ISRAEL JOURNAL OF MATHEMATICS, Vol. 24, No. 1, 1976

NEW K-AUTOMORPHISMS AND A PROBLEM OF KAKUTANI

BY

J. FELDMAN

ABSTRACT

A property is introduced, for 1-1 measure-preserving transformations of probability spaces, called *loose Bernoulliness* (LB), which is invariant under taking factors, inducing, and tower-building. It amounts to replacing, in Ornstein's definition of *very weak Bernoulli*, the Hamming distance on strings by a coarser metric. The main result is the construction of a transformation T_0 which is ergodic and of entropy 0 but *not* LB. On the other hand, any irrational rotation *is* LB. Consequently, the equivalence relation generated by inducing and tower-building (which I call *Kakutani equivalence*, and the Russians call *monotone equivalence*) has at least two distinct equivalence classes among the ergodic entropy zero transformations. A similar situation exists for ergodic positive-entropy transformations: on the one hand, any Bernoulli shift is LB, while on the other hand a non LB K-automorphism $\hat{T}_0$ can be made by skewing T_0 over a Bernoulli base.

1. Introduction

Let S be a Bernoulli automorphism of the probability space $(X, \mathcal{A}, \mu)$ with independent generating partition $\mathcal{P} = \{P_0, P_1\}$, $\mu(P_i) = \frac{1}{2}$. The term "automorphism" will here always mean bimeasurable measure-preserving bijection. For any ergodic automorphism with finite entropy, call it T, on $(Y, \mathcal{B}, \nu)$, let $\hat{T}$ be the skew product defined by

$$\hat{T}(x, y) = \begin{cases} (Sx, y), & x \in P_0 \\ (Sx, Ty), & x \in P_1 \end{cases}.$$

I. Meilijson [7] has shown that $\hat{T}$ is always a K-automorphism. This was noticed independently but later by myself and Ken Berg.

The question leaps to mind, will $\hat{T}$ always be a *Bernoulli* automorphism? If T is either an irrational rotation of the circle or itself a Bernoulli automorphism

Received November 3, 1975

then $\hat{T}$ is indeed Bernoulli (Theorem 2). The more interesting problem is to find some T for which $\hat{T}$ is *not* Bernoulli.

Such T do in fact exist! Specifically: I introduce a property, *loosely Bernoulli* (LB), which must be possessed by any T for which $\hat{T}$ is Bernoulli (Corollary 2 to Theorem 3), and then construct an ergodic T_0 of entropy zero which is *not* LB (Theorem 4). But actually the method gives much more. The condition LB is preserved under taking factors, inducing and building towers of finite measure (Theorem 3). Thus, in particular, LB is invariant under Kakutani-equivalence [5]: no single ergodic automorphism can induce both an LB automorphism and a non LB automorphism. Equivalently, no special flow built over an LB base can be isomorphic to a flow built over a non LB base. The non LB automorphism T_0 is thus not Kakutani-equivalent to any irrational rotation (Corollary 1 to Theorem 4), even though both have zero entropy. It is further argued, in Corollary 2 to Theorem 4, that $\hat{T}_0$ is also not LB, and consequently provides an example of a K-automorphism which is not Kakutani-equivalent to a Bernoulli automorphism.

To the best of my knowledge, all previously known non-Bernoulli K-automorphisms have been constructed by an infinite sequence of independent cuttings and stackings, plus occasional addition of new material; cf. Ornstein [8]. Since this technique always gives rise to automorphisms which induce Bernoullis, as was essentially shown by L. Swanson [11], it follows that $\hat{T}_0$ is not isomorphic, or even Kakutani-equivalent, to any of these.

It is further possible (Corollary 3 to Theorem 4), using results of Gurevič [4], to build a K-flow over $\hat{T}_0$, thus providing an example of a K-flow which cannot be time-changed to a Bernoulli flow.

Some of these techniques may be extended to study skew products over $(S, \mathcal{P})$ where $\mathcal{P}$ is not independent. This will be carried out elsewhere; in the present paper just enough technique is developed to provide the examples mentioned.

One of the crucial ideas is to substitute into the definition "very weak Bernoulli" (VWB) given by D. Ornstein [8], a certain notion of distance between strings of symbols which is different from the "Hamming distance" used by Ornstein. This notion has also been encountered by S. Ulam [12] in a biological context, and I feel that it is a natural notion for Information Theory.

Essential use is made of the recent D. Ornstein–B. Weiss result [10] that every finite partition for a Bernoulli automorphism is VWB.

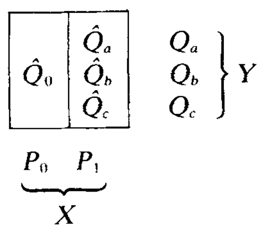
This work was carried out mainly during the 1975 Symposium on Ergodic Theory at the University of Warwick, while on sabbatical leave from the University of California, Berkeley, and with partial support of NSF Grant MPS-

75-05576. I am grateful to Ken Berg, Jean-Paul Thouvenot, and especially Harry Furstenburg for valuable suggestions and discussions; also to Don Ornstein for a suggestion which simplified the proof of Step V in Theorem 4.

The referee has brought to my attention a note by A. Katok (Dokl. Akad. Nauk SSSR, **223** (1975), 784–792). The results announced there have some things in common with the present paper. To be exact: My definition LB, in the case of entropy zero transformations, is precisely the condition which appears in Katok's theorem 4, part 3. Thus, his announced results cover the zero entropy case of my Theorem 3.

2. *K*-automorphisms and Bernoulli automorphisms

Let S be a Bernoulli with independent generator $\mathcal{P} = \{P_0, P_1\}$, $\mu(P_0) = \mu(P_1) = \frac{1}{2}$. Let T have finite generator $\mathcal{Q} = \{Q_a : a \in \mathbf{I}\}$ (although doubtless much of what follows does not really require this assumption). Let $\hat{\mathcal{Q}}$ be the partition $\{\hat{Q}_a : a \in \hat{\mathbf{I}}\}$ of $X \times Y$, where $\hat{\mathbf{I}} = \mathbf{I} \cup \{0\}$, $\hat{Q}_0 = P_0 \times Y$, and $\hat{Q}_a = P_1 \times Q_a$ for $a \in \mathbf{I}$.



If $\mathcal{Q}$ is a generator under T then, as may easily be verified, $\hat{\mathcal{Q}}$ is a generator under $\hat{T}$. Let $\{\xi_n\}$ be the stationary process defined by $\xi_n(x) = j$ if $S^n x \in P_j$. Let $\eta_n(y) = a$ if $T^n y \in Q_a$. Let $\hat{\eta}_n(x, y) = a$ if $\hat{T}^n y \in \hat{Q}_a$. Thus $\hat{\eta}_n(x, y) = \eta_{\sigma_n(x)}(y)$ if $\xi_n = 1$, otherwise $\hat{\eta}_n(x, y) = 0$, where the σ_n are partial sums of the ξ_n, defined by $\sigma_0 = 0$ and $\sigma_{n+1} = \sigma_n + \xi_{n+1}$. The ξ_n, η_n and σ_n may also be regarded, in an obvious way, as functions on $X \times Y$. The processes may be viewed graphically thus:

$$\begin{array}{ccccc|ccccc} & \cdots\eta_{-2} & & \eta_{-1} & & & \eta_0 & & \eta_1\ \eta_2 \cdots & \\ \cdots 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1\cdots \end{array}.$$

That is, the η_n are substituted for the 1's in the ξ-sequence in order, η_0 being substituted in the first non-negative place j where $\xi_j = 1$, to obtain the $\hat{\eta}$-sequence.

THEOREM 1. *$\hat{T}$ is a K-automorphism.*

REMARKS. This is a special case of [7]. The general idea of the proof is to utilize the fact that the probabilities $\mu\{\sigma_N = j\}$ are "asymptotically flat", in the sense that $\lim_{N\to\infty}\Sigma_j|\mu\{\sigma_N = j\} - \mu\{\sigma_N = j+1\}| = 0$ which may be shown directly, or by applying a theorem in [9]. This enables the process $\{\eta_{\sigma_n}\}$ to "forget" where the process $\{\eta_j\}$ is.

THEOREM 2. *If T is either an irrational rotation or a Bernoulli automorphism then $\hat{T}$ is a Bernoulli automorphism.*

PROOF. The case of an irrational rotation follows from Adler and Shields [1]. The case of a Bernoulli transformation is obvious, because if $\mathcal{Q}$ is an independent generator for T then $\hat{\mathcal{Q}}$ is an independent generator for $\hat{T}$.

REMARK. It may be shown directly that for any ergodic rotation T of a compact abelian group, or any T which is direct product of such a rotation with a Bernoulli automorphism, $\hat{T}$ is Bernoulli. However, I'll refrain from presenting that argument here.

3. Loosely Bernoulli processes

Let T be an automorphism on $(Y, \mathcal{B}, \nu)$, and $\mathcal{Q} = \{Q_a : a \in \mathbf{I}\}$ a finite partition. By $\mathcal{Q}_j^k$ is meant the partition generated by $\{T^{-i}\mathcal{Q} : j \leqq i \leqq k\}$. Let $\{\eta_i\}$ be the corresponding $\mathbf{I}$-valued process, defined by $T^i y \in Q_{\eta_i(y)}$.

Now I define "very weakly Bernoulli" for a partition. This will differ from the original definition of [8] in two respects. One is minor: the use of measures $n_{A,B}$ on $\mathbf{I}^N \times \mathbf{I}^N$, rather than enlarging to a continuous probability space. This is easily seen to amount to the same thing. The second is more substantive: instead of demanding "$\exists N$" in the definition below, I demand "for *all* sufficiently large N". That this apparently more stringent requirement is in fact equivalent to the original definition was pointed out to me by Don Ornstein. It may be seen by examining the proof of the lemma in [10]. I could have worked with the original definition throughout, making a corresponding change in the definition of LB, but the present form simplifies checking the basic example in Section 5.

$\mathcal{Q}$ is called *very weakly Bernoulli* (VWB) for T if for every $\varepsilon > 0$ it is the case that for every sufficiently large integer N and for each $M > 0$ the following holds: there exists a collection $\mathcal{G}$ of "good" atoms of $\mathcal{Q}_{-M}^0$ whose union has measure $> 1 - \varepsilon$, and so that for each pair A, B of atoms in $\mathcal{G}$ the conditional distributions $\nu(\cdot \mid A)$ and $\nu(\cdot \mid B)$, restricted to $\mathcal{Q}_1^N$, are "close" in this precise sense: there is a measure $n_{A,B}$ on $\mathbf{I}^N \times \mathbf{I}^N$ such that

a) $n_{A,B}(\alpha \times \mathbf{I}^N) = \nu((\eta_1, \cdots, \eta_N) = \alpha \mid A)$, and
$n_{A,B}(\mathbf{I}^N \times \beta) = \nu((\eta_1, \cdots, \eta_N) = \beta \mid B)$,

and also

b) $n_{A,B}(\{(\alpha, \beta)$: α differs from β in at least εN places$\}) < \varepsilon$.

It was shown by Ornstein [8] that if T has a VWB finite generator then T is a Bernoulli automorphism; and by Ornstein and Weiss [10] that, conversely, for a Bernoulli automorphism every finite generator is VWB.

Now I introduce another property of a finite partition $\mathscr{Q}$: that of being *loosely Bernoulli* (LB) with respect to T. The only difference between VWB and LB wil occur in (b), where the notion of the closeness of the strings used there will be replaced by a weaker notion. First, for two strings of symbols $\alpha = (a_1, \cdots, a_N)$ and $\beta = (b_1, \cdots, b_N)$, set $d(\alpha, \beta) = (1/N)|\{j: a_j \neq b_j\}|$. Then condition (b) in the definition of VWB may be replaced by

$$n_{A,B}(\{(\alpha, \beta): d(\alpha, \beta) \geqq \varepsilon\}) < \varepsilon.$$

Now: for two strings $\alpha = (a_1, \cdots, a_M)$ and $\beta = (b_1, \cdots, b_N)$, no longer necessarily the same length, define a *match* π of α with β as an order-preserving bijection from a subset $\mathscr{D}(\pi) \subset \{1, \cdots, M\}$ onto a subset $\mathscr{R}(\pi) \subset \{1, \cdots, N\}$ such that $b_{\pi(j)} = a_j$. The *fit* of the match, $|\pi|$, is the number $(|\mathscr{D}(\pi)| + |\mathscr{R}(\pi)|)/(M + N)$. Of course $|\mathscr{D}(\pi)| = |\mathscr{R}(\pi)|$, and if $M = N$ we just ge $|\mathscr{D}(\pi)|/M$. The *distance* $\delta(\alpha, \beta)$ is defined to be $1 - \sup\{|\pi|: \pi$ any match of α with $\beta\}$. Notice that if π is a match of α with β then π^{-1} is a match of β with α and $\delta(\pi) = \delta(\pi^{-1})$, so $\delta(\alpha, \beta) = \delta(\beta, \alpha)$. Also, $\delta(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$. On string of equal length, the triangle equality also holds, so that δ is a distance; on string of varying length this is not the case, although minor changes in the definitio could be made to fix this if it were needed. Now say $\mathscr{Q}$ is *loosely Bernoulli* (LB for S if it satisfies the previous definition of VWB with the following change instead of (b), I demand the weaker

b′) $n_{A,B}\{(\alpha, \beta): \delta(\alpha, \beta) \geqq \varepsilon\} < \varepsilon$.

Here is an argument which shows that a certain apparent weakening of th definitions actually leads to the same thing.

PROPOSITION 1. *Let (a_{ij}), $i, j = 1, \cdots, N$, be a matrix with nonnegative entrie adding up to 1. Let $b_i = \Sigma_j a_{ij}$ and $c_j = \Sigma_i a_{ij}$. Let b'_i, $i = 1, \cdots, N$, and c'_j, $j = 1, \cdots, N$, be probability vectors with $\Sigma_i |b_i - b'_i| < \varepsilon$ and $\Sigma_j |c_j - c'_j| < \varepsilon$. Then there is another matrix (a'_{ij}) with nonnegative entries adding up to 1 such tha $\Sigma_j a'_{ij} = b'_i$, $\Sigma_i a'_{ij} = c'_j$, and $\Sigma_{ij} |a_{ij} - a'_{ij}| < 2\varepsilon$.*

PROOF. It will suffice to show the case $c_j = c'_j$, provided that $\Sigma_{ij} |a_{ij} - a'_{ij}| < \varepsilon$; for then one can first fix the b_j leaving the c_j alone, and then fix the c_j leavin the b_j alone. So let us assume $c_j = c'_j$. Now: take each row i for which $b_i > b'_i$

look along it for terms a_{ij} which are >0, remove enough mass from these a_{ij} to bring b_i down to b'_i, and transfer the mass removed from a_{ij} to some $a_{i'j}$ for which $b_{i'} < b'_{i'}$. This can certainly be done, since $\Sigma_i b_i = \Sigma_i b'_i$. Call the new matrix a'_{ij}. The sum $\Sigma_i a'_{ij}$ is still equal to c_j for each j, while $\Sigma_j a'_{ij} = b'_i$ as desired. Finally,

$$\sum_{i,j} | a_{ij} - a'_{ij} | = 2 \sum_{i,j} \{ a_{ij} - a'_{ij} : a_{ij} > a'_{ij} \}$$

$$= 2 \sum_i \{ b_i - b'_i : b_i > b'_i \} = \sum_i | b_i - b'_i | < \varepsilon. \qquad \square$$

REMARK. The result easily generalizes to the context of marginals of probability measures on product spaces.

COROLLARY 1. *In* (a) *of the definition of VWB and LB, it suffices to demand that the marginals of* $n_{A,B}$ *approximate to within* ε.

DEFINITION. This weaker demand I will call (a′).

COROLLARY 2. *In the definition of VWB and LB it suffices, instead of saying "for all* $M > 0$*", to say "for all **sufficiently large*** $M > 0$*".*

PROOF. Suppose the definition is applied for ε^2 and some M (for fixed ε and N). I show that, for any positive integer $M' < M$, the definition also works.

Let $\mathscr{G}$ be a set of $\mathscr{Q}^0_{-M}$ atoms of measure $> 1 - \varepsilon^2$ satisfying the definition: so that associated with each A, B in $\mathscr{G}$ there is a measure $n_{A,B}$ on $\mathbf{I}^N \times \mathbf{I}^N$ satisfying (a), and (b) or (b′). Those atoms of $\mathscr{Q}^0_{-M'}$ which contain a proportion $1 - \varepsilon$ of $\mathscr{Q}^0_{-M}$ atoms in $\mathscr{G}$, we call $\mathscr{G}'$. The remaining $\mathscr{Q}^0_{-M'}$ atoms form a set of measure $< \varepsilon$.

For $A' \in \mathscr{G}'$ set $\mathscr{G}(A') = \{A \in \mathscr{G} : A \subset A'\}$. For $A \in \mathscr{G}(A')$ set $\rho(A) = \nu(A)/\Sigma_{B \in \mathscr{G}(A')} \nu(B)$. For $A', B' \in \mathscr{G}'$ set

$$n_{A',B'} = \sum \rho(A) \rho(B) n_{A,B},$$

where A is summed over $\mathscr{G}(A')$ and B over $\mathscr{G}(B')$. Then (b) or (b′) is satisfied for $n_{A',B'}$, since it is an average of measures satisfying the condition. As for (a):

$$n_{A',B'}(\alpha \times \mathbf{I}^N) = \sum \rho(A)\rho(B) n_{A,B}(\alpha \times \mathbf{I}^N)$$

$$= \sum \rho(A)\rho(B) \nu((\eta_1, \cdots, \eta_N) = \alpha \mid A)$$

$$= \sum \rho(A) \nu((\eta_1, \cdots, \eta_N) = \alpha \mid A)$$

$$= \frac{\nu(A')}{\sum_{B \in \mathscr{G}(A')} \nu(B)} \sum \nu(A \mid A') \nu((\eta_1, \cdots, \eta_N) = \alpha \mid A).$$

The factor in front is between 1 and $1/(1-\varepsilon)$, and the summation reduces to $\nu((\eta_1,\cdots,\eta_N)=\alpha \mid A')$. A similar argument works for $n_{A',B'}(\mathbf{I}^N\times\beta)$. Thus, after a shrewder choice of ε, one gets the weaker looking (a') of Corollary 1. □

REMARK. If the entropy $h(T,\mathcal{Q})=0$, then the definition LB can be recast in a considerably more manageable form: namely, that for every $\varepsilon>0$ and every sufficiently large N there exists a set $\mathcal{H}$ of $\mathcal{Q}_1^N$ atoms of total measure $\geqq 1-\varepsilon$, such that any two of the names can be matched better than $1-\varepsilon$. The proof of this remark is easy and is left to the reader.

In the remainder of this section, and in the next section there will be proven several functorial properties of "LB". It should be remarked that all of these may be proven much more simply in the entropy zero case, by using the above simplified definition, and the reader may want initially to carry out for himself this version of the arguments. However, in order to obtain all of Corollary 2 to Theorem 4, which is one of the main results of this paper, I will need the positive entropy case.

LEMMA 1. *If $\mathcal{Q}$ is LB for T, and $\mathcal{R}\subset\mathcal{Q}$, then $\mathcal{R}$ is LB for T.*

PROOF. Choose $\varepsilon>0$ and a sufficiently large N for $\mathcal{Q}$ and ε^2. Choose $M>0$ and a corresponding set of $\mathcal{Q}^0_{-M}$-atoms. Let $\mathcal{G}'$ be those $\mathcal{P}^0_{M}$-atoms which are made up, to within a proportion $1-\varepsilon$, of members of $\mathcal{G}$; then the total measure of the members of $\mathcal{G}'$ is $\geqq 1-\varepsilon$.

Continue exactly as in the proof of the preceding Corollary 2, to obtain for each A', B' in $\mathcal{G}'$ a measure $n_{A'B'}$ on $\mathbf{I}^N\times\mathbf{I}^N$ which satisfies (a') and (b'). Here $\mathbf{I}$ is of course the index set of $\mathcal{Q}$. However, what is needed is a measure on $\mathbf{J}^N\times\mathbf{J}^N$ where $\mathbf{J}$ is the index set of $\mathcal{P}$. But $\mathbf{J}$ can be regarded as a certain collection of subsets of $\mathbf{I}$, and by restricting $n_{A',B'}$ to the subsets of $\mathbf{I}^N\times\mathbf{I}^N$ corresponding to $\mathbf{J}^N\times\mathbf{J}^N$ one obtains a measure on $\mathbf{J}^N\times\mathbf{J}^N$ satisfying (a') and (b') for $\mathcal{P}$. □

LEMMA 2. *If $\mathcal{Q}$ is LB for T, then so is $\mathcal{Q}_0^k$.*

PROOF. It suffices to consider the case $k=1$, because then iteration gives the general case.

Pick $\varepsilon>0$. Since $\mathcal{Q}$ is LB, pick N for δ, where δ will be chosen later. Suppose α and β are in $\mathbf{I}^N$ and are matched better than $1-\delta$. Let $\alpha=(a_1,\cdots,a_N)$ $\beta=(b_1,\cdots,b_N)$. Let $\tilde{\alpha}=(\tilde{a}_2,\cdots,\tilde{a}_N)$, $\tilde{\beta}=(\tilde{b}_2,\cdots,\tilde{b}_N)$, where $\tilde{a}_{j+1}$ is the pair (a_j,a_{j+1}) and $\tilde{b}_{j+1}$ is (b_j,b_{j+1}). For all pairs (a_j,a_{j+1}) which are paired with (b_j,b_{j+1}) we get $\tilde{a}_{j+1}$ paired with $\tilde{b}_{j+1}$. If at least $(1-\delta)N$ of the α-sequence are paired, so that no more than δN are unpaired, then at least $(1-2\delta)N-1$ of the terms of $(a_2,\cdots,a_N)$ are paired and have the previous term paired as well. Consider the images of such terms in β via the match. Restricting to $(b_2,\cdots,b_N)$ leaves at least $(1-2\delta)N-2$ terms; and since no more than δN terms of β were originally unpaired, at least $(1-3\delta)N-2$ of these images will have the previous term paired. Thus there are at least $(1-3\delta)N-2$ pairings induced between $(\tilde{\alpha}_2,\cdots,\tilde{\alpha}_N)$ and $(\tilde{\beta}_2,\cdots,\tilde{\beta}_N)$. Pick $\delta=\varepsilon/4$ and N at least $8/\varepsilon$. Then a match as good as $1-\varepsilon$ for α and β gives a match as good as $1-\varepsilon$ for $(\tilde{a}_1,\cdots,\tilde{a}_N)$ and $(\tilde{b}_1,\cdots,\tilde{b}_N)$, where a_0 and b_0 are chosen in any way whatsoever. Now: an atom of $(T^{-1}\mathcal{Q}\vee\mathcal{Q})^0_{-M}$ is also an atom of $\mathcal{Q}^0_{-M-1}$. Let $n_{A,B}$ be the measure on $\mathbf{I}^N\times\mathbf{I}^N$ given by LB of $\mathcal{Q}$. Extend this measure to allow for a 0-component in the first factor and the second factor, which takes on with certainty the value a_0 (from A) in the first factor and b_0 (from B) in the second factor. This induces a measure $\tilde{n}_{A,B}$ on $\tilde{\mathbf{I}}^N\times\tilde{\mathbf{I}}^N$, where $\tilde{\mathbf{I}}=\mathbf{I}\times\mathbf{I}$, using the map $(a_0,\cdots,a_N)\mapsto(\tilde{a}_1,\tilde{a}_2,\cdots,\tilde{a}_N)$ and $(b_0,\cdots,b_N)\mapsto(\tilde{b}_1,\cdots,\tilde{b}_N)$ to transfer the measure. Then $n_{A,B}$ will have the desired properties. □

PROPOSITION 2. *If $\mathcal{Q}$ is LB with respect to T, and $\mathcal{P}$ is a finite partition with $\mathcal{P}\subset\mathcal{Q}^0_{-\infty}$, then $\mathcal{P}$ is also LB. (Here $\mathcal{Q}^0_{-\infty}$ is the σ-field generated by $\{\mathcal{Q}^0_{-k}:\ k\leqq 0\}$).*

PROOF. Pick $\varepsilon>0$. Pick δ (whose dependence on ε will be specified later). Pick L so $\mathcal{P}\subset_\delta\mathcal{Q}^0_{-L}$. Pick N as in the definition of LB, for $\mathcal{Q}^0_{-L}$ and δ. Then for each $M>0\ \exists\,\mathcal{G}\subset(\mathcal{Q}^0_{-L})^0_{-M}=\mathcal{Q}^0_{-(L+M)}$ so that if $A,B\in\mathcal{G}$ then $\exists\, n_{A,B}$ on $\mathbf{I}^{(L+1)M}\times\mathbf{I}^{(L+1)M}$ satisfying (a) and (b′) for δ and $\mathcal{Q}^0_{-L}$. Let $\mathcal{P}=\{P_a:\ a\in\mathbf{J}\}$ and $\tilde{\mathcal{P}}=\{\tilde{P}_a:\ a\in\mathbf{J}\}$, with $\tilde{\mathcal{P}}\subset\mathcal{Q}^0_{-L}$ and $|\tilde{\mathcal{P}}-\mathcal{P}|<\delta$. Then $\mathbf{J}$ can be regarded as certain subsets of $\mathbf{I}^{N+1}$, and a map is induced from $\mathbf{I}^{(L+1)M}\times\mathbf{I}^{(L+1)M}$ to $\mathbf{J}\times\mathbf{J}$, call it ϕ.

Now fix $M_1>0$, and choose M so large that each $\mathcal{P}^0_{-M_1}$-atom consists, up to $1-\delta$, of a union of $\mathcal{Q}^0_{-(L+M)}$-atoms. Then, as in previous arguments, most $\mathcal{P}^0_{-M}$-atoms are made up mostly of atoms of $\mathcal{G}$, call these $\mathcal{P}^0_{-M}$-atoms $\mathcal{G}_1$, and for A_1,B_1 in $\mathcal{G}_1$ the measure n_{A_1,B_1} is defined as $\Sigma_{A,B}n_{A,B}\circ\phi^{-1}\nu(A\mid A_1)\nu(B\mid B_1)$, the sum being extended over those A,B in $\mathcal{G}$ which are mostly contained in A_1 and B_1 respectively. Sufficiently small choice of δ guarantees that (a′) and (b′) are satisfied for $\mathcal{P}$ and ε. □

COROLLARY. *If T is of entropy 0 and $\mathcal{Q}$ is a finite generator which is LB with respect to T, then **every** finite partition is LB with respect to T.*

PROOF. If T has entropy 0 and $\mathcal{Q}^\infty_{-\infty}=\mathcal{B}$ then already $\mathcal{Q}^0_{-\infty}=\mathcal{B}$, and Proposition 2 may be applied. □

4. Loosely Bernoulli automorphisms

DEFINITION. The ergodic automorphism T on $(Y, \mathcal{B}, \nu)$ will be called loosely Bernoulli (LB) if *every* finite partition of Y is LB with respect to T.

THEOREM 3.

a) *If T is LB then so is every factor.*
b) *If T is LB and $\nu(E)>0$ then T_E is LB.*
c) *Conversely, if T_E is LB then so is T.*

PROOF OF (a). This is immediate from the definition. □

PROOF OF (b). Assume that T is LB. Let $\mathcal{P} = \{P_a : a \in \mathbf{I}\}$ be a finite partition of E. The object is to show that $\mathcal{P}$ is LB for T_E. Let $\mathcal{P}' = \mathcal{P} \cup \{Y - E\} = \{P_a : a \in \mathbf{I}'\}$, where $\mathbf{I}' = \mathbf{I} \cup \{0\}$ and $Y - E = P_0$. Let ν_E be $\nu(\cdot)/\nu(E)$. Let $\{\xi_j\}$ be the $\mathbf{I}$-valued stochastic process on $(E, \mathcal{B}/E, \nu_E)$ defined by $T_E^j y = P_{\xi_j(y)}$. Let $\{\xi'_j\}$ be the $\mathbf{I}'$-valued process on $(Y, \mathcal{B}, \nu)$ defined by $T_y^j = P_{\xi'_j(y)}$. Let t_n be the nth arrival time in E, counting forward from 1 for $n>0$ and backward from 0 for $n \leqq 0$. Thus, $\xi'_{t_n} \mid E = \xi_n$.

Now pick $\varepsilon>0$. Pick $\varepsilon'>0$, to depend on ε in a manner which will be specified later, but make sure $2\varepsilon' < \nu(E)$. $\mathcal{P}'$ is LB with respect to T, so N' may be chosen "big enough" for T, $\mathcal{P}'$, ε'. Also choose it so large that

$$\nu(E) - \varepsilon' < \frac{1}{N'} \sum_{j=1}^{N'} 1_E(T_y^j) \leqq \nu(E) + \delta'$$

on a set F' of measure $\geqq 1 - \varepsilon'$. F' is obviously $\mathcal{P}_1'^{N'}$-measurable, corresponding to a set $\mathbf{F}' \subset \mathbf{I}'^{N'}$. Choose N between $(\nu(E) - \varepsilon')N'$ and $(\nu(E)+\varepsilon')N'$; since N can be chosen arbitrarily once it gets big enough, so can N.

Now choose a nonnegative integer M. Choose M' so large that for each $A \in \mathcal{P}_{-M}^0$, $t_{-M} \leqq M'$ on a subset of A whose ν_E-measure is at least a proportion $(1-\varepsilon')$ of that of A. Let $\mathcal{G}'$ and $n_{A',B'}$ satisfy (a), (b′) for this M'. Also let $\mathcal{G}$ be the set of $\mathcal{P}_{-M}^0$-atoms A such that $\nu_E(F' \mid A)$ is $\geqq 1 - \sqrt{\varepsilon'}$; since $\nu_E(F') \geqq 1 - (\varepsilon'/\nu(E))$, the total measure of $\mathcal{G}$ (for ν_E) is $\geqq 1 - (\sqrt{\varepsilon'}/\nu(E))$. Now for A, B in $\mathcal{G}$, define

$$\tilde{n}_{A,B}(\alpha, \beta) = \sum_{\alpha', \beta', A', B'} n_{A',B'}(\alpha', \beta') \nu(A' \mid A) \nu(B' \mid B),$$

where α' ranges over all elements of $\mathbf{F}'$ containing α as the first N symbols in $\mathbf{I}$ and β' has a similar range relative to β, $A' \in \mathcal{G}'$ and $\subset A$, while $B' \in \mathcal{G}'$ and $\subset B$

Then (without repeating explicitly the range of the summation at each stage) we have

$$\tilde{n}_{A,B}(\alpha, \mathbf{I}^N) = \sum_{\beta \in \mathbf{I}^N} \sum_{\alpha', \beta', A', B'} n_{A',B'}(\alpha', \beta')\nu(A' \mid A)\nu(B' \mid B)$$

$$= \sum_{\alpha', A', B'} n_{A',B'}(\alpha', \mathbf{F}')\nu(A' \mid A)\nu(B' \mid B)$$

$$\leqq \sum_{\alpha', A', B'} n_{A',B'}(\alpha', \mathbf{I}'^{N'})\nu(A' \mid A)\nu(B' \mid B)$$

$$= \sum_{\alpha', A', B'} \nu((\xi'_1, \cdots, \xi'_{N'}) = \alpha' \mid A')\nu(A' \mid A)\nu(B' \mid B)$$

$$\leqq \sum_{\alpha'} \nu((\xi'_1, \cdots, \xi'_{N'}) = \alpha' \mid A) \leqq \nu((\xi'_{t_1}, \cdots, \xi'_{t_N}) = \alpha \mid A)$$

$$= \nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A).$$

Similarly, $\tilde{n}_{A,B}(\mathbf{I}^N, \beta) \leqq \nu_E((\xi_1, \cdots, \xi_N) = \beta \mid B)$.

How much has been omitted in the above inequalities? The total mass of the deficit in $\tilde{n}_{A,B}(\cdot, \mathbf{I}^N)$ is dominated by the sum of the following four terms:

$$\sum_{A', B'} n_{A',B'}(\mathbf{F}', \mathbf{I}'^{N'} - \mathbf{F}')\nu(A' \mid A)\nu(B' \mid B)$$

1) $$\leqq \nu((\xi'_1, \cdots, \xi'_{N'}) \in \mathbf{I}'^{N'} - \mathbf{F}' \mid A)$$

$$= \nu(Y - F' \mid A) = \nu_E(Y - F' \mid A) \leqq \sqrt{\varepsilon'},$$

because $A \in \mathscr{G}$, and similarly,

2) $$\sum_{A', B'} n_{A',B'}(\mathbf{I}'^{N'} - \mathbf{F}', \mathbf{F}')\nu(A' \mid A)\nu(B' \mid B) \leqq \sqrt{\varepsilon'},$$

because $B \in \mathscr{G}$.

3) $\Sigma_{A'} \nu(A' \mid A)$ summed over all A' whose names have fewer than M terms in $\mathbf{I}$; this is $< \varepsilon'$ by choice of M'.

4) Similarly $\Sigma_{B'} \nu(B' \mid B)$ over all B' whose names have fewer than M terms in $\mathbf{I}$, which gives $< \varepsilon'$.

Thus, $\tilde{n}_{A,B}(\cdot, \mathbf{I}^N)$ approximates the appropriate marginals to within $2(\varepsilon' + \sqrt{\varepsilon'})$ in total variation. It is not a probability measure, but normalizing it clearly only changes it a little.

Finally, if α' and β' in $\mathbf{F}'$ have α and β respectively as the first N $\mathbf{I}$-terms, and

$\delta(\alpha',\beta') < \varepsilon'$, then, since $(\nu(E)-\varepsilon')N' < N < (\nu(E)+\varepsilon')N'$, it follows from the definition of δ that $\delta(\alpha,\beta) < \varepsilon'/(\nu(E)+\varepsilon')$.

Since $\tilde{n}_{A,B}$ is a sub-convex average of measures which give mass $< \varepsilon'$ to pairs (α',β') such that $\delta(\alpha',\beta') \geqq \varepsilon'$, it follows that $\tilde{n}_{A,B}$ gives mass $< \varepsilon'$ to pairs (α,β) with $\delta(\alpha,\beta) \geqq \varepsilon'/(\nu(E)+\varepsilon')$. Normalization of $\tilde{n}$ does not change this much.

So, finally, choosing ε' very small—$\varepsilon^2/100$ will surely be enough—causes $\mathcal{P}$ to satisfy (a′) and (b′) with respect to T and ε. □

For the proof of (c) some lemmas will be needed.

LEMMA 1. *Let E be a set of positive measure in Y, and let* $E_n = \{y \in E \colon T^n y \in E, T^j y \notin E \text{ for } 0 < j < n\}$. *Then* $-\Sigma_n \nu(E_n)\log \nu(E_n) < \infty$.

PROOF. Let $q_n = 2^{-n}$, $n = 1,2,\cdots$, and $p_n = \nu(E_n)/\nu(E)$. Then $\{q_n\}$ and $\{p_n\}$ are both probability distributions, and $-\Sigma_n p_n \log q_n = -\Sigma_n(\nu(E_n)/\nu(E))n = 1/\nu(E)$. But $-\Sigma_n p_n \log q_n \geqq -\Sigma_n p_n \log p_n$; see Billingsley [2]. □

LEMMA 2. *If $\mathcal{R}$ is a countable partition, $\mathcal{R} = \{R_1, R_2, \cdots\}$, and $-\Sigma_n \nu(R_n)\log\nu(R_n) < \infty$, then there is a* **finite** *partition $\mathcal{Q}$ such that $\mathcal{Q}^0_{-\infty} \subset \mathcal{R}$.*

PROOF. This is implicit in the way Krieger [6] constructs his finite partition. □

LEMMA 3. *Let R be a finite partition of Y, and $\nu(E) > 0$. Then there is a finite partition $\mathcal{Q}$ of E such that, setting $\mathcal{Q}' = \mathcal{Q} \cup \{Y - E\}$ as before, we have $\mathcal{Q}'^0_{-\infty} \supset \mathcal{R}$.*

PROOF. It will suffice to consider the special case where $\mathcal{R} = \{E, R_0, R_1\}$ with $\nu(E) \geqq \nu(R_0)$.

Write F for R_0 and set

$$F_1 = \{y \in F \colon \exists n > 0, T^n y \in E, T^j y \notin F \text{ if } 0 < j < n\},$$

$$E_1 = \{y \in E \colon \exists n > 0, T^{-n} y \in F, T^{-j} y \notin E \text{ if } 0 < j < n\}.$$

A measure-preserving bijection from F_1 onto E_1 may be defined by sending each y in F_1 to its first image in E_1.

Inductively, set $F_{k+1} =$

$$\left\{y \in F - \bigcup_{i=1}^{k} F_i \colon \exists n > 0, T^n y \in E - \bigcup_{i=1}^{k} E_i, T^j y \notin F - \bigcup_{i=1}^{k} F_i \text{ if } 0 < j < n\right\},$$

and set $E_{k+1} =$

$$\left\{y \in E - \bigcup_{i=1}^{k} E_i \colon \exists n > 0, T^{-n} y \in F - \bigcup_{i=1}^{k} F_i, T^{-j} y \notin E - \bigcup_{i=1}^{k} E_i \text{ if } 0 < j < n\right\}.$$

A measure-preserving bijection from F_{k+1} onto E_{k+1} may be defined by sending each y in F_{k+1} to its first image in E_{k+1}.

I claim that the set $F - \bigcup_{n=1}^{\infty} F_k = F_\infty$ has measure zero. For suppose $\nu(F_\infty) > 0$. Then also $\nu(E_\infty) > 0$ where $E_\infty = E - \bigcup_{k=1}^{\infty} E_k$, since $\nu(F) \leqq \nu(E)$ while $\nu(F_k) = \nu(E_k)$. Choose any $y \in F_\infty$. Then for every k, $y \notin F_{k+1}$, no images of y reach $F - \bigcup_{i=1}^{k} F_i$ before reaching $E - \bigcup_{i=1}^{k} E_i$, and a fortiori before reaching E_∞. It follows that no image of y can be in E_∞. This contradicts ergodicity.

Now set $Q_0 = \bigcup_{k=1}^{\infty} E_k$, $Q_1 = E - Q_0$. Clearly each F_k is in $\mathcal{Q}'^{0}_{-\infty}$, so R_0 is. But so is E, which completes the argument. □

LEMMA 4. *In order to prove that a particular partition $\mathcal{P}$ is LB for a transformation T, it suffices to find for each $\varepsilon > 0$ some partition $\mathcal{Q} \supset \mathcal{P}$ so that for all sufficiently big N there exist arbitrarily big M for which the LB definition holds in the form* (a′), (b′) *for that particular N, M, ε.*

PROOF. It suffices to show that $\mathcal{P}$ satisfies the LB definition in the form (a′), (b′) for every $\varepsilon > 0$. To get this: an examination of the proofs of Lemma 1 in Section 3 and Corollary 7 of Proposition 1 in Section 3 shows that it suffices to get $\mathcal{Q}$ satisfying LB in the form (a), (b′) for $(\varepsilon/2)^4$. Replacing $(\varepsilon/2)^4$ by $(\varepsilon/2)^8$ enables us to replace "all M" by "arbitrarily large M" in the definition; again see the proof of Corollary 2 of Proposition 1 in Section 3. Finally, using $(1/3)(\varepsilon/2)^8$ instead makes it possible to replace (a) by the weaker condition (a′). □

PROOF OF PART (c) OF THEOREM 3. In view of Proposition 2 of Section 3 and Lemma 3 above, it need only be shown that if T_E is LB and $\mathcal{P}$ is a finite partition of E then $\mathcal{P}' = \mathcal{P} \cup \{Y - E\}$ is LB. Applying Lemmata 1 and 2 above, it may further be assumed that $\mathcal{P}^0_{-\infty}$ contains the return-time partition $\{E_1, E_2, \cdots\}$ of E. In view of Lemma 4, it will suffice to find, for each $\varepsilon' > 0$, some $\mathcal{Q} \supset \mathcal{P}$ so that $\mathcal{Q}'$ satisfies (a′) and (b′) for ε', in the weakened form described there, where only "large" M' need be considered.

Let ε' be given. Let ε be chosen > 0, depending on ε in a manner to be specified later. Since $\Sigma_{l=1}^{\infty} l\nu(E_l) = 1$, an integer K may be chosen so large that $\Sigma_{l>K} l\nu(E_l) < \varepsilon$. Let $\mathcal{Q} = \mathcal{P} \vee \{E_1, \cdots, E_K, E^K\}$ where $E^K = \bigcup_{l>K} E_l$. This $\mathcal{Q}$ will be shown to do the trick, provided ε is sufficiently small. As before, let $\mathcal{Q} = \{Q_a : a \in \mathbf{I}\}$ and $\mathcal{Q}' = \{Q_a : a \in \mathbf{I}'\}$, where $\mathbf{I}' = \mathbf{I} \cup \{0\}$ and $Q_0 = Y - E$.

Again using the fact that $\Sigma_{l=1}^{\infty} l\nu(E_l) = 1$, there is for every $\varepsilon > 0$ some $\delta(\varepsilon)$, which for convenience will be taken $< \varepsilon$, so that if $A \subset E$ and $\nu(A) < \delta$ then $\Sigma_{l>L} l\nu(E_l \cap A) < \varepsilon$. Define $p: Y \to E$ by $py = T^{-j}y$ where j is the smallest nonnegative integer with $T^{-j}y \in E$; thus $p^{-1}(y) = \{y, Ty, \cdots, T_y^{l-1}\}$ if $y \in E_l$. Then $\nu(A) < \delta(\varepsilon) \Rightarrow \nu(p^{-1}A) < \varepsilon$.

Also let $\{\xi_j\}$, $\{\xi'_j\}$, $\{t_j\}$ be defined as in the proof of part (b), so $\xi'_{t_j} \mid E = \xi_j$.

Let $\lambda = \inf\{\nu(E_l)\colon \nu(E_l) > 0, l \leqq K\}$. Let $\delta = (\frac{1}{2}\lambda\delta(\varepsilon))^4$. Since $\mathcal{Q}$ is LB for T_E, there is some "big enough" N_0 for $\mathcal{Q}$ and δ; also make $N_0 > K/\varepsilon$; and finally, make N_0 so large that if $N \geqq N_0$ then on a set F' of measure $> 1 - \delta$, setting $\alpha' = (\xi'_1, \cdots, \xi'_N)$, the symbols of α' have a "typical" distribution, in that there are at least $(1-\delta)\nu(E_l)N$ symbols in α' which come from $\mathcal{Q}'$-atoms contained in E_l, for $l = 1, \cdots, K$. For convenience, let us call such a symbol a *symbol of height* l; that is, $a \in \mathbf{I}'$ is a symbol of height l if $Q_a \subset E_l$, $l = 1, \cdots, K$. Thus there may be chosen a substring $\bar{\alpha}$ of α' consisting of N_l symbols of height l, where N_l is the smallest integer $\geqq (1-\delta)(E_l)N$. F' is obviously a $\mathcal{Q}_1'^N$-measurable set; let $\mathbf{F}'$ be the corresponding subset of $\mathbf{I}'^N$: that is, those strings $\alpha' \in \mathcal{Q}_1'^N$ which have the described property.

Having chosen some $N \geqq N_0$, choose M_0 large enough that $\nu(t_0 \geqq -M_0) \geqq 1 - \delta$. Then it will be shown that if ε is sufficiently small, $\mathcal{Q}'$ will satisfy (a') and (b') for ε', all $N \geqq N_0$, and all $M' \geqq M_0$.

Now let L be chosen so large that $\nu(G) \geqq 1 - \delta$, where $G = \{\xi_j \in \bigcup_{i \leqq L} E_i$ for $-M' \leqq j \leqq N\}$. Let $\mathcal{E}$ be the partition $\{E_1, \cdots, E_L, E^L\}$, where again $E^L = \bigcup_{l > L} E_l$. By a *controlled* atom of $\mathcal{E}_j^k$ will be meant one with no $T^{-i}E^L$ factor.

Since $\mathcal{Q}_{-\infty}^0 \supset \{E_1, E_2, \cdots\}$, M may be chosen so large that the following holds: there is a set of atoms A of $\mathcal{Q}_{-M}^0$, of total measure $> 1 - \delta$, and for each such atom A a collection $\mathbf{H}(A)$ of sequences $\alpha \in \mathbf{I}^N$, with $\nu((\xi_1, \cdots, \xi_N) \in \mathbf{H}(A) \mid A) > 1 - \delta$, and such that

1) A has a proportion $\geqq 1 - \delta$ of its measure contained in some controlled atom B of $\mathcal{E}_{-M'-1}^0$. Write B as $\bigcap_{j=-M'-1}^0 T^{-j}E_{h_j(A)}$, where $1 \leqq h_j(A) \leqq L$.

2) $\{(\xi_1, \cdots, \xi_N) = \alpha\}$ has a proportion $\geqq 1 - \delta$ of its $\nu(\cdot \mid A)$ measure in some controlled atom C of $\mathcal{E}_{-M'-1}^N$ with $C \subset B$; C then may be written as $\bigcap_{j=-M'-1}^N T^{-j}E_{h_j(A,\alpha)}$ where $1 \leqq h_j(A, \alpha) \leqq L$ and $h_j(A, \alpha) = h_j(A)$ for $j \leqq 0$.

Let $\mathcal{G}$ be those "good" atoms of $\mathcal{Q}_{-M}^0$, with respect to N and δ, which also satisfy (1) and (2). Thus $\mathcal{G}$ has total measure $\geqq 1 - 2\delta$.

An atom A of $\mathcal{Q}_{-M}^0$ will be said to *induce* A' in $\mathcal{Q}'^0_{-M'}$ if the sequence of $\mathbf{I}$-symbols in the name of A' are precisely the rightmost of the symbols in the name of A; that is, writing j for the largest integer such that $t_j \geqq -M'$ on all of A', the (constant) value of $(\xi'_{t_j}, \cdots, \xi'_{t_0})$ on A' is precisely the (constant) value of $(\xi_j, \cdots, \xi_0)$ on A. If also $\alpha \in \mathbf{I}^N$ and $\alpha' \in \mathbf{I}'^N$, say (A, α) induces (A', α') if furthermore the subsequence of $\mathbf{I}$-names in α' are precisely the leftmost elements in α.

Let A' be an atom of $\mathcal{Q}_{-M'}^{0}$ on which $-M' \leqq t_0$. If $A \in \mathcal{G}$, then say A induces A' *properly* if A induces A', and the **I**-terms in the name of A' appear according to the rule $t_0 \geqq -h_0(A)$, $t_1 = t_0 - h_1(A), \cdots, t_j = t_{j-1} - h_j(A)$, $t_{j+1} < -M'$ everywhere on A', and $t_j + M' < h_{j+1}(A)$; that is, in a manner consistent with the controlled $\mathcal{E}_{-M'-1}^{0}$ atom which contains most of A. Similarly, if $A \in \mathcal{G}$ and $\alpha \in \mathbf{H}(A)$, say (A, α) induces (A', α') properly if (A, α) induces (A', α'), and the gaps are consistent with the controlled $\mathcal{E}_{-M'-1}^{N}$ atom which contains most of $\{(\xi_1, \cdots, \xi_N) = \alpha, A\}$, so that the first **I**-symbol in α' appears in the place $h_0(A) + t_0$, where t_0 is the place of the last **I**-symbol in the name of A, and subsequent **I**-symbols in α' have gaps $h_1(A, \alpha) - 1$, $h_2(A, \alpha) - 1, \cdots$ between them.

Notice that an A in $\mathcal{G}$ induces properly exactly $h_0(A)$ different $A' \in \mathcal{Q}_{-M'}^{0}$ corresponding to different values of t_0 on A'; and given one of these A', and given $\alpha \in \mathbf{H}(A)$, there is then exactly one α' in $\mathbf{I}'^{N}$ such that (A, α) induces (A', α') properly. Notice also that if A induces A' properly, then $\nu(A) \geqq \nu(p^{-1}A \cap A') \geqq (1-\delta)\nu(A)$, and if (A, α) induces (A', α') properly then

$$\begin{aligned}\nu((\xi_1, \cdots, \xi_N) = \alpha, A) &\geqq \nu((\xi'_{t_1}, \cdots, \xi'_{t_N}) = \alpha, p^{-1}A \cap A') \\ &\geqq (1-\delta)\nu((\xi_1, \cdots, \xi_N) = \alpha, A).\end{aligned}$$

Here is the argument: p is 1-1 on A' because $t_0 \geqq -M'$, and clearly measure-preserving, so the left hand inequalities are obvious in both cases. As for the right hand inequalities: let j be the largest integer for which $t_j \geqq -M'$ on A'. Then $pA' \supset A \cap D$, where D is an $\mathcal{E}_{-j-1}^{0}$-measurable set, which, by definition of proper inducing, must contain the $\mathcal{E}_{-M'-1}^{0}$ atom in which A mainly lives; thus

$$\nu(A \cap pA') \geqq (1-\delta)\nu(A);$$

and a similar argument holds for the other case.

Next, let $\mathcal{G}'$ be those atoms A' in $Q_{-M'}^{0}$ which have a proportion $\geqq 1 - \sqrt{\delta}$ of their measures filled up by $\cup\{p^{-1}A : A$ induces A' properly$\}$, and for which $\nu((\xi'_1, \cdots, \xi'_N) \in \mathbf{F}' \mid A') \geqq 1 - \sqrt{\delta}$, and for which also $t_0 \geqq -M'$. Then $\mathcal{G}'$ has total measure $\geqq 1 - \sqrt{\delta}$.

Now, for A', B' in $\mathcal{G}'$, a measure $\tilde{n}_{A', B'}$ on $\mathbf{I}'^{N} \times \mathbf{I}'^{N}$ will be defined by the formula

$$\tilde{n}_{A', B'}(\alpha', \beta') = \sum_{A, B, \alpha, \beta} n_{A, B}(\alpha, \beta)\nu(p^{-1}A \mid A')\nu(p^{-1}B \mid B'),$$

where the $n_{A, B}$ come from the LB definition for $\mathcal{Q}$, and where the range of the summation is as follows: $A \in \mathcal{G}$, $\alpha \in \mathbf{H}(A)$, (A, α) inducing (A', α') properly,

$B \in \mathcal{G}$, $\beta \in \mathbf{H}(B)$, (B, β) inducing (B', β') properly. It will be shown that if ε is chosen small enough then the $\tilde{n}_{A', B'}$, after normalization to total mass 1, will satisfy (a') and (b'). This will be done as was done in the proof of part (b), by showing that they satisfy appropriate inequalities, and that the error in total variation introduced in these inequalities is sufficiently small.

If $\tilde{n}_{A', B'}(\alpha', \beta')$ is summed over all β', then since each (B, β) can induce (B', β') for only *one* β', any given β only appears at most once in the expanded sum, so

(i)
$$\tilde{n}_{A', B'}(\alpha' \times \mathbf{I}'^{N}) \leqq \sum_{A, B, \alpha} n_{A, B}(\alpha \times \mathbf{I}'^{N}) \nu(p^{-1}A \mid A') \nu(p^{-1}B \mid B')$$
$$= \sum_{A, B, \alpha} \nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A) \nu(p^{-1}A \mid A') \nu(p^{-1}B \mid B').$$

The range of summation on A and α is as before, while now B ranges over $\mathcal{G}$, subject to B inducing B' properly. The B are disjoint, so the $p^{-1}B$ are, and

(ii)
$$\sum_{A, B, \alpha} \nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A) \nu(p^{-1}A \mid A') \nu(p^{-1}B \mid B')$$
$$\sum_{A, \alpha} \nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A) \nu(p^{-1}A \mid A').$$

An estimate of three paragraphs back gives:

$$\nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A) = \frac{\nu_E((\xi_1, \cdots, \xi_N) = \alpha, A)}{\nu_E(A)}$$
$$= \frac{\nu((\xi_1, \cdots, \xi_N) = \alpha, A)}{\nu(A)} \leqq \frac{1}{1 - \delta} \frac{\nu((\xi'_{i_1}, \cdots, \xi'_{i_N}) = \alpha, p^{-1}A \cap A')}{\nu(p^{-1}A \cap A')}.$$

So one may write

(iii)
$$\sum_{A, \alpha} \nu_E((\xi_1, \cdots, \xi_N) = \alpha \mid A) \nu(p^{-1}A \mid A)$$
$$\leqq \frac{1}{1 - \delta} \sum_{A, \alpha} \nu((\xi'_{i_1}, \cdots, \xi'_{i_N}) = \alpha, p^{-1}A \mid A').$$

The range of summation on A and α is what it was to begin with. If however this is enlarged a bit, so that all $A \in \mathcal{Q}^0_{-M}$ and all $\alpha \in \mathbf{I}^N$ for which (A, α) induces (A', α') are permitted, then one gets an inequality for the last expression:

(iv)
$$\frac{1}{1-\delta}\sum_{A,\alpha}\nu((\xi'_{i_1},\cdots,\xi'_{i_N})=\alpha,p^{-1}A\mid A')$$
$$\leqq\frac{1}{1-\delta}\nu((\xi'_1,\cdots,\xi'_N)=\alpha'\mid A').$$

Thus:

$$\tilde{n}_{A',B'}(\alpha'\times\mathbf{I}'^N)\leqq\frac{1}{1-\delta}\nu((\xi'_1,\cdots,\xi'_N)=\alpha'\mid A')$$

and similarly

$$\tilde{n}_{A',B'}(\mathbf{I}'^N\times\beta')\leqq\frac{1}{1-\delta}\nu(\xi'_1,\cdots,\xi'_N)=\beta'\mid B').$$

In order to see that the excess is not large, it is necessary to see what mass has been thrown in at each of the inequalities (i) $\cdots$ (iv).

INEQUALITY (i). The total mass of the omission is precisely

$$\sum_{\alpha'}\sum_{A,B,\alpha,\beta}n_{A,B}(\alpha,\beta)\nu(p^{-1}A\mid A')\nu(p^{-1}B\mid B').$$

The inside summation is over $A\in\mathcal{G}$, $\alpha\in\mathbf{I}(A)$, (A,α) inducing (A',α') properly, $B\in\mathcal{G}$, B inducing B' properly, and $\beta\notin\mathbf{H}(B)$. This is dominated (using estimates like before) by

$$\sum_{A,B}\sum_{\beta\notin\mathbf{H}(B)}\nu((\xi_1,\cdots,\xi_N)=\beta\mid B)\nu(p^{-1}A\mid A')\nu(p^{-1}B\mid B').$$

Since $\nu((\xi_1,\cdots,\xi_{N'})\notin\mathbf{H}(B)\mid B)<\delta$, the entire sum is less than δ.

INEQUALITY (ii). The excess mass is precisely

$$\nu(B'-\cup^{-1}B\mid B')\Big(\sum_{A,\alpha}\nu_E((\xi_1,\cdots,\xi_N)=\alpha\mid A)\nu(p^{-1}A\mid A')\Big)$$

$$\leqq\nu(B-\cup p^{-1}B\mid B')\frac{1}{1-\delta}\nu((\xi'_1,\cdots,\xi'_N)=\alpha'\mid A')\leqq\nu(B'-\cup p^{-1}B\mid B'),$$

using previous inequalities.

B ranges over all atoms of $\mathcal{G}$ which induce B' properly. Thus, since $B'\in\mathcal{G}'$, it follows that $\nu(B'-\cup p^{-1}B\mid B')\leqq 1-\sqrt{\delta}$.

INEQUALITY (iii). The total mass of the difference may be dominated by $1/(1-\delta)$ times

$$1 - \nu(\cup\{p^{-1}B : B \in \mathcal{G} \text{ and induces } B \text{ properly}\} \mid B'),$$

which is less than $\sqrt{\delta}$.

INEQUALITY (iv). The difference may be dominated by $1/(1-\delta)$ times the sum of two terms:

1) $\nu(A' - \cup\{p^{-1}A : A \in \mathcal{G}, A \text{ induces } A' \text{ properly}\} \mid A')$, which was already seen to be $< \delta$.

2) $$\sum_{A \in \mathcal{G}} \nu((\xi'_{t_1}, \cdots, \xi'_{t_N}) \notin \mathbf{H}(A), p^{-1}A \mid A')$$
$$\leqq \frac{1}{1-\delta} \sum_{A \in \mathcal{G}} \nu((\xi_{t_1}, \cdots, \xi_{t_N}) \notin \mathbf{H}(A) \mid A)\nu(p^{-1}A \mid A') \leqq \frac{\delta}{1-\delta}.$$

Sufficiently small choice of ε will now guarantee that the normalized version $n_{A',B'}$ of $\tilde{n}_{A',B'}$ will have property (a′) for T, $\mathcal{Q}'$, N, M' and ε'.

Finally, it is necessary to estimate $\tilde{n}_{A',B'}(\{(\alpha', \beta') : \delta(\alpha', \beta') \geqq \varepsilon'\})$. This may be estimated from above by the sum of three terms:

$$n_{A',B'}(\{(\alpha', \beta') : \alpha', \beta' \in \mathbf{F}', \delta(\alpha', \beta') \geqq \varepsilon'\})$$
$$+ \tilde{n}_{A',B'}(\mathbf{I}'^N \times (\mathbf{I}'^N - \mathbf{F}'))$$
$$+ \tilde{n}_{A',B'}((\mathbf{I}'^N - \mathbf{F}') \times \mathbf{I}'^N).$$

The last two terms are dominated by $(1/(1-\delta))\nu(\mathbf{I}'^N - \mathbf{F}' \mid B')$ and $(1/(1-\delta))\nu(\mathbf{I}'^N - \mathbf{F}' \mid A')$ respectively from the inequalities already noted.

As for the first term: if α' and β' are in $\mathbf{F}'$ and α and β are in $\mathcal{Q}_1^N$ and induce α' and β' respectively, then for each $l = 1, \cdots$, α and β each contain $\geqq N$ symbols of height l. Here, as defined earlier, N_l is the smallest integer $\geqq (1-\delta)\nu(E_j)N$. Suppose now that α and β can be matched well enough that there are no more than δN unmatched symbols in each string. Then there are no more than $\varepsilon\nu(E_j)N$ unmatched symbols among those N_l symbols of α and β which correspond to the chosen N_l symbols of height l in α' and β', $j = 1, \cdots, K$. Thus there is a match induced from α' to β' which matches at least $(1-\delta-\varepsilon)\nu(E_j)N$ symbols of $\mathcal{Q}'$ of height l, $l = 1, \cdots, K$. Now, if a symbol of $\mathcal{Q}'$ in α has height l, and α' is a string which can really occur, then the next $j-1$ symbols are zeros. Thus the match from α' to β' may be extended to a total of at least $(1-\delta-\varepsilon)\sum_{j=1}^{K} j\nu(E_j)N - (L-1)$ pairs from (α', β'). (The reason for subtracting $L-1$ is that the rightmost pair which is matched might not have enough space left before the ends of the α' and β' strings to get in any more pairings.) So α and β' may be matched better than

$$(1-\delta-\varepsilon)\sum_{j=1}^{k} j\nu(E_j)-\frac{K-1}{N}\geqq(1-\delta-\varepsilon)(1-\varepsilon)-\frac{K-1}{N_0}\geqq(1-2\varepsilon)(1-\varepsilon)-\varepsilon.$$

Summarizing the last paragraph: if α' and β' are in $\mathbf{F}'$, and $\delta(\alpha,\beta)\leqq\delta$, then $\delta(\alpha',\beta')\leqq 1-(1-2\varepsilon)(1-\varepsilon)-\varepsilon$, which of course can be made less than ε' by choosing ε small. So

$$\tilde{n}_{A',B'}(\{(\alpha',\beta')\colon \alpha',\beta'\in\mathbf{F}', \delta(\alpha',\beta')\geqq\varepsilon'\})$$
$$\leqq \sum_{A,B,\alpha,\beta} n_{A,B}(\{(\alpha,\beta)\colon \delta(\alpha,\beta)\geqq\delta\})\nu(p^{-1}A\mid A')\nu(p^{-1}B\mid B')\leqq\delta.$$

Now a sufficiently small choice of ε (and consequently of δ) completes the argument. □

COROLLARY 1. *If $\hat{T}$ is LB then so is T.*

PROOF. $(\hat{T})_{P_1\times Y}$ is isomorphic in an obvious way to $S_{P_1}\times T$, which has T as a factor. □

COROLLARY 2. *If $\hat{T}$ is Bernoulli, then T is LB.*

PROOF. $\hat{T}$ Bernoulli $\Rightarrow$ $\hat{T}$ LB $\Rightarrow$ T LB by the previous corollary. □

5. Examples of non LB automorphisms

THEOREM 4. *There exists an ergodic automorphism T_0 of entropy zero which is not LB.*

PROOF. The method will be to make longer and longer strings of symbols, putting them together by the stacking method (cf. Friedman [3]). Underlying the construction is the observation that no two of the following strings can be matched very well:

$$a\ b\ a\ b\ a\ b\ a\ b,$$
$$a\ a\ b\ b\ a\ a\ b\ b,$$
$$a\ a\ a\ a\ b\ b\ b\ b.$$

Step I. Description of the strings. Let $\mathbf{I}_1$ be an alphabet of $N(1)$ symbols $\{a_{1,1},\cdots,a_{1,N(1)}\}$. I inductively construct sets $\mathbf{I}_{n+1}=\{a_{n+1,1},\cdots,a_{n+1,N(n+1)}\}$ of strings of length $L(n+1)$, formed by stringing together members of $\mathbf{I}_n$. The strings $a_{n,j}$ are called n-symbols. The construction of $(n+1)$-symbols goes like this:

$$a_{n+1,1}=(a_{n,1}^{N(n)^2}\cdots a_{n,N(n)}^{N(n)^2})^{N(n)^{2N(n+1)}}$$

$$a_{n+1,2}=(a_{n,1}^{N(n)^4}\cdots a_{n,N(n)}^{N(n)^4})^{N(n)^{2(N(n+1)-1)}}$$

$$\vdots$$

$$a_{n+1,N(n+1)}=(a_{n,1}^{N(n)^{2N(n+1)}}\cdots a_{n,N(n)}^{N(n)^{2N(n+1)}})^{N(n)^2}.$$

Thus $L(1)=1$, and $L(n+1)=L(n)N(n)^{2(N(n+1)+1)}$.

LEMMA. *If $N(n)>200\times 2^{n+2}$, then for all positive integers r and s and all n, $a^r_{n,i}$ and $a^s_{n,j}$ cannot be fitted better than* $(1/8)\ (1-(1/2^n))$ *if $i<j$.*

The argument is by induction, and is obvious for $n=1$. Suppose, then, that we have a match between $a^r_{n+1,i}$ and $a^s_{n+1,j}$. Write $j=i+k$, with $k>0$. Write α_h for $a^{N(n)^{2i}}_{n,h}$. Then

$$a^r_{n+1,i}=(\alpha_1\cdots\alpha_{N(n)})^{N(n)^{2(N(n+1)-i)}r}$$

and

$$a^s_{n+1,j}=(\alpha_1^{N(n)^{2k}}\cdots\alpha_{N(n)}^{N(n)^{2k}})^{N(n)^{2(N(n+1)-j)}s}.$$

Now: $a^s_{n+1,j}$ may be factored into $N(n)^{2(N(n+1)-j)}s$ strings of the form $\alpha_h^{N(n)^{2k}}$. The given match between $a^r_{n+1,i}$ and $a^s_{n+1,j}$ may be split into matches between the successive $\alpha_h^{N(n)^{2k}}$ and $N(n)\times N(n)^{2(N(n+1)-k)}s$ certain disjoint substrings of $a^r_{n+1,i}$ whose union is all of $a^r_{n+1,i}$. If we can show that each of these induced matches has a fit of no better than $(1/8)\ (1-(1/2^{n+1}))$, then the original match likewise will have a fit of no better than $(1/8)\ (1-(1/2^{n+1}))$. Now, substrings of $a^r_{n+1,i}$ are of the form $\beta(\alpha_1\cdots\alpha_{N(n)})^t\gamma$, where β and γ are respectively final and initial substrings of the string $\alpha_1\cdots\alpha_{N(n)}$ and t is a nonnegative integer. If we "complete" β and γ, thus enlarging the substring $\beta(\alpha_1\cdots\alpha_{N(n)})^t\gamma$ to $(\alpha_1\cdots\alpha_{N(n)})^{t+2}$, the total length $|\beta(\alpha_1\cdots\alpha_{N(n)})^t\gamma|+|\alpha^{N(n)2k}|$ will be changed to $|(\alpha_1\cdots\alpha_{N(n)})^{t+2}|+|\alpha_h^{N(n)^{2k}}|$. The added length is less than $2JN(n)$, where J is the common length of all the α_h. So, if we regard our match between $\alpha_h^{N(n)^{2k}}$ and $\beta(\alpha_1\cdots\alpha_{N(n)})^t\gamma$ as rather a match between $\alpha_h^{N(n)^{2k}}$ and $(\alpha_1\cdots\alpha_{N(n)})^{t+2}$, the fit is decreased, but by a factor of no less than

$$1-\frac{2JN(n)}{|\alpha_h^{N(n)^{2k}}|}>1-\frac{2}{N(n)}>1-\frac{1}{100}\frac{1}{2^{n+1}}.$$

So it will suffice to show that any match between an $\alpha_h^{N(n)^{2k}}$ and $(\alpha_1\cdots\alpha_{N(n)})^{t+2}$, $t\geqq 0$, fits no better than

$$\frac{1}{8}\left(1-\frac{1}{2^{n+1}}\right)\left(1-\frac{1}{100}\frac{1}{2^{n+1}}\right).$$

Repeat the above trick: $(\alpha_1\cdots\alpha_{N(n)})^{t+2}$ is a product of $N(n)(t+2)$ strings of the form a_l, l varying. Any match between $\alpha_h^{N(n)^{2k}}$ and $(\alpha_1\cdots\alpha_{N(n)})^{t+2}$ may be decomposed into matches between these $N(n)(t+2)$ successive α_l and $N(n)\cdot(t+2)$ certain disjoint substrings of $\alpha_h^{N(n)^{2k}}$ whose union is all of $\alpha_h^{N(n)^{2k}}$. Any substring of $\alpha_h^{N(n)^{2k}}$ has the form $b\,a^u_{n,h}c$, where b is a final substring of $a_{n,h}$ and c an initial substring of $a_{n,h}$. If we "complete" the string $b\,a^u_{n,h}c$ to obtain the string $a_{n,h}^{u+2}$, then the total length $|ba_{n,h}^{u}c| + |\alpha_l|$ changes by less than $2L(n)$, and if we regard our given match between α_l and $ba_{n,h}^{u}c$ as one between α_l and $a_{n,h}^{u+2}$, the fit is decreased, but by a factor of no less than

$$1-\frac{2K}{|\alpha_l|}=1-\frac{2K}{KN(n)^{2i}}>1-\frac{1}{100}\frac{1}{2^{n+1}};$$

here K is the common length of the $a_{n,h}$.

Now, any match between α_l and a string of the form $a_{n,h}^{u+2}$ with $h \neq l$ fits no better than $(1/8)\,(1-(1/2^n))$, by induction hypothesis. The strings α_l for which $l = h$ form a fraction $1/N(n)$ of the total length of the string $(\alpha_1 \cdots \alpha_{N(n)})^{t+2}$. So we are guaranteed that the total match between $(\alpha_1 \cdots \alpha_{N(n)})^{t+2}$ and $\alpha_h^{N(n)^{2k}}$ has fit no better than

$$\frac{1}{8}\left(1-\frac{1}{2^n}\right)+\frac{2}{100}\frac{1}{2^{n+1}},$$

which is less than

$$\frac{1}{8}\left(1-\frac{1}{2^{n+1}}\right)\left(1-\frac{1}{100}\frac{1}{2^{n+1}}\right)^2. \qquad \square$$

Step II. Construction of T_0. Take the unit interval, which we call F_1, and partition it into $N(1)$ equal intervals: $\mathcal{Q} = \{Q_a : a \in \mathbf{I}_1\}$. Inductively, suppose we have constructed intervals F_j and partially defined automorphisms T_j, $j = 1, \cdots, n$ satisfying

1) $F_{j+1} \subset F_j$ and $T_{j+1} \supset T_j$,

2) T_j maps F_j piecewise-linearly and disjointly for $L(n)-1$ steps, with

$$\bigcup_{m=0}^{L(j)-1} T_j^m F_j = F_1, \text{ and } T_j \text{ undefined on } T_j^{L(n)-1}F_j,$$

3) the $L(j)$-names of points in F_j are, with equal probability, the $N(j)$ j-symbols in $\mathbf{I}_j$.

Divide F_n into $N(n)$ equal intervals $A^1, \cdots, A^{N(n)}$ according to which n-symbol the $L(n)$-name takes on. Divide each A^i into $M(n) = N(n)^{2N(n+1)}$ equal intervals $A_1^i, \cdots, A_{M(n)}^i$, and further divide each A_j^i into $N(n+1)$ equal intervals $A_{j,1}^i, \cdots, A_{j,N(n+1)}^i$. $M(n)$ is exactly the number of times each n-symbol occurs in each $(n+1)$-symbol.

Set $B_{j,k}^i = T_n^{L(n)-1}A_{j,k}^i$: the portion of the "roof" of the nth stack which lies above the portion $A_{j,k}^i$ of the floor F_n. Set $F_{n+1} = \bigcup_{k=1}^{N(n+1)} A_{1,k}^1 = A_1^1$. T_{n+1} will be defined on F_{n+1} for $L(n+1)-1$ steps, agreeing with T_n wherever that has already been defined. Here is the prescription: send $A^1_{1,k}$ "upward" via T_n for $L(n)-1$ steps, arriving at $T_n^{L(n)-1}A^1_{1,k}=B^1_{1,k}$. Map $B^1_{1,k}$ linearly onto $A^1_{2,k}$. Send $A^1_{2,k}$ "upward" again via T_n for $L(n)-1$ more steps, to $B^1_{2,k}$; in general: having started at $A^1_{1,k}$, and having arrived at $B^i_{j,k}$, after some number—say l—of trips from the floor to the roof, so that l n-symbols have been traversed in the process, send $B^i_{j,k}$ linearly onto $A^{i'}_{j',k'}$, where i' is the subscript of the $(l+1)$th n-symbol in the $(n+1)$-symbol $\mathcal{A}_{n+1,k}$, and where $j'-1$ of the previously traversed n-symbols have had the subscript i'. In this manner, eventually one arrives at $B^{N(n)}_{M(n),k}$, having in the process traversed, for each i and j, the column from $A^i_{j,k}$ to $B^i_{j,k}$. It takes $L(n+1)-1$ steps to go from $A^1_{1,k}$ to $B^{N(n)}_{M(n),k}$. Doing this for each k gives a stack of height $L(n+1)$ with base F_{n+1}.

It may easily be checked that the hypotheses (1), (2), (3) still hold. The common extension T_0 of the T_n is defined on a set of measure 1, and T_0 clearly has the property that for each n the interval F_n moves disjointly for $L(n)-1$ steps, and F_n splits into $N(n)$ equal subintervals, the points of which take on with equal probability the n-symbols as their $L(n)$-names.

Step III. T_0 has zero entropy. There are at most $L(n)N(n)^2$ different strings of length $L(n)$; some may be repetitions of others. This is clear from looking at the tower based on F_n and considering various starting points. So

$$h(T)\leqq \lim_{n\to\infty}\frac{1}{L(n)}\log L(n)N(n)^2$$

$$=\lim_{n\to\infty}\left(\frac{\log L(n)}{L(n)}+\frac{2\log N(n)}{L(n)}\right).$$

But since

$$\frac{N(n)}{L(n)}\to 0,\quad h(T)=0.$$

Step IV. T_0 is ergodic. This may be seen directly, by counting k-blocks in n-strings.

Step V. T_0 is not LB. If T_0 were LB, then for some n there would be a subset A of the unit interval, of measure $>99/100$, such that for any y_1, y_2 in A, the $L(n)$-name of y_1 and y_2 match better than 99/100. Now, an $L(n)$-name consists of the tail of an n-symbol followed by the beginning of a (possibly different) n-symbol. It is easy to see, from the magnitudes involved, that y_1 and y_2 in A may be so chosen that the one or two n-symbols making up the name of y_1 are different from those making up the name of y_2; then, from the lemma, the two $L(n)$-names cannot be matched better than 1/4. ∎

COROLLARY 1. *T_0, although of entropy zero, is not Kakutani-equivalent to a* *irrational rotation; in other words, no flow built over T_0 can be isomorphic to a flow built over an irrational rotation.*

PROOF. Immediate from Theorems 3 and 4. □

COROLLARY 2. *$\hat{T}_0$ is a K-automorphism which is not LB. Consequently it is not Kakutani-equivalent to a Bernoulli transformation, and no flow over it can be time-changed to a Bernoulli flow.*

PROOF. Immediate from Theorem 4 and Corollary 1 to Theorem 2. □

COROLLARY 3. *There exists a K-flow which cannot be time-changed to a Bernoulli flow.*

PROOF. In view of Corollary 2, it suffices to show that there exists a K-flow over $\hat{T}_0$. But this will follow from the next observation, which is perhaps worth isolating as a Proposition.

PROPOSITION. *For any ergodic automorphism T, there exists a K-flow over $\hat{T}$.*

PROOF. $\hat{T}$ is a K-automorphism, and we let $\mathcal{S}$ be the "K-partition" generated by the past of $\hat{\mathcal{Q}}$ (see Section 2). Let $f(x,y) = p$ if $x \in P_0$ and q if $x \in P_1$, where p and q are incommensurable positive real numbers. Then the results of Gurevič [4] may be applied to show that the flow over $\hat{T}$ under f is a K-flow. □

SOME OPEN PROBLEMS

A) Corollary 2 to Theorem 3 says that if $\hat{T}$ is Bernoulli then T is LB. Is the converse true?

B) It is my feeling that "most" ergodic automorphisms are *not* LB. It would be interesting to formulate and prove a precise result along these lines, or to find some "natural", i.e., algebraic or geometric examples of non LB automorphisms.

C) In connection with the last proposition: is it the case that every K-automorphism, or perhaps even every ergodic automorphism of entropy > 0, can be the base of some K-flow? It *is* known, for example, that every ergodic automorphism can serve as the base for some flow with continuous spectrum.

Note added in proof, February 21, 1976. In connection with Problem C, D. Ornstein and M. Smorodinsky have now shown that in fact any ergodic automorphism of positive entropy induces a K-automorphism, and any ergodic flow of positive entropy may be time-changed to a K-flow.

References

1. R. Adler and P. Shields, *Skew products of Bernoulli shifts with rotations II*, Israel J. Math. **19** (1974), 228–236.

2. P. Billingsly, *Ergodic Theory and Information*, Wiley, New York, 1965.

3. N. Friedman, *Introduction to Ergodic Theory*, Van Nostrand Reinhold, New York, 1970.

4. B. M. Gurevič, *Some existence conditions for K-decompositions for special flows*, Trans. Moscow Math. Soc. **17** (1967), 99–128.

5. S. Kakutani, *Induced measure-preserving transformations*, Proc. Imp. Acad. Tokyo **19** (1943), 635–641.

6. W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. **199** (1970), 453–464; *Erratum*, ibid. **168** (1972), 549.

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

8. D. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems*, Yale University Press, New Haven, 1970.

9. D. Ornstein and L. Sucheston, *An operator theorem on L_1 convergence to zero with applications to Markov kernels*, Ann. Math. Statist. **41** (1970), 1631–1639.

10. D. Ornstein and B. Weiss, *Finitely determined implies very weak Bernoulli*, Israel J. Math. **17** (1974), 95–104.

11. L. Swanson, *Induced automorphisms and factors in ergodic theory*, Ph.D. thesis, U. C. Berkeley, 1975 (to be submitted to Advances in Math., under title *Induced automorphisms and Bernoulli shifts*).

12. S. Ulam, *Some combinatorial problems studied experimentally on computing machines*, in *Applications of Number Theory to Numerical Analysis* (S. K. Zaremba, ed.), Academic Press, New York, 1972.

Department of Mathematics
University of California
Berkeley, California 94720 U. S. A.