

Standardness of sequences of σ -fields given by certain endomorphisms

by

Jacob Feldman (Berkeley, Calif.) and
Daniel J. Rudolph (College Park, Md.)

Abstract. Let E be an ergodic endomorphism of the Lebesgue probability space $\{X, \mathcal{F}, \mu\}$. It gives rise to a decreasing sequence of σ -fields $\mathcal{F}, E^{-1}\mathcal{F}, E^{-2}\mathcal{F}, \dots$. A central example is the one-sided shift σ on $X = \{0, 1\}^{\mathbb{N}}$ with $\frac{1}{2}, \frac{1}{2}$ product measure. Now let T be an ergodic automorphism of zero entropy on (Y, ν) . The $[I|T]$ endomorphism is defined on $(X \times Y, \mu \times \nu)$ by $(x, y) \mapsto (\sigma(x), T^{x(1)}(y))$. Here \mathcal{F} is the σ -field of $\mu \times \nu$ -measurable sets. Each field is a two-point extension of the one beneath it. Vershik has defined as “standard” any decreasing sequence of σ -fields isomorphic to that generated by σ . Our main results are:

THEOREM 2.1. *If T is rank-1 then the sequence of σ -fields given by $[I|T]$ is standard.*

COROLLARY 2.2. *If T is of pure point spectrum, and in particular if it is an irrational rotation of the circle, then the σ -fields generated by $[I|T]$ are standard.*

COROLLARY 2.3. *There exists an exact dyadic endomorphism with a finite generating partition which gives a standard sequence of σ -fields, while its natural two-sided extension is not conjugate to a Bernoulli shift.*

1. Background and examples. Consider a decreasing sequence of σ -fields

$$\mathcal{F} = \mathcal{F}_0 \supset \mathcal{F}_1 \supset \dots$$

on the Lebesgue probability space (X, \mathcal{F}, μ) where $\bigcap_n \mathcal{F}_n$ is the trivial algebra of null and co-null sets. The term “ σ -field” will implicitly be taken to mean complete with respect to the obvious measure, so that we always get Lebesgue spaces (on appropriate quotient spaces and apart from the pos-

1991 *Mathematics Subject Classification*: Primary 28D05; Secondary 60G.

Feldman supported in part by NSF Grant DMS 9500803 and by a Fulbright Research Grant.

Rudolph supported in part by the Institute for Advanced Studies, Hebrew Univ. and NSF Grant DMS 9706829.

sible presence of atoms). The sequence is called *dyadic* if $\mu(\mathcal{F}_n|\mathcal{F}_{n+1})$ almost surely consists of two atoms of equal probability. It is called *standard* if there exist independent partitions $\mathcal{C}_1, \mathcal{C}_2, \dots$ for which $\mathcal{F}_n = \bigvee_{m>n} \mathcal{C}_m$. It is easy to see that \mathcal{C}_n is an independent complement to \mathcal{F}_n in \mathcal{F}_{n-1} . Obviously each \mathcal{C}_n consists of two sets of equal measure. If we label these two sets 0 and 1 then we get an i.i.d. sequence of $\{0, 1\}$ -valued random variables which generates all measurable sets, enabling us to identify X with $\{0, 1\}^{\mathbb{N}}$ with $\frac{1}{2}, \frac{1}{2}$ product measure. For σ the shift map on $\{0, 1\}^{\mathbb{N}}$ we can now identify \mathcal{F}_n as $\sigma^{-n}(\mathcal{F})$.

The introduction of the study of sequences of σ -fields as above and the introduction of standardness (in general, not just the dyadic case as described above) are due to A. M. Vershik. An overview of his work on this subject over the last 25–30 years is given in [19]. Here we have restricted ourselves to the dyadic case, but the p -adic case (where the positive integer p replaces 2) can be treated in the same way at the price of more notation.

Let E be an ergodic measure-preserving map of a Lebesgue probability space into itself, i.e. an *endomorphism* of (X, \mathcal{F}, μ) . Recall that such a map can always be “extended” to be 1-1. That is, there is a Lebesgue probability space $(\bar{X}, \bar{\mathcal{F}}, \bar{\mu})$, an *automorphism* \bar{E} of this space, and a measurable $\pi : \bar{X} \rightarrow X$ sending $\bar{\mu}$ to μ with $\pi\bar{E} = E\pi$. It may be made unique in the obvious sense by requiring that $\bigcup_{n=1}^{\infty} \bar{E}^n(\pi^{-1}\mathcal{F})$ be dense in \bar{X} . \bar{E} is called the *natural extension* of E . The endomorphism E is called *exact* if $\bigcap_n E^{-n}\mathcal{F}$ is trivial. The natural extension will likewise be ergodic. Hereafter when we write \bar{E} we will always mean the natural extension of E . Exactness of E implies that \bar{E} is *K-mixing*.

We shall be especially interested in the case when the endomorphism E has a *finite generator*; that is, there exists a finite measurable partition \mathcal{P} of X for which the sequence of partitions $\mathcal{P}, E^{-1}\mathcal{P}, E^{-2}\mathcal{P}, \dots$ generates \mathcal{F} . This property has an interpretation in terms of entropy: the existence of a finite generator is precisely equivalent to the equality of the dynamical entropy $h(\bar{E})$ and $H(\mathcal{F}|E^{-1}\mathcal{F})$. Also, if there is a finite generator for E , then E will be exact whenever \bar{E} is *K-mixing*. We will only be considering endomorphisms for which there is a finite generator. Hence for our purposes exactness of E and *K-mixing* of \bar{E} will be equivalent.

E is called *dyadic* if $\mu(\mathcal{F}|E^{-1}\mathcal{F})$ consists almost surely of two atoms of equal measure. To say E is dyadic is then equivalent to saying the sequence of σ -fields $\mathcal{F}_n = E^{-n}\mathcal{F}$ is dyadic as defined above. If an ergodic E is dyadic and has a finite generator then its entropy must be just $\log_2 2 = 1$.

Here are a few examples of exact dyadic endomorphisms.

1. *The canonical example.* Set $X = \{0, 1\}^{\mathbb{N}}$ with $\frac{1}{2}, \frac{1}{2}$ product measure, and σ the shift. The natural extension $\bar{\sigma}$ is the full-shift on $\{0, 1\}^{\mathbb{Z}}$ and σ

clearly has a finite generator. σ is isomorphic to the map $\lambda \mapsto 2\lambda \pmod{1}$ on $[0, 1)$ with Lebesgue measure.

2. *Toral endomorphisms.* If A is a $d \times d$ matrix with integral entries and $\det A = 2$ then A gives rise to an algebraic endomorphism of \mathbb{T}^d and this to E_A , a measure-preserving endomorphism of \mathbb{T}^d with normalized Haar measure. It is evident that E_A is dyadic. It will be ergodic if and only if A has no eigenvalues which are roots of unity (as an easy Fourier series argument shows). \bar{E}_A may be realized as an ergodic automorphism of a compact group (the inverse limit under A of \mathbb{T}^d) and hence is a Bernoulli automorphism ([12], [14], [16]). The entropy of E_A is the log of the product of the norms of all eigenvalues of A of norm ≥ 1 . But the product of *all* the eigenvalues is the determinant of A , namely 2. Then $h(\bar{E}_A) = 1$ if and only if all eigenvalues have norm ≥ 1 and this—in addition to the aforementioned ergodicity condition—is precisely what is required for E_A to have a finite generator. Thus, under these assumptions, E_A is exact.

3. *The $[T^{-1}|T]$ endomorphism.* This is defined on the space $(X \times Y, \mathcal{F}, \mu)$ where X is the one-sided two-shift as above with $\frac{1}{2}, \frac{1}{2}$ product measure and T is a measure-preserving automorphism of (Y, ν) . The endomorphism $E = [T|T^{-1}]$ is defined by the formula

$$(x, y) \mapsto (\sigma x, T^{2x(1)-1}y),$$

σ being the shift on X . Note that when T is the identity on a one-point space then $[T|T^{-1}]$ is isomorphic to σ . Ergodicity of E is equivalent to ergodicity of T^2 . E is clearly dyadic and has a finite generating partition. The natural extension \bar{E} of E is just the $[T|T^{-1}]$ automorphism, where σ is replaced by the shift $\bar{\sigma}$ on $\{0, 1\}^{\mathbb{Z}}$. In the ergodic case \bar{E} was shown in [15] to be a K -automorphism and consequently E is exact. Of special interest is the case where T is of positive entropy. It was shown in [8] that then the $[T|T^{-1}]$ automorphism is not Bernoulli (or even “loosely Bernoulli”; see [2] or [17] for definitions and details).

4. *The $[I|T]$ endomorphism.* This is similar to the previous example. Both have the form of a “cocycle extension” of the one-sided two-shift. In the previous example one applied the automorphism T or T^{-1} to the second coordinate depending upon the current value of the first coordinate. Here we apply either the automorphism T or I , the identity. More precisely, E is defined by the formula

$$(x, y) \mapsto (\sigma x, T^{x(1)}y).$$

Again if T is the identity on a one-point space then E is isomorphic to σ . For $E = [I|T]$, ergodicity of E is equivalent to that of T . Again E is clearly dyadic. Its natural extension is just the $[I|T]$ automorphism, replacing σ

by $\bar{\sigma}$ as above, and is again a K -automorphism by [15]. If T is of zero entropy then E has a finite generating partition. Thus it is also exact. It is shown in [2] that when T is an irrational rotation then \bar{E} is Bernoulli, but that it is also possible to construct T of zero entropy for which \bar{E} is *not* Bernoulli (or even loosely Bernoulli). This leads one to ask if there is any connection between Bernoullicity of \bar{E} and standardness of the sequence of σ -fields $\mathcal{F}, E^{-1}\mathcal{F}, E^{-2}\mathcal{F}, \dots$. It has been speculated that for proper dyadic endomorphisms E the corresponding decreasing sequence of σ -fields will be standard if and only if \bar{E} is Bernoulli. In the present paper we will give a counterexample to the “only if” direction.

The authors are grateful for the hospitality of the Mathematics Institute and the Institute for Advanced Studies of Hebrew University while this work was being done.

2. Discussion of results

THEOREM 2.1. *If T is rank-1 then the sequence of σ -fields given by $[I|T]$ is standard.*

COROLLARY 2.2. *If T is of pure point spectrum, and in particular if it is an irrational rotation of the circle, then the σ -fields generated by $[I|T]$ are standard.*

COROLLARY 2.3. *There exists an exact dyadic endomorphism with a finite generating partition which gives a standard sequence of σ -fields, while its natural two-sided extension is **not** conjugate to a Bernoulli shift.*

Corollary 2.2 simply uses the fact that any automorphism of pure point spectrum is of rank 1 (see [7]). Corollary 2.2 has already been stated by Vershik, who—in conversation with one of the authors—outlined an alternate argument. Corollary 2.3 follows from the construction by R. Burton (in his PhD thesis [1]) of a rank-1 automorphism for which the $[I|T]$ automorphism is *not* Bernoulli.

Some related results have recently been found:

1. C. Hoffman [6] has given an example of a zero entropy T for which $[I|T]$ gives a *nonstandard* sequence of σ -fields.

2. D. Heicklen and C. Hoffman [4] have shown that if T has positive entropy then the sequence of σ -fields given by the $[T|T^{-1}]$ endomorphism is *not* standard.

3. D. Heicklen, C. Hoffman, and D. J. Rudolph [5] have shown that if $[T|T^{-1}]$ and $[S|S^{-1}]$ give isomorphic sequences of σ -fields then S and T are of equal entropy.

Probabilistic significance may be attached to Corollary 2.3 in connection with coding of stationary processes from i.i.d. processes. Questions about

such coding were raised and studied by P. Lévy [13], N. Wiener [11], and M. Rosenblatt [18]. (The work of Kallianpur and Wiener to our knowledge remains unpublished, but in [10] there is a discussion of the work.) Corollary 2.3 implies the following: there exists a stationary stochastic process ξ_1, ξ_2, \dots with finite state space such that

(a) There are i.i.d. random variables η_n so that both

$$\xi_n = f_n(\eta_n, \eta_{n+1}, \dots) \quad \text{and} \quad \eta_n = g_n(\xi_n, \xi_{n+1}, \dots).$$

(b) On the other hand, there is no way of choosing i.i.d. random variables η'_n and a measurable function f so that

$$\xi_n = f(\eta'_n, \eta'_{n+1}, \dots).$$

3. Proof of Theorem 2.1. There is a basic criterion, due to A. Vershik [19], for standardness of a dyadic decreasing sequence of σ -fields. First note that an atom of \mathcal{F}_n consists of 2^n points arranged as the branches of a binary tree of height n . This arrangement goes as follows. Each atom of \mathcal{F}_1 consists of two points, the branches of a binary tree of height 1. If the atoms of \mathcal{F}_{n-1} are taken to be binary trees of height $n-1$, two of these combine to form an atom of \mathcal{F}_n and we construct the corresponding tree by attaching their two roots together to form a binary tree of height n . In these terms, Vershik's criterion says—roughly—that standardness of a dyadic sequence of σ -fields is equivalent to the following property: for any $\varepsilon > 0$ if n is large enough, then for all but ε in measure of the atoms of \mathcal{F}_n , the finite sets of points they consist of can be matched in such a way that all but ε in measure of the matched points are within ε of each other (in X as a metric space), with the match being obtained via a tree automorphism of the binary tree. We now discuss this with more precision for the case of σ -fields arising from an $[I|T]$ endomorphism. Further discussion of Vershik's criterion is in [3].

For $(x, y) \in X \times Y$, the two inverse images under $[I|T]$ of (x, y) are $(0x, y)$ and $(1x, T^{-1}y)$. For $p \in \{0, 1\}^n$ set

$$\bar{p} = \sum_{i=1}^n p(i).$$

More generally, $[I|T]^{-n}(x, y)$ consists of the 2^n points $\{(px, T^{-\bar{p}}y) : p \in \{0, 1\}^n\}$. These sets then are the atoms of \mathcal{F}_n . The tree structure on such an atom is then the natural tree structure on $\{0, 1\}^n$. In this tree the collection of branches (read “elements of $\{0, 1\}^n$ ”) which meet at a common node at height k in the tree are precisely the words that agree on their *rightmost* k indices. An automorphism α of this tree structure is then a bijection of $\{0, 1\}^n$ with the property that if two words agree on their rightmost k indices, the same is true of their images.

We can put metrics on these bijections in a consistent fashion by setting $D(\alpha, \alpha') = 2^{-k}$ if for all $p \in \{0, 1\}^n$, $\alpha(p)$ and $\alpha'(p)$ agree on their first (i.e. leftmost) k indices and k is the largest such value. (If $k = n$, i.e. $\alpha = \alpha'$ then of course we set $D(\alpha, \alpha) = 0$.) Here is a simple way to view this collection of metrics. Any map from $\{0, 1\}^n$ can be extended to all of $\{0, 1\}^{\mathbb{N}}$ by setting $\alpha(px) = \alpha(p)x$, i.e. just extend the map as the identity on the rest of the coordinates. Now $\{0, 1\}^{\mathbb{N}}$ is a compact metric space and these extensions are all continuous endomorphisms. Any of our tree automorphisms will actually extend to a homeomorphism that is a tree automorphism for the infinite tree. On the space of continuous maps from $\{0, 1\}^{\mathbb{N}}$ we have the topology of uniform convergence. A natural metric for these continuous maps is of course $D(\phi, \phi') = 2^{-k}$ if for all x , $\phi(x)$ and $\phi'(x)$ agree on their first k indices and k is the largest such value. Our metrics are simply this single metric on the extensions to $\{0, 1\}^{\mathbb{N}}$ of the tree automorphisms.

Suppose the automorphism T has a finite generator $Q = \{Q_a : a \in A\}$ where A is some labeling set for the elements of Q . The space Y can be taken to be $A^{\mathbb{Z}}$, a representation as a compact metric space. Of course two points in this space are close together iff they agree on a long interval of coordinates $[-N, N]$. The topology on $X \times Y$ is then simply the product of the two “product” topologies. Referring back to our brief description of Vershik’s criterion, notice that for an automorphism of one tree of inverse images to match most of the points in it to some other tree the automorphism is not necessarily close in D to the identity, but on “most” branches it must preserve lots of indices at the *left* end of a word x . Since Vershik’s criterion only specifies that most points be close, we could then actually replace this automorphism with one which is close in D to the identity and still have most points matched closely. Knowing the automorphism is close in D to the identity now, in order for most points to be close comes down to just asking that most points in one set are matched to points in the other that belong to the same element of Q . If the tree automorphism is D close to the identity and most points match to points in the same element of Q then most (this “most” is of course smaller than the previous one) points will be matched to points which have the same T, Q -name across a long segment $[-N, N]$. That is to say, the two inverse images will have been matched in the sense Vershik’s criterion demands. Stated with more precision, Vershik’s criterion for the endomorphism $[I|T]$ is then equivalent to the following.

VERSHIK’S STANDARDNESS CRITERION STATED FOR $[I|T]$. For some generating partition $\mathcal{Q} = \{Q_a : a \in A\}$ for T and for each $\varepsilon > 0$ there is some positive integer N , some sequence (a_0, \dots, a_N) of members of A (which we call \mathcal{Q} -symbols, or just symbols), a set $Y_0 \subseteq Y$ of measure $> 1 - \varepsilon$, and

for each $y \in Y_0$ a tree automorphism α of $\{0, 1\}^N$ such that

$$(1) \quad \frac{1}{2^N} \sum_{p \in \{0, 1\}^N} d(\mathcal{Q}(T^{-\alpha(p)}y), a_{\bar{p}}) + D(\alpha, \text{id}) < \varepsilon.$$

Here $\mathcal{Q}(y)$ is the subscript in A of the partition element in \mathcal{Q} containing y and d is the discrete 0, 1 valued metric on the finite space A .

To help understand this criterion, notice that for any point y the 2^N inverse images of a point (x, y) each have the form (x', y') where y' is one of $y, T^{-1}y, \dots, T^{-N}y$ and the values x' all agree with x after the N th binary coordinate (where they are all just x). The number of times a particular value $T^{-k}(y)$ occurs in this list is simply $\binom{N}{k}$, once for each x' with precisely k ones in its first N positions. This list of $N + 1$ points in Y gives a list of $N + 1$ labels in A , $\mathcal{Q}(y), \mathcal{Q}(T^{-1}y), \dots, \mathcal{Q}(T^{-N}y)$. Such a list can be thought of as putting a label at the tip of each branch of the tree. If $x' = px$ then the label for this branch is just $\mathcal{Q}(T^{-\bar{p}}y)$, the label of the y' where the branch “ends”.

Any list of $N + 1$ values from A will give rise to a labeling of the branches of the tree. In particular, the list (a_0, a_1, \dots, a_N) in the description of the criterion gives rise to such a labeling. Notice that these labelings are not arbitrary labelings of branches. All branches with the same value for \bar{p} yield the same label. A tree automorphism α , or for that matter any bijection of the branches $\{0, 1\}^N$ will permute a labeling to a new labeling of the branches. Of course it would be quite unusual for this new labeling to still arise from a list of N symbols; two branches “ending” at the same symbol in a list may not be mapped to branches ending at the same symbol. Given two lists (a_0, a_1, \dots, a_N) and $(a'_0, a'_1, \dots, a'_N)$ one can use them to produce two labelings of the tree $\{0, 1\}^N$ and then see how closely the two labelings can be matched by a tree automorphism. That is to say, for $p \in \{0, 1\}^N$ let $a(p)$ and $a'(p)$ be the labels the two lists attach to p . Now for any tree automorphism α one can calculate

$$\frac{1}{2^N} \sum_{p \in \{0, 1\}^N} d(a(p), a'(\alpha(p))) = \frac{\#\{p \in \{0, 1\}^N : a(p) \neq a'(\alpha(p))\}}{2^N}.$$

We also need to pay a price for how badly the automorphism itself disrupts the tree, i.e. we add in the distance D from the automorphism α to the identity automorphism. The minimum value this sum takes as α varies over the tree automorphisms gives a pseudometric on the space of all names. This bears a strong resemblance to \bar{f} and other similar constructions. What Vershik’s criterion says is that there is essentially only one name in this pseudometric, that is to say once N is large enough, “most” (read “all but ε ”) of the values y yield lists of labels close to some fixed list in this pseudomet-

ric. Here is an observation which provides a useful notational simplification. Notice that if in (1) one replaces y with $T^{-N}(y)$ one gets the identical statement, except that the minus sign in front of $\overline{\alpha(p)}$ disappears. Hereafter we will use this equivalent but simpler looking version of (1).

Rank-1 transformations. The proof of Theorem 2.1 will consist in verifying this criterion when T is rank-1. The definition of rank-1 which will be used is the following. First we will have a sequence of finite “words” in two symbols $\{a, b\}$. W_1 will just be the word a . Inductively the words W_n will have a concatenated structure of the form

$$W_{n+1} = W_n f_{n,1} W_n F_{n,2} \dots f_{n,l_n} W_n,$$

where each $F_{n,j}$ is a sequence of b 's of length $f_{n,j} \geq 1$ (yes 1, not 0). We refer to the b 's in a substrings $f_{n,j}$ as n -*filler*. The number of terms of W_n in n -filler is $f_n = \sum_{j=1}^{l_n} f_{n,j}$ and the length of W_n is w_n . We assume, in order to have a finite measure space, that $l_n > 1$ and $\sum_n f_n/w_n < \infty$. By dropping to a subsequence of words we may assume that $w_n^2/w_{n+1} \rightarrow 0$.

We say T is a *rank-1 transformation* of the Lebesgue space (Y, ν) if there is a list of words W_n as described above, and T has a two-set generating partition $\mathcal{Q} = \{Q_a, Q_b\}$ so that for ν -a.e. y , if n is sufficiently large, there is some value k between 0 and w_n so that

$$\mathcal{Q}(T^{-k}y) \dots \mathcal{Q}(T^{-1}y) \mathcal{Q}(y) \dots \mathcal{Q}(T^{w_n-k}y) = W_n.$$

Here $\mathcal{Q}(x) = a$ or b according as $x \in Q_a$ or $x \in Q_b$.

It is a fact that such a T must be ergodic and of zero entropy, and that the positions of the W_n 's in the \mathcal{Q} -name of y are uniquely determined (see [9]).

To verify Vershik's criterion in our case we will show that using tree automorphisms to rearrange \mathcal{Q} -names we can modify most names to be close to some single name. We can now tell you what that single name will be. For each value n set \mathcal{N}_n to be the infinite concatenation of copies of W_n . More precisely,

$$\mathcal{N}_n(j) = W_n(j \bmod w_n).$$

Proof of Theorem 2.1. To prove Theorem 2.1 we must show how to use tree automorphisms to “shift” most occurrences of W_n in the names of most y 's to lie at the uniformly spaced positions they have in \mathcal{N}_n .

Here are some preliminaries. On $\{0, 1\}^J$ denote by Pr fair-coin-tossing measure, i.e. uniform mass 2^{-J} on each element of $\{0, 1\}^J$. As before, for any $p \in \{0, 1\}^J$ we set $\bar{p} = \sum_{j=1}^J p(j)$ (the number of 1's in p).

LEMMA 3.1 (localization). *Given $\varepsilon > 0$ there exists a C such that for all sufficiently large J ,*

$$\text{Pr}\{p \in \{0, 1\}^J : |\bar{p} - J/2| < C\sqrt{J}\} \geq 1 - \varepsilon.$$

LEMMA 3.2 (scaling). *Given $\varepsilon > 0$ and C there exists D such that for all $J \in \mathbb{N}$, if S is a set of integers in $(J/2 - C\sqrt{J}, J/2 + C\sqrt{J})$ and $\#S < \frac{C}{D}\sqrt{J}\varepsilon$ then*

$$\Pr\{p \in \{0, 1\}^J : \bar{p} \in S\} < \varepsilon.$$

LEMMA 3.3 (shifting). *Given $\varepsilon > 0$ there exists $\sigma > 0$ so that for all sufficiently large J , if s is an integer with $|s| < \sigma\sqrt{J}$, then there is a tree automorphism α of $\{0, 1\}^J$ with $D(\alpha, \text{id}) < \varepsilon$ so that*

$$\Pr\{\overline{p} - \overline{\alpha(p)} = s\} \geq 1 - \varepsilon.$$

Proof. The first two of these three lemmas are fairly obvious from the Central Limit Theorem. The third may be seen as follows. We assume s is nonnegative. The nonpositive case follows directly: solve the problem for $-s$ and use α^{-1} . We will in fact only use the nonpositive case. On the probability space $\{0, 1\}^J$ we construct J steps of a random walk: $X_n(p) = 2p(n) - 1$. Thus the X_n are i.i.d. on $\{-1, 1\}$ and

$$S_n(p) = X_1(p) + \dots + X_n(p) = 2 \sum_{j=1}^n p(j) - n, \quad 0 \leq n \leq J,$$

is a standard symmetric random walk. Let $A = \{S_n \geq s \text{ for some } n, 0 \leq n \leq J\}$. Then

$$\Pr\{A\} \geq \Pr\{A, S_J \geq s\} + \Pr\{A, S_J \leq s\} - \Pr\{S_J = s\}.$$

Symmetry and independence of the X_n 's give

$$\Pr\{A, S_J \geq s\} = \Pr\{A, S_J \leq s\}$$

(reflect the portion of the path beginning from the last time $S_n = s$ across the horizontal line of height s). Since $A \supset \{S_J \geq s\}$, we have

$$\begin{aligned} \Pr\{A\} &\geq 2\Pr\{S_J \geq s\} - \Pr\{S_J = s\} \\ &\geq \Pr\{|S_J| \geq s\} - \Pr\{S_J = s\} \\ &\geq \Pr\{|S_J| \geq \sigma\sqrt{J}\} - \Pr\{S_J = s\}. \end{aligned}$$

Now

$$\Pr\{|S_J| \geq \sigma\sqrt{J}\} \rightarrow 1 - \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} e^{-\lambda^2/2} d\lambda$$

as $J \rightarrow \infty$, while $\Pr\{S_J = s\}$ is either $\binom{J}{(J-s)/2} / 2^J$ or 0 depending on whether J and s are or are not of the same parity. This latter tends to zero as $J \rightarrow \infty$ and the integral of course tends to 0 as $\sigma \rightarrow 0$. Thus once J is large enough we can be sure $\Pr\{A\} > 1 - \varepsilon$.

We now define an auxiliary map α' . If $p \in A$ and n is the first integer for which $S(n) = s$, then set $\alpha'(p)(j) = 1 - p(j)$ for all $j < n$, that is to say interchange 0 and 1 along this string of n indices. For such a $p \in A$ and

$j \geq n$ set $\alpha'(p)(j) = p(j)$; that is, do not change the value at the rest of the indices. For such a p it is an easy calculation that

$$\bar{p} - \overline{\alpha'(p)} = \sum_{j=1}^n (p(j) - (1 - p(j))) = 2 \sum_{j=1}^n p(j) - n = s.$$

For $p \notin \Lambda$, define $\alpha'(p)(j) = 1 - p(j)$ for all j , that is to say switch all indices.

The construction of α' can be envisioned in the following dynamic fashion. To find $\alpha'(p)$ we start switching the 0, 1 values of p index by index. We are looking for a time when the difference between the running sum of p and its switched version becomes precisely s . If and when this happens, we stop switching. If it does not ever happen, all the values get switched. Notice that this constructs a map on $\{0, 1\}^J$ with the property that the images of any two names that agree on their leftmost k indices still agree on their leftmost k -indices. This is symmetric to what we want of a tree automorphism. Define α to be α' conjugated by the flip $\phi(p)(j) = p(n - j + 1)$. This now makes α a tree automorphism. Since ϕ leaves the number of 1's in the string unchanged, we still obtain $\bar{p} - \overline{\alpha(p)} = s$, provided $\phi(p) \in \Lambda$. Vershik's criterion (1) also requires that $D(\alpha, \text{id}) < \varepsilon$. That is to say, for all p , p and $\alpha(p)$ should agree in their leftmost $[-\log(\varepsilon)]$ indices. To obtain this, modify the construction of α' so that if after $J - [-\log(\varepsilon)]$ indices it is still switching, then stop switching and start fixing values. ■

Now let T be a rank-1 map as described above.

A finite string W of symbols from $\{a, b\}$ is said to *appear in position* $[k, l]$ in the name of y if $\mathcal{Q}(T^k(y)), \dots, \mathcal{Q}(T^l(y)) = W$. Thus for our rank-1 map T and for a.e. y , if n is sufficiently large, then for some $k = 0, 1, \dots, w_n$, the word W_n appears in position $[-k, w_n - k]$ in the name of y .

LEMMA 3.4. *Given $\varepsilon > 0$ there is some $e > 0$ so that for all n large enough and all y in a set of measure $\geq 1 - \varepsilon$ there is a substring of W_{n+1} which appears in position $[-ew_{n+1}, ew_{n+1}]$ in the name of y (of course $e < 1$).*

PROOF. As remarked earlier, the breakup of the name of a point y into copies of W_{n+1} is unique. In this breakup the origin is either at some position k between 0 and $w_{n+1} - 1$ in W_{n+1} or is in filler. Once n is large enough the set of points y with the origin in filler will have measure less than $\varepsilon/3$. Among the points y with the origin inside a copy of W_{n+1} , the probabilities of the origin being at various values k are all equal. Hence those within a fraction e of either end have probability at most $2e$. Thus $e = \varepsilon/3$ will do. ■

By $[[j]]$ we denote the smallest *even* integer $\geq j$.

LEMMA 3.5. Let C be as in Lemma 3.1 and

$$K(n) = \lceil \lceil ew_{n+1}/C \rceil \rceil^2.$$

If n is sufficiently large then

$$\Pr\{p : |\bar{p} - K(n)/2| < ew_{n+1}\} < \varepsilon.$$

Proof. Immediate from Lemma 3.1. ■

To verify Vershik’s criterion we need to construct some tree automorphisms. These will be built in stages. We will write $\{0, 1\}^N$ as $\{0, 1\}^M \times \{0, 1\}^L \times \{0, 1\}^K$ where $N = K + L + M$, writing a string of length N as a concatenation of three strings rqp . Our automorphisms will be built by choosing for each p a tree automorphism $\alpha(p)$ of $\{0, 1\}^L$ and for each qp a tree automorphism $\beta(qp)$ of $\{0, 1\}^M$. Given these, we can define a tree automorphism

$$\gamma(rqp) = (\beta(qp)(r))(\alpha(p)(q))p.$$

(This brief description has omitted the dependence our choices of α ’s and β ’s must have on the \mathcal{Q} -name of y .)

To begin the proof choose a (large) integer H and then a (small) positive value δ . Also choose even positive integers $L(n)$ and $M(n)$ for the purpose of shifting names. We require

$$\left(\frac{2}{\sigma}w_n\right)^2 < L(n) < \left(\frac{\varepsilon}{2CD}Hw_n\right)^2$$

(so H had to be chosen appreciably larger than $4CD/(\sigma\varepsilon)$) and

$$\left(\frac{\delta Hw_{n+1}}{\sigma}\right)^2 < M(n) < \left(\frac{\varepsilon}{2CD}w_n\right)^2$$

(so δ needs to be chosen appreciably smaller than $\sigma\varepsilon/(2DCH)$). Here C , D , and σ are the values considered in Lemmas 3.1–3.3.

Let $N(n)$ denote $K(n) + L(n) + M(n)$. Mark off intervals of length Hw_n starting from 0 in \mathbb{Z} . Such an interval will look like this:

$$[kHw_n, \dots, (k + 1)Hw_n - 1].$$

For H fixed, n can be chosen so large that

$$\frac{2Hw_n}{ew_{n+1}} < \frac{\varepsilon}{D},$$

where D is again from Lemma 3.2. It will help to keep in mind as we continue our construction that we regard the “origin” in a name as sitting at position $N(n)/2$, the mean position of the random walk. We call the interval $[-ew_{n+1}, ew_{n+1}]$ the n -center. It will contain many n -blocks. We fix a “border” on the n -center consisting of two intervals of length Hw_n , i.e. the union

of the two intervals $[-ew_{n+1}, -ew_{n+1} + Hw_n]$ and $[ew_{n+1} - \beta w_n, ew_{n+1}]$. By Lemma 3.2,

$$\Pr\{p \in \{0, 1\}^{K(n)} : \bar{p} - N(n)/2 \in \text{this border}\} < \varepsilon.$$

In each n -block B we make a border of width

$$\frac{\varepsilon}{2D} \#(B) = \frac{\varepsilon}{2D} Hw_n.$$

If this is done for all n -blocks B that lie entirely in the n -center, then Lemma 3.2 implies that

$$\Pr\{p : \bar{p} \in \text{the border of some } n\text{-block} \subseteq n\text{-center}\} < \varepsilon.$$

The part of an n -block B not contained in this border will be called the *interior* of B .

To get a sense of relative sizes at this point, $K(n)$ is on the order of $((\text{small constant}) \times w_{n+1})^2$ and hence $\sqrt{K(n)}$ grows more rapidly in n than w_n^2 . $L(n)$ is on the order of $((\text{large constant}) \times w_n)^2$, hence grows less rapidly than $\sqrt{K(n)}$. As both C and D can be assumed very large, $M(n)$ is on the order of $(\text{very small constant}) \times w_n$. Thus $N(n)$ is essentially just $K(n)$.

We will be able to work with those y for which the interval $I_n = [K(n)/2 - ew_{n+1}, N(n)/2 + ew_{n+1}]$ in the name of y lies entirely within W_{n+1} . Such a y we will call *n-usable*. Notice $N(n) - K(n) = L(n) + M(n) < (\text{constant}) \times w_n^2$ and $w_n^2/w_{n+1} \rightarrow 0$ in n . Recalling that $e \leq \varepsilon/2$, we see that the length of I_n divided by w_{n+1} converges in n to $2e \leq \varepsilon$. Now Lemma 3.4, where the value of e was set, tells us that once n is large enough the set of *n-usable* y will have measure $\geq 1 - 2\varepsilon$.

Let y be *n-usable*. An n -block B is called *good for y* if it is entirely in the n -center and the portion of the name of $T^{N(n)/2}y$ in B contains a fraction $< \delta$ of n -filler. If n is sufficiently large, then for all *n-usable* y , a proportion more than $1 - \varepsilon/D$ of the n -blocks in the n -center are good for y . Then by Lemma 3.2, for any *n-usable* y ,

$$\Pr\{p \in \{0, 1\}^{K(n)} : \bar{p} + L(n)/2 + M(n)/2 \text{ is in a } y\text{-good } n\text{-block}\} \geq 1 - 2\varepsilon.$$

For an n -block B which is good for y define $s(B, y)$ to be the number of symbols to the left of the leftmost W_n in the name of $T^{N(n)/2}y$ which is fully contained in B . Then

$$s(B, y) < w_n + \delta Hw_n \leq 2w_n.$$

Now we define a tree automorphism $\alpha_y(p)$ of $\{0, 1\}^{L(n)}$. If $\bar{p} + (L(n) + M(n))/2$ is in an n -block B which is good for y , choose $\alpha_y(p)$ by Lemma 3.3 so that

$$\Pr\{\bar{q} - \overline{\alpha_y(p)(q)} = s(B, y)\} \geq 1 - \varepsilon.$$

For other p it does not matter, but for explicitness let $\alpha_y(p)$ be the identity.

Let B be an n -block which is good for y (y is assumed n -usable). Then B is mostly covered by disjoint intervals of length w_n in which the name of $T^{N(n)/2}y$ is a copy of W_n . Call these intervals n -cells. We will label them from left to right as Γ_i . The leftmost n -cell Γ_1 begins after precisely $s(B, y)$ steps, in B , and in general the k th n -cell Γ_k begins after $s(B, y) + (k-1)w_n + t(k, y)$ steps, where we know

$$0 \leq t(k, y) < \delta H w_n.$$

This last is true because we know that the *total* number of indices in B that lie in n -filler in the name of $T^{N(n)/2}y$ is less than $\delta H w_n$. The rightmost n -cell has its right end within $2H w_n$ of the right end of B and so is inside the right border of B .

Now we define tree automorphisms of $\{0, 1\}^{M(n)}$ to implement a shift by $t(k, y)$ for the n -cell Γ_k . If $\bar{p} + \overline{\alpha_y(p)(q)} + M(n)/2$ lies in Γ_k , an n -cell in a good n -block in the name of $T^{N(n)/2}y$, then once more use Lemma 3.3 to choose a tree automorphism $\beta_y(p, q, k)$ of $\{0, 1\}^{M(n)}$ so that

$$\Pr\{\bar{r} - \beta_y(p, q, k) = t(k, y)\} \geq 1 - \varepsilon.$$

Suppress k , defining $\beta_y(qp)(r)$ by defining it as $\beta(p, q, k)$ for r as above. For other p, q or y the effect of the tree automorphism will not matter, and so we extend it to be the identity.

We write a sequence in $\{0, 1\}^{N(n)}$ as rqp where $p \in \{0, 1\}^{K(n)}$, $q \in \{0, 1\}^{L(n)}$, and $r \in \{0, 1\}^{M(n)}$ and now define a tree automorphism on $\{0, 1\}^{N(n)}$ by

$$\gamma_y(rqp) = \beta_y(pq)(r)\alpha_y(p)(q)p$$

(remember that placing two strings of symbols next to one another concatenates them). Furthermore, for most rqp we have

$$\mathcal{Q}(T^{\gamma(rqp)}y) = \mathcal{N}_n(\overline{rqp}).$$

More precisely:

PROPOSITION 3.6. *If n is sufficiently large, then for each n -usable y we have*

$$\Pr\{\mathcal{Q}(T^{\gamma(rqp)}y) = \mathcal{N}_n(\overline{rqp})\} > 1 - 11\varepsilon.$$

Proof. This is just a bit of bookkeeping. Fix some n -usable y . First we consider $\{0, 1\}^{K(n)}$; its typical element we call, as always, p . Then

$$\Pr\{p : |\bar{p} - K(n)/2| < \varepsilon w_{n+1}\} \geq 1 - \varepsilon,$$

$$\Pr\{p : \bar{p} + (L(n) + M(n))/2 \text{ lies in an } n\text{-block entirely in the } n\text{-center}\} \geq 1 - \varepsilon,$$

$$\Pr\{p : \bar{p} + (L(n) + M(n))/2 \text{ is interior to an } n\text{-block}\} \geq 1 - \varepsilon,$$

$$\Pr\{p : \bar{p} + (L(n) + M(n))/2 \text{ lies in an } n\text{-block which is good for } y\} \geq 1 - \varepsilon.$$

So the intersection of these four sets has measure $\geq 1 - 4\varepsilon$.

Next, for each fixed p in the above intersection we examine $\{0, 1\}^{L(n)}$ whose typical element we denote by q . Let B be the good n -block in which $\bar{p} + (L(n) + M(n))/2$ lies. From Lemma 3.1,

$$\Pr\{q : \bar{p} + \bar{q} + M(n) \in B\} \geq 1 - \varepsilon.$$

From the fact that tree automorphisms are measure-preserving,

$$\Pr\{q : \bar{p} + \overline{\alpha_y(p)(q)} + M(n) \in B\} \geq 1 - \varepsilon.$$

By assumption on p ,

$$\Pr\{q : \bar{q} - \overline{\alpha_y(p)(q)} = s(B, y)\} \geq 1 - \varepsilon.$$

From Lemma 3.2,

$$\Pr\{q : \bar{q} - s(B, y) + M(n) \text{ is interior to a cell in } B\} \geq 1 - \varepsilon.$$

Thus for each such fixed p we have a set of q of measure $\geq 1 - 4\varepsilon$ in all four of these sets.

Finally, we examine $r \in \{0, 1\}^{M(n)}$. Fix p and q satisfying the above eight conditions (four on each). Assume $\bar{p} + \overline{\alpha_y(p)} + M(n)/2$ is in an n -block B good for y , and interior to an n -cell Γ_k . Then

$$\Pr\{r : \bar{p} + \overline{\alpha_y(p)(q)} + \bar{r} \in \Gamma_k\} \geq 1 - \varepsilon,$$

$$\Pr\{r : \bar{p} + \overline{\alpha_y(p)(q)} + \overline{\beta_y(qp)(r)} \in \Gamma_k\} \geq 1 - \varepsilon,$$

$$\Pr\{r : \bar{r} - \overline{\beta_y(qp)(r)} = t(k, y)\} \geq 1 - \varepsilon.$$

For each given such p and q the set of all r in these three sets has measure $\geq 1 - 3\varepsilon$. Thus the set of rqp satisfying all eleven of the conditions has probability $\geq 1 - 11\varepsilon$. If all these conditions are satisfied then

$$\mathcal{Q}(T^{\overline{(rqp)}}y) \text{ is precisely } \mathcal{N}_n(\overline{rqp}).$$

Replacing ε by $\varepsilon/11$ we see that our version of Vershik's criterion is satisfied. ■

References

- [1] R. Burton, *A non-Bernoulli skew product which is loosely Bernoulli*, Israel J. Math. 35 (1980), 339–348.
- [2] J. Feldman, *New K-automorphisms and a problem of Kakutani*, *ibid.* 24 (1976), 16–38.
- [3] D. Heicklen, *Entropy and r equivalence*, *Ergodic Theory Dynam. Systems*, to appear.
- [4] D. Heicklen and C. Hoffman, *T, T^{-1} is not standard*, *ibid.* to appear.

- [5] D. Heicklen, C. Hoffman and D. Rudolph, *Entropy and dyadic equivalence of random walks on a random scenery*, preprint.
- [6] C. Hoffman, *A zero entropy (T, Id) endomorphism that is not standard*, Ergodic Theory Dynam. Systems, to appear.
- [7] A. del Junco, *Transformations with discrete spectrum are stacking transformations*, Canad. J. Math. 28 (1976), 836–839.
- [8] S. Kalikow, *T, T^{-1} transformation is not loosely Bernoulli*, Ann. of Math. 115 (1982), 393–409.
- [9] —, *Twofold mixing implies threefold mixing for rank one transformations*, Ergodic Theory Dynam. Systems 4 (1984), 237–259.
- [10] G. Kallianpur, *Some ramifications of Wiener's ideas on nonlinear prediction*, in: Norbert Wiener, Collected Works (P. Masani, ed.), Vol. III, MIT Press, 1981, 402–423.
- [11] G. Kallianpur and N. Wiener, 1956, Tech. report No. 1, Office of Naval Research CU-2-56 NONR-266 (39), CIRMIP project NR-047-015.
- [12] Y. Katznelson, *Ergodic automorphisms of T^n are Bernoulli shifts*, Israel J. Math. 10 (1971), 186–195.
- [13] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1954.
- [14] D. Lind, *The structure of skew products with ergodic group automorphisms*, Israel J. Math. 28 (1977), 205–248.
- [15] I. Meilijson, *Mixing properties of a class of skew products*, *ibid.* 19 (1974), 266–270.
- [16] G. Miles and K. Thomas, *Generalized torus automorphisms are Bernoullian*, in: Adv. Math. Suppl. Stud. 2, Academic Press, 1978, 231–249.
- [17] D. S. Ornstein, D. Rudolph and B. Weiss, *Equivalence of measure preserving transformations*, Mem. Amer. Math. Soc. 262 (1982).
- [18] M. Rosenblatt, *Stationary processes as shifts of functions of independent random variables*, J. Math. Mech. 8 (1959), 665–681.
- [19] A. Vershik, *The theory of decreasing sequences of measurable partitions*, St. Petersburg Math. J. 6 (1995), 705–761.

Department of Mathematics
 University of California at Berkeley
 Berkeley, California 94720
 U.S.A.
 E-mail: feldman@math.berkeley.edu

Department of Mathematics
 University of Maryland
 College Park, Maryland 20742
 U.S.A.
 E-mail: djr@math.umd.edu

*Received 9 October 1997;
 in revised form 18 February 1998*