Ergod. Th. & Dynam. Sys. (1999), **19**, 447–473
Printed in the United Kingdom © *1999 Cambridge University Press*

Partitions with independent iterates along IP-Sets

ANDRES DEL JUNCO†, KARIN REINHOLD‡ and BENJAMIN WEISS§

† *Department of Mathematics, University of Toronto, Toronto, Canada M5S 1A1*
(e-mail: deljunco@math.toronto.edu)
‡ *Department of Mathematics, SUNY Albany, Albany NY 12222, USA*
(e-mail: reinhold@math.albany.edu)
§ *Department of Mathematics, Hebrew University, Jerusalem, Israel*
(e-mail: weiss@humus.huji.ac.il)

(*Received* 30 *January* 1997 *and accepted in revised form* 20 *August* 1997)

Abstract. We prove a collection of results inspired by Krengel's theorem on the existence of partitions with infinitely many independent iterates in any weakly mixing measure-preserving dynamical system. Our approach avoids Krengel's use of two-fold mixing thereby obtaining stronger results, including characterizations of mild and strong mixing, as well as weak mixing. We also obtain results for non-weakly mixing systems and for more general group actions.

0. *Introduction*

The results in this paper were inspired by Krengel's work on weakly independent partitions in [**6**]. These notes present a new approach to Krengel's result as well as extensions of it.

Given a probability space (X, β, m) and $T : X \mapsto X$ an invertible measure-preserving transformation, we are concerned with the problem of finding a partition P and a sequence $\{n_k\}$ such that the family of partitions $\{T^{n_k} P\}$ is mutually independent, that is, $T^{-n_{k+1}} P \perp P \vee T^{-n_1} P \vee \cdots \vee T^{-n_k} P$ for all $k \geq 1$. In this case, the partition P is called weakly independent.

Krengel showed that T is weakly mixing if and only if the family of weakly independent partitions is dense in the family of all finite partitions. More precisely, given any non-trivial finite probability vector π and letting Π denote the family of all partitions with distribution π, his proof shows that T is weakly mixing precisely if the family of weakly independent partitions in Π is dense in Π.

THEOREM 1. (Krengel)

(a) *If T is a weakly mixing transformation, then the weakly independent partitions are dense in the set of finite partitions.*

(b) *If for every partition $P = \{A, A^c\}$ and any $\epsilon > 0$ there exists $P' = \{A', A'^c\}$ with $m(A \Delta A') < \epsilon$, and three integers n_1, n_2, n_3 such that $T^{n_1} P'$, $T^{n_2} P'$, $T^{n_3} P'$ are independent, then T is weakly mixing.*

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Krengel's proof of this result was very combinatorial in nature, showing exactly how to cut the atoms of the partition to achieve the independence result. In these notes, we present two different proofs of this theorem which, in fact, yield stronger results: weak independence along IP-sets and extensions to group actions.

An IP-set Γ is the set of finite sums with no repetitions generated by a sequence $\{n_k\}_{k=1}^{\infty}$, that is, Γ consists of the elements of the form $n_{i_1} + \cdots + n_{i_k}$, $k \geq 1$, $i_1 < i_2 < \cdots < i_k$.

THEOREM 2. *Let T be a weakly mixing transformation on a probability space (X, β, m). Then, the weakly independent partitions along IP-sets are dense in the set of finite partitions.*

Before discussing Theorem 2, we would like to mention that the problem of finding weakly independent partitions has a counterpart in the Hilbert space setting. Let H designate a Hilbert space, U an isometry on it and $f \in H$ a vector. We say that f is a *weakly wandering vector* if there exists an infinite sequence $\{n_k\}_{k=1}^{\infty}$ of positive integers such that $\{U^{n_k} f\}_{k=1}^{\infty}$ are pairwise orthogonal vectors.

THEOREM 3. (Krengel) *An isometry U has no discrete spectrum if and only if the weakly wandering vectors are dense in H.*

This theorem has recently been generalized to weakly wandering vectors along IP-sequences. Bergelson, Cornfield and Mityagin used an ingenious fixed point argument to prove the following.

THEOREM 4. ([**1**]) *If U is an isometry with no discrete spectrum then the weakly wandering vectors along IP-sequences are dense in H.*

The first proof of Theorem 2 (Theorem 5 below) is fundamentally different from Krengel's combinatorial argument. It depends on the existence, in certain situations, of a shift invariant extension to $\Omega^{\mathbb{Z}}$ of a shift-invariant measure on Ω^I, where I is a finite subset of $\mathbb{Z}$ (Proposition 1.4 followed by Markov extension). Moreover, it does not use the two-fold mixing property which is essential to Krengel's construction.

The second proof of Theorem 2 turns out to be similar in spirit to Krengel's argument in that it makes use of the two-fold mixing property of the action. However, our argument extends to free actions of discrete abelian groups (Theorem 10 below) whose action is two-fold weakly mixing.

In §1 we establish the necessary results on extensions of measures and in §2 we present the first version of Theorem 2 under a weaker hypothesis, as well as corollaries of it.

Given a measure-preserving transformation T on (X, β, m), a sequence $\{m_k\}_{k=1}^{\infty}$ is mixing for T if for any pair of sets $A, B \in \beta$, $\lim_{k\to\infty} m(A \cap T^{-m_k} B) = m(A)m(B)$.

THEOREM 5. *Let $S = \{m_k\}_{k=1}^{\infty}$ be a mixing sequence for T. Then, the weakly independent partitions along IP-sets with generators in S are dense in the set of all finite partitions.*

Clearly, Theorem 2 follows from Theorem 5. Moreover, this version has applications to weakly mixing, mildly mixing and mixing transformations.

COROLLARY 6. *If T is weakly mixing, given any set Σ with positive density, the weakly independent partitions along IP-sets with generators in Σ are dense in the set of all finite partitions.*

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

COROLLARY 7. *If T is mildly mixing, given any IP-set Σ, the weakly independent partitions along IP-sets with generators in Σ are dense in the set of all finite partitions.*

COROLLARY 8. *If T is mixing, given any infinite set Σ, the weakly independent partitions along IP-sets with generators in Σ are dense in the set of all finite partitions.*

Krengel's theorem provides a characterization of weakly mixing transformations. The corollaries above provide the corresponding characterization for mild mixing and strong mixing respectively. We also include in the following proposition a characterization of the K-property. Although this characterization is in the same spirit, its proof (which is a simple application of the fundamental lemma of the isomorphism theory) has nothing to do with the ideas in this paper.

PROPOSITION 9.

(a) *T is a mild mixing transformation if and only if given any IP-set Σ, the weakly independent partitions along subsequences of Σ are dense in the set of all finite partitions.*

(b) *T is a mixing transformation if and only if given any sequence $\{n_k\}_{k=1}^{\infty}$, the weakly independent partitions along subsequences of $\{n_k\}_{k=1}^{\infty}$ are dense in the set of all finite partitions.*

(c) *T is a K-automorphism if and only if the partitions which have independent iterates along some subgroup of $\mathbb{Z}$ are dense in the set of all finite partitions.*

Section 3 concerns the extension of Theorem 2 to free actions of discrete abelian groups. In the spirit of Krengel's original proof, we must require the mixing sequence to be two-fold mixing as well. There are two reasons for this. In the first place there is no obvious analogue of Markov extension of stationary measures when the acting group is not $\mathbb{Z}$. Secondly, we have not been able to get an analogue of Proposition 1.4 for more general groups. Indeed, both of these questions are, as far as we know open even for $\mathbb{Z}^d$ and seem to be interesting and natural in their own right.

A sequence of maps $\{T_i\}_{i=1}^{\infty}$ on (X, β, m) is called *mixing* if for any pair $A, B \in \beta$, $\lim_{n\to\infty} m(A \cap T_n(B)) = m(A)m(B)$. The sequence is called *two-fold mixing* if for any triple $A, B, C \in \beta$, $\lim_{n\to\infty} m(T_n^{-1}(A) \cap B \cap T_n(C)) = m(A)m(B)m(C)$.

We remark that any two-fold weakly mixing G-action contains a two-fold mixing sequence (see §3.1). Moreover, if $G/2G$ is finite then any weakly mixing action is two-fold weakly mixing. Thus our results apply, for example, to $\mathbb{Z}^d$ actions.

Let G be a discrete abelian group and $\{g_i\}_{\in I}$ a collection of elements of G. The *IP-set* they generate is the set of finite sums $\Gamma = \{g_{i_1} + g_{i_2} + \cdots + g_{i_k} : i_1 < i_2 < \cdots < i_k, k \geq 1\}$, and we denote it by $\Gamma = \mathrm{FS}(\{g_k\}_{k=1}^{\infty})$. (FS = 'finite sums'.)

THEOREM 10. *Let G be a discrete abelian group acting freely on X. If $S \subset G$ is a two-fold mixing sequence, then the set of weakly independent partitions along IP-sets with generators in S is dense in the set of all finite partitions.*

If G is, for example, the countable direct sum of copies of $\mathbb{Z}_2$ then every element has order two, so there are no two-fold mixing sequences. Nonetheless for any G in which all elements have finite order the proof of Theorem 10 can be modified to give the same conclusion assuming only that S is mixing. We leave the details to the interested reader.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

It is a very interesting question whether the hypothesis of two-fold weak mixing is needed at all in Theorem 10. An interesting group to look at is the countable direct sum of copies of $\mathbb{Z}$, in which case there are weakly mixing but not two-fold weakly mixing actions and we do not know whether there is a Krengel-type theorem. More precisely, the question is: given a weakly mixing action of this group must there be lots of partitions which move independently along some infinite sequence?

An extension of Krengel's result to group actions was already present in the literature. Bergelson and Rudolph [**3**] extended the result for actions of groups of the form $G = \oplus_{i=1}^{\infty} F_i \equiv F^{\infty}$, where $F_i = F$ is a fixed finite field.

THEOREM 11. ([**3**]) *Let F be a finite field. Let $\{T_g\}_{g\in F^{\infty}}$ be an action of F^{∞} on a Lebesgue space (X, β, m). If the action is weakly mixing, then there is an infinite subgroup H and an increasing sequence of finite partitions $\{P_i\}_{i=1}^{\infty}$ such that for every i, $\{T_g P_i\}_{g\in H}$ is a mutually independent family, and $\vee_{i=1}^{\infty} \vee_{g\in H} T_g P_i = \beta$.*

This result is stronger than Theorem 10 but its strength relies on the particular structure of the group.

For groups without finite torsion it is hopeless to prove independence along an infinite subgroup. However, one can improve Theorem 10 to get independence along an IP-set with repetitions, that is each element may be repeated say k times, provided we assume that the action has a sufficiently high order of weak mixing, which is automatic when the group is $\mathbb{Z}^d$.

Returning to single transformations T, if T is not weak mixing, it is natural to conjecture that if P is independent of the Kronecker factor than it can be perturbed to move independently along some infinite sequence. In §4 we show that this is the case at least when the transformation is not a finite extension of its Kronecker factor, leaving open the case when the Kronecker factor has finite fibers.

THEOREM 12. *Let T be an ergodic measure preserving transformation. If the fibers of its Kronecker factor are non-atomic then, given $\epsilon > 0$ and P a partition independent of the Kronecker factor, there exists a partition P', also independent of the Kronecker factor, and a sequence $\{n_k\}_{k=1}^{\infty}$ such that $\|P - P'\| < \epsilon$ and $\{T^{n_k} P'\}_{k=1}^{\infty}$ is a mutually independent family.*

1. *Extension of measures*

The proofs of Theorems 5 and 10 are based on the ability to find a measure with given marginals. Although this problem has no solution in general, we are only concerned with the case in which the marginal measures are sufficiently close to product measure.

Let $\Omega = \{1, \ldots, r\}$ and p a probability measure on it. Let A and B be finite sets and μ and ν measures on Ω^A and Ω^B respectively. Then μ and ν are *consistent* measures if their projections onto $\Omega^{A\cap B}$ coincide.

Let $\{A_i\}_{i=1}^n$ be a collection of finite sets and $A = \cup_{i=1}^n A_i$. Suppose $\{\nu_i\}_{i=1}^n$ is a collection of (pairwise) consistent probability measures such that each ν_i is defined on Ω^{A_i} with one-dimensional marginals given by p. Let $\nu_{i_1,i_2,\ldots,i_k}$ denote the common projection

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

of the measures v_{i_j}, $1 \le j \le k$, onto $\Omega^{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}}$. For any finite set B, let

$$M_B = \{\text{signed measures on } \Omega^B \text{ with one-dimensional marginals given by } p\}.$$

Definition 1.1. Let B be a finite set, $v, w \in M_B$ and π the product measure in M_B. We say that $|v - w| < \epsilon\pi$ if for any $y \in \Omega^B$, $|v(y) - w(y)| < \epsilon\pi(y)$. Also, we say that $v \ge w$ if $v(y) \ge w(y)$ for all $y \in \Omega^B$.

Let π_i denote the product measure on Ω^{A_i}, and π the product measure on Ω^A, both with one-dimensional marginals given by p.

PROPOSITION 1.2. *With the above notation, there exists $\eta > 0$ depending only on n, such that, if $|v_i - \pi_i| < \eta\pi_i$ for all $i = 1, \ldots, n$, then there exists a probability measure λ on Ω^A with marginals v_i, i.e.* $\mathrm{proj}_{\Omega^{A_i}}(\lambda) = v_i$, $1 \le i \le n$. *Moreover,* $|\lambda - \pi| < 2^n \eta\pi$.

Proof. The proof is essentially an exercise in linear algebra: the linear algebra consists of showing that if we allow signed measures then the problem has a solution. By continuity, if the marginals are close to independence then so is the extension, and hence the extension will be positive. We proceed as follows.

Let $\pi_{i_1,\ldots,i_k}$ denote the product measure on $\Omega^{A_{i_1} \cap \cdots \cap A_{i_k}}$ and $\pi^c_{i_1,\ldots,i_k}$ the product measure on $\Omega^{A \setminus (A_{i_1} \cap \cdots \cap A_{i_k})}$, both measures with one-dimensional marginals given by p. Define the function

$$F : \{(m_1, \ldots, m_n) \in M_{A_1} \times \cdots \times M_{A_n} : \{m_i\}_{i=1}^n \text{ is a consistent set of measures}\} \to M_A$$

by

$$F(m_1, \ldots, m_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} m_{i_1,\ldots,i_k} \pi^c_{i_1,\ldots,i_k}, \tag{1.1}$$

where $m_{i,\ldots,i_k}$ denotes the projection of m_{i_j} onto $\Omega^{A_{i_1} \cap \cdots \cap A_{i_k}}$, $1 \le j \le k$.

Since $F(v_1, \ldots, v_n)(\Omega^A) = 1$, in order to show that $F(v_1, \ldots, v_n)$ is a probability measure it suffices to show it is positive.

Note that $F(\pi_1, \ldots, \pi_n) = \pi$, thus

$$\begin{aligned} |F(v_1, \ldots, v_n) - \pi| &\le \sum_{k=1}^{n} \sum_{1 \le i_1 < \cdots < i_k \le n} |v_{i_1,\ldots,i_k} - \pi_{i_1,\ldots,i_k}| \pi^c_{i_1,\ldots,i_k} \\ &\le \sum_{k=1}^{n} \sum_{1 \le i_1 < \cdots < i_k \le n} \eta\pi_{i_1,\ldots,i_k} \pi^c_{i_1,\ldots,i_k} \\ &\le 2^n \eta\pi. \end{aligned}$$

Hence, for η small enough, $F(v_1, \ldots, v_n) > 0$.

Let $\lambda = F(v_1, \ldots, v_n)$ and $\lambda_{A_i} = \mathrm{proj}_{\Omega^{A_i}}(\lambda)$. Fixing i, rewrite

$$\begin{aligned} \lambda = v_i \pi_i^c + \sum_{\substack{1 \le j \le n \\ j \ne i}} v_j \pi_j^c + \sum_{k=1}^{n-1} (-1)^k \sum_{\substack{j_1 < \cdots < j_k \\ j_s \ne i}} v_{i,j_1,\ldots,j_k} \pi^c_{i,j_1,\ldots,j_k} \\ + \sum_{k=2}^{n-1} (-1)^{k+1} \sum_{\substack{j_1 < \cdots < j_k \\ j_s \ne i}} v_{j_1,\ldots,j_k} \pi^c_{j_1,\ldots,j_k}. \end{aligned}$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Then, denoting $\pi^{A_i}_{j_1,\dots,j_k}$ the product measure on coordinates $A_i \setminus (A_{j_1} \cap \cdots \cap A_{j_k})$,

$$\begin{aligned}\lambda_{A_i} &= v_i + \sum_{\substack{1\le j\le n\\ j\neq i}} v_{i,j}\pi^{A_i}_j + \sum_{k=1}^{n-1}(-1)^k \sum_{\substack{j_1<\cdots<j_k\\ j_s\neq i}} v_{i,j_1,\dots,j_k}\pi^{A_i}_{i,j_1,\dots,j_k}\\ &\quad + \sum_{k=2}^{n-1}(-1)^{k+1} \sum_{\substack{j_1<\cdots<j_k\\ j_s\neq i}} v_{i,j_1,\dots,j_k}\pi^{A_i}_{i,j_1,\dots,j_k}\\ &= v_i. \end{aligned}$$

□

We can relax the hypothesis of Proposition 1.2 by allowing one of the marginals not to be product measure. In this case, we need all the atoms of that measure not to be of zero measure.

Definition. We say that a probability measure v on Ω^A is *ρ-quasi-independent* if $v(x) \ge \rho$, $\forall x \in \Omega^A$. We say that v is *quasi-independent* (q.i.) if it is ρ-q.i. for some $\rho > 0$.

COROLLARY 1.3. *With the above notation, if v_1 is ρ-q.i., then there exists $\eta > 0$ depending only on ρ, n and the size of A_1, such that, if $|v_i - \pi_i| < \eta\pi_i$ for all $i = 1,\dots,n$, then there exists a probability measure λ on Ω^A with marginals v_i, i.e.* $\mathrm{proj}_{A_i}(\lambda) = v_i$, $1 \le i \le n$.

Proof. The proof is exactly the same as the one above. Following the same notation as in (1.1),

$$\begin{aligned}&F(v_1,\dots,v_n)\\ &\quad = \pi + \sum_{k=1}^{n} \sum_{1\le i_1<\cdots<i_k\le n} (v_{i_1,\dots,i_k} - \pi_{i_1,\dots,i_k})\pi^c_{i_1,\dots,i_k}\\ &\quad \ge \pi + (v_1 - \pi_1)\pi^c_1 - \sum_{k=2}^{n} |v_i - \pi_i|\pi^c_i - \sum_{2\le i_1<\cdots<i_k\le n} |v_{i_1,\dots,i_k} - \pi_{i_1,\dots,i_k}|\pi^c_{i_1,\dots,i_k}\\ &\quad \ge v_1\pi^c_1 - (2^n - 2)\eta\pi \ge [v_1 - 2^n\eta\pi_1]\pi^c_1.\end{aligned}$$

Letting $b = \min_{1\le i\le r} p_i$, since the atoms of v_1 are bigger than ρ, then, choosing η so that $2^n\eta b^{|A_1|} < \rho$, the measure $F(v_1,\dots,v_n)$ is positive. □

We remark that Proposition 1.2 is not valid without the assumption of closeness to product measure. To see this just consider random variables X_1, X_2, X_3 with common distribution $(1/2, 1/2)$ and couple pairs by the conditions $X_1 = X_2$, $X_1 = X_3$, $X_2 \perp X_3$. Evidently there is no three-fold joint distribution for X_1, X_2, X_3 consistent with these conditions.

In §2 we will need a result, of interest in its own right, which allows us to extend a shift-invariant measure on Ω^I, $I \subset \{0,\dots,h\}$ to a shift-invariant measure on $\Omega^{\{0,\dots,h\}}$. Here 'shift-invariant' means as shift-invariant as possible, that is, whenever both J and $J + k$ are contained in I then the marginals of μ on Ω^J and Ω^{J+k} are the shifts of each other.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

PROPOSITION 1.4. *Given $h > 0$ $\exists \eta > 0$ such that whenever $I \subset \{0, \ldots, h\}$ and μ is any shift-invariant measure on Ω^I which is η-close to the product of its one-dimensional marginals then there is a shift-invariant measure on $\Omega^{\{0,\ldots,h\}}$ with marginal μ on Ω^I. Moreover, given $\epsilon > 0$, if η is sufficiently small then λ will be ϵ-close to product measure.*

Proof. Define the length of a set of integers $S = \{a_1, a_2, \ldots, a_i\}$, ordered increasingly, to be $l = a_i - a_1$. That is, the length of the set $\{0, 1\}$ is one, not two. Without loss of generality we can assume that $I = \{0, s_1, s_2, \ldots, s_k\}$ with $0 < s_1 < s_2 < \cdots < s_k$. A subset of I, say $B = \{b_1, \ldots, b_j\}$ (ordered increasingly), is maximal if for any $b_i < m < b_{i+1}$, $1 \leq i < j$, the set $B \cup \{m\}$ is not a subset of I. Let $0 < l_1 < l_2 < \cdots < l_L$ be a listing of all possible lengths of subsets of I. Consider

$$\hat{D}_i = \{\text{all maximal subsets } \sigma \subset I \text{ of length } l_i\}.$$

Translate these sets to the origin, let

$$D_i = \{\sigma - s : \sigma \in \hat{D}_i, s \text{ the smallest element in } \sigma\}.$$

By hypothesis we have that for any $\sigma \in \cup_{i=1}^{L} D_i = D$,

$$|\mu_\sigma - \pi_\sigma| < \eta \pi_\sigma.$$

For simplicity of notation, we identify the set of patterns D with the set

$$\cup_{\sigma \in D} \{n : n \in \sigma\} \subset H.$$

It will be clear from the context which of the two we are referring to. Note that the collection of measures $\{\mu_\sigma\}_{\sigma \in D}$ is consistent because all these measures are marginals of μ, and μ is shift-invariant.

We will inductively construct a shift-invariant measure λ on $\Omega^{\{0,1,\ldots,h\}}$ whose projection onto Ω^σ is μ_σ, for all $\sigma \in D$. We start the induction by letting λ_1 be the measure on coordinates $\{0, 1, \ldots, l_1\}$ defined as the independent joining of μ_σ, $\sigma \in D_1$, and product measure on coordinates 1 through $l_1 - 1$,

$$\lambda_1(i_0, i_1, \ldots, i_{l_1}) = \mu_\sigma(i_0, i_{l_1}) \prod_{k=1}^{l_1 - 1} p(i_k).$$

Recall that since l_1 is the smallest length of subsets in A, D_1 consists of only one set, $\sigma = \{0, l_1\}$. Let λ_1' be the Markov extension (Definition 1.5 below) of λ_1 to coordinates $\{0, 1, \ldots, l_2 - 1\}$.

Now suppose $\lambda_1, \lambda_1', \ldots, \lambda_j, \lambda_j'$ have all been defined, for $j < L$, and all are shift-invariant measures. Define λ_{j+1} as follows. Let ν^l and ν^r be two copies of λ_j', one defined on coordinates $\{0, 1, \ldots, l_{j+1} - 1\}$ and the other on coordinates $\{1, 2, \ldots, l_{j+1}\}$. Let λ_{j+1} be the shift-invariant measure obtained by applying Proposition 1.2 to the measures ν^l, ν^r and μ_σ, $\sigma \in D_{j+1}$. The measures ν^l, ν^r and μ_σ, $\sigma \in D_{j+1}$ form a consistent set of measures. This is due to the fact that ν^l and ν^r are shift-invariant and their projection onto coordinates $v \in D_k (k \leq j)$ is μ_v, and because the collection $\{\mu_\sigma\}_{\sigma \in D}$ is consistent. It is clear that λ_{j+1} is shift-invariant since its projection onto coordinates $\{0, 1, \ldots, l_{j+1} - 1\}$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

and $\{1, 2, \ldots, l_{j+1}\}$ are ν^l and ν^r respectively, both copies of λ'_j. If $j+1 < L$, let λ'_{j+1} be the Markov extension of λ_{j+1} to coordinates $\{0, 1, \ldots, l_{j+2} - 1\}$. Finally, let λ be the Markov extension of λ_L to $\Omega^{\{0,1,\ldots,h}\}$. By construction, λ satisfies $\lambda_J = \rho_J$ for all $J \subset I$. Indeed, if $A \subset \{0, 1, \ldots, l_j\}$, then $\lambda_A = (\lambda_j)_A$.

The construction of these measures can be carried out provided that at each step, the involved measures satisfy the hypothesis of Proposition 1.2. To see that this can be ensured, recall that μ_σ is η-close to product measure, for all $\sigma \in D$. So λ_1 is also η-close to product measure. Since the number of iterations of Proposition 1.2 and the number of Markov extensions made in our construction can be bounded *a priori* once h is known, if η is sufficiently small then all the λ_j and λ'_j will remain close to product measure (see Note 1.6), which guarantees that we can keep applying Proposition 1.2. □

Like Proposition 1.2, Proposition 1.4 is not valid without the assumption of closeness to independence, as we can see by considering two-valued random variables $X_1 = X_2$ and $X_4 \perp X_1$. This is clearly a stationary sequence but if we could complete it to a stationary sequence X_1, X_2, X_3, X_4 we would have $X_1 = X_2 = X_3 = X_4$, contradicting $X_1 \perp X_4$.

It is natural to ask whether η in Proposition 1.4 can be made to depend only on the cardinality of I and not on its extent. We have not been able to answer this question. The following result gives a positive answer in a very special case which will suffice for our purposes in §2. Its proof is just a comment about the proof of Proposition 1.4.

PROPOSITION 1.4′. *Given $I \subset \{0, \ldots, h\}$ $\exists \eta > 0$ such that whenever $k > h$ and μ is a shift-invariant measure on $\Omega^{I \cup I+n}$ such that*

$$|\mu - \pi_{I \cup I+k}| < \eta \pi_{I \cup I+k} \quad \text{and} \quad \mu_I = \pi_I$$

(so $\mu_{I+k} = \pi_{I+k}$ also) then there exists a shift invariant measure λ on $\Omega^{\{0,\ldots,h+k\}}$ such that $\lambda_{I \cup I+k} = \mu$.

Proof. Since $\mu_I = \pi_I$ the inductive construction in Proposition 1.4 can be started with product measure on $\Omega^{\{0,\ldots,k-1\}}$. Then the number of applications of Proposition 1.2 and Markov extensions is a again *a priori* bounded as in Proposition 1.4. □

In §2, we will also need to extend a shift-invariant measure on Ω^N to a shift-invariant measure on a larger space. In this case we do not need to apply Proposition 1.2, rather we simply use the $(N-1)$-step Markov extension of a stationary process $(X_1, X_2, \ldots, X_N)$.

If μ is a measure on Ω^N we write $\mu(\Omega, i_2, \ldots, i_n) = \sum_{j \in \Omega} \mu(j, i_2, \ldots, i_N)$.

Definition 1.5. The one-step *Markov extension* of a shift-invariant measure μ on Ω^N to a shift-invariant measure $\tilde{\mu}$ on Ω^{N+1} is the relative independent joining of μ with its shift, that is

$$\tilde{\mu}(i_1, \ldots, i_{N+1}) = \frac{\mu(i_1, \ldots, i_N)\mu(i_2, \ldots, i_{N+1})}{\mu(\Omega, i_2, \ldots, i_N)}$$

if the denominator is not zero, and zero otherwise.

Note that since μ is shift-invariant, $\tilde{\mu}$ is well defined, and the projection of $\tilde{\mu}$ to coordinates $\{1, 2, \ldots, N\}$ or $\{2, 3, \ldots, N+1\}$ is μ. Repeating this process one step at a time, given a shift-invariant measure μ on Ω^N, we can find a shift-invariant measure on Ω^{N+M} for any $M > 0$ (or $\Omega^{\mathbb{Z}}$), whose projection onto Ω^N is μ. This extension will be called the *Markov extension* of μ.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Note 1.6. Suppose that μ is a probability measure on Ω^N with one-dimensional marginals given by a fixed probability vector p. Let π_N denote the product measure on Ω^N with the same one-dimensional marginals. If μ is δ-close to product measure, i.e. $|\mu - \pi_N| < \delta\pi_N$, then it is easy to see that μ', the one-step Markov extension of μ, is also close to product measure on Ω^{N+1}. Indeed, it is δ'-close to product measure for some δ' depending on δ, and $\delta' \to 0$ as $\delta \to 0$. Therefore, any finite Markov extension enjoys the same property. Of course, δ' depends on both δ and the length of the extension.

There is a result analogous to Proposition 1.4 in the Hilbert space setting, extending a finite collection of stationary random vectors which are pairwise almost orthogonal to an infinite stationary process. We thank Peter Schwartz for providing the following simplified proof.

PROPOSITION 1.7. *Suppose $(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$ is a stationary sequence of vectors in a Hilbert space, that is, $\|X_{i_j}\| = 1$ and for any difference that occurs more than once, say $i_j - i_k = i_{j'} - i_{k'}$, then $(X_{i_j}, X_{i_k}) = (X_{i_{j'}}, X_{i_{k'}})$. Then there exists η, depending only on m, such that, if $|(X_{i_j}, X_{i_k})| < \eta$ for all $1 \le j \ne k \le m$, then there exists an infinite stationary sequence $\{Y_n : n \in \mathbb{Z}\}$ such that for all j, k, $(X_{i_j}, X_{i_k}) = (Y_{i_j}, Y_{i_k})$.*

Proof. Let $D = \{i_j - i_k : 1 \le j < k \le m\}$ be the set of differences. The infinite process $(Y_n : n \in \mathbb{Z})$ exists if and only if (Y_0, Y_n) is a positive definite sequence. Then the problem is to interpolate the values on $i_j - i_k$, namely $(X_{i_j}, X_{i_k}) = \rho(i_j - i_k)$ to a positive definite sequence. We will achieve this by considering an appropriate trigonometric polynomial.

Consider the polynomials

$$\tilde{R}(t) = \sum_{d \in D} \rho(d) e^{2\pi i d t} + \sum_{d \in D} \overline{\rho(d)} e^{-2\pi i d t} \quad \text{and} \quad R(t) = 1 + \tilde{R}(t).$$

Then, $\tilde{R}(t)$ is real for all t and we can estimate

$$|\tilde{R}(t)| < 2|D|\eta, \quad \text{where } \eta = \max_{d \in D} |\rho(d)|.$$

Thus, $R(t) \ge 0$ if $2|D|\eta < 1$. Clearly, if $\hat{R}$ denotes the Fourier transform,

$$\hat{R}(d) = \rho(d), \quad d \in D \quad \text{and} \quad \hat{R}(0) = 1,$$

as desired.

Note that η depends only on the size of D, a number that can be bounded by a constant depending only on m. Indeed, the set D may depend on the sequence n_k, but it has at most $\binom{m}{2}$ elements. □

Again, closeness to orthogonality is needed in Proposition 1.7, as the example after Proposition 1.4 shows. Note also that, unlike Proposition 1.4, in Proposition 1.7, η depends only on the cardinality of the set $\{i_1, \ldots, i_m\}$ and not on its extent.

2. *Krengel's theorem for IP sequences*

We introduce some notation that will be used in the rest of the paper.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

By a partition we understand a finite measurable partition of the space (X, β, m). There is a natural pseudometric on the space of partitions with r-atoms, namely, if $P = \{p_1, \ldots, p_r\}$ and $Q = \{q_1, \ldots, q_r\}$, then

$$\|P - Q\| = \inf_{\sigma} \sum_{i=1}^{r} m(p_i \Delta q_{\sigma(i)}),$$

where the infimum is taken over all possible orderings σ of r elements. The set of finite partitions (modulo sets of zero measure), becomes a complete metric space under this metric. Given two partitions P and Q, $P \vee Q$ denotes the common refinement partition consisting of the intersections of the atoms of P with those of Q.

Definition 2.1. Given a partition Q, we say that Q *moves independently along* F if $\{T_g Q\}_{g \in F}$ is a mutually independent family. We say that Q *moves* ϵ*-independently along* F if for any choice $\{A_g\}_{g \in F}$ of atoms of Q,

$$\left| m(\cap_{g \in F} T_g A_g) - \prod_{g \in F} m(A_g) \right| < \epsilon \prod_{g \in F} m(A_g).$$

Let $r \in N$ be fixed, and let $\Omega = \{1, 2, \ldots, r\}$ be a set of r symbols. Let p be a probability measure on Ω such that $p(i) > 0$ for all $i = 1, \ldots, r$. We will use P, P', Q, Q' to denote partitions with r atoms and distribution p. That is, if $A_1, \ldots, A_r$ are the atoms of one of them, then $m(A_i) = p(i)$. Assume $Q = \{A_1, \ldots, A_r\}$. Given a finite set S of integers, we can code the distribution of the partition $\vee_{n \in S} T^n Q$ by associating it with a probability measure on Ω^S, $\mu = \text{distr}(\vee_{n \in S} T^n Q)$, defined as

$$\mu(\{a_n\}_{n \in S}) = m(\cap_{n \in S} T^n A_{a_n}). \tag{2.1}$$

For the proof of Theorem 5 we will use the following three lemmas, which all say that if a partition moves almost independently along a certain subset then it can be perturbed to improve the independence.

LEMMA 2.2. *Let I be a finite set of positive integers and Q a finite partition. Given $\epsilon > 0$ there exists $\eta > 0$ such that if Q moves η-independently along I, then, $\forall \eta' > 0$ there is a partition Q' such that $\|Q - Q'\| < \epsilon$ and Q' moves η'-independently along I.*

Proof. See below.

LEMMA 2.3. *Let I be a finite set of positive integers and Q a finite partition. Given $\epsilon > 0$ there exists $\eta > 0$ such that if Q moves η-independently along I, then there is a partition Q' such that $\|Q - Q'\| < \epsilon$ and Q' moves independently along I.*

Proof. Iterate Lemma 2.2. □

LEMMA 2.4. *Given $I \subset \{0, \ldots, h\}$ and $\epsilon > 0$ $\exists \delta = \delta(I, \epsilon) > 0$ such that if P moves exactly independently along I and δ-independently along $I \cup I + k$ for some $k > h$, then $\forall \eta > 0$ $\exists P'$ such that $\|P' - P\| < \epsilon$ and P' moves η-independently along $I \cup I + k$.*

Proof. See below.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Notation: if A is an index set, v is a measure on Ω^A and $B \subset A$, then v_B denotes the projection of v onto the coordinates B.

Proof of Lemma 2.2. Let h be a large integer to be specified later and $H = \{0, 1, 2, \ldots, h-1\}$. Let μ be the measure on Ω^H defined by the distribution of the partition $\vee_{k=0}^{h-1} T^k Q$, as in equation (2.1). Since T is measure preserving, μ is shift-invariant. Since Q moves η-independently along I, μ_I is η-close to product measure.

We define a (signed) measure ρ on Ω^I by the equation

$$\pi_I = (1 - \eta^{1/2})\mu_I + \eta^{1/2}\rho.$$

We have

$$|\rho - \pi_I| < (1 - \eta^{1/2})\eta^{1/2}|\mu_I - \pi_I| < \eta^{1/2}\pi_I.$$

In particular, ρ will in fact be a probability measure, and also shift-invariant, since μ_I and π_I are shift-invariant. The reader should think of ρ as a 'correcting' measure which will be used to adjust the distribution of P closer to independence. If η is sufficiently small, depending only on I, by Proposition 1.4 and Markov extension we get a shift-invariant measure λ on Ω^H with $\lambda_I = \rho$, so we have

$$\pi_I = (1 - \eta^{1/2})\mu_I + \eta^{1/2}\lambda_I.$$

Now construct a Rohlin tower of height $h + 2l$ (l = length of I) with base B independent of the partition $\Phi = \vee_{k=0}^{h+2l-1} T^{-k}(\vee_{n\in I} T^n Q)$. Thus, every level in the tower is independent of $\vee_{n\in I} T^n Q$. Next select a subset $B_0 \subset B$ with $B_0 \perp \Phi$ and $m(B_0) = \alpha m(B)$, where α is to be specified later. Set $B_1 = T^l(B_0)$. On the tower of height h over B_1, $E = \cup_{k=0}^{h-1} T^k B_1$, find a partition $R = \{R_1, R_2, \ldots, R_r\}$ such that

$$\frac{m(\cap_{j=1}^{h} T^{-j} R_{i_j} \cap B_1)}{m(B_1)} = \lambda(i_0, i_1, \ldots, i_{h-1}).$$

This is done in the standard way by 'painting names on the tower': choose a partition Γ of B_1, indexed by Ω^h, such that $\mu(\Gamma_{\mathbf{i}})/\mu(B_1) = \lambda(\mathbf{i})\ \forall \mathbf{i} \in \Omega^h$ and decree that $T^j(\Gamma_{\mathbf{i}}) \subset R_{i_j}$ for each $\mathbf{i} = (i_0, \ldots, i_{h-1}) \in \Omega^h$ and $j \in \{0, \ldots, h-1\}$. Also let

$$E_1 = \cup_{j=0}^{h+2l-1} T^j(B_0) = \cup_{j=-l}^{l} T^j E \quad \text{and} \quad E_0 = \cup_{j=l}^{h-l-1} T^j(B_1).$$

That is,

$$E_0 \subset E \subset E_1.$$

Now let $Q' = \{A'_1, A'_2, \ldots, A'_r\}$ be the partition defined by

$$A'_i = (A_i \cap E^c) \cup R_i, \quad i = 1, 2, \ldots, r.$$

Then

$$m(A'_i) = m(A_i \cap E^c) + m(R_i) = m(A_i)m(E^c) + p(i)m(E) = m(A_i),$$

by the independence of E from the partition Q.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

$$E_0 \subset E \subset E_1.$$

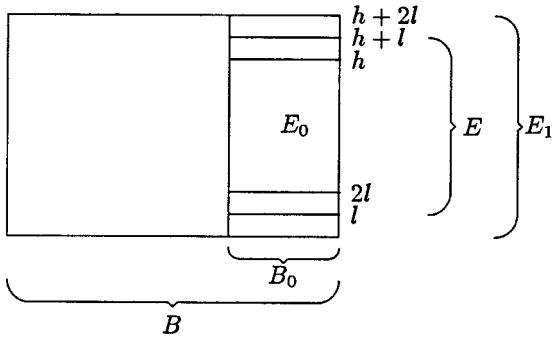


FIGURE 1. Tower over E.

Now consider

$$\begin{aligned}\operatorname{distr}(\vee_{n\in I} T^n Q') = {} & m(E_1^c)\operatorname{distr}(\vee_{n\in I} T^n Q' \mid E_1^c) \\ & + m(E_0)\operatorname{distr}(\vee_{n\in I} T^n Q' \mid E_0) \\ & + m(E_1\backslash E_0)\operatorname{distr}(\vee_{n\in I} T^n Q' \mid (E_1\backslash E_0)).\end{aligned}$$

For the first term, we have that, for any $i = 1, \ldots, r$, $A_i' \cap E_1^c = A_1 \cap E_1^c$ and $T^n A_i' \cap E_1^c = T^n A_i \cap E_1^c$, for all $n \in I$. Therefore, by the independence of the levels from the partition $\vee_{n\in I} T^n Q$,

$$\operatorname{distr}(\vee_{n\in I} T^n Q \mid E_1^c) = \mu_I.$$

For the second term we have that, for any $i = 1, \ldots, r$, $A_i' \cap E_0 = R_i \cap E_0$ and $T^n A_i' \cap E_0 = T^n R_i \cap E_0$, for all $n \in I$. Therefore

$$\operatorname{distr}(\vee_{n\in I} T^n Q' \mid E_0) = \lambda_I.$$

Finally, we do not have control over the third term. Denote it by

$$\operatorname{distr}(\vee_{n\in I} T^n Q' \mid (E_1\backslash E_0)) = \nu.$$

Noting that $m(E_0) = (h - 2l)\alpha m(B)$, choose α so that $(h - 2l)\alpha m(B) = \sqrt{\eta}$. This choice makes $\alpha < 1$ provided that η is sufficiently small. (This condition on η depends on the size of I.) Thus $m(E_1) = \sqrt{\eta} + 2l\alpha m(B)$ and

$$\operatorname{distr}(\vee_{n\in I} T^n Q') = (1 - \sqrt{\eta})\mu_I + \sqrt{\eta}\lambda_I + 2l\alpha m(B)(\nu - \mu_I).$$

Since $(1 - \sqrt{\eta})\mu_I + \sqrt{\eta}\lambda_I = \pi_I$,

$$\operatorname{distr}(\vee_{n\in I} T^n Q') = \pi_I + 2l\alpha m(B)(\nu - \mu_I). \tag{2.4}$$

Choosing h sufficiently large, we can control the last term. Indeed,

$$2l\alpha m(B) = \frac{2l\sqrt{\eta}}{h - 2l} < \eta' \frac{b^{|I|}}{2}, \tag{2.5}$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

for h large enough. Since $b = \min\{p(i) : i = 1, 2, \ldots, r\}$, we obtain

$$|\operatorname{distr}(\vee_{n\in I} T^n Q') - \pi_I| < \eta' \pi_I.$$

Therefore, by further requiring that $\eta < \epsilon^2/9$, we obtain a partition Q' such that
(a) Q' moves η'-independently along I,
(b) $\|Q - Q'\| < 2m(E) = 2h\alpha m(B) = 2h\sqrt{\eta}/(h - 2l) < 3\sqrt{\eta} < \epsilon$,
thus finishing the proof of Lemma 2.2. □

Proof of Lemma 2.4. The proof is exactly analogous to the proof of Lemma 2.2, using Proposition 1.4′ in place of Proposition 1.4 to construct the desired correcting measure ρ. □

Proof of Theorem 5. Let P be a finite partition and $\epsilon > 0$. We will construct a sequence $\{n_k\}_{k=1}^{\infty}$ together with a sequence of partitions $\{P_k\}_{k=1}^{\infty}$ such that $P_0 = P$ and for all $i \geq 1$,
(a) $\|P_i - P_{i-1}\| < \epsilon_i = \epsilon/2^i$ and
(b) P_i moves independently along Γ_i, where $\Gamma_i = \mathrm{FS}(0, n_1, \ldots, n_i)$.
Letting $P' = \lim P_i$ we will then have that $\|P' - P\| < \epsilon$ and P' moves independently along the IP-set generated by $\{n_i\}$.

Suppose that $n_1, \ldots, n_k$ and $P_1, \ldots, P_k$ have already been defined. Let $\delta = \delta(I, \epsilon/2^{i+2})$ as in Lemma 2.5 and choose $n_{k+1} \in \{m_j\}_{j=1}^{\infty}$ such that

$$n_{k+1} > (n_1 + \cdots + n_k) = \max \Gamma_k \tag{2.6}$$

and

$$\vee_{n\in\Gamma_k} T^n P_k \text{ moves } \delta\text{-independently along } \{0, n_{k+1}\}. \tag{2.7}$$

These two conditions can be met simultaneously because $\{m_j\}_{j=1}^{\infty}$ is a mixing sequence. The choice of n_{k+1} guarantees that the size of Γ_{k+1} is twice that of Γ_k so the cardinality of Γ_j is 2^j, $j = 1, \ldots, k+1$.

By Lemma 2.5 given $\eta > 0$, $\exists P_i'$ such that $\|P_i' - P_i\| < \epsilon/2^{i+2}$ and P_i' moves η-independently along $\Gamma_k \cup \Gamma_k + n_k = \Gamma_{k+1}$. If η is sufficiently small then Lemma 2.3 yields P_{i+1} such that $\|P_{i+1} - P_i'\| < \epsilon/2^{i+2}$ and P_{i+1} moves independently along Γ_{k+1}. Moreover, $\|P_{i+1} - P_i\| < \epsilon/2^{i+1}$ as desired. □

Proof of Corollaries 6, 7 and 8. For Corollary 6 it suffices to observe that if T is weakly mixing then it has a mixing sequence of density one. Since this set intersects any set E of density $a > 0$ in a set of density a again, E contains a mixing sequence so Theorem 5 applies.

For Corollary 7 recall that T is mildly mixing if and only if $\forall A, B$ and $\epsilon > 0$,

$$\{n \geq 0 : |\mu(T^n A \cap B) - \mu(A)\mu(B)| < \epsilon\}$$

is an IP* set, that is, it intersects every IP subset of $\mathbf{N}$. Since the intersection of IP* sets is IP* [**5**, Lemma 9.5] it follows that for any finite partitions P, Q and $\epsilon > 0$,

$$\{n \geq 0 : T^n P \perp^{\epsilon} Q\}$$

is IP*. By taking successively finer partitions and letting $\epsilon \to 0$ it follows easily that every IP-set contains a mixing sequence of T and Corollary 7 then follows from Theorem 5.

Finally, the proof of Corollary 8 is immediate from Theorem 5 since here every infinite sequence is mixing. □

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Proof of Proposition 9. The sufficiency for both parts (a) and (b) is a consequence of Corollary 7 and 8 respectively. It remains to prove the necessity.

(a) If T is not mild mixing then there is a measurable set A with $0 < m(A) < 1$, such that, for any $\delta > 0$, the set of integers for which

$$|m(T^n A \cap A) - m(A)| < \delta$$

contains an IP-set (see Furstenberg [**5**]), call it Σ. By hypothesis, given $\epsilon > 0$, there exists a partition $Q = \{B, B^c\}$ and a subset $\Gamma \subset \Sigma$ such that

$$m(B) = m(A), \quad m(A \Delta B) < \epsilon \quad \text{and} \quad Q \text{ moves independently along } \Gamma.$$

Then, for $n \in \Gamma$ we have

$$\begin{aligned} |m(T^n A \cap A) - m(A)^2| &= |m(T^n A \cap A) - m(T^n B \cap B)| \\ &\leq 2\epsilon. \end{aligned}$$

On the other hand, since $n \in \Sigma$,

$$\begin{aligned} |m(T^n A \cap A) - m(A)^2| &> |m(A)^2 - m(A)| - |m(T^n A \cap A) - m(A)| \\ &> m(A)(1 - m(A)) - \delta. \end{aligned}$$

Choosing δ and ϵ sufficiently small, we arrive at a contradiction.

(b) Suppose T is not mixing. Then there exists a measurable set A with $0 < m(A) < 1$, and a sequence $\{n_k\}$ such that

$$|m(T^{n_k} A \cap A) - m(A)^2| > a$$

for some $a > 0$ and all $k \geq 1$.

By hypothesis, for any $\epsilon > 0$, we can find a subsequence $\{m_k\} \subset \{n_j\}$ and a partition $Q = \{B, B^c\}$ such that

$$m(B) = m(A), m(A \Delta B) < \epsilon \quad \text{and} \quad Q \text{ moves independently along } \{m_k\}.$$

Then we have

$$\begin{aligned} |m(T^{m_k} A \cap A) - m(A)^2| &= |m(T^{m_k} A \cap A) - m(T^{m_k} B \cap B)| \\ &\leq cm(A \Delta B) < c\epsilon, \end{aligned}$$

which gives a contradiction by choosing ϵ appropriately because $\{m_k\}$ is a subsequence of $\{n_j\}$.

(c) Suppose T is K, P is a finite partition and $\epsilon > 0$. By the K-mixing property of the process (P, T) given $\eta > 0$ there is an integer $d > 0$ such that

$$T^d P \perp^{\eta} P_{-\infty}^0.$$

It follows that the sequence

$$\{T^{dn} P : n \in \mathbb{Z}\}$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

is η-independent. If η is sufficiently small the fundamental lemma of the isomorphism theory [**12**, Chapter 9] then yields a P' such that $\|P' - P\| < \epsilon$ and

$$\{T^{dn}P' : n \in \mathbb{Z}\}$$

is independent.

Conversely suppose that T satisfies the condition at hand and P is a non-trivial finite partition. Then, given $\eta > 0$ there is a P' and $d > 0$ such that $\|P' - P\| < \eta$ and $\{T^{dn}P'\}$ is an independent sequence, that is $h(T^d, P') = h(P')$. If η is sufficiently small, depending only on the distribution of P it follows [**12**, Lemma 8.2] that

$$h(T^d, P) > h(P)/2 > 0.$$

Since $h(T^d, P) \leq dh(T, P)$ we get $h(T, P) > 0$. Thus T has completely positive entropy, showing that it is a K-automorphism. □

3. *Krengel's theorem for group actions*

3.1. *Notation.* Let G be a countable, discrete abelian group acting on a non-atomic probability space (X, β, m) by measure preserving transformations, say $\{T_g\}_{g\in G}$. Assume also that the action of G on X is free, i.e. $m(\{x \in X : T_g x = x\}) = 0$ for all $g \in G$, $g \neq e$.

For any set $A \subset G$, $|A|$ denotes the cardinality of A. A Følner sequence is a collection of subsets $\{F_n\}$ of G which are asymptotically invariant, i.e. for all $g \in G$, $\lim_{n\to\infty} |F_n \cap (g + F_n)|/|F_n| = 1$. Let

$$C + F = \{c + f : c \in C, f \in F\}.$$

A finite set F is *ϵ-invariant* under a finite subset $C \subset G$, if

$$|F \cap (C + F)|/|F| > 1 - \epsilon.$$

Then the action of G is *weakly mixing* if for any Følner sequence $\{F_n\}$ in G,

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} |m(A \cap T_g B) - m(A)m(B)| = 0$$

for all $A, B \in \beta$ of positive measure.

Definition 3.1.1. The action of G is k-fold weakly mixing if, for any Følner sequence $\{F_n\}$,

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} |m(A_0 \cap T_g A_1 \cap T_{2g} A_2 \cap \cdots \cap T_{kg} A_k) - m(A_0)m(A_1)\ldots m(A_k)| = 0$$

for any choice of sets $A_0, \ldots, A_k \in \beta$ of positive measure.

Remark 3.1.2. It follows from this theorem that any two-fold weakly mixing G-action, contains a two-fold mixing sequence.

Any weakly mixing transformation is k-fold weakly mixing for all k's (Furstenberg [**5**]). But this is not the case for weakly mixing group actions. The following theorem of Bergelson and Rosenblatt [**2**], characterizes groups with this property.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

THEOREM 3.1.3. ([**2**]) *Let $\{T_g^{(1)}\}_{g\in G},\dots,\{T_g^{(k)}\}_{g\in G}$ be commuting actions of G such that $\{T_g^{(1)}\times\cdots\times T_g^{(k)}\}_{g\in G}$ is an ergodic action on X^k and such that $\{T_g^{(i)}T_{-g}^{(j)}\}_{g\in G}$ is an ergodic action on X for all $i\neq j$. Then, for any $f_1,\dots,f_k\in L^\infty(X)$,*

$$\lim_{n\to\infty}\frac{1}{|F_n|}\sum_{g\in F_n}T_g^{(1)}(f_1)\dots T_g^{(k)}(f_k)=\int f_1\,dm\cdots\int f_k\,dm\quad \text{in } L^2.$$

An immediate consequence is the following.

COROLLARY 3.1.4. *If the action of G is weakly mixing on X and the action of $2G$ is ergodic, then the action of G is two-fold weakly mixing.*

Remarks 3.1.5. (1) If the action of G is weakly mixing and $G/2G$ is compact, then the action of $2G$ is weakly mixing and consequently, the action of G is two-fold weakly mixing.

(2) If $G=\oplus_{i\in I}G_i$ such that for every $i\in I$, G_i is a field with characteristic not equal to two, then if the action of G is weakly mixing, it is also two-fold weakly mixing because, in this case $2G=G$.

(3) There exists free $G=\oplus_{i=1}^{\infty}\mathbb{Z}$ actions which are weakly mixing but not two-fold weakly mixing. For example, let $H=\oplus_{i=1}^{\infty}\mathbb{Z}_2$, $\{S_g\}_{g\in G}$ a weakly mixing action on X, and $\{R_h\}_{h\in H}$ a weakly mixing action on Y. Define a G action on $X\times Y$ by $T_g=S_g\times R_g$, where if $g=(n_1,n_2,\dots)R_g=R_h$ with $h=(n_1 \bmod(2),n_2 \bmod(2),\dots)$. This action is weakly mixing on $X\times Y$ but the action of $2G$ is not ergodic on $X\times Y$. Indeed, for any $g\in G$, $T_{2g}=S_{2g}\times \mathrm{Id}_Y$. Thus, any set of the form $X\times A$ with $A\in Y$ of positive measure is invariant under the action of $2G$.

3.2. *Rohlin tower.* Before proving Theorem 10 we need a technical lemma. Given $g_1,\dots,g_k\in G$ and a partition P of X, we will find a set $E\subset X$ such that $\{T_gE\}_{g\in \mathrm{FS}(\pm g_1,\dots,\pm g_k)}$ are pairwise disjoint, each independent of P, and such that the measure of E is large enough and does not depend on the elements in consideration. We use a technique due to Conze [**4**] in his construction of Rohlin towers for $\mathbb{Z}^2$ actions.

Definition 3.2.1. Given a finite set $F\subset G$, a *Rohlin tower (R-tower) along F* consists of a base set $B\subset X$, together with all its translates along F such that $\{T_gB\}_{g\in F}$ are pairwise disjoint.

LEMMA 3.2.2. *Suppose $F=\{f_1,\dots,f_n\}$ is a finite subset of G, $P=\{P_1,\dots,P_n\}$ is a partition and $\gamma=\min\{m(P_i):i=1,\dots,n\}$. Then there is an R-tower along F with base E such that $E\perp P$ and $\mu(E)\geq\gamma n^{-2}$.*

Proof. Freeness implies that for any finite set $\Gamma\subset G$ and $A\subset X$ with $m(A)>0$ there is a subset $E\subset A$ with $m(E)>0$ and $gE\cap E=0$ for $g\in\Gamma$. If $\Gamma=F-F$ this is equivalent to $\{fE:f\in F\}$ being an R-tower. Now choose $E\subset X$ maximal with the property that $\{fE:f\in F\}$ being an R-tower and $E\perp P$. If $m(E)<\gamma n^{-2}$ setting $\Gamma=F-F$ we get

$$m(P_i-\cup\{gE:g\in\Gamma\})>0\ \forall i.$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

We can then choose in succession non-null subsets $E_1, E_2, \ldots, E_n$ of $P_1, P_2, \ldots, P_n$ such that $gE_i \cap E_i = 0$ for $g \in \Gamma$ and

$$E_i \subset P_i - \cup\{gE_j : g \in \Gamma, j = 1, \ldots, i-1\}.$$

If $E_1, \ldots, E_{i-1}$ are chosen sufficiently small there will always be room to choose E_i. Now choosing $E_i' \subset E_i$ such that $m(E_i')/m(P_i)$ is independent of i and setting

$$E' = E \cup \bigcup\{E_i' : i = 1, \ldots, k\},$$

we will have $E' \perp P$ and $\{fE' : f \in F\}$ will be an R-tower, contradicting maximality of E. □

3.3. *Proof of Theorem 10.* The proof of this theorem follows a similar structure as that of Theorem 5 in that the new partition is constructed by the painting names on towers technique. However, in this case, instead of using a regular Rohlin tower, we use an R-tower restricted to the appropriate times.

Let F be a finite subset of G and Q a finite partition.

Definition 3.3.1. We say that Q *moves quasi-independently (q.i.) along* F if the atoms of the partition $\vee_{g\in F} T_g Q$ all have non-zero measure. We say that Q *moves* ϵ*-quasi-independently* (ϵ*-q.i.) along* F if each atom of the partition $\vee_{g\in F} T_g Q$ has measure at least ϵ.

Let $r \in N$ be fixed and let $\Omega = \{1, 2, \ldots, r\}$ be a set of r symbols. Let p be a probability measure on Ω such that $p(i) > 0$ for all $i = 1, \ldots, r$. As in §2, P, P', Q, Q' denote partitions with r atoms and distribution p. Assume $Q = \{A_1, \ldots, A_r\}$. Given a finite set $S \subset G$, we can code the distribution of the partition $\vee_{g\in S} T_g Q$ by associating it with a probability measure on Ω^S, $\mathrm{distr}(\vee_{g\in S} T_g Q) = \mu$, defined as

$$\mu(\{a_g\}_{g\in S}) = m(\cap_{g\in S} T_g A_{a_g}). \tag{3.1}$$

Suppose Γ is a finite set in G. Let $\Gamma^* = \Gamma - \Gamma$, $h \notin \Gamma^*$, $\Gamma_1 = \Gamma \cup (h + \Gamma)$, $\Gamma_1^* = \Gamma_1 - \Gamma_1$. Also let

$$\Phi(Q) = \vee_{g\in\Gamma^*} T_g Q \quad \text{and} \quad \Phi_1(Q) = \vee_{g\in\Gamma_1^*} T_g Q. \tag{3.2}$$

Theorem 10 follows from the following proposition.

PROPOSITION 3.3.2. *Let Q be a finite partition such that Q moves independently along Γ and q.i. along Γ^*. Given $\epsilon > 0$ there exists $\delta > 0$ such that if $\Phi(Q)$ moves δ independently along $\{-h, 0, h\}$, then there exists a partition Q' such that $\|Q' - Q\| < \epsilon$ and Q' is independent along Γ_1 and q.i. along Γ_1^*.*

Proof of Theorem 10. The partition P' and the sequence $\{g_n\}_{n=1}^{\infty} \subset S$ are constructed inductively. Let $\epsilon_k = \epsilon/2^{k+1}$. Together with the sequence $\{g_k\}_{k=1}^{\infty}$ we will construct a sequence of partitions $\{P_k\}_{k=1}^{\infty}$. Let $\Gamma_j = \mathrm{FS}(0, g_1, \ldots, g_j)$ and $\Gamma_j^* = \mathrm{FS}(0, \pm g_1, \ldots, \pm g_j)$.

The sequences are constructed in such a way that $P_0 = P$ and for every $k > 0$:

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

(a) the sets $\Gamma^*_{k-1} - g_k$, Γ^*_{k-1} and $\Gamma^*_{k-1} + g_k$ are pairwise disjoint;
(b) $\|P_k - P_{k-1}\| < \epsilon_k$;
(c) P_k moves independently along Γ_k;
(d) P_k is q.i. along Γ^*_k.

Once this is done, P' is the limiting partition, $P' = \lim_{k\to\infty} P_k$, and $\Gamma = \mathrm{FS}(\{g_k\})$ is the IP-set generated by the sequence $\{g_k\}_{k=1}^{\infty}$. P' then satisfies that $\|P - P'\| < \epsilon$ and the family $\{T_g P'\}_{g\in\Gamma}$ is mutually independent.

Assume that $g_1, \ldots, g_k$ and $P_1, \ldots, P_k$ have already been chosen. Then P_k is η-q.i. along Γ^*_k for some $\eta > 0$. By Proposition 3.3.2 there exits a $\delta > 0$ depending on η, the size of Γ_k and ϵ_k such that if g satisfies (a) and $\Phi(P_k)$ moves independently along $\{-g, 0, g\}$, then we can find a partition Q, ϵ_k-close to P_k, such that Q moves independently along $\Gamma' = \Gamma_k \cup \Gamma + g$ and is q.i. along $(\Gamma')^*$. Since the sequence S is two-fold mixing we can choose g to be an element of S, call it g_{k+1}. Then $P_{k+1} = Q$ has the desired properties. □

To prove Proposition 3.3.2, we need an inductive procedure by which we achieve independence along subsets of Γ_1 of a certain size, once we have achieved independence of the subsets of smaller size. Moreover, at each step, the partition also moves independently along Γ. We do the induction via the help of a technical lemma which carries the main idea behind Theorem 10.

Let us call a subset of Γ_1 a proper subset if it is not a translate of a subset of Γ. Let τ'_k be the collection of all proper subsets of Γ_1 of cardinality k. Now let

$$\tau_k = \{\sigma - s : \sigma \in \tau'_k, s \in \sigma\}.$$

Then $0 \in \sigma$ and $\sigma \subset \Gamma_1 - \Gamma_1 = \Gamma^*_1$, for all $\sigma \in \tau_k$. The size of τ_k is at most

$$|\tau_k| = \sum_{j=1}^{k-1} \binom{|\Gamma|}{j}\binom{|\Gamma|}{k-j} k \le k\binom{2|\Gamma|}{k} \le 4^{|\Gamma|}.$$

Hence it does not depend on the elements of Γ_1, only on its size.

Proposition 3.3.2 will be proved by induction on the sizes of the proper subsets of Γ_1. Note that $|\Gamma_1| = 2|\Gamma|$, Let $N = 4^{|\Gamma|} + 1$.

LEMMA 3.3.3. *Let Γ and Q be as in Proposition 3.3.2 and assume h satisfies that the sets $\Gamma^* - h$, Γ^* and $\Gamma^* + h$ are pairwise disjoint. Given $\epsilon > 0$ and $\eta > 0$ there exists $\delta > 0$ such that if:*

(a) *Q moves independently along Γ;*
(b) *Q moves independently along σ for all $\sigma \in \tau_k$;*
(c) *Q is η-q.i. along Γ^*;*
(d) *$\Phi(Q)$ moves δ-independently along $\{-h, 0, h\}$,*

then, there exists Q' such that $\|Q - Q'\| < \epsilon$ and:

(a′) *Q' moves independently along Γ;*
(b′) *Q' moves independently along σ for all $\sigma \in \tau_{k+1}$;*
(c′) *Q' is η'-q.i. along Γ^*;*
(d′) *$\Phi(Q')$ moves δ'-independently along $\{-h, 0, h\}$,*

where $\eta' = (1 - |\Gamma^*|\sqrt{\delta})\eta$ and $\delta' = \delta + 6|\Gamma^*|\sqrt{\delta}/\eta^3$.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Proof of Proposition 3.3.2. By iterating Lemma 3.3.3 with $\epsilon/|\Gamma_1|$ instead of ϵ, we can find partitions Q_k, and parameters $\delta_k, \eta_k > 0$, $1 \le k \le |\Gamma_1|$, such that $Q_1 = Q$, $\delta_1 = \delta$, $\eta_1 = \eta$ and for all $k \ge 1$:

(a) Q_k moves independently along Γ;
(b) Q_k moves independently along σ for all $\sigma \in \tau_k$;
(c) Q_k is η_k-q.i. along Γ^*;
(d) $\Phi(Q_k)$ moves δ_k-independently along $\{-h, 0, h\}$;
(e) $\|Q_k - Q_{k+1}\| < \epsilon/|\Gamma_1|$.

By Lemma 3.3.3, the parameters δ_k and η_k depend continuously on δ and η because they satisfy the recurrence relation in (d′) above. Thus, given any $\eta_k > 0$, we can find $\delta_k > 0$ so small (and independent from h) such that for each η_k, δ_k is small enough that Lemma 3.3.3 can be applied to Q_k. Taking $Q' = Q_{|\Gamma_1|}$, we have $\|Q - Q'\| < \epsilon$, Q' moves independently along Γ_1 and q.i. along Γ_1^* as desired. □

We now proceed to prove the main lemma.

Proof of Lemma 3.3.3. Using Lemma 3.3.2, construct an R-tower over Γ_1^* with base set E such that

$$E \perp \Phi_1(Q) \quad \text{and} \quad m(E) \ge \varphi,$$

where

$$\varphi = \frac{1}{|\Gamma_1^*|} \min\{m(A) : A \in \Phi_1(Q)\} \ge \frac{1}{|\Gamma_1^*|}(1-\delta)\eta^3,$$

by (c) and (d).

We will change the partition Q on a small portion of the set E, in such a way that it will correct the distributions of all proper subsets of Γ_1 of size $k+1$ while preserving the distribution of proper subsets of smaller cardinality. By hypotheses (a) and (b), Q moves independently along Γ and σ for each $\sigma \in \tau_k$. Hence, the distributions of

$$\vee_{g\in\Gamma} T_g Q \quad \text{and} \quad \vee_{g\in\sigma} T_g Q, \quad \text{for all } \sigma \in \tau_k,$$

are product measure. Also by hypothesis (d), the distribution of $\vee_{g\in\sigma} T_g Q$ is δ-close to product measure, for all $\sigma \in \tau_{k+1}$. Recall that Q is a partition with r atoms, $Q = \{A_1, \ldots, A_r\}$, and distribution p.

Let μ be the distribution of the partition $\Phi_1(Q)$ defined by equation (3.1). Let π_B denote the product measure on coordinates $B \subset \Gamma_1^*$ with one-dimensional marginals given by p, and μ_B the projection of μ onto those coordinates. For short, denote $\mu_{\Gamma_1^*\setminus\{0\}}$ by $\tilde{\mu}$. By hypothesis, $\mu_\Gamma = \pi_\Gamma$ and $\mu_\nu = \pi_\nu$, for all $\nu \in \tau_k$. Also, μ_σ is δ-close to product measure, for all $\sigma \in \tau_{k+1}$. Notice that since $0 \in \sigma$ for all $\sigma \in \tau_{k+1}$, $\sigma \cap (\Gamma_1^*\setminus\{0\})$ is a pattern of smaller cardinality than σ. Thus, the projection of $\tilde{\mu}$ onto those coordinates is product measure. Also, since the set of measures $\tilde{\mu}$ and $\{\mu_\sigma\}_{\sigma\in\tau_{k+1}}$ are all marginal measures of μ, they form a consistent set of measures.

For each $\sigma \in \tau_{k+1}$, define the correcting measure ρ_σ on Ω^σ by

$$\pi_\sigma = (1-\beta)\mu_\sigma + \beta\rho_\sigma,$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

where $\beta = \sqrt{\delta}$. Since the measure μ_σ is close to product measure, so is ρ_σ, Indeed,

$$|\rho_\sigma - \pi_\sigma| = \frac{(1-\beta)}{\beta}|\mu_\sigma - \pi_\sigma| \leq \frac{1}{\sqrt{\delta}}\delta\pi_\sigma = \sqrt{\delta}\pi_\sigma. \tag{3.3}$$

Thus, ρ_σ is a probability measure on Ω^σ.

By construction and the above remark the measures $\tilde{\mu}$ and $\{\rho_\sigma\}_{\sigma\in\tau_{k+1}}$ form a collection of consistent measures. Applying Corollary 1.3 to this set of measures, we obtain a measure λ in $\Omega^{\Gamma_1^*}$ with marginals

$$\lambda_{\Gamma_1^*\setminus\{0\}} = \tilde{\mu} \quad \text{and} \quad \lambda_\sigma = \rho_\sigma \quad \text{for all } \sigma \in \tau_{k+1}.$$

To check that Corollary 1.3 can be used here notice that, by (3.3), each of the ρ_σ is $\sqrt{\delta}$-close to product measure, and the atoms of $\tilde{\mu}$ are large enough. In fact, by hypothesis (d) and the fact that

$$\Gamma_1^* = (\Gamma^* - h) \cup \Gamma^* \cup (\Gamma^* + h),$$

the atoms of $\tilde{\mu}$ (and μ) are at least $(1-\delta)\,\alpha^3$ large, where α is the minimum size of atoms of Γ^*. Thus, from hypothesis (c) the atoms of $\tilde{\mu}$ (and μ) are then at least $(1-\delta)\,\eta^3$ large. Thus, by Corollary 1.3, we can choose δ so small and depending only on η, the size of Γ_1^* and N (upper bound on the total number of measures employed in the construction of λ (see (3.4))), such that the extending measure λ constructed in Corollary 1.3 is a probability measure.

The measure λ has the added property that if F is Γ or a proper subset of Γ_1 of cardinality $\leq k$, then if $F + l \subset \Gamma_1^*$,

$$\lambda_{F+l} = \pi_{F+l}. \tag{3.4}$$

This property holds because the projection onto common coordinates of any pair of measures in the collection $\{\hat{\mu}, \{\rho_\sigma\}_{\sigma\in\tau_{k+1}}\}$ is product measure, λ has the form

$$\lambda = \hat{\mu}\pi_{\{0\}} + \sum_{\sigma\in\tau_{k+1}} \rho_\sigma \pi_\sigma^c - C\pi_{\Gamma_1^*} \tag{3.5}$$

(see (1.1)), where C is the cardinality of τ_{k+1}, and π_σ^c is the product measure on coordinates $\Gamma_i^*\setminus\sigma$.

Let $\alpha = \sqrt{\delta}/((k+1)m(E))$. Then

$$\alpha \leq \sqrt{\delta}/\varphi < |\Gamma_1^*|\sqrt{\delta}/(1-\delta)\eta^3 < 1, \tag{3.6}$$

if δ is chosen sufficiently small. Select a set $E_0 \subset E$ such that

$$m(E_0) = \alpha m(E) \quad \text{and} \quad T_g E_0 \perp \Phi_1(Q)$$

for all $g \in \Gamma_{k+1}$, where $\phi_1(Q)$ is the partition defined in (3.2).

We now want to perturb Q, *only on* E_0, to obtain a partition Q' of X such that

$$\frac{m(\vee_{g\in\Gamma_1} T_g^{-1} Q'_{\mathbf{i}} \cap E_0)}{m(E_0)} = \lambda(\mathbf{i})\ \forall \mathbf{i} \in \Omega^{\Gamma_1}.$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

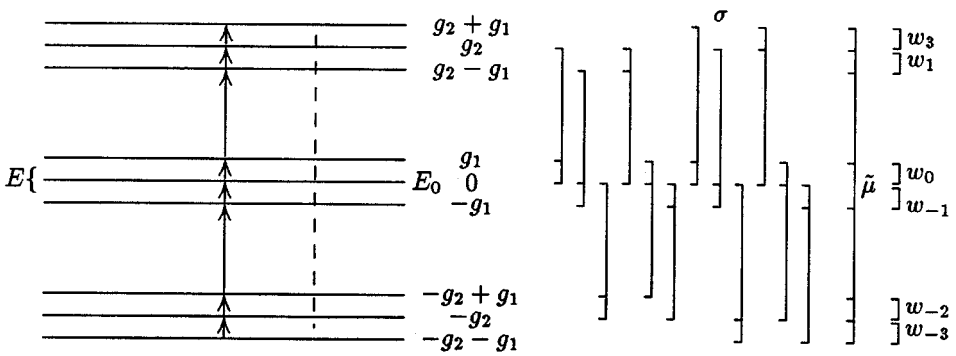


FIGURE 2. Representation of marginals in τ_3.

This is possible because the projection of λ onto $\Omega^{\Gamma_1}\backslash\{0\}$ is $\tilde{\mu}$. Namely, for each $\mathbf{i} \in \Omega^{\Gamma_1\backslash\{0\}}$ perturb Q to Q' on

$$E_{0,\mathbf{i}} = \bigcap\{Q_{\mathbf{i}_g} : g \in \Gamma_1\backslash 0\} \cap E_0$$

so that for each $j \in \Omega$,

$$m\left(Q'_j \cap \bigcap\{Q_{\mathbf{i}_g} : g \in \Gamma_1\backslash 0\} \cap E_0\right) = \lambda(\mathbf{i}^j),$$

where $\mathbf{i}^j$ denotes the extension of $\mathbf{i}$ to Ω^{Γ_1} such that $\mathbf{i}_0 = j$. It follows that

$$\|Q' - Q\| < 2\alpha m(E) = 2\frac{\sqrt{\delta}}{k+1} < 2\sqrt{\delta} < \epsilon. \tag{3.7}$$

The equations (3.6) and (3.7) determine the choice of δ. That is, δ depends only on ϵ, η and the size of Γ_1 (see in (3.3) the definition of N and φ). Hence the choice of δ is independent of h.

For example, when $\Gamma = \{0, g_1\}$, $h = g_2$, $k = 2$, we have Figure 2, where the w_i's denote the projections of μ onto the translations of Γ that fit inside Γ_1^*.

Let B be a subset of Γ_1 and $F = \cup_{g\in B} T_g E_0$. Then

$$\begin{aligned}\operatorname{distr}(\vee_{g\in B} T_g Q') = (1 - |B|m(E_0))\operatorname{distr}(\vee_{g\in B} T_g Q' \mid F^c) \\ + |B|m(E_0)\operatorname{distr}(\vee_{g\in B} T_g Q' \mid F).\end{aligned} \tag{3.8}$$

For the first term of this sum, notice that if $g \in B$, $T_g A'_i \cap F^c = T_g A_i \cap F^c$ for any $1 \le i \le r$. Since $F \perp \vee_{g\in B} T_g Q$, we obtain

$$\operatorname{distr}(\vee_{g\in B} T_g Q' \mid F^c) = \operatorname{distr}(\vee_{g\in B} T_g Q) = \mu_B.$$

For the second term, notice that for any $1 \le i \le r$,

$$\begin{aligned} T_g A'_i \cap T_{g'} E_0 &= T_g A_i \cap T_{g'} E_0 \quad \text{if } g \ne g', \\ T_h A'_i \cap T_{g'} E_0 &= T_{g'} R_i \cap T_{g'} E_0. \end{aligned}$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Thus, since $B - g' \subset \Gamma_1^*$, $\text{distr}(\vee_{g\in B} T_g Q' \mid T_{g'} E_0) = \text{distr}(\vee_{g\in B-g'} T_g Q' \mid E_0)$ equals the projection of λ onto the coordinates $B - g'$. Hence

$$\begin{aligned}\text{distr}(\vee_{g\in B} T_g Q' \mid F) &= \frac{1}{|B|}\sum_{g'\in B} \text{distr}(\vee_{g\in B} T_g Q' \mid T_{g'} E_0)\\ &= \frac{1}{|B|}\sum_{g'\in B} \text{distr}(\vee_{g\in B-g'} T_g Q' \mid E_0)\\ &= \frac{1}{|B|}\sum_{g'\in B} \lambda_{B-g'}.\end{aligned}$$

If $B = \Gamma$ or $B = \sigma$ for some $\sigma \in \tau_k$, then λ_B is product measure by (3.6). Therefore, in both cases we obtain

$$\text{distr}(\vee_{g\in B} T_g Q') = \pi_B.$$

Moreover, if $B = \sigma$ for some $\sigma \in \tau_{k+1}$, then $\lambda_\sigma = \rho_\sigma$. Thus, by shift-invariance,

$$\text{distr}(\vee_{g\in\sigma} T_g Q' \mid F) \equiv \rho_\sigma.$$

Since

$$|\sigma| m(E_0) = (k+1)\alpha m(E) = \sqrt{\delta} = \beta,$$

we have, by definition of ρ_σ,

$$\text{distr}(\vee_{g\in B} T_g Q') = (1-\beta)\mu_\sigma + \beta\rho_\sigma = \pi_\sigma.$$

Therefore Q' satisfies properties (a′) and (b′) of the lemma.

Let $\Gamma^* E_0$ denote $\cup_{g\in\Gamma^*} T_g E_0$. If B is an atom of $\Phi(Q')$ (recall Definition 3.2), then $B \supset A \cap (\Gamma^* E_0)^c$. Since $\Gamma^* E_0 \perp \Phi_1(Q)$,

$$\begin{aligned}\min\{m(B) : B \text{ atom of } \Phi(Q')\} &\geq (1 - |\Gamma^*| m(E_0)) \min\{m(A) : A \text{ atom of } \Phi(Q)\}\\ &= \left(1 - |\Gamma^*|\frac{\sqrt{\delta}}{(k+1)}\right)\eta\\ &= \eta',\end{aligned}$$

proving (c′).

Lastly, to see (d′), let $\beta = |\Gamma_1^*| m(E_0)$. Then, as in (3.10),

$$\text{distr}(\vee_{g\in\Gamma_1^*} T_g Q') = (1-\beta)\mu_{\Gamma_1^*} + \beta\gamma$$

for some measure γ.

Let A, B, C be atoms of $\Phi(Q')$. There exists elements a, b, c in Ω^{Γ^*} corresponding to A, B, C respectively, and an element in $\Omega^{\Gamma_1^*}$, that for simplicity we denote by $a \vee b \vee c$, corresponding to $T_{-h}A \cap B \cap T_h C$. Then

$$\begin{aligned}&|m(T_{-h}A \cap B \cap T_h C) - m(A)m(B)m(C)|\\ &\quad= (1-\beta)|\mu(a\vee b\vee c) - \mu_{\Gamma^*}(a)\mu_{\Gamma^*}(b)\mu_{\Gamma^*}(c)|\\ &\qquad+ \beta|\gamma(a\vee b\vee c) - \mu_{\Gamma^*}(a)\mu_{\Gamma^*}(b)\mu_{\Gamma^*}(c)|\\ &\quad\leq (1-\beta)\delta m(A)m(B)m(C) + 2\beta.\end{aligned}$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Since $\beta = 3|\Gamma^*|\sqrt{\delta}/(k+1) < 3|\Gamma^*|\sqrt{\delta}$ and η' is smaller that the size of the atoms of $\Phi(Q')$,

$$|m(T_{-h}A \cap B \cap T_h C) - m(A)m(B)m(C)| \leq \left(\delta + \frac{6|\Gamma^*|\sqrt{\delta}}{\eta^3}\right) m(A)m(B)m(C)$$
$$= \delta' m(A)m(B)m(C)$$

for all A, B, C atoms of $\Phi(Q')$.

This finishes the proof of Lemma 3.3.3. □

4. *Krengel's theorem for general ergodic systems*

A probability space (X, β, m) together with a measure preserving transformation on it is called a *measure preserving system* (m.p.s.) or *dynamical system*. A m.p.s. $(Y, \mathcal{G}, \nu, S)$ is a *factor* of (X, β, m, T) if there is a measurable map $\varphi : X \mapsto Y$ which is measure preserving (i.e. $A \in \mathcal{G} \Rightarrow \varphi^{-1}A \in \beta$ and $m(\varphi^{-1}A) = \nu(A)$); and $S\varphi = \varphi T m$-a.e. In this case, φ is called the *factor map*. For short, we will denote (X, β, m, T) by X and $(Y, \mathcal{G}, \nu, S)$ by Y.

Given a factor Y of X with factor map φ, $A \in \beta$ is called a *cylinder set* if $A = \varphi^{-1}(C)$ for some measurable set $C \in Y$. A factor Y is called *measurable* if there exists a countable collection I of cylinder sets that separates points in Y, i.e. for all $y, y' \in Y$ there exists $A \in I$ such that either $\varphi^{-1}(y) \in A$ and $\varphi^{-1}(y') \notin A$ or $\varphi^{-1}(y) \notin A$ and $\varphi^{-1}(y') \in A$.

We will further assume that (X, β) is a Lebesgue space and $(Y, \mathcal{G})$ is a measurable factor. Under these hypotheses, Y is a Lebesgue space and the σ-algebra $\mathcal{G}$ has the property that not only the inverse image of every measurable set is measurable, but also the image of every measurable set is measurable, i.e. $\varphi(B) \in \mathcal{G}$ for all $B \in \beta$. We refer the reader to Rohlin [**9**] for a treatment of Lebesgue spaces and their factors.

These regularity conditions make it possible top disintegrate the measure m over the factor. Moreover, if the system X is ergodic, then (Rohlin [**9**]) X is isomorphic to a skew product of Y with some space $(Z, \mathcal{E}, w)$.

Suppose the m.p.s. X is ergodic but not weakly mixing. Then, the system has a non-trivial eigenfunction, which defines a rotation factor of X which is either isomorphic to (S^1, Borel sets, λ, R_α), (where λ is the normalized Lebesgue measure and R_α is the rotation by α on S^1) or to ($\mathbb{Z}_n$, all sets, ν, R_α) (where ν= counting measure and R_α is addition by some $0 < p < n$ $(n : p) = 1$) according to whether α is a root of unity or not,

The *Kronecker factor*, $\mathrm{K} = (K, \mathcal{G}, \lambda, R_g)$, is the maximal group rotation factor of the m.p.s. system. It has the property that one can associate to it a sub-sigma algebra $\mathcal{G}$ of β such that K is identified with $(X, \mathcal{G}, m, T)$ and $L^2(X, \mathcal{G}, m)$ is the subspace generated by the eigenfunctions of T. We will denote it simply by $L^2(K)$.

LEMMA 4.1. *For $f \in L^2(K)^\perp$, $g \in L^2(X, \beta, m)$ we have*

$$1/N \sum_{n=0}^{N-1} |\langle T^n f, g\rangle| \to 0.$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Proof. Since $f \in L^2(K)^\perp$ the restriction of T to the cyclic subspace generated by f is a unitary operator with continuous spectrum. The lemma then follows by the standard spectral argument (see for example [**7**, Theorem 3.4]). □

From now on we assume that X is a Lebesgue space and any factor we consider is a measurable factor. Moreover, since the action on X commutes with that of Y, abusing the notation, we denote both of them by T.

Definition 4.2. We say that a set A is *independent of the factor* Y if $m_y(A) = m(A)$ for a.e. $y \in Y$. A finite partition P of X is independent of Y if each of its atoms is independent of the factor.

Note. $\lim_{n\to\infty}(1/N)\sum_{n=1}^{N}|a_n| = 0$ if and only if D-$\lim_{n\to\infty}|a_n| = 0$, where D-lim means the limit taken along a sequence of density one (Furstenberg [**5**]).

LEMMA 4.3. *Let A be a measurable set independent of the Kronecker factor. Then D-$\lim_{n\to\infty} m(T^nA \cap B) = m(A)m(B)$ for any $B \in \beta$.*

Proof. Let $f = \chi A - m(A)$. Since $A \perp$ K, it follows that $f \in H$, where H is the subspace in Lemma 4.1. Given $B \in \beta$, let $g = \chi_B$. Then $\lim_{N\to\infty}(1/N)\sum_{n=1}^{N}|\langle T^n f, g\rangle| = 0$ which implies D-$\lim_{N\to\infty}|\langle T^n f, g\rangle| = 0$, finishing the lemma. □

The proof of Theorem 12 is based on two inductive lemmas, which are the counterpart of Lemmas 2.2 and 2.3. In fact, the proof is essentially the same. However, one has to be more careful in this case because each time one modifies a partition independent of the Kronecker factor, one must make sure that the new partition is also independent of K. To this end, we need first a strong Rohlin lemma over the Kronecker factor. Because of technicalities, we do not exactly get a strong Rohlin lemma but an ϵ-strong one instead.

A factor Y of a m.p.s. X is non-atomic if the induced measure on Y is a non-atomic measure.

LEMMA 4.4. *Let X be an ergodic system and Y a non-atomic factor, with factor map φ. Given a partition P of X, $\epsilon > 0$ and h a positive integer, there exists a cylinder set B such that $\{T^jB\}_{j=0}^{h-1}$ is a Rohlin tower with $m(\cup_{j=0}^{h-1}T^jB) > 1-\epsilon$ and B is ϵ-independent of the partition P.*

Proof. First suppose that P has the form $P_1 \vee P_2$ where P_1 is Y-measurable and P_2 is independent of Y. By the strong Rohlin lemma in Y we obtain a Rohlin tower with base B independent of P_1 and hence automatically independent of $P_1 \vee P_2$. Since an arbitrary P can be approximated by partitions of the form $P_1 \vee P_2$ the lemma follows immediately. □

We assume now that X is an ergodic system such that the fibers over K are a.e. non-atomic. The notation $Q \perp$ K means that the partition Q is independent of the Kronecker factor.

LEMMA 4.5. *Let I be a set of positive integers and Q a finite partition such that $Q \perp K$. Given $\epsilon > 0$ there exists $\eta > 0$ such that if Q moves η-independently along I, then given any $\eta' > 0$ there is a partition $Q' \perp K$ such that $\|Q - Q'\| < \epsilon$ and Q' moves η'-independently along I.*

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

LEMMA 4.6. *Let I be a finite set of positive integers and Q a finite partition with $Q \perp K$. Given $\epsilon > 0$ there exists $\eta > 0$ such that if Q moves η-independently along I, then, there is a partition $Q' \perp K$ such that $\|Q - Q'\| < \epsilon$ and Q' moves η'-independently along I.*

Proof. Iterate Lemma 4.5. □

Proof of Lemma 4.5. This is the analogous version of Lemma 2.2 and its proof follows the same ideas. However, in this case one should pay closer attention to the construction of the partition Q' to ensure that it will be independent of the Kronecker factor.

Assume Q is a partition with r atoms, say $Q = \{A_1, \ldots, A_r\}$. Let $\Omega = \{1, \ldots, r\}$ be the symbol space and p the probability on Ω defined by the distribution of Q, $p_i = m(A_i)$, $1 \leq i \leq r$. Let $n = \max\{i : i \in I\}$. Let $h \gg n$ an integer to be determined and μ the measure on $\Omega^{\{0,1,\ldots,h\}}$ defined by $\text{distr}(\vee_{j=0}^{h-1} T^j Q)$. For any subset J of the interval $\{0, 1, \ldots, h\}$, let μ_J denote the projection of μ onto Ω^J and π_J the product measure on Ω^J with one-dimensional marginals given by p.

If Q moves η-independently along I, then μ_I is η-close to product measure. Thus, by Lemma 2.4, there exists an η such that if Q moves η-independently along I, we can construct a shift-invariant probability measure λ on $\Omega^{\{0,1,\ldots,h\}}$ such that

$$\pi_J = (1 - \sqrt{\eta})\mu_J + \sqrt{\eta}\lambda_J$$

for all $J \subset I$. With the measure λ and the painting names on towers technique, we can construct the new partition. We have two cases: K is non-atomic or purely atomic.

Set $H = h + 2n$ and let Φ be the partition $\Phi = \vee_{j=0}^{H-1} T^{-j}(\vee_{i \in I} T^i Q)$. α is a quantity to be chosen appropriate for the construction of Q' and depends on η, h and I. Moreover, choosing η small enough guarantees that $\alpha < 1$ (see Lemma 2.2).

Case 1. K is non-atomic. By Lemma 4.4, given δ arbitrarily small, we can find an R-tower with base B and height H such that B is a cylinder set, $m(\cup_{j=0}^{H-1} T^j B) > 1 - \delta$ and $B \perp^\delta \Phi$ (hence, $T^n B \perp^\delta \vee_{i \in I} T^i q$). Since the factor is non-atomic we can find a subset $B_0 \subset B$ such that B_0 is also a cylinder set, $B_0 \perp^\delta \Phi$ and $m(B_0) = \alpha m(B)$. Since B_0 is a cylinder set, $B_0 = \varphi^{-1}(E_0)$ where $E_0 \subset Y$ is the base of an R-tower on Y of height H.

We use the measure λ to find a partition $R = \{C_1, \ldots, C_r\}$ of the tower $\{T^j B_0\}_{j=n}^{h+n-1}$, such that

$$m_y(\cap_{u=0}^{h-1} T^{-u} C_{i_u} \cap T^n B_0) = \lambda(i_0, i_1, \ldots, i_{h-1}) \tag{4.1}$$

for a.e. $y \in T^n E_0$. Since the fibers are a.e. non-atomic, this construction can be carried out by dividing $T^n B_0$ into pieces $\{W_\gamma\}_{\gamma \in \Omega^{\{0,1,\ldots,h-1\}}}$, such that $m_y(W_\gamma) = \lambda(\gamma)$ for a.e. $y \in K$ (Rohlin [**8**]) and using the painting names on towers technique.

Now let $D = \cup_{j=n}^{h+n-1} T^j B_0$ and define $Q' = \{A'_1, \ldots, A'_r\}$ by letting $A'_i = (A_i \setminus D) \cup C_i$, $1 \leq i \leq r$. Since the one-dimensional marginals of λ have the same distribution as Q, Q' is also independent of the Kronecker factor. Indeed, $\varphi(D) = \cup_{j=n}^{h+n-1} T^j E_0$. Then, for a.e. $y \in \varphi(D)^c$,

$$m_y(A'_i) = m_y(A_i) = p_i \quad \text{because} \quad A'_i \cap D^c = A_i \cap D^c.$$

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

Since $A_i' \cap D = C_i \cap D$, for a.e. $y \in T^{j+n}E_0$, we have by (4.1),

$$m_y(A_i) = m_y(C_i) = m_z(T^{-j}C_i) = \lambda_j(i) = p_i$$

where $T^j(z) = y$, $z \in T^n E_0$ and λ_j denotes the projection of λ into its jth coordinate.

Case 2. K is atomic. In this case since K is ergodic it is simply a rotation on a finite number m of points. Take h such that $H = h + n$ is a multiple of m. Let F denote the partition of X defined by the fibers over K, i.e.

$$F = \{\varphi^{-1}\{i\} : i : 1, \ldots, m\}.$$

Construct an R-tower with base $B \perp (\Phi \vee F)$ and $m(\cup_{j=0}^{H-1} T^j B) > 1 - \delta$. Then select $B_0 \subset B$ such that $\{B_0, B_0^c\} \perp (\Phi \vee F)$ and $m(B_0) = \alpha m(B)$. The rest of the argument follows as in Case 1 but the estimation of the distribution of $\vee_{j\in I} T^j Q'$ presents a slight problem because the Rohlin tower is not independent of the partition ϕ.

Following the notation of Lemma 2.2, let

$$F_1 = \cup_{j=0}^{h+2n-1} T^j(B_0) \quad \text{and} \quad F_0 = \cup_{j=2n}^{h-1} T^j(B_0).$$

Then

$$\begin{aligned}\operatorname{distr}(\vee_{j\in I} T^j Q') = {} & m(F_1^c)\operatorname{distr}(\vee_{j\in I} T^j Q' \mid F_1^c) \\ & + m(F_0)\operatorname{distr}(\vee_{j\in I} T^j Q' \mid F_0) \\ & + m(F_1 \backslash F_0)\operatorname{distr}(\vee_{j\in I} T^j Q' \mid F_1 \backslash F_0)).\end{aligned}$$

Estimating the distributions as in Lemma 2.2, we have

$$\operatorname{distr}(\vee_{j\in I} T^j Q \mid F_1^c) = \tilde{\mu}_I, \quad \operatorname{distr}(\vee_{j\in I} T^j Q' \mid F_0) = \lambda_I$$
$$\text{and} \quad \operatorname{distr}(\vee_{j\in I} T^j Q' \mid (F_1 \backslash F_0)) = \nu,$$

where the last measure ν is unknown (but the measure of $F_1 \backslash F_0$ can be made very small), and since $F_1^c \perp^{\delta} \vee_{j\in I} T^j Q$,

$$\|\tilde{\mu}_I - \mu_I\| < \delta \pi_I.$$

Then, as in Lemma 2.2, since we can choose δ arbitrarily small and h large enough we can make $\operatorname{distr}(\vee_{j\in I} T^j Q')$ arbitrarily close to product measure. □

Proof of Theorem 12. The sequence $\{n_k\}_{k=1}^{\infty}$ is constructed inductively and together with it, a sequence of partitions $\{P_k\}_{k=1}^{\infty}$ is constructed such that $P_0 = P$ and for all $i \geq 1$:

(a) P_i is independent of the Kronecker factor;

(b) $\|P_i - P_{i-1}\| < \epsilon_i = \epsilon/2^i$; and

(c) P_i moves independently along Γ_i, where $\Gamma_i = \{0, n_1, \ldots, n_i\}$.

Suppose that $n_1, \ldots, n_k$ and $P_1, \ldots, P_k$ have already been defined. Let $\delta = \delta(k)$ to be determined and choose $n_{k+1} > n_k$ such that

$$\vee_{j=0}^{k} T^{n_j} P_k \perp^{\delta} T^{n_{k+1}} P_k.$$

This condition is satisfied by applying Lemma 4.3 to each atom of P_k.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press

By Lemma 4.6, there exists $\eta = \eta(I)$ such that if Q is a partition independent of K that moves η-independently along I, there is a partition P_{k+1} independent of K such that $\|P_{k+1} - Q\| < \epsilon_{k+1}/2$ and P_{k+1} moves independently along I. By Lemma 4.5, there exists δ such that if P_k moves δ-independently along I, there is a partition Q independent of K such that $\|P_k - Q\| < \epsilon_{k+1}/2$ and Q moves η-independently along I. Theorem 12 would be finished if we could show that δ was independent of n_{k+1}. By a straight application of Lemma 4.5, δ depends on the set I. However, since P_k already moves independently along I, δ is independent of n_{k+1} (see proof of Theorem 5). □

REFERENCES

[1] V. Bergelson, I. Cornfeld and B. Mityagin. Unitary $\mathbb{Z}^d$ actions with continuous spectrum. *Proc. Amer. Math. Soc.* **119** (1993), 1127–1134.

[2] V. Bergelson and J. Rosenblatt. Joint ergodicity for group actions. *Ergod. Th. & Dynam. Sys.* **8** (1988), 351–365.

[3] V. Bergelson and D. Rudolph. Weakly mixing actions of F^∞ have infinite subgroup actions which are Bernoulli. *Dynamical Systems (Proc. Univ. Maryland, 1986–87) (Lecture Notes in Mathematics, 1342)*. Ed. Alexander. Springer, Berlin, 1987, pp. 7–22.

[4] J. P. Conze. Entropie d'un groupe abélien de transformations. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **25** (1973), 103–108.

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

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

[7] U. Krengel. *Ergodic Theorems.* de Gruyter, Berlin, 1985.

[8] V. A. Rohlin. On the fundamental ideas of measure theory. *Math. Sb. (N.S.)* 25 (67) **1** (1949), 107–150. Engl. transl. *Amer. Math. Soc. Translations, Ser. 1*, **10** (1962), 1–54.

[9] V. A. Rohlin. Selected topics from the metric theory of dynamical systems. *Amer. Math. Soc. Translations, Ser. 2*, **49** (1966), 171–240.

[10] W. Rudin. *Functional Analysis*. McGraw-Hill, 1973.

[11] D. Rudolph. *Fundamentals of Measurable Dynamics, Ergodic Theory on Lebesgue Spaces*. Clarendon Press, Oxford, 1990.

[12] P. Shields. *The Theory of Bernoulli Shifts*. University of Chicago Press, Chicago, 1973.

https://doi.org/10.1017/S0143385799126580 Published online by Cambridge University Press