\boldsymbol{b}_j * \boldsymbol{a}_1 * \boldsymbol{c}_1 * \boldsymbol{a}_2 * \boldsymbol{c}_2 * \cdots * \boldsymbol{c}_L * \boldsymbol{a}_{L+1} * \boldsymbol{b}_{f(n)-j}$$

where j is a random integer distributed as μ_n and the $\boldsymbol{c}_i$ are (S^*, P^*) $(n-1)$-blocks chosen independently.

It is clear that we may choose $f(n)$ so that we get a finite measure space, hence a well-defined process (S^*, P^*). Further by using more 'randomizing' symbols $P_{f_1}, \cdots, P_{f_m}$ and varying the distributions μ_n we can get exponentially as many names as we like so we can prescribe and entropy for (S^*, P^*).

The most well-known example of such a process is the K-automorphism that is not Bernoulli defined by Ornstein and Shields [5, 6]. Their proof that this example is a K-process also shows the following general

PROPOSITION 1. *(S^*, P^*) defined as above is K-process for any rank one (S, P) and any choice of μ_n, $f(n)$ so $\mu_n(j) > 0$ for $j = 1, \cdots, f(n) - 1$ and that give a finite measure to the system (S^*, P^*).*

We have also

PROPOSITION 2. *(S^*, P^*) is loosely Bernoulli.*

PROOF. The proof is similar to the zero entropy case: strings of future distributions (conditioned on past atoms) can be $\bar{f}$ matched by deleting the space between n-blocks, shifting one string so n-blocks have the same initial indices (all n-blocks have the same length) and matching n-blocks independently. ■

We will say any process of one of the above forms is of *rank one*.

We may now state our theorem.

THEOREM. *For any finite $\alpha \geqq 0$ there is a rank one process, (S, P) of entropy α that is not pre-Bernoulli, i.e., $(\hat{S}, \hat{P})$ is not Bernoulli.*

We will prove the Theorem by constructing an example for the zero entropy case first. Then we will show increasing the entropy by adding randomizing symbols as above will not affect the proof.

3. Constructing an example

Let $N > 10^6$ be a large enough integer where the meaning of 'large enough' will be made clear as we go along. We define a rank one process by describing n-blocks for $n \geqq N$. Let

$$h(N-1) = 2^{(N-1)^3/3} \qquad \text{and} \qquad h(N) = 2^{N^3/3 + N^2/2 + N/6} = \prod_{k=1}^{N} 2^{k^2}$$

and set $h(n) = 2^{n^2}[h(n-1) + h(n-2)]$ for $n > N$. $h(n)$ will denote the length of an n-block in our construction. Since $h(n-1) > h(n-2)$ we have

$$(2) \qquad 2^{n^2} h(n-1) < h(n) < 2^{n^2+1} h(n-1).$$

So

$$(3) \qquad 2^{n^3/3} < h(n) < 2^{n^3/3 + P(n)}, \qquad n \geqq N$$

where $P(n)$ is a degree two polynomial that is always nonnegative for $n \geqq N$.

Let $s_n = \{h(n)^{7/8}\}$ where $\{x\}$ is the least integer greater than x. We assume N is large enough so that

$$2^{2n^2} s_n < h(n-2) \tag{4}$$

and

$$(1/8)s_{n-1} > s_{n-2} > h(n)^{3/4}. \tag{5}$$

This is possible by (3).

We define an N-block to be a string of length $h(N)$ of P_2-symbols. For $n > N$ each n-block is composed of 2^{n^2} n-boxes. Each n-box is a string of symbols $h(n-1)+h(n-2)$ long. The kth n-box in an n-block beginning with $k^2 s_n$ P_3-symbols, followed by an $(n-1)$-block, followed by $h(n-2)-k^2 s_n$ P_s-symbols. This is well defined by (4). Note that between the kth $(n-1)$-block and the $(k+1)$st $(n-1)$-block are $(2k+1)s_n + h(n-2)$ P_3-symbols. Figure 1 shows a schematic diagram of an n-block.

PROPOSITION 3. *There is a unique rank one process (S, P, X) with n-blocks as defined above.*

PROOF. The usual cutting and stacking of P-labeled Rochlin towers will give us a transformation S acting on a measure space X. We need only show $m(X) < \infty$. Let $g_n = m\{x \mid x \in n\text{-block}\}$ for $n \geqq N$, normalized so $g_N = 1$. By looking at the corresponding Rochlin towers we see

$$g_{n+1} = g_n[1 + h(n-1)/h(n)] \qquad \text{for } n > N.$$

Continuing we see (since $g_N = 1$) $g_n = \Pi_{k=N}^{n}[1 + h(k-1/h(k)]$. So

$$m(X) = \lim_{n\to\infty} g_n = \prod_{k=N}^{\infty}\left[1 + \frac{h(k-1)}{h(k)}\right]. \tag{6}$$

Now by (2) $2^{-k^2} > h(k-1)/h(k)$ so (6) converges to a finite value and we can renormalize so $m(X) = 1$. ■

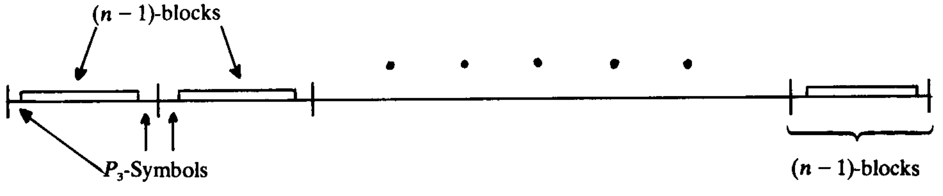

Fig. 1. Constructing an n-block.

By construction (S, P) is rank one so it is ergodic, zero entropy, and loosely Bernoulli. Before showing (S, P) is not pre-Bernoulli we need a proposition about the spacing of n-blocks in P-names.

Let $\boldsymbol{a}$ be an n-block in the P-name of $x \in X$. The [initial] *index* of a is the least i such that $T^{-i}x \in \boldsymbol{a}$. Two n-blocks $\boldsymbol{a}$, $\boldsymbol{b}$ in the names of $x, y \in X$ have distance k if their indices differ by k. If $\boldsymbol{a}'$ is the kth $(n-1)$-subblock of $\boldsymbol{a}$ we say k is the *order* of $\boldsymbol{a}'$.

PROPOSITION 4. *Let $\boldsymbol{a}$, $\boldsymbol{b}$ be n-blocks in the names of $x, y \in X$. If $\boldsymbol{a}$, $\boldsymbol{b}$ have distance greater than s_n then at most one $(n-1)$-subblock of $\boldsymbol{a}$ is closer than s_n to an $(n-1)$-subblock of $\boldsymbol{b}$.*

PROOF. Suppose two $(n-1)$-subblocks of $\boldsymbol{a}$, $\boldsymbol{b}$ are closer than s_n. These subblocks must have different order in $\boldsymbol{a}$, $\boldsymbol{b}$ respectively. From then on overlapping $(n-1)$-subblocks have indices differing by at least $2s_n$ and the maximum shift is $2^{2n^2}s_n < h(n-1)$ so these corresponding subblocks continue to overlap and have distance greater than s_n. ■

4. Proof of Theorem

Here we prove (S, P) is not pre-Bernoulli. The obvious $\bar{f}$ matches between two future names of (S, P) would arise from deleting the indices between n-blocks for some large n and make a shift in one name so n-blocks are lined up. In view of Proposition 4 this would require adjustments of more than $s_n > h(n)^{3/4}$. The law of the iterated logarithm says we cannot, using an independent process, make adjustments this large very often. The proof will work by showing all $\bar{f}$ matches between almost all P-names must make too many adjustments of this sort.

We need a series of lemmas. The first two are facts from probability.

LEMMA 1. *Let X_n, X'_n for $n = 1, 2, \cdots$ be independent, binary, equi-probable random variables with values in $\{0, 1\}$. Let $S_n = X_1 + \cdots + X_n$ and $S'_n = X'_1 + \cdots + X'_n$ and $d_n = |n - 2S_n|$ = the absolute value of the difference between the number of occurrences of 0's and 1's. Then for $\varepsilon = 1/N^2$ small enough we can remove ε of the space so that on the remaining part we have for $n \geqq N$*

$$d_n < (S_n)^{3/4}, \tag{7}$$

$$|S_n - S'_n| < (S_n)^{3/4}. \tag{8}$$

PROOF. (7) implies (8) so we only show (7). By the law of the iterated logarithm for N large enough we have on $1 - 1/N^3$ of the space

$$d_n < 8\sqrt{\frac{1}{2}\,n\log\log n} < 1/3n^{3/4}.$$

By the strong law of large numbers $n < 3S_n$ for all $n \geqq N$ large enough on $1 - 1/N^6$ of the space. So $n \geqq N$ implies $d_n < \frac{1}{3}(3S_n)^{3/4} < (S_n)^{3/4}$ on at least $1 - 1/N^2$ of the space. Note that we could get much sharper error estimates. ∎

Lemma 2 below is a consequence of a well-known geometric approximation to the binomial distribution which we state without proof. See e.g. [2].

LEMMA 2. *Let X_n, $n = 1, 2, \cdots$ be independent, identically distributed, binary random variables with* $\mathrm{Prob}(1) = p$ *and* $\mathrm{Prob}(0) = 1 - p$. *We will denote* $X_1 + X_2 + \cdots + X_n$ *by* $\mathrm{Bi}(n, p)$. *Then* $\mathrm{Prob}(\mathrm{Bi}(n,p) \leqq 2np) \geqq 1/2np$.

Now we turn our attention to the skew product $(\hat{S}, \hat{P})$. Note that $(\hat{S}, \hat{P})$ inherits a block structure from (S, P) by setting $(x, y) \in n$-block if $x \in n$-block.

We need a strong version of Lemma 2. We state the following definitions and Lemma 3 in terms of $\hat{P}$-names but we are really making statements about independent $\hat{Q}$-names.

Let $\boldsymbol{d}$ be a string in the $\hat{P}$-name of $(x, y) \in X \times Y$. Let m = number of $\hat{P}_0$ occurrences in $\boldsymbol{d}$ and n-number of occurrences of $\hat{P}_1$ (i.e., of $\hat{P}_2$ or $\hat{P}_3$). The *excess* of $\boldsymbol{d}$ is defined to be $|n - m|$. $\boldsymbol{d}$ is good if the excess is less than $n^{3/4}$, i.e., the string is *good* if it satisfies (7) above. We say an n-block is *good* if all initial substrings of length greater than 2^n are good. We say an n-block is *very good* if (a) it is good, and (b) by induction for $n > N$, at least $1 - 2^{-n}$ of the $(n-1)$-subblocks are very good.

LEMMA 3. *Given $\varepsilon > 0$ if N is large enough then*

$$m\{(x, y) \in \textit{very good } n\textit{-block, for all } n \leqq N\} > 1 - \varepsilon.$$

PROOF. Assume first N is large enough so that $m\{(x, y) \in N\text{-block}\} > 1 - \varepsilon/2$ and condition on this set. Call the conditional measure m_1. Let $A_n = \{(x, y) \in \text{very good } n\text{-block}\}$ and $B_n = \{(x, y) \in \text{good } n\text{-block}\}$. We show if N was chosen large enough and $n \geqq N$,

$$m_1(A_n) = m_1(A_n/B_n)m_1(B_n) \geqq 1 - 2^{-n}.$$

By induction assume $m_1(A_{n-1}) > 1 - 2^{-n+1}$. By Lemma 1 we may assume $m_1(B_n) \geqq 1 - 2^{-2n}$. $m_1(A_n^c/B_n) = m_1$ (more than 2^{-n} of the $(n-1)$-subblocks are not very good) $\leqq \mathrm{Prob}(\mathrm{Bi}(2^{n^2}, 2^{-n+1}) \geqq 2^{-n}) \leqq 2(2^{n^2} \cdot 2^{-n+1})^{-1} = 2^{-n^2+n} \leqq 2^{-2n}$. So $m_1(A_n/B_n) \geqq 1 - 2^{-2n}$ so $m_1(A_n) \geqq (1 - 2^{-2n})^2 \geqq 1 - 2^{-n}$. This means for a large enough N we need throw out less than $\Sigma_{n \geqq N} 2^{-n} = 2^{-N+1}$ of the space and the lemma will be satisfied on the remainder. ∎

Take $\varepsilon = 1/10$ and assume N is large enough so that Lemma 3 is satisfied for this ε.

We say two n-blocks in the $\hat{P}$-names are *close* if their initial indices differ by less than s_n.

Lemma 4 is the skewed analog of Proposition 4.

LEMMA 4. *If two very good n-blocks in the names of (x, y) and (x', y') are not close then at most one pair of $(n-1)$-subblocks are close.*

PROOF. Assume an $(n-1)$-subblock in the (x, y) name is close to an $(n-1)$-subblock in the (x', y') name. Then by Proposition 4 we have to adjust (using the independent Bernoulli process) by at least $s_n - s_{n-1} > \frac{1}{2} s_n$ indices to line up any further $(n-1)$-subblocks. This means one of the n-blocks must have a substring with excess greater than $\frac{1}{4} s_n$ so it has an initial substring with excess greater than $(1/8)s_n > [h(n)]^{3/4}$ and of length greater than 2^n. This is impossible since the n-blocks are very good. ∎

The next lemma insures $(n-1)$-subblocks of very good n-blocks occur at regular enough intervals for our purposes.

LEMMA 5. *Let $\boldsymbol{a}$ be a very good n-block in the name of (x, y) that only overlaps very good n-blocks in the (x', y') name. Then $\boldsymbol{a}$ overlaps at most two n-blocks in the (x', y') name and each $(n-1)$-subblock overlaps at most two $(n-1)$-blocks in the (x', y') name.*

PROOF. Since $\boldsymbol{a}$ is good any initial substring of length greater than 2^n has excess less than $h(n)^{3/4} < s_{n-2}$, so any substring has excess less than $2s_{n-2}$. This says the length of any good n-block is strictly between the values $2h(n) \pm 2s_{n-2}$. If more than two n-blocks overlap $\boldsymbol{a}$ then one must be entirely overlapped by $\boldsymbol{a}$ and the maximal number of left over indices is $4s_{n-2} < s_{n-1}$. But more P_3-symbols than that separate all n-blocks giving us a contradiction. The proof is similar for $(n-1)$-subblocks. ∎

Recall if (x, y), $(x', y') \in X \times Y$ and $M \geqq N$. Then

$$\bar{d}(x, y)_N^M, (x', y')_N^M =$$

$$\frac{1}{M - N + 1} \operatorname{card}\{n \mid N \leqq n \leqq M \text{ and } (x, y)(n) \neq (x', y')(n)\},$$

i.e. the fraction of indices where the $\hat{P}$-names disagree. Below is the key lemma that makes the proof work. It is the analogue of lemma 3 in [5].

LEMMA 6. *There is an $\bar{\varepsilon} > 0$ so that if we have any n-block $\boldsymbol{a}$ of (x, y) together with (x', y') satisfying the conditions of Lemma 5 then either*

(a) *there is an n-block $\boldsymbol{b}$ in the (x', y') name close to $\boldsymbol{a}$ or*

(b) *$\boldsymbol{a}$ has d distance greater than $\bar{\varepsilon}$ from the corresponding string in the (x', y') name.*

PROOF. Let ε_n be the $\bar{\varepsilon}$ we would get from the statement of the Lemma for n-blocks If $\boldsymbol{a}$ is not close to an n-block then by Lemmas 4 and 5 at most two $(n-1)$-subblocks of $\boldsymbol{a}$ are close to $(n-1)$-blocks in the (x', y') name. At most $2^{-n} \cdot 2^{n^2}$ of the $(n-1)$-blocks of each very good n-blocks are not very good. So at most $2 \cdot 2^{-n} \cdot 2^{n^2}$ of the $(n-1)$-subblocks of $\boldsymbol{a}$ do not only overlap very good $(n-1)$-blocks in the (x', y') name. So even assuming perfect $\bar{d}$ matches except on the remaining very good $(n-1)$-subblocks of $\boldsymbol{a}$ we get

$$\varepsilon_n \geqq (1 - 2/2^{n^2} - 2/2^n - h(n-2)/h(n-1))\varepsilon_{n-1} \geqq (1 - 2^{-n+3})\varepsilon_{n-1}.$$

Continuing we get $\varepsilon_n \geqq \varepsilon_N \prod_{k=N+1}^{n} (1 - 2^{-k+4}) > \alpha > 0$ independent of n. Take $\bar{\varepsilon} = \inf(\varepsilon_n) > 0$. ■

Proposition 7 below may be proven exactly as in [5] which we assume the reader is familiar with. We only sketch the proof.

PROPOSITION 7. *(S, P) is loosely Bernoulli but not pre-Bernoulli.*

PROOF. We have only to show $(\hat{S}, \hat{P})$ is not very weak Bernoulli. Choose $\varepsilon < (\bar{\varepsilon})^2/4$. $(\hat{S}, \hat{P})$ very weak Bernoulli would imply an $M > 0$ and atoms of the past of $(\hat{S}, \hat{P})$, $\hat{\mathcal{P}}_1$, $\hat{\mathcal{P}}_2$ so that the conditional future distributions $\bigvee_{n=0}^{M-1} S^n P / p_i$, $i = 1, 2$ have $\bar{d}$ distance less than ε. We can assume $\hat{\mathcal{P}}_1$, $\hat{\mathcal{P}}_2$ are chosen so we would be forced to match strings of length M in $\hat{P}$-names contained in very good $(m+1)$-blocks that are not close, containing at least 20 very good m-blocks and so that at least 9/10 of the indices of the first M-name belong to very good m-blocks that only overlap very good m-blocks in the other name. Further we may assume these names have $\bar{d}$-distance less than $\sqrt{\varepsilon} < \bar{\varepsilon}/2$. But by Lemma 6 the $\bar{d}$-distance is greater than $\bar{\varepsilon}(9/10 - 1/10) > \bar{\varepsilon}/2 > \varepsilon$. This contradiction proves the result. ■

To finish the proof of the Theorem recall we can construct a K-process (S^*, P^*) of arbitrary entropy that contains much of the structure of (S, P). The above proof also applies word for word to (S^*, P^*) if $f(n)$ grows slowly enough so (S^*, P^*) also is not pre-Bernoulli.

ACKNOWLEDGEMENT

I would like to offer my sincere appreciation to Professors Donald Ornstein and Jacob Feldman who introduced me to this subject and who gave me much help and encouragement.

REFERENCES

1. J. Feldman, *New K-automorphisms and a problem of Kakutani*, Israel J. Math. **24** (1976), 16–38.

2. W. Feller, *An Introduction to Probability and its Applications*, Vol. 1, New York, Wiley, 1957.

3. N. A. Friedman, *Introduction to Ergodic Theory*, Van Nostrand Reinhold Math. Studies #29, York, (1970).

4. I. Meilijson, *Mixing properties of a class of skew-products*, Israel J. Math. **19** (1974), 266–270.

5. D. S. Ornstein, *Ergodic Theory, Randomness and Dynamical Systems*, Yale University Press, 1974.

6. D. S. Ornstein and P. Shields, *An uncountable family of K-automorphisms*, Advances in Math. **10** (1973), 1.

7. P. C. Shields, *Cutting and independent stacking of intervals*, Math. Systems Theory **7** (1973).

8. B. Weiss, *Equivalence of measure preserving transformations*, Lecture Notes, The Institute for Advanced Studies of the Hebrew University of Jerusalem, 1975.

DEPARTMENT OF MATHEMATICS
OREGON STATE UNIVERSITY
CORVALLIS, OR 97330 USA