ISRAEL JOURNAL OF MATHEMATICS 76 (1991), 289–298

MIXING OF ALL ORDERS AND PAIRWISE INDEPENDENT JOININGS OF SYSTEMS WITH SINGULAR SPECTRUM

BY

BERNARD HOST

Département de Mathématiques
C.S.P., Université de Paris-Nord, 93430 Villetaneuse, France
host@lumimath.univ-mrs.fr

ABSTRACT

We prove that every pairwise independent joining of weakly mixing systems with purely singular spectrum is independent; it follows that every mixing system with purely singular spectrum is mixing of all orders.

1. Introduction

1.1 RESULTS. The property of mixing of all orders has been introduced by Rokhlin ([R]), and Halmos ([H]) asked whether it holds for every mixing system. A similar question, for the weak mixing property, has been resolved by Furstenberg [F]. But, until now, only little progress has been done for the higher order mixing problem. Recently, Kalikow [K] showed that the answer is positive for rank one systems.

Solving the problem for zero entropy systems would give the general solution. Moreover, the property for a system to have an absolutely continuous spectrum looks like a very strong mixing property. The main result of this paper restricts significantly the class of systems where counterexamples are likely to be found:

Received February 10, 1991

THEOREM 1: *Every mixing system with purely singular spectrum is mixing of all orders.* (The complete definitions are given in section 1.3.)

A related problem is to find under what conditions every pairwise independent self-joining of a system is independent; such systems are called "pairwise independently determined" in [dJ–R]; they are weakly mixing and have zero entropy. Our Theorem 1 is actually a direct consequence of:

THEOREM 2: *Every pairwise independent joining of $r \geq 3$ weakly mixing systems with purely singular spectrum is independent.*

1.2 In section 2, we define 3-correlation measures for a pairwise independent joining of three systems. They are finite complex measures on $\mathbf{T}^3(\mathbf{T} = \mathbf{R}/\mathbf{Z})$, concentrated on the closed subgroup $H = \{(s,t,u) \in \mathbf{T}^3; s+t+u = 0\}$, with each of their natural projections on $\mathbf{T}^2$ absolutely continuous with respect to some product measure. Such measures are studied in sections 3 and 4; we show that their projections on $\mathbf{T}$ have no continuous singular part (Theorem 5). This question seems to be beyond the scope of Ergodic Theory, but the proof uses concepts of non-singular dynamics, namely the properties of the measures on $\mathbf{T}$ which are ergodic under some countable group of rotations. In section 5, we prove Theorem 2 by applying Theorem 5 to the correlation measures introduced in section 2, and Theorem 1 is deduced from Theorem 2.

1.3 DEFINITIONS. Given an integer $r \geq 2$ a **joining** of r dynamical systems $(X_i, \mathcal{B}_i, T_i, \mu_i), 1 \leq i \leq r$, is a probability measure ω on the product space $\prod_i (X_i, \mathcal{B}_i)$ which is invariant under the diagonal transformation $\prod_i T_i$ and whose projection (marginal) on each X_i is equal to μ_i; the joining ω is **pairwise independent** if its projection on $X_i \times X_j$ is equal to $\mu_i \times \mu_j$ for all $i \neq j$; it is **independent** if it is the product measure.

For $r \geq 2$, a system $(X, \mathcal{B}, T, \mu)$ is **mixing of order** r if, for every $A_i \in \mathcal{B}$ $(1 \leq i \leq r), \mu(A_1 \cap T^{n_2} A_2 \cap \cdots \cap T^{n_r} A_r)$ converges $\prod_i \mu(A_i)$ when $n_2 \to +\infty$ and $n_{i+1} - n_i \to +\infty$ for $2 \leq i < r$. Mixing of order to 2 is the ordinary mixing.

1.4 GENERAL HYPOTHESIS AND NOTATIONS. For all dynamical systems $(X, \mathcal{B}, T, \mu)$ we assume that $(X, \mathcal{B})$ is a standard Borel space, T an invertible bimeasurable transformation of X, and μ a T-invariant probability measure on $(X, \mathcal{B})$. By the spectral maximal type σ of this system, we mean the reduced one: a positive measure is absolutely continuous with respect to σ if and only if it is equal to measure σ_f of some $f \in L^2(\mu)$ with $\int f d\mu = 0$. In particular, if the system is ergodic, σ has no point mass at 0.

$\mathbf{T}$ is the quotient group $\mathbf{R}/\mathbf{Z}, m$ the Lebesgue measure of $\mathbf{T}$.

The Fourier transform $\hat{\sigma}$ of a finite complex measure σ on $\mathbf{T}$ is given by $\hat{\sigma}(n) = \int e^{2\pi int} d\sigma(t)$ for $n \in \mathbf{Z}$.

Except when the contrary is explicitly stated, all measures are supposed to be positive and finite. If μ and ν are measures on the same Borel space, the notations "$\mu \perp \nu$", "$\mu \not\perp \nu$", "$\mu \approx \nu$" mean "μ and ν are mutually singular", "μ and ν are not mutually singular" and "μ and ν are equivalent", that is "$\mu \ll \nu$" and "$\nu \ll \mu$". If $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ is a Borel map and μ is a measure on $(X, \mathcal{A}), f\mu$ denotes the image measure of μ by f.

2. 3-Correlation Measures of a Pairwise Independent Joining

2.1 We begin with some classical facts about correlation measures of the product of two systems $(X_i, b_i, T_i, \mu_i), 1 \leq i \leq 2$. For every $m, n \in \mathbf{Z}$, the measure $\mu_1 \times \mu_2$ is invariant under $T_1^m \times T_2^n$; these transformations define an action of $\mathbf{Z}^2$ on $L^2(\mu_1 \times \mu_2)$. The correlation measure $\sigma_{\phi,\psi}$ of $\phi, \psi \in L^2(\mu_1 \times \mu_2)$ is the finite complex measure on $\mathbf{T}^2$ with Fourier transform given by

$$\hat{\sigma}_{\phi,\psi}(m,n) = \int \phi(T_1^m x_1, T_2^n x_2)\overline{\psi(x_1,x_2)} d\mu_1(x_1) d\mu_2(x_2) \qquad (m, n \in \mathbf{Z}).$$

For $\phi \in L^2(\mu_1 \times \mu_2)$, we note σ_ϕ instead of $\sigma_{\phi,\phi}$; σ_ϕ is non-negative and $\sigma_{\phi,\psi} \ll \sigma_\phi$ for all $\psi \in L^2(\mu_1 \times \mu_2)$. For $f_i \in L^2(\mu_i), 1 \leq i \leq 2$, the correlation measure of the function $f_1 \otimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2)$ is $\sigma_{f_1} \times \sigma_{f_2}$.

2.2 THEOREM 3: *Let ω be a pairwise independent joining of the systems $(X_i, b_i, T_i, \mu_i), 1 \leq i \leq 3$, and $f_i \in L^2(\mu_i)$. Then*

(i) *There exists a (unique) finite complex measure τ on $\mathbf{T}^3$ with Fourier transform given by*

$$\hat{\tau}(m,n,p) = \int f_1(T_1^m x_1) f_2(T_2^n x_2) f_3(T_3^p x_3) d\omega(x_1, x_2, x_3) \qquad (m, n, p \in \mathbf{Z}).$$

(ii) *τ is concentrated on the closed subgroup*

$$H = \{(x_1, x_2, x_3) \in \mathbf{T}^3 ; x_1 + x_2 + x_3 = 0\}.$$

(iii) *The images of* $|\tau|$ *by the natural projections* $\pi_{1,2}, \pi_{1,3}$ *and* $\pi_{2,3}$ *of* $\mathbf{T}^3$ *on* $\mathbf{T}^2$ *are absolutely continuous with respect to* $\sigma_{f_1} \times \sigma_{f_2}, \sigma_{f_1} \times \sigma_{f_3}$ *and* $\sigma_{f_2} \times \sigma_{f_3}$*, respectively.*

($\pi_{1,2}$ *is the projection* $(x_1, x_2, x_3) \to (x_1, x_2)$*;* $\pi_{1,3}$ *and* $\pi_{2,3}$ *are defined analogously.)*

Proof: Let $G(x_1, x_2)$ be the conditional expectation of $f_3(x_3)$ given x_1 and x_2 : G is the element of $L^2(\mu_1 \times \mu_2)$ characterized by

$$\int G(x_1, x_2)F(x_1, x_2)d\mu_1(x_1)d\mu_2(x_2) = \int f_3(x_3)F(x_1, x_2)d\omega(x_1, x_2, x_3)$$

for every $F \in L^2(\mu_1 \times \mu_2)$. We denote by λ the correlation measure $\sigma_{f_1 \otimes f_2, \bar{G}}$, and τ the image measure of λ by the mapping $(x_1, x_2) \to (x_1, x_2, -x_1 - x_2)$ of $\mathbf{T}^2$ on $\mathbf{T}^3$. τ is concentrated on H and, for $m, n, p \in \mathbf{Z}$,

$$\begin{aligned}\hat{\tau}(m, n, p) &= \hat{\lambda}(m - p, n - p) \\ &= \int f_1(T_1^{m-p}x_1)f_2(T_2^{n-p}x_2)G(x_1, x_2)d\mu_1(x_1)d\mu_2(x_2) \\ &= \int f_1(T_1^{m-p}x_1)f_2(T_2^{n-p}x_2)f_3(x_3)d\omega(x_1, x_2, x_3) \\ &= \int f_1(T_1^{m}x_1)f_2(T_2^{n}x_2)f_3(T_3^{p}x_3)d\omega(x_1, x_2, x_3).\end{aligned}$$

This proves (i) and (ii).

λ is absolutely continuous with respect to $\sigma_{f_1 \otimes f_2} = \sigma_{f_1} \times \sigma_{f_2}$ and is the image by the projection $\pi_{1,2}$ of the measure τ which is concentrated on H; as the restriction to H of $\pi_{1,2}$ is one-to-one, $\pi_{1,2}|\tau|$ is equal to $|\lambda|$ and thus absolutely continuous with respect to $\sigma_{f_1} \times \sigma_{f_2}$. Similarly, $\pi_{1,3}|\tau| \ll \sigma_{f_1} \times \sigma_{f_3}$ and $\pi_{2,3}|\tau| \ll \sigma_{f_2} \times \sigma_{f_3}$.

3. Rotations and Measures on the Circle

3.1 Notations. For every $t \in \mathbf{T}$, let us denote by R_t the rotation $x \to x + t$ of $\mathbf{T}$. If D is a countable subgroup of $\mathbf{T}$, we say that a subset of $\mathbf{T}$ is D-invariant if it is invariant under R_t for every $t \in D$.

3.2 Group Associated with a Measure. Lemma: *Let* μ *be a measure on* $\mathbf{T}$*. There exists a countable subgroup* D *of* $\mathbf{T}$ *such that,*

for every $\lambda \ll \mu$ *and every* $t \in \mathbf{T}$ *with* $R_t\lambda \ll \mu$,

(1) $$R_t\lambda(A) = \lambda(A) \quad for \quad every\ D - invariant\ Borel\ set\ A.$$

Proof: We identify $L^1(\mu)$ with the space of finite complex measures which are absolutely continuous with respect to μ. For $\lambda \in L^1(\mu)$ and $t \in \mathbf{T}$, let $L_t(\lambda)$ denote the part of $R_t\lambda$ which is absolutely continuous with respect to μ. L_t is a linear operator on $L^1(\mu)$, of norm ≤ 1. As $L^1(\mu)$ is separable, there exists a countable set $J \subset \mathbf{T}$ such that $\{L_t; t \in J\}$ is dense in $\{L_t; t \in \mathbf{T}\}$ for the strong operator topology. Let D denote the subgroup of $\mathbf{T}$ spanned by J.

Let λ, t be as above, and A a D-invariant Borel set. For every $d \in D$, $L_d\lambda(A) \leq R_d\lambda(A) = \lambda(A)$ since A is D-invariant; by density, $L_t\lambda(A) \leq \lambda(A)$. Moreover, as $R_t\lambda$ is absolutely continuous with respect to μ, $R_t\lambda = L_t\lambda$ and $R_t\lambda(A) \leq \lambda(A)$. Substituting $\mathbf{T}\setminus A$ for A gives the reverse inequality, and the result follows. ∎

Remark: There are several groups with the property of the lemma. However, it can be easily checked that the σ-algebra $\mathcal{F}$ spanned by the D-invariant and the μ-null Borel sets does not depend of the choice of D: a Borel set A belongs to $\mathcal{F}$ if it satisfies (1). This property is equivalent to:

(2) $$\text{For every } t \in \mathbf{T}, \quad R_t(1_A \cdot \mu) \perp 1_{\mathbf{T}\setminus A} \cdot \mu.$$

3.3 D-ERGODIC MEASURES. Given a countable subgroup D of $\mathbf{T}$, a probability measure μ on $\mathbf{T}$ is D-ergodic if:

$$\mu(A) = 0 \text{ or } 1 \text{ for every } D\text{-invariant Borel set } A.$$

Note that there is no assumption of quasi-invariance for μ. In the next section, we shall use the following theorem of Brown and Moran:

THEOREM 4 ([B–M]; see [G–McG], Chap. 6): *Let D be a countable subgroup of $\mathbf{T}$ and μ a D-ergodic probability measure. If $\mu * \mu \not\perp \mu$, then μ is discrete or absolutely continuous with respect to the Lebesgue measure.*

The original proof of this theorem uses deep results of the theory of convolution measure algebras of J. L. Taylor. As the result seems interesting for ergodic theory, the author and F. Parreau will publish a short and elementary proof later on.

4. The Key Result

4.1 THEOREM 5: *Let ρ be a measure on $\mathbf{T}^3$, concentrated on*

$$H = \{(s,t,u) \in \mathbf{T}^3; s+t+u=0\},$$

and suppose that each of its natural projections on $\mathbf{T}^2$ is absolutely continuous with respect to some product measure. Then each projection of ρ on $\mathbf{T}$ is the sum of a discrete measure and an absolutely continuous measure (with respect to Lebesgue measure).

4.2 REDUCTION OF THE PROBLEM. We denote by S the symmetry $s \to -s$ of $\mathbf{T}$; P the mapping $(s,t) \to s+t$ of $\mathbf{T}^2$ on $\mathbf{T}$; and V the mapping $(s,t) \to (s,-s-t)$ of $\mathbf{T}^2$ on itself.

Each of the measures arising from ρ by coordinate permutation satisfies the same properties as ρ does, and the average of these measures too, since a sum of product measures is absolutely continuous with respect to some product measure; if the conclusion of the theorem holds for this average measure, it holds for ρ. We can thus restrict ourselves to the case where ρ is invariant by coordinate permutation. Similarly, we can assume that ρ is invariant by the symmetry $S \times S \times S$ of $\mathbf{T}^3$. We denote by λ the image of ρ by the projection $\pi_{1,2}; (s,t,u) \to (s,t)$ of $\mathbf{T}^3$ onto $\mathbf{T}^2$ and σ the image of ρ by the projection $(s,t,u) \to s$ of $\mathbf{T}^3$ onto $\mathbf{T}$. σ is invariant under S. We have to prove that σ is the sum of some discrete and some absolutely continuous measure.

The restriction to H of the projection $\pi_{1,2}$ is one-to-one, and its inverse is $(s,t) \to (s,t,-s-t)$; as ρ is concentrated on H, it is the image of λ by this mapping; ρ being invariant by coordinate permutation, it is the image of λ by $(s,t) \to (s,-s-t,t)$; therefore λ is the image of λ by V:

$$V\lambda = \lambda. \tag{1}$$

Looking at the projections of these two measures on $\mathbf{T}$ we get $SP\lambda = \sigma$, thus $P\lambda = S\sigma = \sigma$:

$$P\lambda = \sigma. \tag{2}$$

4.3. As λ is absolutely continuous with respect to some product measure, it is absolutely continuous with respect to the product of its projections: $\lambda \ll \sigma \times \sigma$. Let $F(\cdot,\cdot)$ be the Radon–Nikodym derivative of λ with respect to $\sigma \times \sigma$. For every $s \in \mathbf{T}$, let σ_s be the measure on $\mathbf{T}$ with density $F(s,\cdot)$ with respect to σ. For σ-almost every s, σ_s is a positive finite measure, absolutely continuous with respect to σ. For every bounded Borel function ϕ on $\mathbf{T}^2$,

$$\int \phi(s,t)d\lambda(s,t) = \int \left(\int \phi(x,t)d\sigma_s(t)\right) d\sigma(s) \tag{3}$$

and

$$\begin{aligned}\int \phi(s,t)dV\lambda(s,t) &= \int \phi(s,-s-t)d\lambda(s,t)\\ &= \int \left(\int \phi(s,-s-t)d\sigma_s(t)\right) d\sigma(s)\\ &= \int \left(\int \phi(s,t)dSR_s\sigma_s(t)\right) d\sigma(s).\end{aligned}$$

As $V\lambda = \lambda$ it follows that, for σ-almost $s, \sigma_s = SR_s\sigma_s$ and $R_s\sigma_s = S\sigma_s \ll S\sigma = \sigma$. Therefore for σ-almost every s,

$$R_s\sigma_s \ll \sigma. \tag{4}$$

4.4. Let D be a countable subgroup of $\mathbf{T}$, associated with σ as in the lemma. Let us denote by $\mathcal{D}$ the σ-algebra of D-invariant Borel sets.

Let $B \in \mathcal{D}$. For σ-almost every s, $\mathbf{1}_B \cdot \sigma_s$ is absolutely continuous with respect to σ; moreover, $R_s(\mathbf{1}_B \cdot \sigma_s)$ is absolutely continuous with respect to $R_s\sigma_s$, thus with respect to σ by (4). By the property of the group D, $R_s(\mathbf{1}_B \cdot \sigma_s)(\mathbf{T} \setminus B) = (\mathbf{1}_B \cdot \sigma_s)(\mathbf{T} \setminus B) = 0$, that is $\int \mathbf{1}_{\mathbf{T}\setminus B}(s+t)\mathbf{1}_B(t)d\sigma_s(t) = 0$. Integrating with respect to σ we get $\int \mathbf{1}_{\mathbf{T}\setminus B}(s+t)\mathbf{1}_B(t)d\lambda(s,t) = 0$, that is $P(\mathbf{1}_{\mathbf{T}\times B}\cdot\lambda)(\mathbf{T}\setminus B) = 0$. Therefore

$$\text{for every } B \in \mathcal{D}, \quad P(\mathbf{1}_{\mathbf{T}\times B} \cdot \lambda) \quad \text{is concentrated on } B. \tag{5}$$

Similarly, for every $A \in \mathcal{D}, P(\mathbf{1}_{A\times\mathbf{T}} \cdot \lambda)$ is concentrated on A, and $P(\mathbf{1}_{A\times B} \cdot \lambda)$ is concentrated on $A \cap B$. Therefore

$$\text{for every } A, B \in \mathcal{D} \text{ such that } A \cap B = \emptyset, \quad \lambda(A \times B) = 0. \tag{6}$$

Let $B \in \mathcal{D}$. The image by P of the measure $\mathbf{1}_{\mathbf{T}\times B} \cdot \lambda$ is absolutely continuous with respect to σ by (2), and concentrated on B by (5), thus absolutely continuous with respect to $\mathbf{1}_B \cdot \sigma$. Moreover, by (6), $\mathbf{1}_{\mathbf{T}\times B} \cdot \lambda = \mathbf{1}_{B\times B} \cdot \lambda$, which is absolutely continuous with respect to $(\mathbf{1}_B \cdot \sigma)\times (\mathbf{1}_B \cdot \sigma)$; the image by P of this last measure is $(\mathbf{1}_B \cdot \sigma) * (\mathbf{1}_B \cdot \sigma)$ by definition of the convolution product. Thus:

$$\text{(7)} \qquad \text{for every} \quad B \in \mathcal{D} \quad \text{with} \quad \sigma(B) \neq 0, \quad \mathbf{1}_B \cdot \sigma \not\perp (\mathbf{1}_B \cdot \sigma)^*(\mathbf{1}_B \cdot \sigma).$$

4.5 We claim that the restriction of σ to the algebra $\mathcal{D}$ is atomic. If not, there would exist $A \in \mathcal{D}$ with $\sigma(A) > 0$ and, for every $\varepsilon > 0$, a finite partition $(A_i; i \in I)$ of A such that $A_i \in \mathcal{D}$ and $\sigma(A_i) < \varepsilon$ for all i; then, by (5) and (6), $\mathbf{1}_{A\times\mathbf{T}} \cdot \lambda = \mathbf{1}_{A\times A} \cdot \lambda$ would be concentrated on the set $K = \cup_i A_i \times A_i$, with $\sigma \times \sigma(K) < \varepsilon$; thus $\mathbf{1}_{A\times\mathbf{T}} \cdot \lambda$ would be concentrated on a set of $\sigma \times \sigma$ measure 0. As λ is absolutely continuous with respect to $\sigma \times \sigma$, $\sigma(A) = \lambda(A \times \mathbf{T}) = 0$, and a contradiction would follow.

There exists therefore a finite or countable partition $(E_i; i \in I)$ of $\mathbf{T}$ such that, for all $i, \sigma(E_i) > 0$ and $\sigma(A \cap E_i) = \sigma(E_i)$ or 0 for all $A \in \mathcal{D}$. Each of the probability measures

$$\sigma_i = \frac{1}{\sigma(E_i)} \mathbf{1}_{E_i} \cdot \sigma$$

is D-ergodic. For each i, by (7), $\sigma_i \not\perp \sigma_i * \sigma_i$ and, by Theorem 4, σ_i is discrete or absolutely continuous, and the result follows.

5. Proof of Theorems 2 and 1

5.1 Proof of Theorem 2. Let ω be a pairwise independent joining of the systems $(X_i; \mathcal{B}_i, T_i, \mu_i), 1 \leq i \leq r$, and suppose that each of them is weakly mixing and has a purely singular spectrum. We have to prove that ω is the product measure $\prod_i \mu_i$.

Let us begin with $r = 3$, but assuming only that X_1 is weakly mixing and has a purely singular spectrum. For $1 \leq r \leq 3$, let $f_i \in L^2(\mu_i)$, and τ the correlation measure of these functions as constructed in section 2. Suppose that $\int f_1 d\mu_1 = 0$. Each projection of $\rho = |\tau|$ on $\mathbf{T}^2$ is absolutely continuous with respect to some product measure, thus, by Theorem 5, the first projection ρ_1 of ρ on $\mathbf{T}$ is the sum of some discrete measure and some absolutely continuous measure. But ρ_1 is absolutely continuous with respect to correlation measure σ_{f_1}, which is absolutely continuous with respect to the maximal spectral type of X_1 since f_1 has zero integral. Thus ρ_1 is continuous and purely singular. Therefore $\rho_1 = 0$, thus $\rho = 0, \tau = 0$, and $\int f_1(x_1)f_2(x_2)f_3(x_3)d\omega(x_1,x_2,x_3) = \hat{\tau}(0) = 0$. As ω is pairwise independent, it follows that ω is the product measure.

We turn now to the general case, and take $r > 3$. Proceeding by induction, we assume that the theorem is valid for $r-1$. The projection of ω on $X_1 \times \cdots \times X_{r-1}$ is a pairwise independent joining of these systems, and they are thus independent. *A fortiori*, $X_2, \ldots, X_{r-1}$ are independent, and X_1 is independent of $Y = X_2 \times \cdots \times X_{r-1}$. Similarly, X_r is independent of $Y : \omega$ can be viewed as a pairwise independent joining of X_1, Y and X_r. X_1 is weakly mixing and has a purely singular spectrum, thus, by the first part of the proof, this joining is independent and $\omega = \prod_i \mu_i$. ∎

Remark: The conclusion remains valid if only $r-2$ of the systems are weakly mixing and have a purely singular spectrum. ∎

5.2 PROOF OF THEOREM 1. We shall deduce Theorem 1 from Theorem 2. As the method is classical, we omit the details and restrict ourselves to prove the mixing of order 3.

Let $(X, \mathcal{B}, T, \mu)$ be a mixing system with purely singular spectrum, and suppose that it is not mixing of order 3. There exist $A, B, C \in \mathcal{B}$ and a sequence (m_i, n_i) in $\mathbf{Z}^2$ with $m_i \to +\infty$ and $n_i - m_i \to +\infty$, such that $\mu(A \cap T^{m_i}B \cap T^{n_i}C)$ does not converge to $\mu(A)\mu(B)\mu(C)$. Substituting a subsequence for (m_i, n_i), we can suppose that $\mu(A \cap T^{m_i}B \cap T^{n_i}C)$ converges to a limit $c \neq \mu(A)\mu(B)\mu(C)$. As $(X, \mathcal{B})$ is standard, we can identify it with $\{0,1\}^{\mathbf{N}}$ endowed with its Borel σ-algebra. Substituting a subsequence for (m_i, n_i), we can assume that the sequence $\mu(D \cap T^{m_i}E \cap T^{n_i}F)$ converges for all cylinder sets D, E, F; there exists a (unique) probability measure ω on X^3 such that $\omega(D \times E \times F) = \operatorname{Lim} \mu(D \cap T^{m_i}E \cap T^{n_i}F)$ for all cylinder sets D, E, F. Moreover, for every $D, E, F, D', E', F' \in \mathcal{B}$ and every $m, n \in \mathbf{Z}$,

$$|\mu(D \cap T^m E \cap T^n F) - \mu(D' \cap T^m E' \cap T^n F')| \leq \mu(D \Delta D') + \mu(E \Delta E') + \mu(F \Delta F').$$

It follows that the sequence $\mu(D \cap T^{m_i}E \cap T^{n_i}F)$ converges to $\omega(D \times E \times F)$ for all $D, E, F \in \mathcal{B}$. The measure ω is invariant under $T \times T \times T$, and each of its projections on X is $\mu : \omega$ is a self-joining of X; it is pairwise independent because X is mixing, and $\omega(A \times B \times C) = c \neq \mu(A)\mu(B)\mu(C)$, thus it is not independent, and Theorem 2 provides a contradiction. ∎

References

[B–M] G. Brown and W. Moran, *A dichitomy for infinite convolution products of discrete measures*, Proc. Cambridge Phil. Soc. **73** (1973), 307–316.

[F] H. Furstenberg, *Recurrence in Ergodic Theory and in Combinatorial Number Theory*, Princeton University Press, Princeton, 1981.

[G–Mc G] C. Graham and C. McGehee, *Essays in Commutative Harmonic Analysis*, Springer-Verlag, New York, Heidelberg, Berlin, 1979.

[H] P. Halmos, *Lectures on Ergodic Theory*, Chelsea Publishing Co., New York, 1956.

[dJ–R] A. Del Junco and D. Rudolf, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory and Dynamical Systems **7** (1987), 531–557.

[K] S. Kalikow, *Twofold mixing implies threefold mixing for rank one transformations*, Ergodic Theory and Dynamical Systems **4** (1984), 237–259.

[R] V Rokhlin, *On endomorphisms of compact commutative groups*, Izv. Akad. Nauk SSSR **13** (1949), 329–340.