MIXING k -FOLD INDEPENDENT PROCESSES OF ZERO ENTROPY

L. FLAMINIO

(Communicated by Charles Pugh)

ABSTRACT. For each fixed k we give a simple construction of a stochastic process (X_n) which is mixing, has zero entropy, and such that any k -tuple of the X_n are independent.

The motivation for this note came from the following question raised by Steve Kalikov some time ago. "Does there exist a process of zero entropy which is two-fold independent?" The answer to this question is positive and we are going to exhibit a very simple class of examples of such processes.

Pairwise and more generally k -fold independent stationary processes have received attention lately especially in relation to the failure to satisfy the Central Limit Theorem (CLT). The examples by Janson [Ja] and Bradley [Br] are pairwise independent and fail to satisfy the CLT. Our basic example in the next section is a simple modification of Janson's which makes the process ergodic. It is described here for the sake of completeness. Our goal in this note is to show how, using techniques of ergodic theory, one can modify this example to produce processes which satisfy strong mixing properties while keeping the zero-entropy and the k -fold independence constraint. One important open question is whether these examples satisfy the CLT.

Definitions. A transformation is for brevity a probability space $(\Omega, \mathcal{B}, \mu)$ endowed with a bimeasurable map $T: \Omega \rightarrow \Omega$ which is measure preserving, i.e., $\forall A \in \mathcal{B}$ we have $\mu(T^{-1}A) = \mu(A)$. A transformation is ergodic if $\mu(T^{-1}A \Delta A) = 0$ implies that $\mu A = 0$ or $\mu A = 1$.

Every measurable partition $\mathcal{P} = (P_1, \dots, P_l)$ of Ω defines a stationary random process $(X_n^{\mathcal{P}})_{n \in \mathbb{Z}} = (\mathcal{P}, T)$ in the following way. For $\omega \in \Omega$, let $X_n^{\mathcal{P}}(\omega) = k$ if $T^n \omega \in P_k$, $k \in [1, l]$. Then

$$\Pr\{X_{s_1} = k_1, X_{s_2} = k_2, \dots, X_{s_j} = k_j\} = \mu(T^{-s_1}P_{k_1} \cap T^{-s_2}P_{k_2} \cap \dots \cap T^{-s_j}P_{k_j}).$$

A process $(X_n)_{n \in \mathbb{Z}}$ is said to be *m-fold independent* if the r.v.'s $X_{s_1}, X_{s_2}, \dots, X_{s_m}$ are jointly independent for every m -tuple of distinct integers $(s_1, \dots, s_m)$. It is said to be *mixing* if $\Pr\{X_{t_1+s} = k_1, \dots, X_{t_m+s} = k_m | X_{s_1} = j_1, \dots, X_{s_n} = j_n\}$ converges to $\Pr\{X_{t_1} = k_1, \dots, X_{t_m} = k_m\}$ as s tends to infinity. It is said to be *weakly mixing* if $\Pr\{X_{t_1+s} = k_1, \dots, X_{t_m+s} = k_m | X_{s_1} = j_1, \dots, X_{s_n} = j_n\}$

Received by the editors December 30, 1991.

1991 *Mathematics Subject Classification.* Primary 60G10, 28D05.

© 1993 American Mathematical Society
 0002-9939/93 \$1.00 + \$.25 per page

converges to $\Pr\{X_{t_1} = k_1, \dots, X_{t_m} = k_m\}$ as s tends to infinity avoiding a set of integers of density zero.

Finally a stationary process $(X_n)_{n \in \mathbb{Z}}$ with a finite number of states has zero entropy if $\frac{1}{n} \log \Pr\{X_1 = X_1(\omega), X_2 = X_2(\omega), \dots, X_n = X_n(\omega)\}$ converges to zero almost surely.

Given a transformation T on $(\Omega, \mathcal{B}, \mu)$ we say that T is *mixing* (resp. *weakly mixing*) or that it has *zero entropy* if for all partitions $\mathcal{P} = (P_1, \dots, P_l)$ of Ω the stationary random process $(\mathcal{P}, T)$ is mixing (resp. weakly mixing), or it has zero entropy.

The definitions of mixing, weakly mixing, and zero entropy given above are not quite the standard definitions as given, for example, in [Pe]. Nevertheless, they are equivalent to the standard definitions (see [Pe]) and more useful for our purposes.

The goal of this note is to show the following.

Theorem. *There are mixing processes that are k -fold independent and have zero entropy.*

Basic construction. Let $\mathbf{T}^k$ be the k -torus considered as the additive group of k real numbers modulo 1, with the normalized Lebesgue measure μ defined on the σ -algebra of Borel measurable sets $\mathcal{B}$.

Let α be an irrational number. We define a transformation on S on $\mathbf{T}^k$ by setting

$$S(t_1, \dots, t_k) = (\alpha + t_1, t_1 + t_2, \dots, t_{k-1} + t_k).$$

It is easily seen that S preserves the measure μ . The n th iterate of S can be written explicitly as

$$\begin{aligned} S^n(t_1, \dots, t_k) \\ = \left(n\alpha + t_1, \binom{n}{2}\alpha + nt_1 + t_2, \dots, \binom{n}{k}\alpha + \binom{n}{k-1}t_1 + \dots + nt_{k-1} + t_k \right). \end{aligned}$$

It is well known and easy to see that the transformation S on $(\mathbf{T}^k, \mathcal{B}, \mu)$ is an ergodic transformation of zero entropy.

Consider a partition $\tilde{\mathcal{P}} = (\tilde{P}_0, \tilde{P}_1, \dots, \tilde{P}_l)$ of $\mathbf{T}^1$, and let $\mathcal{P} = (P_0, \dots, P_l)$ be the partition of $\mathbf{T}^k$ defined by

$$P_i = \mathbf{T}^{k-1} \times \tilde{P}_i = \{(t_1, \dots, t_k) \in \mathbf{T}^k \mid t_k \in \tilde{P}_i\}.$$

Proposition 1. *The process $(\mathcal{P}, S)$ has zero entropy and is k -fold independent.*

Proof. The nullity of entropy follows from the fact that S has zero entropy. Thus we need to show that for any choice of k distinct integers $i_1, \dots, i_k$ the partitions $S^{-i_1}\mathcal{P}, \dots, S^{-i_k}\mathcal{P}$ are independent. Let p be the projection $(t_1, \dots, t_k) \mapsto t_k$. Our goal will be achieved if we show that the random variables $p \circ S^{i_1}, \dots, p \circ S^{i_k}$ are independent for any choice of distinct integers $i_1, \dots, i_k$.

Since

$$p \circ S^{i_s}(t_1, \dots, t_k) = \left(\binom{i_s}{k}\alpha + \binom{i_s}{k-1}t_1 + \dots + i_st_{k-1} + t_k \right),$$

this is tantamount to showing that the map $(p \circ S^{i_1}, p \circ S^{i_2}, \dots, p \circ S^{i_k}): \mathbf{T}^k \rightarrow \mathbf{T}^k$ given by

$$(t_1, \dots, t_k) \mapsto \left(\binom{i_1}{k} \alpha + \binom{i_1}{k-1} t_1 + \dots + i_1 t_{k-1} + t_k, \dots, \right. \\ \left. \binom{i_k}{k} \alpha + \binom{i_k}{k-1} t_1 + \dots + i_k t_{k-1} + t_k \right)$$

is measure preserving. This follows immediately from the fact that

$$\det \begin{pmatrix} \binom{i_1}{k-1} & \binom{i_1}{k-2} & \dots & i_1 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \binom{i_k}{k-1} & \binom{i_k}{k-2} & \dots & i_k & 1 \end{pmatrix} = \frac{\prod_{u < v} (i_u - i_v)}{1!2!3! \dots (k-1)!}$$

is nonzero for all distinct choices of $i_1, \dots, i_k$ (to see the last equality, notice that the above is a VanderMonde determinant up to the factor $\frac{1}{1!2!3! \dots (k-1)!}$). $\square$

The example that we have been considering is not mixing. In fact, it is not even weakly mixing. To construct a mixing example one can proceed in the following way that has been suggested to us by B. Weiss. It uses ideas and proofs of [GIWe].

Mixing examples. Let $S_0: X \rightarrow X$ be a mixing transformation of zero entropy of a probability space X . Examples of such transformations have been known for a while [Or, Ma]. Let $S_1: \mathbf{T}^k \rightarrow \mathbf{T}^k$ be the previously considered transformation

$$(t_1, \dots, t_k) \mapsto (\alpha + t_1, t_1 + t_2, \dots, t_{k-1} + t_k)$$

where α is an irrational number and $k \geq 2$. Consider the product S of S_0 and S_1 acting on $Y = X \times \mathbf{T}^k$ endowed with the product measure ν on the product σ -algebra. Then S is an ergodic transformation of zero entropy.

For every measurable function $f: x \in X \mapsto (f_1(x), \dots, f_k(x)) \in \mathbf{T}^k$ let us consider the measure-preserving transformations

$$F_f: (x, t_1, \dots, t_k) \in Y \mapsto (x, t_1 + f_1(x), \dots, t_k + f_k(x)) \in Y.$$

Let $\mathcal{S}$ be the set of transformations of Y of the type $F_f^{-1} S F_f$. It is easily seen that

$$(1) \quad F_f^{-1} S F_f(x, t_1, \dots, t_k) = (S_0 x, \alpha + t_1 + f_1(x) - f_1(S_0 x), \\ t_1 + t_2 + f_1(x) + f_2(x) - f_2(S_0 x), \dots, \\ t_{k-1} + t_k + f_{k-1}(x) + f_k(x) - f_k(S_0 x)).$$

Denote by $\overline{\mathcal{S}}$ the closure of $\mathcal{S}$ in the space of all transformations with respect to the topology

$$T_\alpha \rightarrow T \iff (T_\alpha f, g) \rightarrow (T f, g)$$

for all $f, g \in L^2(Y)$.

Every transformation $R \in \overline{\mathcal{S}}$ has the form

$$(2) \quad (x, t_1, \dots, t_k) \mapsto (S_0(x), t_1 + g_1(x), t_2 + t_1 + g_2(x), \dots, t_{k-1} + t_k + g_k(x)).$$

Such transformations are known as “skew-products”.

Lemma 1. *Let $p: Y \mapsto \mathbb{T}^1$ be the projection on the last coordinate. Then, for $R \in \overline{\mathcal{P}}$, the random variables*

$$X_n(x, t_1, \dots, t_k) = p \circ R^n(x, t_1, \dots, t_k)$$

are k -fold independent.

Proof. From (2) above we have that

$$X_n(x, t_1, \dots, t_k) = \binom{n}{k-1} t_1 + \dots + n t_{k-1} + t_k + H(n, x)$$

for some measurable $H(n, x)$. Now Proposition 1 allows us to conclude. $\square$

We also have

Lemma 2. *Every $T \in \overline{\mathcal{P}}$ has entropy zero.*

The proof of Lemma 2 is immediate. See for example [Pe] for the computation of the entropy of a skew-product.

We now proceed to state the main result.

Theorem. *There exists a dense set of mixing transformation in $\overline{\mathcal{P}}$.*

Proof. One need only observe that the transformations in $\overline{\mathcal{P}}$ are “compact group extensions” of the mixing transformation S_0 . By a theorem of Rudolph [Ru] weakly mixing compact group extensions of a mixing transformation are mixing and we are done if we prove the following proposition. $\square$

Proposition 2. *There exists a dense G_δ subset of $\overline{\mathcal{P}}$ of weakly mixing transformations.*

We will need a few lemmata to prove the above proposition.

It is well known (see, e.g., [Pe]) that a measure-preserving transformation R on a probability space (Y, ν) is weakly mixing if and only if the associated unitary operator on $L^2(Y, \nu)$ given by $\varphi \mapsto \varphi \circ R$ has no eigenfunctions other than the constants. Of course if φ is an eigenfunction of unit norm then $|\langle \varphi \circ R^n, \varphi \rangle| = 1$ for all integers n . We are about to state an approximate version of this fact. With abuse of notation, denote also by R the unitary operator on $L^2(Y, \nu)$ associated with the transformation R . Thus $R\varphi = \varphi \circ R$. Denote by $L_0^2(Y, \nu)$ the subspace of $L^2(Y, \nu)$ orthogonal to the constants.

Lemma 3. *A measure-preserving transformation R on a probability space (Y, ν) is weakly mixing if and only if there exists an η , $0 < \eta < 1$, such that for all φ in a dense subset of $L_0^2(Y, \nu)$ there exists an integer n with $|\langle R^n \varphi, \varphi \rangle| < (1 - \eta) \|\varphi\|^2$.*

The proof follows easily from the definition and from the spectral characterization of weak mixing.

We say that a function $\varphi \in L_0^2(Y, \nu)$ is an η -quasi-eigenfunction for R if for all integers n we have that $|\langle R^n \varphi, \varphi \rangle| \geq (1 - \eta) \|\varphi\|^2$.

Then the previous lemma could be roughly stated by saying that R is weakly mixing if and only if there are no η -quasi-eigenfunctions.

The next lemma states that if φ is an η -quasi-eigenfunction for S then φ is close to an actual eigenfunction for S . This is a general fact about unitary operators on a Hilbert space.

Lemma 4 [GlWe]. *Let U be a unitary operator on a Hilbert space $\mathcal{H}$, and let $\varphi \in \mathcal{H}$ be an η -quasi-eigenfunction for U . Then there exists an eigenfunction φ_0 of U satisfying $\|\varphi - \varphi_0\| \leq \sqrt{2\eta}\|\varphi\|$ and $\|\varphi_0\| = \|\varphi\|$.*

Proof. By rescaling we may assume that $\|\varphi\| = 1$. Let μ_φ be the spectral measure for U associated with φ . Then, denoting by $\hat{\mu}_\varphi(n)$ the Fourier coefficients of μ_φ , we have

$$\langle \varphi, U^n \varphi \rangle = \int_{\mathbf{T}} \exp(2\pi i n t) d\mu_\varphi(t) = \overline{\hat{\mu}_\varphi(n)}.$$

By hypothesis, we have $|\hat{\mu}_\varphi(n)| > 1 - \eta$. Wiener's Theorem [Ka] states that

$$\sum_{t \in \mathbf{T}} \mu_\varphi(\{t\})^2 = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |\hat{\mu}_\varphi(n)|^2.$$

In our case this entails

$$\sum_{t \in \mathbf{T}} \mu_\varphi(\{t\})^2 \geq (1 - \eta)^2.$$

If we had $\mu_\varphi(\{t\}) < (1 - \eta)^2$ for all $t \in \mathbf{T}$ we would also have

$$\begin{aligned} (1 - \eta)^2 &\leq \sum_{t \in \mathbf{T}} \mu_\varphi(\{t\})^2 < \sum_{t \in \mathbf{T}} (1 - \eta)^2 \mu_\varphi(\{t\}) \\ &\leq (1 - \eta)^2 \int_{\mathbf{T}} d\mu_\varphi(t) = (1 - \eta)^2, \end{aligned}$$

reaching a contradiction. Hence there is a $t \in \mathbf{T}$ for which $\mu_\varphi(\{t\}) \geq (1 - \eta)^2$. Recall that $\mu_\varphi(\{t\})$ is the modulus squared of the the projection of φ of the eigenspace E_t of U of eigenvalue $\exp(2\pi i t)$. The claim now follows immediately by taking as φ_0 the normalized projection of φ on E_t . $\square$

Note that the eigenfunctions for S are exactly given by the functions

$$e_l(x, t_1, \dots, t_k) := \exp(2\pi i l t_1).$$

Let

$$\mathcal{M}_\varphi(\eta) = \{R \in \overline{\mathcal{S}} \mid \exists n \in \mathbf{N} \text{ such that } |(R^n \varphi, \varphi)| < (1 - \eta)\|\varphi\|^2\}.$$

The set $\mathcal{M}_\varphi(\eta)$ will be called *the set of transformations that mix φ up to η* .

Next we show that given any eigenfunction φ for S we can find close to S a transformation in $\mathcal{S}$ that mixes φ up to $\frac{1}{2}$.

Lemma 5. *Given $l \neq 0$ and $\varepsilon > 0$ there exists a transformation $R \in \mathcal{S}$ such that*

- (i) *R agrees with S outside a set of measure less than ε , and*
- (ii) *R mixes e_l up to $\frac{1}{2}$.*

Proof. Let $R \in \mathcal{S}$ be given by $F_f^{-1} S F_f$, where $f: x \in X \mapsto (f_1(x), \dots, f_k(x)) \in \mathbf{T}^k$ is a measurable map.

From formula (1) we see that if the terms $f_1(x) - f_1(S_0 x), \dots, f_{k-1}(x) + f_k(x) - f_k(S_0 x)$ vanish outside a set of measure less than ε then condition (i) is satisfied.

Furthermore, since $|\langle F_f^{-1} S^n F_f e_l, e_l \rangle| = |\langle S^n F_f e_l, F_f e_l \rangle|$, the transformation $F_f^{-1} S F_f$ mixes e_l up to $\frac{1}{2}$ if and only if S mixes $F_f e_l$ up to $\frac{1}{2}$. By the contrapositive of Lemma 4, S mixes $F_f e_l$ up to $\frac{1}{2}$ if $\|F_f \circ e_l - e_j\| > 1$ for all $j \in \mathbf{Z}$. We can achieve this by making $F_f e_l$ orthogonal to all the eigenfunctions e_j . Finally, notice that $F_f e_l(x, t_1, \dots, t_k) = \exp[2\pi i l(t_1 + f_1(x))]$. Thus $\langle F_f e_l, e_j \rangle = \delta_{jl} \int_X \exp[-2\pi i l f_1(x)] dx$. Thus if the latter integral vanishes we obtain that $F_f^{-1} S F_f$ mixes e_l up to $\frac{1}{2}$. Here is how we can define $f_1, \dots, f_k$ to achieve both conditions at once.

By Rohlin's lemma [Pe], for any positive integer N , there is a set $B \subset X$ such that $B, S_0 B, \dots, S_0^{N-1} B$ are mutually disjoint measurable sets and $\bigcup_{s=0}^{N-1} S_0^s B$ has measure larger than $\varepsilon/2$. We take $N > 2/\varepsilon$. Let $\{B_1, B_2\}$ be a partition of B in two sets of equal measure. Let also $\{C_1, C_2\}$ be a partition of $C := X - \bigcup_{s=0}^{N-1} S_0^s B$ in two sets of equal measure. Define $f_1(x) = 0$ for $x \in C_1 \cup \bigcup_{s=0}^{N-1} S_0^s B_1$ and $f_1(x) = (2l)^{-1}$ for $x \in C_2 \cup \bigcup_{s=0}^{N-1} S_0^s B_2$. Then we have

$$\int_X \exp[-2\pi i l f_1(x)] dx = 0,$$

and by the above discussion we have achieved condition (ii).

Set $f_i(x) = 0$, $i = 2, \dots, k$, for $x \in B \cup C$, and define inductively f_i , $i = 2, \dots, k$, so that for $x \in \bigcup_{s=0}^{N-2} S_0^s B$ we have $f_i(S_0 x) = f_i(x) + f_{i-1}(x)$. If we let $f(x) = (f_1(x), \dots, f_k(x))$, then $F_f^{-1} S F_f$ agrees with S on $x \in \bigcup_{s=0}^{N-2} S_0^s B$ which has measure larger than $1 - \varepsilon$ thus achieving condition (i). $\square$

Finally using the previous two lemmata we show that given any φ orthogonal to the constants we can find close to S a transformation in $\mathcal{S}$ which mixes φ up to η .

Lemma 6. *For all $\eta < 1/200$ and for all $\varepsilon > 0$ and $\varphi \in L_0^2(Y, \nu)$ there exists a transformation $R \in \mathcal{S}$ such that*

- (i) *R agrees with S outside a set of measure less than ε , and*
- (ii) *R mixes φ up to η , i.e., $R \in \mathcal{M}_\varphi(\eta)$.*

Proof. If $S \in \mathcal{M}_\varphi(\eta)$, we are done. Thus we can assume that S does not mix φ up to η . By rescaling, we can also assume $\|\varphi\| = 1$. By Lemma 4 applied to $\mathcal{H} = L_0^2(Y, \nu)$ there exists a (unique) *nonconstant* eigenfunction e_l for S with

$$\|\varphi - e_l\| \leq \sqrt{2\eta}.$$

By Lemma 5 there exists a transformation $R \in \mathcal{S}$ satisfying (i) and such that $|\langle R^n e_l, e_l \rangle| \leq \frac{1}{2}$ for some n .

Then for the same n one has

$$|\langle R^n \varphi, \varphi \rangle| \leq |\langle R^n e_l, e_l \rangle| + 3\|\varphi - e_l\| \leq \frac{1}{2} + 3\sqrt{2\eta} < 1 - \eta.$$

Hence we have $R \in \mathcal{M}_\varphi(\eta)$, and the lemma is proved. $\square$

Now we can conclude with the proof of Proposition 2.

Proof of Proposition 2. By Lemma 6, $S \in \overline{\mathcal{M}_\varphi(\eta)}$ for $\eta < 1/200$ and all $\varphi \in L_0^2(Y, \nu)$. Thus for any $R = F_f^{-1} S F_f \in \mathcal{S}$, where $f : x \in X \mapsto (f_1(x), \dots, f_k(x)) \in \mathbf{T}^k$ is a measurable map, we have that $R \in F_f^{-1} \overline{\mathcal{M}_\varphi(\eta)} F_f =$

$\overline{F_f^{-1}\mathcal{M}_\varphi(\eta)F_f} = \overline{\mathcal{M}_\psi(\eta)}$, where $\psi = \varphi \circ F_f^{-1}$. Thus $\mathcal{S} \subset \overline{\mathcal{M}_\varphi(\eta)}$ for $\eta < \eta_0$ and all $\varphi \in L_0^2(Y, \nu)$. It follows that $\mathcal{M}_\varphi(\eta)$ is dense in $\overline{\mathcal{S}}$.

Let Σ be a countable dense subset of $L_0^2(Y, \nu)$. Then $\bigcap_{\varphi \in \Sigma} \mathcal{M}_\varphi(\eta)$ is a dense G_δ because $\overline{\mathcal{S}}$ is a complete metric space and the sets $\mathcal{M}_\varphi(\eta)$ are obviously open. By Lemma 3 every $R \in \bigcap_{\varphi \in \Sigma} \mathcal{M}_\varphi(\eta)$ is weakly mixing. This concludes the proof of Proposition 2. $\square$

REFERENCES

- [Br] R. Bradley, *A stationary, pairwise independent, absolutely regular sequence for which the Central Limit Theorem fails*, Probab. Theory Related Fields **81** (1989), 1–10.
- [GIWe] S. Glasner and B. Weiss, *A weakly mixing upside-down tower of isometric extensions*, Ergodic Theory Dynamical Systems **1** (1981), 151–157.
- [Ja] S. Janson, *Some pairwise independent sequences for which the Central Limit Theorem fails*, Stochastics **23** (1988), 439–448.
- [Ka] Y. Katznelson, *An introduction to harmonic analysis*, Dover, New York, 1968.
- [Ma] B. Marcus, *The horocycle flow is mixing of all degrees*, Invent. Math. **46** (1978), 201–209.
- [Or] D. S. Ornstein, *On the root problem in ergodic theory*, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley, 1970, pp. 347–356.
- [Pe] K. Petersen, *Ergodic theory*, Cambridge Univ. Press, Cambridge, New York, Port Chester, Melbourne, and Sydney, 1983.
- [Ru] D. J. Rudolph, *k-Fold mixing lifts to weakly mixing isometric extensions*, Ergodic Theory Dynamical Systems **5** (1985), 445–447.

DEPARTMENT OF MATHEMATICS, 201 WALKER HALL, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611

E-mail address: flaminio@math.ufl.edu