Ergod. Th. & Dynam. Sys. (2010), **30**, 361–377
doi:10.1017/S0143385709000212
© *2009 Cambridge University Press*
Printed in the United Kingdom

Partitions with independent iterates in random dynamical systems

BORIS BEGUN† and ANDRÉS DEL JUNCO‡

† *Institute of Mathematics, Hebrew University, 91904 Jerusalem, Israel*
(e-mail: begun@math.huji.ac.il)

‡ *Department of Mathematics, University of Toronto, Toronto ON, M5S 2E4, Canada*
(e-mail: deljunco@math.toronto.edu)

(*Received* 4 *April* 2008 *and accepted in revised form* 7 *October* 2008)

Abstract. Krengel characterized weakly mixing actions (X, T) as those measure-preserving actions having a dense set of partitions of X with infinitely many jointly independent images under iterates of T. Using the tools developed in later papers—one by del Junco, Reinhold and Weiss, another by del Junco and Begun—we prove analogues of these results for weakly mixing random dynamical systems (in other words, relatively weakly mixing systems).

1. *Introduction*

For measure-preserving transformations, the notion of weak mixing is one of the notions stronger than ergodicity. Along with the notions of mixing, K-property, etc, it belongs to the hierarchy of statistical properties of transformations in ergodic theory. This notion complements the notion of discrete spectrum and has several equivalent characterizations mostly going back to the paper by Koopman and von Neumann [**KN**] from the early 1930s. Another unexpected characterization of weak mixing was given in the 1970s by Krengel [**K**]. In this paper we prove two analogues of Krengel's result for systems that are weakly mixing relative to a factor. Our claim can be alternatively stated in the language of random dynamical systems.

Before presenting the original theorem and its relativization, we provide necessary definitions concerning partitions of probability spaces (for us, a partition is a representation of the space as a finite disjoint union of measurable sets), and recall the notion of weak mixing in terms convenient for our purposes.

A family of measurable sets in a probability space (X, μ) is *independent* if for every finite sub-family $A_1, A_2, \ldots, A_k$ we have

$$\mu\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} \mu(A_i).$$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

A family of measurable partitions of X is independent if every family of sets of the partitions—one set from each partition—is independent. If $P = \{A_1, \ldots, A_n\}$ is an ordered partition then $\text{dist}_\mu P$, the distribution of P, is a vector in $\mathbb{R}^n$ defined by $(\text{dist}_\mu P)_i = \mu(A_i)$. Consequently, if $P = \{A_1, \ldots, A_k\}$ and $Q = \{B_1, \ldots, B_k\}$ are two partitions with the same number of atoms, then we will say that P and Q have the same distribution if $\mu(A_i) = \mu(B_i)$ for all i. The set of all (ordered) partitions into n atoms can be turned into a complete metric space with the metric $d_\mu(P, Q) = \frac{1}{2}\sum_{i=1}^n \mu(A_i \triangle B_i)$. Whenever there is no ambiguity we will omit the subscripts and use plain dist and d.

By a measure-preserving system we mean a quadruple $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, where $(X, \mathcal{B}, \mu)$ is a probability space and T is an automorphism of $(X, \mathcal{B}, \mu)$, that is T is an invertible μ-preserving map. A sequence $\{n_i\} \subset \mathbb{Z}$ is called a *mixing sequence* for T if for all $A, B \in \mathcal{B}$

$$\mu(A \cap T^{-n_i} B) \to \mu(A)\mu(B) \quad \text{as } i \to \infty.$$

One characterization of *weak mixing* of T is that the product system

$$\mathbf{X} \times \mathbf{X} = (X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$$

is ergodic. An equivalent one is that there exists a mixing sequence $\{n_i\} \subset \mathbb{N}$ for T having density one in $\mathbb{N}$, or equivalently there exists any mixing sequence at all.

THEOREM (KRENGEL). *$(X, \mathcal{B}, \mu, T)$ is a weakly mixing system if and only if, for every finite measurable partition P of X and $\epsilon > 0$, there is an infinite sequence $0 = n_0 < n_1 < \cdots$ and a partition Q with the same distribution as P such that $d(P, Q) < \epsilon$ and the partitions $\{T^{-n_i} Q\}_{i=0}^\infty$ are jointly independent.*

Krengel's very technical proof was substantially simplified in 1999 by del Junco *et al* [**JRW**], paving the way for generalizations and refinements. Among other things it was shown there that the sequence in Krengel's theorem can be chosen inside any prescribed mixing sequence for T. The key tool for this was an extension result for stationary measures on $A^{\mathbb{Z}}$, where A is a finite set. In the subsequent paper [**BJ**] the extension theorem was sharpened and generalized to A^G, where G is any discrete amenable group, permitting the Krengel theorem to be lifted to free actions of discrete amenable groups.

We now review some general facts about factors and skew-products, referring the reader to [**G**, §§2.1, 2.2 and 3.3] or [**F**, Ch. 5] for more detail. We restrict our attention to systems $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ such that X is a complete separable metric space and $\mathcal{B}$ is the μ-completion of the Borel σ-algebra of X. We will call such systems standard. This assumption involves no loss of generality for theorems of the kind we are concerned with here. We will be studying the properties of $\mathbf{X}$ relative to a distinguished T-invariant factor algebra $\mathcal{F}$. There is no loss of generality in assuming that $\mathcal{F}$ is given by $\mathcal{F} = \pi^{-1}\bar{\mathcal{F}}$, where

$$\pi : \mathbf{X} \to \mathbf{\Omega} = (\Omega, \bar{\mathcal{F}}, \mathbb{P}, \theta)$$

is a factor map from $\mathbf{X}$ to another standard system $\mathbf{\Omega}$. The sets $\pi^{-1}(\omega)$ for $\omega \in \Omega$ are the fibers of π. The measure μ disintegrates over the factor as

$$\mu = \int_\Omega \mu_\omega \, d\mathbb{P}(\omega),$$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

where μ_ω is supported on the fiber $\pi^{-1}(\omega)$. The μ-invariance of T yields $T\mu_\omega = \mu_{\theta\omega}$, or $\mu_\omega A = \mu_{\theta\omega}(TA)$ for all $A \in \mathcal{F}$ and $\mathbb{P}$-almost every (a.e.) ω.

By a theorem of Rokhlin (see for example [**G**, Theorem 3.18]), when the factor system is ergodic we can represent $\mathbf{X}$ as a skew-product over $\mathbf{\Omega}$:

$$(X, \mathcal{B}, \mu) = (\Omega, \bar{\mathcal{F}}, \mathbb{P}) \times (Y, \mathcal{C}, \nu),$$
$$\mathcal{F} = \bar{\mathcal{F}} \times Y,$$
$$T(\omega, y) = (\theta\omega, T_\omega y),$$

where T_ω is an automorphism of a probability space $(Y, \mathcal{C}, \nu)$ which depends measurably on ω. The pair $(\theta, \{T_\omega\})$ is also called a random dynamical system. From now on we will identify $\mathcal{F}$ with $\bar{\mathcal{F}}$ and write $\mathcal{F}$ for both.

One definition of *relative weak mixing* uses the relative product of $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ with itself, otherwise called its relatively independent joining with itself. This is the system

$$\mathbf{X} \times_{\mathbf{\Omega}} \mathbf{X} = (X \times X, \mathcal{B} \times \mathcal{B}, \lambda, T \times T),$$

where the relative product measure $\lambda = \mu \times_{\mathbf{\Omega}} \mu$ is defined by

$$\lambda = \int_\Omega \mu_\omega \times \mu_\omega \, d\mathbb{P}(\omega)$$

(see [**F**, §5.5] or [**G**, §6.1]). Note that λ is supported on the fibered product

$$\bigcup_{\omega \in \Omega} (\pi^{-1}(\omega) \times \pi^{-1}(\omega)) \subset X \times X.$$

Consequently, $\mathbf{\Omega}$ is a factor of $\mathbf{X} \times_{\mathbf{\Omega}} \mathbf{X}$ in a canonical way.

By definition $\mathbf{X}$ (or simply T) is *weakly mixing relative to* $\mathcal{F}$ if $\mathbf{X} \times_{\mathbf{\Omega}} \mathbf{X}$ is ergodic relative to $\mathcal{F}$. What this means is that any invariant function for the relative product is $\mathcal{F}$-measurable. There is an equivalent characterization as a (relative) mixing condition: T has a *relatively mixing sequence* $\{n_i\}$, which means that for any μ-measurable sets A and B

$$\mu_\omega(A \cap T^{-n_i} B) - \mu_\omega(A)\mu_\omega(T^{-n_i} B)$$

converges to zero in measure $\mathbb{P}$ (as a function of ω). Equivalently, T has a relative mixing sequence of density one in $\mathbb{N}$. The existence of a relative mixing sequence of density one in the presence of relative weak mixing follows easily from [**F**, Lemma 7.6]. The converse fact, that a relative mixing sequence implies relative weak mixing, is easy since a relative mixing sequence for $\mathbf{X}$ is also a relative mixing sequence for $\mathbf{X} \times_{\mathbf{\Omega}} \mathbf{X}$ and the existence of a relative mixing sequence certainly forces relative ergodicity.

We will present two versions of Krengel's theorem for relative weakly mixing systems. For convenience we will state and prove the results in the case of ergodic systems. Theorem 1 remains valid without this assumption and so does Theorem 2, with a minor modification. The proofs are not much more difficult. After proving the results in §4 we will make some comments about how to remove the ergodicity assumption. We remark that when T is weakly mixing relative to $\mathcal{F}$, ergodicity of T is equivalent to ergodicity of θ.

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

The first theorem says that by a uniformly small perturbation of a given partition it is possible to obtain one with the same distribution on almost all fibers so that it has infinitely many independent iterates on a set of fibers of probability arbitrarily close to 1. 'Probability' here means the measure on the factor, which is the customary terminology of the random dynamical systems theory. Consequently, 'almost surely (a.s.)' will mean 'for $\mathbb{P}$-a.e. $\omega \in \Omega$'.

THEOREM 1. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic system with a factor $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ where μ decomposes as $\mu = \int \mu_\omega \, d\mathbb{P}$. Assume that $\{n_i\}_{i=0}^{\infty}$ is a relatively mixing sequence for T, and let P be a finite partition of X for which there is an $\alpha > 0$ such that*

$$\text{for all } p \in P : \mu_\omega(p) > \alpha \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Then for every $\epsilon > 0$ there exist a partition Q of X, a set $E \subset \Omega$, $\mathbb{P}(E) < \epsilon$, and a subsequence $\{m_i\}$ of $\{n_i\}$, such that

$$\operatorname{dist}_{\mu_\omega} Q = \operatorname{dist}_{\mu_\omega} P \quad \text{for } \mathbb{P}\text{-a.e. } \omega, \tag{1}$$

$$d_{\mu_\omega}(P, Q) < \epsilon \quad \text{for } \mathbb{P}\text{-a.e. } \omega, \tag{2}$$

$$\{T^{-m_i} Q\} \text{ is jointly independent with respect to } \mu_\omega \quad \text{for all } \omega \notin E. \tag{3}$$

The independence claim holds for a set of ω of measure (probability) arbitrarily close to 1, not $\mathbb{P}$-almost everywhere. This is not a drawback of the method—if we require that Q be uniformly close to P then we cannot obtain independence on almost all fibers. For a counterexample consider the so-called (T, T^{-1}) transformation (in our notation (S, S^{-1})) in its simplest version. Let $\Omega = Y = \{-1, 1\}^{\mathbb{Z}}$ with the standard product measure $\mathbb{P}$ coming from the $(1/2, 1/2)$ distribution of probabilities on $\{-1, 1\}$, and let S be the Bernoulli shift on both Y and Ω: for $\{a_n\} \in \{-1, 1\}^{\mathbb{Z}}$ define $(Sa)_n = a_{n+1}$. Let $\theta = S$ and define the (S, S^{-1}) map as the following skew-product transformation T on $(\Omega \times Y, \mathbb{P} \times \mathbb{P})$:

$$T(\omega, y) = (S\omega, S^{\omega_0} y).$$

The family $\{S^{\omega_0}\}$ contains only two distinct transformations—the shifts to the left and to the right. The system is weakly mixing with respect to the base (it is even strongly mixing in the L^1 sense—see [**R**] for the definitions and the proof), but the independence a.s. cannot be achieved even along a set of just two numbers.

To see this first observe that if $T^{-n_1} Q$ were μ_ω-independent of $T^{-n_2} Q$ for almost all (a.a.) ω then, for $n = n_1 - n_2$, we would also have $T^{-n} Q$ μ_ω-independent of Q for a.a. ω. We may assume $n > 0$. Choose a partition R of Y such that

$$\operatorname{dist}_{\mathbb{P}} R = (\tfrac{1}{2}, \tfrac{1}{2}) \quad \text{and} \quad d_{\mathbb{P}}(R, SR) < \tfrac{1}{100}.$$

For example, viewing R as a map into the index set $\{-1, 1\}$ let $R(y) = \operatorname{sgn}(\sum_{i=1}^{10^6+1} y_i)$. (Note that this sum cannot be 0.) Define P on X by $P(\omega, y) = R(y)$. Observe that

$$T^n(\omega, y) = (S^n \omega, S^{\phi(n,\omega)} y),$$

so

$$T^{-n} P(\omega, y) = P(T^n(\omega, y)) = R(S^{\phi(n,\omega)} y),$$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

where

$$\phi(n, \omega) = \sum_{i=0}^{n-1} \omega_i.$$

Let $\delta = 0$ or 1 according to whether n is even or odd. Then the set $E = \{\omega \mid \phi(n, \omega) = \delta\}$ has positive measure and for $\omega \in E$ we will have

$$d_{\mu_\omega}(P, T^{-n}P) = d_{\mathbb{P}}(R, S^{\delta}R) < \tfrac{1}{100}.$$

Now suppose Q is a partition of X such that for a.a. ω

$$\operatorname{dist}_{\mu_\omega} Q = (\tfrac{1}{2}, \tfrac{1}{2}) \quad \text{and} \quad d_{\mu_\omega}(Q, P) < \tfrac{1}{100}.$$

Then it is easy to see that for $\omega \in E$

$$d_{\mu_\omega}(Q, T^{-n}Q) < \tfrac{3}{100}.$$

On the other hand, if Q and $T^{-n}Q$ were independent with respect to μ_ω a.s. then we would have $d_{\mu_\omega}(Q, T^{-n}Q) = \frac{1}{2}$ a.s., a contradiction.

In our second result, by dropping the requirement that the perturbation be uniformly small we are able to obtain the desired independence on almost all fibers, rather than just a large set of fibers.

THEOREM 2. *With the same hypotheses as in Theorem 1 one can find a partition Q and a subsequence $\{m_i\}$ of $\{n_i\}$ such that* $\operatorname{dist}_{\mu_\omega} Q = \operatorname{dist}_{\mu_\omega} P$ *for $\mathbb{P}$-a.e. ω, $d_\mu(P, Q) < \epsilon$ and the partitions $\{T^{-m_i}Q\}$ are jointly independent with respect to μ_ω for $\mathbb{P}$-a.e. ω.*

Remark. Of course, to say that $d_\mu(P, Q)$ is small is just to say that $d_{\mu_\omega}(P, Q)$ is small for all but a small set of ω.

Our presentation conforms to the following plan. After a section on preliminaries and notation we prove in §3 a result (Proposition 1) on the existence of measures with prescribed marginals—a non-stationary generalization of the stationary extension result, [**BJ**, Theorem 2]. Its proof mimics closely the proof of the stationary result. Theorems 1 and 2 are proved in §4.

2. *Some definitions and notation*

In this section we introduce some notation and definitions that pertain to both §§3 and 4. Additional notation will be introduced later as needed.

First we adopt a more formal definition of an ordered partition. A (finite) partition of a measurable space $(X, \mathcal{B})$ is a measurable map P from X to a finite index set A (so by definition we are dealing with ordered partitions). The sets $p = P^{-1}(a)$, $a \in A$, are the atoms of P and we write $p \in P$. If $(X, \mathcal{B})$ carries a probability measure μ then dist P (or $\operatorname{dist}_\mu P$ if the measure needs to be emphasized) denotes the measure $\mu \circ P^{-1}$ on A. If $B \in \mathcal{B}$, $\mu(B) \neq 0$, then $\operatorname{dist}_\mu (P|B)$ refers to the restriction $P|_B$ and the normalized measure μ_B. We write $P \prec Q$ (Q refines P) if each atom of P is a union of atoms of Q. For two partitions P and Q indexed by the same alphabet A, the distance $d(P, Q)$, or $d_\mu(P, Q)$, between them is

$$d(P, Q) = \mu\{x \in X \mid P(x) \neq Q(x)\} = \frac{1}{2} \sum_{a \in A} \mu(P^{-1}(a) \triangle Q^{-1}(a)).$$

For a fixed alphabet A this metric defines a complete metric space of partitions.

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

Two partitions of a probability space (X, μ) are independent if every atom p of the first one is independent of every atom q of the second one: $\mu(p \cap q) = \mu(p)\mu(q)$. Joint independence of any finite number of partitions is defined similarly. Joint independence of infinitely many partitions means the joint independence of any finite subset of them. If P and Q are finite partitions of a probability space (X, μ) then we will say P is δ-independent of Q whenever

$$\|\text{dist}(P|q) - \text{ dist } P\|_\infty \leq \delta \quad \text{ for all } q \in Q.$$

(We are identifying a measure σ on a finite set A with the vector $\{\sigma(a)\}_{a\in A}$.) It is implicit in the definition that $\mu(q) > 0$ for all $q \in Q$. When P is δ-independent of Q we will write $P \perp_\delta Q$ (or $P \perp_{\delta,\mu} Q$). Note that this definition of approximate independence differs from a more standard one which uses the l_1-norm and allows for a small exceptional set of q. In fact any reasonable definition of approximate independence could be used in the statements of our results but the one we have chosen is convenient for the proofs. We will make use of the fact that if $P \perp_\delta Q$ and $Q \succ R$ then $P \perp_\delta R$.

3. *Building a measure with prescribed marginals*

This section is static—no dynamics is involved. The object of study is measures on A^Σ, where A is a finite set (alphabet) and Σ is countable. We prove a result on the existence of a probability measure with prescribed projections on A^K for $K \in \mathcal{F}$, a family of finite subsets of Σ.

We will be working with measures on A^H for various subsets $H \subset \Sigma$ and we use the usual Borel structure on A^H. Mostly H will be finite. $H = \phi$ is allowed, in this case A^H is just a singleton. Whenever $K \subset L \subset \Sigma$ we will denote by Pr^K the projection map from A^L to A^K, which is a finite partition of A^L when K is finite. Pr^K induces the projection (or restriction) mapping π_K of measures: if m is a measure on A^L then $\pi_K m = m \circ (Pr^K)^{-1}$. $\pi_K m$ is then a marginal of m, and m is an extension of $\pi_K m$. We identify a word $x \in A^K$ with the atom $(Pr^K)^{-1}x$ of Pr^K, which is a subset in A^L, and $m(x)$ will mean the same as $\pi_K m(x)$.

If $K_1, K_2 \subset L$ and $x \in A^{K_1}$, $y \in A^{K_2}$, $m(y) > 0$, we will use the notation

$$m(x|y) = m((Pr^{K_1})^{-1}x \cap (Pr^{K_2})^{-1}y)/m(y).$$

For H finite we will say that a probability measure m on A^H is δ-independent whenever there is an indexing $H = \{h_1, \ldots, h_n\}$ such that

$$Pr^{\{h_i\}} \perp_{\delta,m} Pr^{\{h_1,\ldots,h_{i-1}\}} \quad \text{for all } i = 2, \ldots, n.$$

One easily sees that if H is non-empty then δ-independence of m implies that

$$\|m - p_H\|_\infty \leq (|H| - 1)\delta, \tag{1}$$

where p_H is the product measure on A^H that has the same one-dimensional marginals as m. Note also that if m is δ-independent then so is $\pi_I m$ for any $I \subset H$.

Two measures on A^{K_1} and A^{K_2} are called consistent if their marginals on $A^{K_1 \cap K_2}$ coincide. Consistency of a family of measures means their pairwise consistency. And

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

the last piece of notation: we write $x \overset{\delta}{\sim} y$ whenever x and y are two elements of a space which are not more than δ apart with respect to some metric or norm which is clear from the context.

PROPOSITION 1. *For every $N \in \mathbb{N}$ and $\alpha > 0$ there is a $\delta = \delta(\alpha, N) > 0$ such that the following assumptions imply the conclusion below.*
Assumptions. A is a finite alphabet, $\mathcal{F}$ is a family of finite subsets of Σ and for every set $K \in \mathcal{F}$ a probability measure μ_K on A^K is given. These data satisfy:
(F) *for all $n \in \Sigma$ we have $|\bigcup\{K \in \mathcal{F} \mid n \in K\}| \leq N$;*
(M1) *the measures μ_K are consistent with each other;*
(M2) *every atom of each one-dimensional marginal of each measure μ_K has size $\geq \alpha$;*
(M3) *every measure μ_K is δ-independent.*
Conclusion. The measures $\{\mu_K\}_{K \in \mathcal{F}}$ have a common extension to A^Σ.

By the classical Kolmogorov extension theorem the proposition is an immediate consequence of the following claim that concerns measures on A^L for L finite.

CLAIM A. *For every $N \in \mathbb{N}$ and $\alpha > 0$ there is a $\delta = \delta(\alpha, N) > 0$ and a $\beta = \beta(\alpha, N) > 0$ such that the following implication holds. Suppose that the assumptions of Proposition 1 are valid. Suppose further that L, L' are finite subsets of Σ, $L \subset L'$, $|L' \backslash L| = 1$ and λ is a measure on A^L that is consistent with every μ_K and β-independent. Then there is a measure λ' on $A^{L'}$ that extends λ, and is again consistent with every μ_K and β-independent.*

Remark. Clearly $\delta \leq \beta$. Typically $\delta \ll \beta \ll 1$. Explicitly, one can take $\beta = \alpha^N/2N$, as will be seen from the proof. The expression for $\delta = \delta(\alpha, N)$ can also be developed in principle but we will not need it.

Proof. To prove Claim A we will need two lemmas. These are Lemmas 2.1 and 2.2 from [**BJ**], the first of which appeared already in [**JRW**] as part of Proposition 1.2. To make our presentation self-contained, we provide sketches of the proofs here and refer the reader to the previous papers for more details.

The first lemma deals with *signed* measures, for which the notions of projection, extension and consistency remain meaningful.

LEMMA 1. *Let K be a finite set and $\{K_i\}$ a family of subsets of K whose union is K. Any consistent family of signed measures on A^{K_i} has a common extension—a signed measure on A^K.*

Sketch of the proof. We are given the measures μ_{K_i}, $i = 1, \ldots, n$. Pick an arbitrary *probability* measure q on A^K. For every $\mathcal{J} \subset \{1, \ldots, n\}$ let $K_\mathcal{J} = \cap\{K_j \mid j \in \mathcal{J}\}$ and denote by $\mu_\mathcal{J}$ the common projection of μ_{K_j}, $j \in \mathcal{J}$, onto $A^{K_\mathcal{J}}$, multiplied by $\pi_{K \backslash K_\mathcal{J}} q$ and thus extended to A^K. We claim that the measure

$$\mu = \sum_{\varnothing \neq \mathcal{J} \subset \{1, \ldots, n\}} (-1)^{|\mathcal{J}|+1} \mu_\mathcal{J}$$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

is a joint extension of the measures in the family. This expression is inspired by the inclusion–exclusion formula in combinatorics and the proof is based on the same idea—when the expression for μ is projected onto any of the A^{K_i}, the sum over all $\Phi \neq \mathcal{J} \subsetneqq \{1, \ldots, n\}$ can be divided into pairs ($\mathcal{J}_1 = \mathcal{J}_2 \cup \{i\}$) that cancel out. □

LEMMA 2. *Suppose that V and W are finite-dimensional real normed vector spaces and $\Pi\colon V \to W$ is any surjective linear map. Then there is a constant $C = C(\Pi)$ with the following property: if $v \in V$ and $w = \Pi(v) \neq 0$ then Π has a right inverse $B(\Pi B = \mathrm{id}_W)$ such that $Bw = v$ and $\|B\| \leq C\|v\|/\|w\|$.*

Sketch of the proof. Using the equivalence of any two norms in a finite-dimensional space it is not hard to see that there is no limitation of generality in the assumption that V is Euclidean and Π is the orthogonal projection onto its subspace W. In this case the result is immediate ($C = 1$). □

Continuing with the proof of Claim A, we shall assume, as we may, that $\bigcup \mathcal{F} = \Sigma$. Let $L' = L \cup \{n\}$, $\mathcal{F}_n = \{K \in \mathcal{F} \mid n \in K\}$ and consider the following two families of subsets of Σ:

$$\mathcal{S} = \{K \cap L' \mid K \in \mathcal{F}_n\},$$
$$\mathcal{R} = \{K \cap L \mid K \in \mathcal{F}_n\} = \{S\backslash\{n\} \mid S \in \mathcal{S}\}.$$

$\mathcal{S}$ and $\mathcal{R}$ are finite families of sets. All sets in $\mathcal{S}$ contain n. Let

$$\bar{S} = \bigcup \mathcal{S} \quad \text{and} \quad \bar{R} = \bigcup \mathcal{R}.$$

Clearly $\bar{R} = \bar{S}\backslash\{n\}$, $\bar{S} \subset L'$, $\bar{R} \subset L$, and by (F) the cardinality of $\bar{S}$ does not exceed N.

We are going first to construct a probability measure σ on $A^{\bar{S}}$—a prospective marginal of λ'. It will satisfy three conditions derived from the corresponding conditions imposed on λ' in Claim A, namely

$$\pi_{\bar{R}}\sigma = \pi_{\bar{R}}\lambda, \tag{2}$$
$$\pi_{K\cap L'}\sigma = \pi_{K\cap L'}\mu_K \quad \text{for every } K \in \mathcal{F}_n, \tag{3}$$
$$Pr^{\{n\}} \perp_{\beta,\sigma} Pr^{\bar{R}}. \tag{4}$$

Let us assume for the moment that we have already constructed σ satisfying (2), (3) and (4). Because of (2) we can then define λ' on $A^{L'}$ to be the relative product measure $\lambda' = \lambda \times_{[A^{\bar{R}}]} \sigma$, that is

$$\lambda'(x) = \frac{\lambda(Pr^I x)\sigma(Pr^{\bar{S}} x)}{\lambda(Pr^{\bar{R}} x)} \quad \text{for all } x \in A^{L'}.$$

λ' is certainly an extension of λ and of σ. λ' is consistent with every μ_K, $K \in \mathcal{F}$. If $K \ni n$ then λ' inherits this property from σ, otherwise—from λ.

Moreover the definition of λ' together with (4) implies that for all $x \in A^{\{n\}}$, $y \in A^{\bar{R}}$ and $z \in A^{L\backslash\bar{R}}$ we have

$$\lambda'(x|y \times z) = \sigma(x|y) \overset{\beta}{\sim} \sigma(x) = \lambda'(x).$$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

This means that $Pr^{\{n\}} \perp_{\beta,\lambda'} Pr^L$, and since λ is β-independent it follows that λ' is also β-independent, completing the proof of Claim A and Proposition 1 under the assumption that a measure σ has been built subject to (2)–(4).

We now proceed to construct σ. For each $S = K \cap L' \in \mathcal{S}$ and for each $R = K \cap L \in \mathcal{R}$, where $K \in \mathcal{F}_n$, let

$$\mu_S = \pi_S \mu_K \quad \text{and} \quad \nu_R = \pi_R \mu_K.$$

By (M1) these requirements unambiguously define μ_S and ν_R even though $K_1 \cap L' = K_2 \cap L'$ for $K_1, K_2 \in \mathcal{F}_n$ does not imply $K_1 = K_2$. Moreover, $\{\nu_R\}_{R\in\mathcal{R}}$ is a consistent family of measures, and so is $\{\mu_S\}_{S\in\mathcal{S}}$. We let V denote the vector space of all signed measures ρ on $A^{\bar{R}}$ and W the space of all consistent families $\{\rho_R\}_{R\in\mathcal{R}}$, where ρ_R is a signed measure on A^R. We regard W as a linear subspace of the direct sum of $|\mathcal{R}|$ spaces, one for each $R \in \mathcal{R}$. All the spaces are endowed with the l_∞-norm (on W the norm is defined by $\|\{\nu_R\}_{R\in\mathcal{R}}\|_\infty = \max_{R\in\mathcal{R}} \|\{\nu_R\}\|_\infty$). We let $\Pi : V \to W$ denote the projection map, that is $\Pi(\rho) = \{\pi_R \rho\}_{R\in\mathcal{R}}$ for each $\rho \in V$, so we have $\Pi(\pi_{\bar{R}}\lambda) = \{\nu_R\}_{R\in\mathcal{R}}$. Lemma 1 asserts that Π is surjective so by Lemma 2 we conclude that Π has a right inverse B such that

$$B(\{\nu_R\}_{R\in\mathcal{R}}) = \pi_{\bar{R}}\lambda \tag{5}$$

and

$$\|B\| \le C\|\pi_{\bar{R}}\lambda\|_\infty / \|\{\nu_R\}_{R\in\mathcal{R}}\|_\infty,$$

where $C = C(\Pi)$ is a constant depending on Π. However, Π is completely determined by specifying the sets $R \in \mathcal{R}$, which are subsets of $\bar{R}$, and by the condition (F) the cardinality of the latter is less than N, so there are less than 2^N subsets. Therefore there are less than $(2^N)^{2^N}$ possibilities for the projection Π. This means that we may take C to be a constant depending only on N. Since $\pi_{\bar{R}}\lambda$ is a probability measure we have $\|\pi_{\bar{R}}\lambda\|_\infty \le 1$. Moreover the hypothesis (M2) ensures that $|A| \le \alpha^{-1}$. Since for each $R \in \mathcal{R}$ there is a $K \in \mathcal{F}_n$, $K \supset R$, we get by (F) that $\|\nu_R\|_\infty > \alpha^N$, and in particular

$$\|\{\nu_R\}_{R\in\mathcal{R}}\|_\infty > \alpha^N,$$

so we obtain an absolute bound $\|B\| < C\alpha^{-N} = C'$.

Recalling that $\bar{R} \cup \{n\} = \bar{S}$, we view each $y \in A^{\bar{S}}$ as a pair $y = (x, a)$, $x \in A^{\bar{R}}$, $a \in A$, and use a similar convention for the sets A^S, $S \in \mathcal{S}$. With this convention, if σ is a signed measure on $A^{\bar{S}}$ and $a \in A$ then $\sigma(\cdot, a)$ is a (signed) measure on $A^{\bar{R}}$. Note that the family $\{\mu_S\}_{S\in\mathcal{S}}$ can be rewritten in the form $\{\mu_{R\cup\{n\}}\}_{R\in\mathcal{R}}$, and if we fix $a \in A$ then $\mu_{R\cup\{n\}}(\cdot, a)$ is a measure on A^R for each $R \in \mathcal{R}$. Clearly

$$\nu_R = \sum_{a\in A} \mu_{R\cup\{n\}}(\cdot, a). \tag{6}$$

We now define σ by specifying that

$$\sigma(\cdot, a) = B(\{\mu_{R\cup\{n\}}(\cdot, a)\}_{R\in\mathcal{R}}) \quad \text{for all } a \in A. \tag{7}$$

(Note that the consistency of the family $\{\mu_{R\cup\{n\}}(\cdot, a)\}_{R\in\mathcal{R}}$ follows from the consistency

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

of the family $\{\mu_{R\cup\{n\}}\}_{R\in\mathcal{R}}$.) Then by (5)–(7) and the linearity of B

$$\begin{aligned}\pi_{\bar{R}}\sigma &= \sum_{a\in A}\sigma(\cdot,a)\\ &= \sum_{a\in A}B(\{\mu_{R\cup\{n\}}(\cdot,a)\}_{R\in\mathcal{R}})\\ &= B\left(\sum_{a\in A}\{\mu_{R\cup\{n\}}(\cdot,a)\}_{R\in\mathcal{R}}\right)\\ &= B(\{\nu_R\}_{R\in\mathcal{R}}) = \pi_{\bar{R}}\lambda,\end{aligned}$$

establishing (2). To check (3) observe that for each $a\in A$ and $R\in\mathcal{R}$ we have

$$(\pi_{R\cup\{n\}}\sigma)(\cdot,a) = \pi_R(\sigma(\cdot,a)) = \pi_R B(\{\mu_{Q\cup\{n\}}(\cdot,a)\}_{Q\in\mathcal{R}}) = \mu_{R\cup\{n\}}(\cdot,a),$$

since $\Pi B = \mathrm{id}_W$.

It remains to check that σ is non-negative and satisfies (4). Note that each of ν_R and $\mu_{R\cup\{n\}}$, $R\in\mathcal{R}$, is δ-independent. Denote generically by p_J the product measure of one-dimensional marginals on $A^{\{j\}}$, $j\in J$. In particular, for $n\in\Sigma$, $p_{\{n\}}$ is the common one-dimensional marginal of all measures in $\mathcal{F}_n$ on $A^{\{n\}}$. Now we obtain for $R\in\mathcal{R}$ by (M3), (1) and (F)

$$\|\mu_{R\cup\{n\}} - p_{R\cup\{n\}}\|_\infty \le (|R\cup\{n\}|-1)\delta = |R|\delta < N\delta.$$

Similarly $\|\nu_R - p_R\|_\infty < N\delta$. Therefore

$$\begin{aligned}\mu_{R\cup\{n\}}(x,a) &\overset{N\delta}{\sim} p_{\{n\}}(a)p_R(x)\\ &\overset{N\delta}{\sim} p_{\{n\}}(a)\nu_R(x) \quad \text{for all } R\in\mathcal{R},\ x\in A^R,\ a\in A.\end{aligned}$$

This means (since we are using l_∞-norms) that for all $a\in A$

$$\{\mu_{R\cup\{n\}}(\cdot,a)\}_{R\in\mathcal{R}} \overset{2N\delta}{\sim} p_{\{n\}}(a)\{\nu_R\}_{R\in\mathcal{R}},$$

so applying the linear operator B of norm $\le C'$ to both sides we get by (5) and (7) that

$$\sigma(\cdot,a) \overset{2C'N\delta}{\sim} p_{\{n\}}(a)B(\{\nu_R\}_{R\in\mathcal{R}}) = p_{\{n\}}(a)\pi_{\bar{R}}\lambda. \tag{8}$$

Since $\pi_{\bar{R}}\lambda$ is β-independent and $|\bar{R}|<N$, for each $x\in A^{\bar{R}}$ we have by (1) and (M2)

$$\pi_{\bar{R}}\lambda(x) \overset{N\beta}{\sim} p_{\bar{R}}(x) > \alpha^N.$$

Taking $\beta = \alpha^N/2N$ it follows that $\pi_{\bar{R}}\lambda(x) > \frac{1}{2}\alpha^N$ for all $x\in A^{\bar{R}}$. Combining this with (8) and taking $\delta = \beta\alpha^N/4C'N$ we get

$$\left|\frac{\sigma(x,a)}{\pi_{\bar{R}}\lambda(x)} - p_{\{n\}}(a)\right| < \frac{4C'N\delta}{\alpha^N} = \beta \quad \text{for all } x\in A^R,\ a\in A. \tag{9}$$

Since $p_{\{n\}}(a)\ge\alpha$ and $\beta\le\alpha/2$, this shows that σ takes only positive values so it is a probability measure. (9) also shows that $Pr^{\{n\}} \perp_{\beta,\sigma} Pr^{\bar{R}}$, establishing (4) and concluding the proof of Claim A and Proposition 1. □

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

4. *Relatively independent iterates of a partition*

If T is an invertible measure-preserving transformation of X and $P: X \to A$ is a partition of X, we define formally $TP = P \circ T^{-1}$, so the atoms of TP are of the form Tp, where p is an atom of P. If K is any subset of $\mathbb{Z}$, then the partition P^K, indexed by A^K, is defined by

$$[P^K(x)](k) = P(T^k x) \quad \text{for } x \in X, k \in K.$$

Informally $P^K(x) = \{P(T^k x)\}_{k\in K}$, the function P evaluated along the K-orbit of x. $P^K(x)$ is called the (P, K)-*name* of x, or just the K-*name* of x when there is no ambiguity. Note that the atoms of P^K are the atoms of the common refinement of the partitions $\{T^{-k}P \mid k \in K\}$. We will say that P is *independent over* K if the partitions $\{T^{-k}P, k \in K\}$ are jointly independent, equivalently dist P^K is the product of its one-dimensional marginals. Similar observations apply to approximate independence. For example, $P^{\{l\}}$ is δ- *independent of* P^K ($P^{\{l\}} \perp_\delta P^K$) if and only if $T^{-l}P$ is δ-independent of $\bigvee_{k\in K} T^{-k}P$. We will use the same terminology in the relative setting, that is with respect to the fiber measures μ_ω.

THEOREM 1. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic system with a factor $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$, where μ decomposes as $\mu = \int \mu_\omega \, d\mathbb{P}$. Assume $\{n_i\}_{i=0}^\infty$ is an $\mathcal{F}$-mixing sequence for T, and P is a finite partition of X for which there is an $\alpha > 0$ such that*

$$\text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega \quad \text{for all } p \in P \quad \text{we have } \mu_\omega(p) > \alpha.$$

Then for every $\epsilon > 0$ there exists a partition Q of X, a set $E \subset \Omega$, $\mathbb{P}(E) < \epsilon$, and a subsequence $\{m_i\}$ of $\{n_i\}$, such that

$$\text{dist}_{\mu_\omega} Q = \text{dist}_{\mu_\omega} P \quad \text{for } \mathbb{P}\text{-a.e. } \omega, \tag{1}$$

$$d_{\mu_\omega}(P, Q) < \epsilon \quad \text{for } \mathbb{P}\text{-a.e. } \omega \tag{2}$$

and

$$Q \text{ is independent over } \{m_i\} \text{ with respect to } \mu_\omega \text{ for all } \omega \notin E. \tag{3}$$

Proof. There are two cases: θ has a set of periodic points of positive measure, and θ is aperiodic. In the first case, by the ergodicity of θ it must be a rotation on a finite number k of points. By dropping to a sub-sequence we may assume that the n_i are all congruent modulo k, say $n_i = n_i' + r$, where the n_i' are multiples of k. The independence we are aiming for over a subsequence $\{n_{i_j}\}$ (which is now on *every* fiber) is then equivalent to independence over n_{i_j}', so we may as well assume that the n_i are themselves multiples of k, say $n_i = q_i k$. For each ω, T^k preserves the measure μ_ω and $\{q_i\}$ is mixing for (T^k, μ_ω). Fix any ω_0 and apply the non-relative version of Theorem 1 ([**BJ**, Theorem 1], also [**JRW**, Theorem 5]) to (T^k, μ_{ω_0}) to modify P by a small amount on the fiber over ω_0, obtaining a new partition and a subsequence of $\{q_i\}$ which give the desired independence with respect to μ_{ω_0}. Now repeat the process in the fiber over $\theta\omega_0$ with the new partition and subsequence to obtain the desired independence with respect to $\mu_{\theta\omega_0}$. Iterating the process until we get to $\theta^{k-1}\omega_0$ concludes the argument in the periodic case.

In the aperiodic case we assume that $\{n_i\}$ is an increasing sequence starting with $n_0 = 0$. The assumption is not restrictive as the sequence $\{n_i\}$ is $\mathcal{F}$-mixing if and only if $\{-n_i\}$ is

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

$\mathcal{F}$-mixing. Any construction or claim regarding a set of $\omega \in \Omega$ should be understood $\mathbb{P}$-almost surely, whether this is stated explicitly or not. We will proceed inductively starting with $Q_0 = P$, $m_0 = 0$, and building the sequences $\{Q_j\}$, $\{m_j\}$ and $\{E_j\}$ so that for all $j \geq 1$ we would have $m_j \in \{n_i\}$, $\mathbb{P}(E_j) < \epsilon/2^j$, and:

(1′) $\text{dist}_{\mu_\omega} Q_j = \text{dist}_{\mu_\omega} Q_{j-1}$ for $\mathbb{P}$-a.e. ω;

(2′) $d_{\mu_\omega}(Q_j, Q_{j-1}) < \epsilon/2^j$ for $\mathbb{P}$-a.e. ω;

(3′) Q_j is independent over $\{m_i\}_{i=0}^{j}$ with respect to μ_ω for all $\omega \notin \bigcup_{i=1}^{j} E_i$.

Assume for a moment that this construction has been carried out. Then define

$$E = \bigcup_{j=1}^{\infty} E_j$$

(so $\mathbb{P}(E) < \epsilon$). Note that the space of partitions indexed by A is complete in the d_{μ_ω} metric for $\mathbb{P}$-a.e. ω. So by (2′) there is a limit μ-measurable partition Q of X such that $Q = \lim Q_j$ with respect to d_{μ_ω} for $\mathbb{P}$-a.e. ω. In this way we get (1), (2) and (3) satisfied.

The proof of the theorem is thereby reduced to the induction step. For convenience we now get rid of unnecessary subscripts and coefficients and restate in equivalent form what remains to be proved. Use the assumptions of the theorem and let K be a subset of $\mathbb{N} \cup \{0\}$ containing 0, so that P is independent over K with respect to μ_ω for all ω outside some $E \in \mathcal{F}$.

Given $\epsilon > 0$, we are looking for $m \in \{n_i\}$, $m > \max K$, a partition Q of X and $\tilde{E} \subset \Omega$, $\mathbb{P}(\tilde{E}) < \epsilon$, such that:

(1″) $\text{dist}_{\mu_\omega} Q = \text{dist}_{\mu_\omega} P$ for $\mathbb{P}$-a.e. ω;

(2″) $d_{\mu_\omega}(P, Q) < \epsilon$ for $\mathbb{P}$-a.e. ω;

(3″) Q is independent over $K \cup \{m\}$ with respect to μ_ω for all $\omega \notin E \cup \tilde{E}$.

Denote $|K| = k$ and let $\eta = \epsilon/10 \cdot \min(\delta, \alpha^{k+1})$ where $\delta = \delta(\alpha, k^2 + 1)$ from Proposition 1 (in §3)—this choice of η will become clear later on. By the hypothesis, for $\omega \in \Omega \backslash E$, P is independent over K with respect to μ_ω. Using the mixing property of the sequence $\{n_i\}$ pick $m \in \{n_i\}$ such that $m > \max K$ and

$$P^{\{m\}} \perp_{\eta, \mu_\omega} P^K \quad \text{for every } \omega \notin \tilde{E}_1, \text{ where } \mathbb{P}(\tilde{E}_1) < \frac{\epsilon}{10}. \tag{4}$$

Let $B_\Omega \subset \Omega$ be the base of a Rokhlin tower R for θ of height M where $M > 10m/\epsilon$, with the error set

$$\tilde{E}_2 = \Omega \backslash \bigcup \{\theta^i B_\Omega \mid 0 \leq i < M\} \text{ of measure } \mathbb{P}(\tilde{E}_2) < \frac{\epsilon}{10}. \tag{5}$$

The tower R can be lifted to X by π^{-1}, where π is the factorization map. This tower in X can informally be regarded as an 'integral of towers' with bases $B_\omega = \pi^{-1}(\omega)$, $\omega \in B_\Omega$, and this observation inspires the construction that follows. Let $\bar{M} = [0, M-1]$. With $\omega \in B_\Omega$ fixed, its $\bar{M}$-orbit $\{\theta^j \omega\}_{j=0}^{M-1}$ gives rise to the T-tower R_ω (in an appropriate space) whose levels are the fibers $\pi^{-1}(\theta^j \omega)$ endowed with the measures $\mu_{\theta^j \omega}$.

Let $\nu_\omega = \text{dist}_{\mu_\omega} P^{\bar{M}}$, a probability measure on $A^{\bar{M}}$. Denote $K' = K \cup \{m\}$. For all j that satisfy $j + K' \subset \bar{M}$ (i.e., $0 \leq j < M - m$) let $\nu_{\omega,j}$ denote the marginal of ν_ω on $A^{j+K'}$. Then $\nu_{\omega,j} = \pi_{j+K'} \nu_\omega = \text{dist}_{\mu_\omega} P^{j+K'}$. For those values of j for which $\theta^j \omega \notin E \cup \tilde{E}_1$, the measure $\nu_{\omega,j}$ on $A^{j+K'}$ is η-independent, by our choice of m and because its marginal on

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

A^{j+K} is exactly independent. We now aim at perturbing P to obtain Q that satisfies $(1'')$ and $(2'')$ and for which $\mathrm{dist}_{\mu_\omega} Q^{j+K'}$ is exactly independent for all combinations of $\omega \in B_\Omega$ and $0 \le j < M - m$ that yield $\theta^j \omega$ outside of a subset of $\Omega \backslash \tilde{E}_2$ only slightly larger than $E \cup \tilde{E}_1$.

Fix an $\omega \in B_\Omega$ and consider the tower R_ω. Denote its base $\pi^{-1}(\omega)$ by B. Partition B into two parts $B = B_0 \cup B_1$, $\mu_\omega(B_0) = \epsilon/10$, $\mu_\omega(B_1) = 1 - (\epsilon/10)$, independently of the $\bar{M}$-names of the points in B (i.e., after proper normalization $\mathrm{dist}_{\mu_\omega} P^{\bar{M}}$ on each of B_0 and B_1 is identical to that on B). This can be done since the fibers of π are non-atomic by the mixing property of T.

The partition P assigns points in X to the letters of the alphabet A. We will construct the required partition Q by re-assigning points on the stack of R_ω with the base B_0 to the letters of A (the technique of 'painting names on towers'), while retaining the original assignment on the stack of R_ω with the base B_1. We now explain how the re-assignment is done. (Recall that ω is fixed for the time being.)

Let $m_{\omega,j}$ denote the marginal of ν_ω on $A^{\{j\}}$, a probability measure on $A^{\{j\}}$ which is the same as $\mathrm{dist}_{\mu_{\theta^j \omega}} P$ (up to identification of $A^{\{j\}}$ with A). Of course, the one-dimensional marginal of $\nu_{\omega,j}$ on $A^{\{j+k\}}$ for any particular $k \in K'$ is $m_{\omega,j+k}$. For all j satisfying

$$0 \le j < M - m \quad \text{and} \quad \theta^j \omega \notin E \cup \tilde{E}_1 \tag{6}$$

we define a 'correcting measure' $\xi_{\omega,j}$ on $A^{j+K'}$ by the following equation:

$$\left(1 - \frac{\epsilon}{10}\right) \cdot \nu_{\omega,j} + \frac{\epsilon}{10} \cdot \xi_{\omega,j} = \prod_{k \in K'} m_{\omega,j+k}. \tag{7}$$

The idea is that blending a small fraction of $\xi_{\omega,j}$ into $\nu_{\omega,j}$ corrects it from approximate to exact independence. *A priori* $\xi_{\omega,j}$ is a signed measure but we shall see shortly that it is in fact positive. Observe that for any $k \in K'$ the marginal of $\xi_{\omega,j}$ on $A^{\{j+k\}}$ is $m_{\omega,j+k}$, since this is the case for both $\nu_{\omega,j}$ and $\prod_{k \in K'} m_{\omega,j+k}$. Moreover, applying the relation (7) to the marginals on A^{j+K} we conclude that $\pi_{A^{j+K}} \xi_{\omega,j}$ is the product of its one-dimensional marginals $m_{\omega,j+k}$, since this is true for $\pi_{A^{j+K}} \nu_{\omega,j}$. In addition, we claim that for all j satisfying (6) the measure $\xi_{\omega,j}$ is positive and δ-independent.

To show this, rewrite the definition of $\xi_{\omega,j}$ as

$$\xi_{\omega,j} - \prod_{k \in K'} m_{\omega,j+k} = \left(\frac{10}{\epsilon} - 1\right)\left(\prod_{k \in K'} m_{\omega,j+k} - \nu_{\omega,j}\right). \tag{7'}$$

To prove positivity we will show that the l_∞-norm of the right-hand side of $(7')$ is bounded by α^{k+1}. Since $\prod_{k \in K'} m_{\omega,j+k}$ on individual elements of $A^{j+K'}$ is greater than α^{k+1}, the positivity of $\xi_{\omega,j}$ will follow. Further on, in view of the joint independence of the first k (out of the total of $k+1$) one-dimensional marginals of this measure, for establishing δ-independence it would be enough to prove that for all $x \in A^{j+K}$ and $a \in A^{\{j+m\}}$ we have

$$\left| \frac{\xi_{\omega,j}(x, a)}{\xi_{\omega,j}(x)} - \xi_{\omega,j}(a) \right| < \delta. \tag{8}$$

(As in §3 we view $y \in A^{j+K'}$ as a pair $y = (x, a)$, $x \in A^{j+K}$, $a \in A^{\{j+m\}}$. Recall also that by our convention from that section the expression $\xi_{\omega,j}(z)$ is meaningful for any $z \in A^L$

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

where $L \subset j + K'$.) This is achieved by our choice of η, as the following calculation shows. Evaluate (7′) at (x, a) and then divide by

$$\xi_{\omega,j}(x) = \nu_{\omega,j}(x) = \left(\prod_{k\in K} m_{\omega,j+k}\right)(x) \tag{9}$$

to obtain

$$\left|\frac{\xi_{\omega,j}(x,a)}{\xi_{\omega,j}(x)} - \xi_{\omega,j}(a)\right| = \left(\frac{10}{\epsilon} - 1\right)\left|\nu_{\omega,j}(a) - \frac{\nu_{\omega,j}(x,a)}{\nu_{\omega,j}(x)}\right| < \frac{10}{\epsilon}\cdot\eta. \tag{10}$$

(The inequality is due to (4).) Now recall that $\eta = \epsilon/10 \cdot \min(\delta, \alpha^{k+1})$. So on the one hand (10) implies (8), and on the other hand multiplying (10) back by the expression (9) we get the required bound for the right-hand side of (7′):

$$\left|\xi_{\omega,j}(x,a) - \left(\prod_{k\in K'} m_{\omega,j+k}\right)(x,a)\right| < \alpha^{k+1}\cdot\left(\prod_{k\in K} m_{\omega,j+k}\right)(x) \le \alpha^{k+1}.$$

Thus the positivity of $\xi_{\omega,j}$ is established and the proof of the δ-independence of $\xi_{\omega,j}$ is complete.

Proposition 1 of §3 can now be applied to the family $\mathcal{F}$ of those shifts $j + K'$ of K' for which j satisfies (6), endowed by the measures $\xi_{\omega,j}$. (F) holds for $N = (k+1)^2$. (M1), the consistency condition for the measures $\xi_{\omega,j}$, is implied by (7) and the consistency of the measures $\nu_{\omega,j}$. (M2) holds because the one-dimensional marginals of $\xi_{\omega,j}$ are the same as those of $\nu_{\omega,j}$. (M3) has just been demonstrated. So there is a measure ξ_ω on $A^{\bar{M}}$ that is a common extension of all the measures in the family. Write the names on the levels of the stack of R_ω with the base B_0 as prescribed by ξ_ω. Explicitly what this means is we choose a partition R of B_0 indexed by $A^{\bar{M}}$ so that $\text{dist}\,(R|B_0) = \xi_\omega$ and then we partition the stack according to the rule: for each $x \in B_0$ and $j \in \bar{M}$ let $Q(T^j x) = (R(x))(j)$. To complete the construction of the partition Q on the whole M-tower above B recall that this partition coincides with P in the stack of R_ω with the base B_1.

The procedure we described depends on the point $\omega \in B_\Omega$. Performing it for every $\omega \in B_\Omega$ we get a partition Q of the whole $\pi^{-1}(\Omega\backslash\tilde{E}_2)$. On $\pi^{-1}(\tilde{E}_2)$ let Q coincide with P as well. One only needs to assure that Q can be chosen μ-measurable. This follows from the fact that for any measurable $A \subset X$ and $0 < t < 1$ there is a set $B \subset X$ such that $\mu_\omega B = t$ for a.a. ω and A and B are independent with respect to μ_ω for a.a. ω. Using this and following every step of the procedure, it is easy to convince oneself that everything can be done in a μ-measurable way.

Having done all this, we obtain a partition Q of X that clearly satisfies (1″) and (2″). (1″) holds because the one-dimensional marginals of the measures $\nu_{\omega,j}$ and $\xi_{\omega,j}$ are the same. (2″) holds since $\mu_\omega(B_0) \equiv \epsilon/10$. By virtue of the construction, (3″) is satisfied for all ω outside of $E \cup \tilde{E}$, where $\tilde{E} = \tilde{E}_1 \cup \tilde{E}_2 \cup \tilde{E}_3$. $\tilde{E}_1$ and $\tilde{E}_2$ were defined in (4) and (5), while $\tilde{E}_3$ is the union of m uppermost levels in the $\bar{M}$-tower for θ with the base B_Ω. So $\mathbb{P}(\tilde{E}_3) \le m/M < \epsilon/10$ by the choice of M. Consequently, $\mathbb{P}(\tilde{E}) < \epsilon$ as required. □

Remark. The proof shows that, as in [**JRW**] and [**BJ**], the independence can be achieved not just along the sequence $\{m_i\}$, but also along the *IP*-set which it generates, namely the set of all finite sums of the form $\sum_{i\in F} m_i$, where F is a finite subset of $\mathbb{N}$.

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

THEOREM 2. *With the same hypotheses as in Theorem 1 there is a Q and a subsequence $\{m_i\}$ such that $\mathrm{dist}_{\mu_\omega} Q = \mathrm{dist}_{\mu_\omega} P$ for a.a. ω, $d(P, Q) < \epsilon$ and Q is independent over $\{m_i\}$ with respect to μ_ω for a.a. ω.*

The proof is a modification of the proof of Theorem 1, using the following lemma.

LEMMA 3. *Suppose K is a finite subset of $\mathbb{N} \cup \{0\}$ and P is a partition which is independent over K with respect to μ_ω for all $\omega \notin E$ where E is some subset of Ω. Then for any $\epsilon > 0$ there is a partition Q such that P and Q agree on all fibers $\pi^{-1}(\omega)$ for ω not in $KE = \bigcup\{\theta^k E \mid k \in K\}$, $\mathrm{dist}_{\mu_\omega} P = \mathrm{dist}_{\mu_\omega} Q$ for a.a. ω and there is a subset $F \subset \Omega$ with $\mathbb{P}(F) < \epsilon$ such that Q is independent over K with respect to μ_ω for all $\omega \notin F$.*

Proof. Let $B \subset \Omega$ be the base of a Rokhlin tower of height $M \gg m = \max K$ with error set of measure smaller than ϵ. Fix $\omega_0 \in B$. As in the proof of Theorem 1 we will describe how to change P on the tower over $\pi^{-1}(\omega_0)$, with the understanding that this should be done simultaneously for each $\omega_0 \in B$, in a measurable varying way.

Fix any i_1 such that $0 \le i_1 < M - m$ and $\theta^{i_1}\omega_0 \in E$. Now change P on each $\pi^{-1}(\theta^{i_1+k}\omega_0)$, $k \in K$, retaining the same distribution on each of these fibers, to obtain a new partition P_1 so that, with respect to μ_{ω_0}, P_1 is independent over $i_1 + K$ and in addition $P_1^{i_1+K}$ is independent of $P_1^{[0,n-1]\setminus(i_1+K)}$. This is possible simply because the fibers are non-atomic.

Of course, we will then have that P_1 is independent over K with respect to $\mu_{\theta^{i_1}\omega_0}$. In addition, for any $i \neq i_1$ such that $0 \le i < M - m$ and $\theta^i\omega_0 \notin E$, we will still have that P_1 is independent over K with respect to $\mu_{\theta^i\omega_0}$. Indeed, with respect to μ_{ω_0}, P_1 is independent over $S_1 = (i + K) \cap (i_1 + K)^c$, since P has this property while P_1 agrees with P over the relevant fibers. Also with respect to μ_{ω_0}, P_1 is independent over $S_2 = (i + K) \cap (i_1 + K)$, since P_1 is actually independent over all of $i_1 + K$ by construction. Finally $P_1^{S_1}$ and $P_1^{S_2}$ are independent with respect to μ_{ω_0}, again by construction of P_1. This shows that P_1 is independent over $i + K$ with respect to μ_{ω_0} which implies that P_1 is independent over K with respect to $\mu_{\theta^i\omega_0}$.

Now replace i_1 with any i_2, $0 \le i_2 < M - m$ such that we still do not have the desired independence with respect to $\mu_{\theta^{i_2}\omega_0}$ and modify P_1 to P_2 as before to achieve the independence with respect to $\mu_{\theta^{i_2}\omega_0}$. Continuing in this way we will arrive at a partition Q which has the desired independence with respect to $\mu_{\theta^i\omega_0}$ for every i, $0 \le i < M - m$. Note also that Q differs from P only on fibers over KE. Doing this for all $\omega_0 \in B$ we obtain the desired independence with respect to μ_ω for all $\omega \in \theta^i B$, $0 \le i < M - m$ and the measure of this set of ω can be made as close to 1 as we please. □

COROLLARY 1. *In the assumptions of Lemma 3, for any $\epsilon > |K| \cdot \mathbb{P}(E)$ there is a partition Q and a set $G \subset \Omega$ with $\mathbb{P}(G) < \epsilon$, such that P and Q agree on all fibers outside G, $\mathrm{dist}_{\mu_\omega} Q = \mathrm{dist}_{\mu_\omega} P$ for a.a. ω, and Q is independent over K with respect to μ_ω for a.a. ω.*

Proof. Iterating Lemma 3 with a sequence of sets E_i in Ω, $E_1 = E$, of total measure at most $\sum_i \mathbb{P}(E_i) < \epsilon/|K|$, produces a sequence of partitions which, by the Borel–Cantelli

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

lemma, eventually stabilizes over ω for a.a. ω. This gives a limiting partition Q with the desired properties. □

Now we can prove Theorem 2: in the proof of Theorem 1 at each stage, after achieving the independence of Q_k over $\{m_1, \ldots, m_k\}$ with respect to μ_ω for a large set of ω, use Corollary 1 to change Q_k over a small set of ω to achieve the desired independence over a.a. ω. This concludes the proof of Theorem 2. □

We now sketch how to remove the ergodicity assumption in Theorems 1 and 2. Let us first observe that it is enough to prove the result in the case when θ is periodic with period $n \in \{1, 2, \ldots, \infty\}$. Indeed, supposing that we have done this, partition Ω into the invariant sets Ω_n, $n = 1, 2, \ldots, \infty$ where θ has period exactly n. We start in Ω_∞ and find a Q_∞ uniformly close to P over Ω_∞ and equal to P over $\Omega \backslash \Omega_\infty$, and a first subsequence $\{m_i\}$ which gives the desired independence in the fibers over most of Ω_∞. Then we work over Ω_1 and find Q_1 close to Q_∞ and a further subsequence which gives the desired independence on most of Ω_1. We continue in this way, perturbing the partition and refining the subsequence, for finitely many steps until we have exhausted all but a small invariant set $E \subset \Omega$ consisting of some tail of the sequence $\{\Omega_n\}$. For Theorem 1 we only need independence on all but a small subset of Ω so we are done. In the case of Theorem 2, if we are willing to relax the requirement that the perturbed partition have the same distribution as the original one on all fibers, we can get the desired independence on fibers over E by simply making the partition trivial on those fibers.

So, we may now assume that θ has constant period k, possibly infinite. In both Theorems 1 and 2 the case $k = \infty$ is identical to the ergodic case since aperiodicity is all that is needed for the Rokhlin lemma. In the case when k is finite, Theorem 1, the same congruence trick used in the ergodic case reduces us to working with T^k so we may assume that $\theta = \mathrm{id}_\Omega$. So we have reduced ourselves to proving the non-ergodic versions of Theorems 1 and 2 in the case when θ is the identity. In the case of Theorem 1 this is just a slightly souped up version of the absolute theorem where instead of having just one weakly mixing system one has a whole measurable field of them. It is quite straightforward to prove this by suitably modifying the proof of the absolute result. To actually carry it out would require us to delve into the proof of the absolute result. This is the reason that we chose to prove Theorems 1 and 2 formally only in the ergodic case. In the case of Theorem 2 we again just make the partition trivial in the small set of fibers where Theorem 1 does not give independence.

The observant reader will have noticed that Theorems 1 and 2 do not quite generalize Krengel's theorem in that we don't insist that the subsequence $\{m_i\}$ start with $m_0 = 0$. The place where our argument would break down if we wanted $m_0 = 0$ is at the congruence trick used to deal with the periodic case. The results may well be true with $m_0 = 0$ and in any case our argument shows that they do hold at least when θ is aperiodic.

Acknowledgements. We would like to thank B. Weiss and Y. Kifer for helpful discussions that led to improvement of the exposition. The first author's work has been supported by the Edmund Landau Minerva Center for Research in Mathematical Analysis and Related Areas, a Niedersachsen grant and ISF grant 1333/04.

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press

REFERENCES

[BJ] B. Begun and A. del Junco. Amenable groups, stationary measures and partitions with independent iterates. *Israel J. Math.* **158** (2007), 41–63.

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

[G] E. Glasner. *Ergodic Theory via Joinings (Mathematical Surveys and Monographs, 101)*. American Mathematical Society, Providence, RI, 2003.

[JRW] A. del Junco, K. Reinhold and B. Weiss. Partitions with independent iterates along IP-sets. *Ergod. Th. & Dynam. Sys.* **19** (1999), 447–473.

[KN] B. O. Koopman and J. V. Neumann. Dynamical systems of continuous spectra. *Proc. Natl. Acad. Sci. USA* **18** (1932), 255–263.

[K] U. Krengel. Weakly wandering vectors and weakly independent partitions. *Trans. Amer. Math. Soc.* **164** (1972), 199–226.

[R] D. J. Rudolph. Pointwise and L^1 mixing relative to a sub-sigma algebra. *Illinois J. Math.* **48** (2004), 505–517.

https://doi.org/10.1017/S0143385709000212 Published online by Cambridge University Press