Ergod. Th. & Dynam. Sys. (2003), **23**, 1275–1304
DOI: 10.1017/S0143385702001657
© *2003 Cambridge University Press*
Printed in the United Kingdom

Multiple mixing and local rank group actions

REEM YASSAWI

Department of Mathematics, Lady Eaton College, Trent University, Peterborough, ON, Canada K9L 1Z6
(e-mail: ryassawi@trentu.ca)

(*Received* 14 *December* 2001 *and accepted in revised form* 17 *December* 2002)

Abstract. Let $r \geq 1$; the Følner sequences $\{F^{n,1}\}_{n=1}^{\infty}, \{F^{n,2}\}_{n=1}^{\infty}, \ldots, \{F^{n,r}\}_{n=1}^{\infty}$ satisfy the *bounded intersection property* if there is a constant p such that, for any $n \in \mathbb{N}$ and $1 \leq i \leq r$, each $F^{n,i}$ can intersect no more than p disjoint translates of $F^{n,1}, F^{n,2}, \ldots, F^{n,r}$. They have *comparable magnitudes* if $0 < \underline{\lim}_n |F^{n,i}|/|F^{n,j}| < \infty$ for $1 \leq i, j \leq r$. Suppose that G is a countable Abelian group with an element of infinite order, and let $\mathcal{X}$ be a mixing finite (or β-local, with $\beta > 1/2$) rank action of G on a probability space. Suppose that the Følner sequences of $\mathcal{X}$ satisfy the bounded intersection property, and have comparable magnitudes. Then $\mathcal{X}$ is mixing of all orders. We follow Ryzhikov's joining technique in our proof: the main theorem follows from showing that any pairwise independent joining of k copies of $\mathcal{X}$ is necessarily product measure.

1. *Introduction*

A mixing group action $(X, \mathcal{B}, \mu, G)$ is not generally multiply mixing [**9**]; however, if any pairwise independent joining of $(X, \mathcal{B}, \mu, G)$ is actually independent, then mixing implies multiple mixing. In this paper we show that this condition holds for a large class of finite rank and local rank group actions. Kalikow [**8**] showed that rank one mixing transformations were 3-mixing, and Host [**4**] proved that mixing transformations with singular spectrum are mixing of all orders. Ryzhikov [**12, 13**] showed that finite rank and β-local ($\beta > 1/2$) mixing transformations are mixing of all orders by showing that pairwise independent self-joinings of the given system are necessarily product measure. Here we generalize Ryzhikov's result to certain finite rank and β-local rank group actions.

2. *Preliminaries*

Throughout this paper G will denote a countable Abelian group. Let $\mathcal{X} = (X, \mathcal{B}, \mu, G)$ and $\mathcal{Y} = (Y, \mathcal{F}, \nu, G)$ be finite measure-preserving G-actions. To each element $g \in G$ there corresponds a measure-preserving transformation $T_g : X \to X$; however, we will mostly use g to denote the element of the group, the measure-preserving transformation it

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

represents, and the unitary operator induced by g on $\mathcal{L}_2(X)$. A sequence $\{F^n\}_{n=1}^{\infty}$ of finite subsets of G is *Følner* if, for all $g \in G$,

$$\lim_{n\to\infty} \frac{|gF^n \triangle F^n|}{|F^n|} = 0.$$

For $i = 1, \ldots, r$, let $\{F^{n,i}\}_{n=1}^{\infty} = F^{1,i}, F^{2,i}, F^{3,i}, \ldots$ be a Følner sequence; we will write $\{F^{n,i}\}$ instead of $\{F^{n,i}\}_{n=1}^{\infty}$. We say that the r Følner sequences $\{F^{n,1}\}, \{F^{n,2}\}, \ldots, \{F^{n,r}\}$ have the bounded intersection property if there exists $p \in \mathbb{N}$ such that, for each $n \in \mathbb{N}$ and $1 \leq j \leq r$, at most p disjoint translates of $F^{n,1}, F^{n,2}, \ldots, F^{n,r}$ can non-trivially intersect $F^{n,j}$. We will assume that the Følner sequence(s) defining our action $(X, \mathcal{B}, \mu, G)$ have the bounded intersection property. We will call such a p an *intersection bound* for the Følner sequences $\{F^{n,1}\}, \{F^{n,2}\}, \ldots, \{F^{n,r}\}$. Examples of groups which have natural Følner sequences satisfying this condition are $\mathbb{Z}^n$, countable direct sums of finite cyclic groups, and direct sums of these two cases.

Furthermore, we will have to assume that the Følner sequences have *comparable magnitude*: i.e. that

$$0 < \underline{\lim}_n \frac{|F^{n,i}|}{|F^{n,j}|} < \infty \tag{1}$$

for $1 \leq i, j \leq r$. Henceforth we will assume that the Følner sequences we are dealing with have the bounded intersection property, and also that (1) is true.

The Følner sequence $\{F^n\}$ satisfies the *Tempel'man condition* if $F^n \subset F^{n+1}$ for each n and there exists $K \in \mathbb{N}$ such that $|F^n(F^n)^{-1}| \leq K|F^n|$ for all $n \in \mathbb{N}$. We mention the Tempel'man condition [**14**] only because it seems to be closely related to the bounded intersection property in that, in all the standard examples we know of, they stand or fall together. However, we do not know if it is possible to derive one from the other.

It is not difficult to find Følner sequences $\{F^n\}$ in $\mathbb{Z}$ which do not have the bounded intersection property. We give an example of a sequence $\{F^n\}$ such that n disjoint translates of F^n, say $F^n+t_1, \ldots, F^n+t_n$, intersect F^n non-trivially. F^n will consist of $n+1$ intervals, all but the last having cardinality n. These intervals will be separated by gaps of length β_i where

$$\beta_1 \gg n^2 \tag{2}$$

and

$$\beta_k > \sum_{i=1}^{k-1} \beta_i. \tag{3}$$

Explicitly, we set $\beta_0 = 0$ and define

$$F^n := \bigcup_{k=0}^{n-1} \left[kn + \sum_{i=0}^{k} \beta_i, kn + \sum_{i=0}^{k} \beta_i + (n-1) \right] \bigcup \left[n^2 + \sum_{i=0}^{n} \beta_i, 2n^2 + \sum_{i=0}^{n} \beta_i \right].$$

Finally, for $k = 1, \ldots, n$ define

$$t_k := n^2 + \sum_{i=k}^{n} \beta_i.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

It can be verified that

$$(F^n + t_k) \cap F^n = \left[(n+k)n + \sum_{i=0}^{n} \beta_i, (n+k)n + \sum_{i=0}^{n} \beta_i + (n-1)\right].$$

Using (2) and (3), one can show that $(F^n + t_j) \cap (F^n + t_k) = \emptyset$ for $1 \leq j \neq k \leq n$.

A group action $(X, \mathcal{B}, \mu, G)$ is *rank* $\leq r$ if there exist r Følner sequences $\{F^{n,i}\}$, where $1 \leq i \leq r$, each indexing a measurable tower

$$X^{n,i} := \bigcup_{g \in F^{n,i}} X_g^{n,i}$$

(if $g \in F^{n,i}$, then $X_g^{n,i}$ is a level of the tower $X^{n,i}$) satisfying: for the sequence of measurable partitions

$$\mathcal{P}_n := \{X_g^{n,i} : g \in F^{n,i}, i = 1, \ldots, r; X \backslash X^n\}$$

of X, where $X^n := \bigcup_{i=1}^{r} X^{n,i}$, we have:

(1) $\mu(X^n) \to 1$;

(2) $hX_g^{n,i} = X_{hg}^{n,i}$ whenever $g \in F^{n,i} \cap h^{-1}F^{n,i}$;

(3) for each measurable set A, and each $n \in \mathbb{N}$ and $1 \leq i \leq r$, there exists a set $A^{n,i}$ which is a union of elements of $\{X_g^{n,i}\}_{g \in F^{n,i}}$ and $\mu((A \cap X^{n,i}) \triangle A^{n,i}) \to_n 0$ for each $i = 1, 2, \ldots, r$.

We let $a^{n,i}$ denote the (common) mass of a level in the ith tower, i.e. $\mu(X_g^{n,i}) = a^{n,i}$ for all $g \in F^{n,i}$.

Property 3 says that the partitions $\mathcal{P}_n$ converge to $\mathcal{B}$ and to denote this we write $\mathcal{P}_n \to_n \epsilon$. We will call $\{F^{n,i}\}_{i=1}^r$ the Følner sequences associated with the rank (at most) r group action $\mathcal{X}$ and write $(\mathcal{X}, \{F^{n,i}\}_{i=1}^r)$ for $\mathcal{X}$ when we wish to specify $\{F^{n,i}\}_{i=1}^r$. We will say that $(\mathcal{X}, \{F^{n,i}\}_{i=1}^r)$ has the bounded intersection property if $\{F^{n,i}\}_{i=1}^r$ have the bounded intersection property. For a general scheme for constructing finite rank group actions, one can generalize the method used in [**5**] to construct rank one group actions. When $G = \mathbb{Z}$, what we call a finite rank action is in fact a generalization of the usual finite rank transformation (which requires that the $F^{n,i}$ be intervals in $\mathbb{Z}$). Ferenczi [**2**] constructed a rank one $\mathbb{Z}$ action which is not loosely Bernoulli [**11**], and therefore not of finite rank. He calls this a *funny* rank one transformation. Ferenczi's example in fact has the bounded intersection property with $p = 4$. We say that the action $(X, \mathcal{B}, \mu, G)$ is *rank* r if the action is rank $\leq r$ and the action is not rank j for $1 \leq j < r$.

Let $0 < \beta \leq 1$. A group action $(X, \mathcal{B}, \mu, G)$ has *local rank* β (or is β-*local*) if there exist a Følner sequence $\{F^n\}$ and a sequence of measurable partitions

$$\mathcal{P}_n := \{X_g^n : g \in F^n; X \backslash X^n\}$$

of X (where $X^n := \bigcup_{g \in F^n} X_g^n$), such that:

(1) $\mu(X^n) \to \beta$;

(2) $hX_g^n = X_{hg}^n$ whenever $g \in F^n \cap h^{-1}F^n$;

(3) for each measurable set A and each $n \in \mathbb{N}$, there exists a set A^n which is a union of elements of $\{X_g^n\}_{g \in F^n}$ and $\mu((A \cap X^n) \triangle A^n) \to 0$.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

In general, if $\{X^n\}$ is a sequence of towers satisfying Property 3, we will say that it is an *approximating* sequence of towers. We will call the Følner sequence used above in the definition of a β-local group action the Følner sequence associated with the β-local group action $\mathcal{X}$ (and sometimes say that $\{F^n\}$ *generates* $X^n = \bigcup_{f\in F^n} X^n_f$) and write $(\mathcal{X}, \{F^n\})$ for the β-local group action. The measure of a level in the nth tower will be denoted by a_n. Ergodic finite rank actions are examples of β-local actions—this can be seen by applying Lemma 1.

If $\{g_j\}$ is a sequence in G we write $g_j \to \infty$ if, whenever $V \subset G$ is finite, then only finitely many of the g_j's belong to V. A group action $\mathcal{X}$ is *2-mixing* if

$$\lim_{j\to\infty} \mu(A_1 \cap g_j A_2) = \mu(A_1)\mu(A_2)$$

for all $A_1, A_2 \in \mathcal{B}$ and for each sequence $g_j \to \infty$. $\mathcal{X}$ is *k-mixing* if

$$\lim_{j\to\infty} \mu(A_1 \cap g_j^1 A_1 \cap g_j^2 A_2 \cap \cdots \cap g_j^{k-1} A_k) = \mu(A_1)\cdots\mu(A_k)$$

where $\lim_{j\to\infty} g_j^i = \infty$ for $i = 1, 2, \ldots, k-1$ and also $\lim_{j\to\infty}(g_j^i)^{-1} g_j^l = \infty$ whenever $i \neq l$. $\mathcal{X}$ is *mixing of all orders* if it is k-mixing for each $k \geq 2$.

Our main theorem is as follows.

THEOREM 1. *Suppose that G is Abelian, countable, and has an element of infinite order. Let $(X, \mathcal{B}, \mu, G)$ be a mixing finite rank action, or a mixing β-local action, with $\beta > 1/2$. In either case, assume that the Følner sequence(s) indexing the tower(s) of the action have the bounded intersection property, and have comparable magnitudes. Then $(X, \mathcal{B}, \mu, G)$ is mixing of all orders.*

To prove Theorem 1, we use the method of joinings, generalizing [**12**].

We say that $(X, \mathcal{B})$ is *regular* if X is compact and $\mathcal{B}$ is the Borel σ-algebra. This implies that $C(X) = \{f : X \to \mathbb{R}, f \text{ is continuous}\}$ is separable and also that any probability measure on $(X, \mathcal{B})$ is regular. In this section $\mathcal{X}$ is assumed regular. A probability measure λ is a *2-self-joining* of $\mathcal{X}$ if $(X \times X, \mathcal{B} \otimes \mathcal{B}, \lambda, G)$ is a measure-preserving group action with the additional conditions that

$$\lambda(A \times X) = \mu(A)$$

and

$$\lambda(X \times A) = \mu(A)$$

for $A \in \mathcal{B}$. We will use the word *coupling* when the measure λ projects onto μ without necessarily preserving the group action. For a detailed account of joinings, see [**6**]. Note that, by regularity, λ is a 2-self-joining if and only if $\lambda(f \otimes 1) = \mu(f)$ and $\lambda(1 \otimes f) = \mu(f)$ for all $f \in C(X)$ (where $f \otimes g(x, y) := f(x)g(y)$).

If λ is a probability measure on $\left(\prod_{i=1}^n X_i, \otimes_{i=1}^n \mathcal{B}_i\right)$ and $J \subset \{1, 2, \ldots, n\}$, then we define $\pi_J\lambda$ to be the projection of λ on $\left(\prod_{j\in J} X_j, \otimes_{j\in J} \mathcal{B}_j\right)$. We use the notation $X^{(n)} := X \times X \times \cdots \times X$ to denote the Cartesian product of n copies of X, and similarly $\mathcal{B}^{(n)}$ for $\mathcal{B} \otimes \cdots \otimes \mathcal{B}$. λ is an *n-self-joining* of $(X, \mathcal{B}, \mu, G)$ if $(X^{(n)}, \mathcal{B}^{(n)}, G, \lambda)$ is a measure-preserving action and $\pi_k\lambda = \mu$ for $k = 1, \ldots, n$. We will often drop the word

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

'self'. For $n > 2$ it is natural to impose stronger conditions on λ: in particular, if λ is an n-self-joining of $(X, \mathcal{B}, \mu, G)$, then we can require that

$$\pi_J \lambda = \mu^k$$

whenever $J \subset \{1, 2, \ldots, n\}$ and $|J| = k$. In this case we write $\lambda \in M(k, n)$ and say that λ is *k-fold independent* (or *pairwise independent* in the case $k = 2$).

Note that $M(k, n) \subset M(1, n)$ for all $k \geq 1$, and that $M(1, n) \subset M(X^{(n)})$, the set of probability measures on $X^{(n)}$. We will work with the weak-$*$ topology on $M(X^{(n)})$. With this (metrizable) topology, $\mu_k \to \mu$ iff $\mu_k(g) \to \mu(g)$ for $g \in C(X^{(n)})$. As the linear span of the family $\{\otimes_{i=1}^n f_i : f_i \in C(X)\}$ is dense in $C(X^{(n)})$, it is sufficient to check convergence on this family. In fact, if $\{\lambda_k\} \subset M(1, n)$, then $\lambda_k \to \lambda$ iff $\lambda_k\left(\prod_{i=1}^n A_i\right) \to_k \lambda\left(\prod_{i=1}^n A_i\right)$ for all measurable rectangles $\left(\prod_{i=1}^n A_i\right)$. We will prove necessity: suppose that $\lambda_n \to \lambda$, and let $\left(\prod_{i=1}^n A_i\right)$ be measurable. Note that λ is a coupling. By regularity there exist compact K_i and open U_i where $K_i \subset A_i \subset U_i$ and $\mu_i(U_i \backslash K_i) < \epsilon/n$ for $i = 1, \ldots, n$. We can then find a continuous function f which is identically 1 on $\prod_{i=1}^n K_i$ and 0 outside $\prod_{i=1}^n U_i$. Thus for each k

$$\left|\lambda_k\left(\prod_{i=1}^n A_i\right) - \lambda_k(f)\right| < \epsilon,$$

and the same is true for λ. A triangle inequality gives the result.

It now follows that, although the maps T_g are not continuous, the set $M(1, n)$ is a closed subset of $M(X^{(n)})$ in the weak-$*$ topology, and hence it is compact.

Theorem 1 will follow from Theorem 2.

THEOREM 2. *Suppose that G is Abelian, countable, and has an element of infinite order. Let $(X, \mathcal{B}, \mu, G)$ be a regular mixing finite rank action, or a regular mixing β-local action, with $\beta > 1/2$. In either case, assume that the Følner sequence(s) indexing the tower(s) of the action have the bounded intersection property, and have comparable magnitudes. Let $\nu \in M(2, k)$. Then $\nu = \mu^k$.*

In proving Theorem 2 we will restrict ourselves to the case $k = 3$. A similar argument shows that $M(k - 1, k) = \{\mu^k\}$ for $k > 2$, so Theorem 2 follows by induction.

Note that regularity is required in the statement of Theorem 2, but not Theorem 1. This is because regularity is essential for all our proofs concerning joinings, but not necessary for the proof of Theorem 1: the finite (or local) rank hypothesis on $\mathcal{X}$ ensures that the measure algebra $\bar{\mathcal{B}}$ of μ-equivalence classes of sets is separable. Standard arguments then show that $\bar{\mathcal{B}}$ is isomorphic to the measure algebra of a regular space and the action of G on $\bar{\mathcal{B}}$ transfers to an action on this regular measure algebra, which can then be realized as a point action. This shows that there is no harm in assuming that $\mathcal{X}$ itself is regular. The concepts of mixing and finite rank are obviously invariant under isomorphism of actions at the level of measure algebras. Throughout the remainder of the paper, except in the proof of Theorem 1, we will assume that the underlying measure space is regular and convergence of measures will always mean weak-$*$ convergence.

Let $(X, \mathcal{B}, \mu, G)$ be mixing and finite rank, and let $\nu \in M(2, 3)$ be ergodic. A sketch of the proof that $\nu = \mu^3$ is as follows. It would be ideal if two things were to occur

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

simultaneously: first, there is a sequence of products of some approximating towers

$$R^n = T^{n,1} \times T^{n,2} \times T^{n,3} \tag{4}$$

with $\underline{\lim}\ \nu(R^n) > 0$; and second that the 'cubes' in R^n do not individually contain too much ν-mass. Ryzhikov would call such cubes ν-light. This ensures that ν is to some extent 'smeared out' over substantially many of the cubes in R^n. The first property would mean that the measure ν_{R^n} 'the measure ν conditioned on R^n', would converge weak-* to ν (Lemma 1). The second property is needed to show that $\nu_{R^n} \to \mu^3$. In the rank one case (see [**7**] or [**12**]) finding R^n is not so hard. The $T^{n,i}$'s will not be the actual rank one towers, rather they will be subsets of the rank one towers. For example, if $G = \mathbb{Z}$, and the Følner sequence is a sequence of intervals: $F^n = [0, h_n)$, then R^n could be some subset of

$$\bigcup_{i=0}^{(1-\delta)h_n} X_i^n \times \bigcup_{i=0}^{(1-\delta)h_n} X_i^n \times \bigcup_{(1-\delta)h_n}^{h_n} X_i^n$$

(or its image under $T^{\delta h_n}$). The facts that ν is pairwise independent and that the rank one tower X^n is most of the space easily yield that

$$\underline{\lim}_n \nu\left(\bigcup_{i=0}^{(1-\delta)h_n} X_i^n \times \bigcup_{i=0}^{(1-\delta)h_n} X_i^n \times \bigcup_{(1-\delta)h_n}^{h_n} X_i^n \right) > 0.$$

Of course, we then have to do a bit of work to show that our chosen R^n has positive ν-mass. The fact that $(X, \mathcal{B}, \mu, G)$ is mixing helps ensure the second requirement (mixing means that no level in X^n can go 'heavily' into any other—Lemma 4), which in turn is needed to guarantee that a certain sequence of probability weights we end up averaging against is uniformly small. Mixing is used a second time to invoke the Blum–Hanson Theorem [**1**], which is where the uniform smallness of the aforementioned weights is needed.

If $(X, \mathcal{B}, \mu, G)$ is finite rank, the difficulty is to get a sequence of R^n's which have both positive ν-mass *and* whose 'cubes' are ν-light at the same time. If $(X, \mathcal{B}, \mu, G)$ has local rank, with towers X^n it is not even clear that $\overline{\lim}_n\ \nu(X^n \times X^n \times X^n) > 0$. In [**13**], Ryzhikov finds two sequences of sets $\{R^n\}$ and $\{S^n\}$ (as in the example above, $S^n \approx h_n R^n$, where $h_n \to \infty$ in G). He then finds a sequence of *induced joinings* $\in M(1, 3)$ satisfying both of the required properties for either R^n or S^n. These joinings are certain projections of relative products built using the disintegration of ν over $(X, \mathcal{B}, \mu, G)$. The sequence of induced joinings $\{\eta_j\}$ is chosen so that, as $j \to \infty$, η_j has an increasingly large pairwise independent component, and we show that this component is mostly product measure, so that $\eta_j = c_j \mu^3$, where $c_j \uparrow 1$. Lemma 3 then shows that $\nu = \mu^3$.

The rest of the paper will proceed in the following order: first, we define induced joinings and show how an appropriate sequence of these joinings will give information about ν (Lemma 3). Next, we construct the sets $\{R^n\}$ and $\{S^n\}$, and then find the sequence of induced joinings (Proposition 3) which give $\{R^n\}$ and $\{S^n\}$ the 'right amount' of mass (Propositions 1 and 2). Finally, we will show that this sequence of measures has a large pairwise independent component (Lemma 6), work with these components, and show that they are mostly product measure. This will allow us (Lemma 3) to conclude that $\nu = \mu^3$.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

3. *Auxiliary results*

If A is a measurable subset of X and μ is a probability measure on X, we define the measure μ_A, the measure μ conditioned on the set A, as

$$\mu_A(F) = \frac{\mu(A \cap F)}{\mu(A)}.$$

If μ, μ_1 are probability measures and $0 < c \leq 1$, then the expression '$\mu \geq c\mu_1$' means that $\mu = c\mu_1 + (1-c)\mu_2$ for some probability measure μ_2.

The next simple lemma is used repeatedly in [**12**].

LEMMA 1. *Let $(X, \mathcal{B}, \nu, G)$ be an ergodic measure-preserving group action, and let I_n be a sequence of measurable sets such that:*

1. $\lim_n \nu(I_n) = c \neq 0$;
2. $\lim_n \nu(gI_n \triangle I_n) = 0$ *for each* $g \in G$.

Then $\lim_n \nu_{I_n} = \nu$.

Proof. Let $\nu_n := \nu_{I_n}$. We know (by passing to a subsequence if necessary) that the sequence $\{\nu_n\}$ converges weak-$*$ to some measure λ. Now $\nu_n = f_n\, d\nu$ where $f_n = \chi_{I_n}/\nu(I_n)$ and, by Property 1, there is a K such that $|f_n| \leq K$ for all large n. We claim that, if $\nu_n = f_n\, d\nu \to \lambda$ and $|f_n| \leq K$, then $\lambda = f\, d\nu$ for some f.

If not, then there is some L with $\lambda(L) \geq \epsilon$ and $\nu(L) = 0$. Find U open containing L such that $\nu(U) < \epsilon/2K$. Then $\nu_n(U) < \epsilon/2$ for all large n, and this is a contradiction since $\liminf_n \nu_n(U) \geq \lambda(U)$ for U open.

Thus, $\nu_n = f_n\, d\nu \to f\, d\nu$. Note that this implies that $\|f_n - f\|_1 \to 0$. Next, we claim that $\lambda \circ g^{-1} = \lambda$ for $g \in G$. To see this, note that

$$\left| \int_A f\, d\nu - \int_{g^{-1}A} f\, d\nu \right| = \left| \int_A (f - f \circ g)\, d\nu \right| \leq 2 \int |f - f_n|\, d\nu + \int_A |f_n - f_n \circ g|\, d\nu.$$

The first summand is small for large n. The second summand is also small since

$$\left| \int_A (f_n - f_n \circ g)\, d\nu \right| \leq \int |f_n - f_n \circ g|\, d\nu = \nu(I_n \Delta g^{-1} I_n) \to 0$$

as $n \to \infty$. Thus $\lambda \ll \nu$ and is G-invariant, and $\nu \geq c\lambda$. The ergodicity of ν forces $\lambda = \nu$. □

We will also need a version of the Blum–Hanson Theorem for group actions [**1**].

THEOREM 3. *Let $\mathcal{X}$ be a mixing action of a countable Abelian group G. Suppose that $\{a^n\}_{n\in\mathbb{N}}$ is a sequence of functions $a^n : G \to [0, \infty)$ satisfying:*

1. $\sum_{g\in G} a^n(g) = 1$, *for all* $n \in \mathbb{N}$;
2. $\lim_{n\to\infty} \sup_{g\in G} a^n(g) = 0$.

Then, for any $\phi \in \mathcal{L}_2(\mu)$,

$$\|A_n(\phi) - \langle \phi, 1 \rangle\|_2 \to 0,$$

where $A_n(\phi) := \sum_{g\in G} a^n(g)(\phi \circ g)$.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Proof. We may assume that $\phi \in \mathcal{L}_2(X)$ has 0 mean and unit norm. Since $\mathcal{X}$ is mixing then, given $\epsilon > 0$, we may choose a finite subset $\mathcal{O}_\epsilon \subset G$ such that $|\langle \phi \circ g, \phi \rangle| < \epsilon/2$ whenever $g \in G \backslash \mathcal{O}_\epsilon$. Next choose N large enough such that

$$\sup_{g \in G} a^n(g) < \frac{\epsilon}{2|\mathcal{O}_\epsilon|}$$

for all $n > N$. Then $\left\langle \sum_{g\in G} a^n(g)(\phi \circ g), \sum_{g\in G} a^n(g)(\phi \circ g) \right\rangle$ can be split up into two summands,

$$\sum_{g\in G} a^n(g) \sum_{\{h: gh^{-1} \in \mathcal{O}_\epsilon\}} a^n(h) \langle \phi \circ (gh^{-1}), \phi \rangle$$

and

$$\sum_{g\in G} a^n(g) \sum_{\{h: gh^{-1} \in G \backslash \mathcal{O}_\epsilon\}} a^n(h) \langle \phi \circ (gh^{-1}), \phi \rangle.$$

The first summand can be bounded by $\epsilon/2|\mathcal{O}_\epsilon| \sum_{g\in G} a^n(g) |\mathcal{O}_\epsilon|$ and the second by $\epsilon/2 \sum_{g\in G} a^n(g) \sum_{h\in G} a^n(h)$ and, using the fact that $\sum_{g\in G} a^n(g) = 1$, the result follows. □

We recall Ramsey's Theorem (see for example [**10**]): given $l \in \mathbb{N}$, there exists $r(l)$ such that if A is any subset of a countable set H of cardinality $r(l)$, with $*$ a 2-colouring (a binary relation) of A, then A contains a set B of cardinality l such that either $b * b'$ for all $b, b' \in B$ or $\neg(b * b')$ for all $b, b' \in B$.

The following lemma will be used in the proof of Proposition 3.

LEMMA 2. *Suppose that $F_1, F_2, \ldots, F_l \in \mathcal{L}_2(\mu)$, $0 \leq F_i \leq 1$ and $\int F_i \, d\mu = u$ for $i = 1, 2, \ldots, l$, with $l \geq 2$. Then there exist distinct i, k such that*

$$\langle F_i, F_k \rangle \geq \frac{u^2 - l^{-1}u}{1 - l^{-1}} \geq u^2 - \frac{1}{l-1}.$$

Proof. Note that, since $0 \leq F_i \leq 1$, then

$$\langle F_i, F_i \rangle \leq \langle F_i, 1 \rangle.$$

If $\phi := \sum_{i=1}^l F_i$ then, by the Cauchy–Schwarz inequality,

$$\langle \phi, \phi \rangle \geq \langle \phi, 1 \rangle^2 = (lu)^2.$$

Hence

$$\begin{aligned} &\max_{i \neq k} \langle F_i, F_k \rangle \\ &\quad \geq \frac{\langle \phi, \phi \rangle - \sum_{i=1}^l \langle F_i, F_i \rangle}{l^2 - l} \\ &\quad \geq \frac{l^2u^2 - lu}{l^2 - l} = \frac{u^2 - l^{-1}u}{1 - l^{-1}} \geq u^2 - \frac{1}{l-1}. \end{aligned}$$

□

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

3.1. *Induced joinings.* Let $\nu \in M(2,3)$ be ergodic for $(Y, \mathcal{F}, G) := (X^3, \mathcal{B}\otimes\mathcal{B}\otimes\mathcal{B}, G)$. Using the disintegration

$$\nu(A\times B\times C) = \int_C \int_B \nu_{x,y}(A)\, d\mu(y)\, d\mu(x) \tag{5}$$

of ν with respect to the factor (X^2, μ^2) (see [**3**]), we can define the family of operators $\{J_x\}_{x\in X}$, where $J_x : \mathcal{L}_2(X) \to \mathcal{L}_2(X)$ is defined as $J_x f(y) := \nu_{x,y}(f)$, so that (5) can also be written as

$$\nu(A\times B\times C) = \int_C \langle J_x \chi_A, \chi_B\rangle\, d\mu(x). \tag{6}$$

Note that

$$J_x(\chi_X) = J_x^*(\chi_X) = 1 \tag{7}$$

and

$$J_x(f) \geq 0 \quad \text{and} \quad J_x^*(f) \geq 0 \quad \text{whenever} \quad f \geq 0. \tag{8}$$

We have

$$\begin{aligned}
\nu(gA\times gB\times gC) &= \int_{gC} \langle J_x \chi_{gA}, \chi_{gB}\rangle\, d\mu(x)\\
&= \int_{gC} \langle J_x \chi_A \circ g^{-1}, \chi_B \circ g^{-1}\rangle\, d\mu(x)\\
&= \int_{gC} \langle g J_x g^{-1} \chi_A, \chi_B\rangle\, d\mu(x)\\
&= \int_{C} \langle g J_{g^{-1}x} g^{-1} \chi_A, \chi_B\rangle\, d\mu(x)
\end{aligned}$$

(where g and g^{-1} act unitarily on $\mathcal{L}_2(X)$) and, since ν is preserved by the action of G, then

$$\int_C \langle g J_{g^{-1}x} g^{-1} \chi_A, \chi_B\rangle\, d\mu(x) = \int_C \langle J_x \chi_A, \chi_B\rangle\, d\mu(x).$$

This is true for all measurable sets $A, B, C \in \mathcal{B}$. Hence

$$g J_{g^{-1}x} g^{-1} \equiv J_x \tag{9}$$

or

$$g J_x g^{-1} \equiv J_{gx}, \tag{10}$$

and this relation is clearly equivalent to the G-invariance of ν. Since $\nu \in M(2,3)$, in particular $\pi_{1,2}\nu = \mu$ and so

$$\int_X \langle J_x \chi_A, \chi_B\rangle\, d\mu(x) = \nu(A\times B\times X) = \mu(A)\mu(B) = \left\langle \int \chi_A, \chi_B \right\rangle$$

or, in the language of operators,

$$\int_X J_x\, d\mu(x) \equiv \int d\mu^2. \tag{11}$$

Conversely, if $\{J_x\}_{x\in X}$ is a measurable family of operators that satisfies (7), (8), (10) and (11), then the measure ν defined by (6) is a pairwise independent joining. If (11) is not satisfied then (7) still gives $\pi_{1,3}\nu = \pi_{2,3}\nu = \mu^2$.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Given $\nu \in M(2,3)$, let $\{J_x\}_{x\in X}$ be the family of operators defined by (6). If $k, l \in G$, then the *induced joining* $\eta_{k,l}$ is the measure defined by

$$\eta_{k,l}(A \times B \times C) := \int_C \langle J_{kx}\chi_A, J_{lx}\chi_B\rangle \, d\mu(x)$$
$$= \int_C \int_X \nu_{(kx,y)}(A)\nu_{(lx,y)}(B)\, d\mu(y)\, d\mu(x). \tag{12}$$

$\eta_{k,l}$ is the projection of a relative product: we will describe this when $l = e$, i.e. for the induced joinings $\eta_k := \eta_{k,e}$. If we label the ith copy of X in Y as X_i for $i = 1, 2, 3$, then we have two factors of Y, namely $\pi_{2,3}Y$ and $\pi_{1,3}Y$, and we identify these two factors using the transformation $\phi : \pi_{2,3}Y \to \pi_{1,3}Y$, which is defined as

$$\phi(x, y) := (kx, y).$$

ϕ induces a joining η'_k on X^4, the relative product over ϕ:

$$\eta'_k(A \times B \times C \times D) := \int_C \int_D \nu_{(kx,y)}(A)\nu_{(x,y)}(B)\, d\mu(y)\, d\mu(x)$$

and from this we see that $\eta_k = \pi_{1,2,3}\eta'_k$.

Note that $\eta_{k,l}$ is G-invariant; this follows using (10). Thus $(Y, \mathcal{F}, \eta_{k,l}, G)$ is a measure-preserving group action. We will only be using induced joinings $\eta_{k,g}$ where $k^{-1}g$ has infinite order.

As an example, we can look at the (well-known) case when ν is a non-trivial three-fold joining of the topological group $X = \{0,1\}^{\mathbb{Z}}$, with the Borel σ-algebra and Haar measure μ. Let $\phi : X^2 \to X^3$ be the map defined by $\phi((x, y)) = (x + y, x, y)$ and let $\nu := \mu^2 \circ \phi^{-1}$. Then ν is supported on the shift-invariant set

$$E = \{(x + y, y, x) : x, y \in X\}.$$

Here

$$J_x f(y) = \nu_{x,y}(f) = f(x + y).$$

Then $\eta_{k,l}$ is the measure

$$\eta_{k,l}(A \times B \times C) = \int_C \int_X \chi_A(\tau^l(x) + y)\chi_B(\tau^k(x) + y)\, d\mu(y)\, d\mu(x),$$

where τ is the shift. $\eta_{k,l}$ is supported on the set

$$\{(\tau^l(x) + y, \tau^k(x) + y, x, y) : x, y \in X\}.$$

In this case $\eta_{k,l}$ is ergodic, but not pairwise independent. So, we cannot assume that $\eta_{k,l} \in M(2,3)$ ($\pi_{1,3}\eta_{k,l}$ and $\pi_{2,3}\eta_{k,l}$ are product measure, but it is not clear that (11) holds); nor in general can we assume that $\eta_{k,l}$ is ergodic.

Let $\{c_j\}$ be a sequence of real numbers increasing to 1. While it is not clear that we can find an appropriate sequence of pairs of towers $\{R^n, S^n\}$ which are given positive mass by ν, we will be able to do this for a sequence of induced joinings $\{\eta_j\}$. Next we will show that this sequence has a large component which is pairwise independent and this will set the stage to prove that $\eta_j \geq c_j\mu^3$. This will imply that $\nu = \mu^3$: we prove this in the next lemma.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

LEMMA 3. *Let $c_j \uparrow_j 1$. Let $(X, \mathcal{B}, \mu, G)$ be a mixing group action, and ν be a pairwise independent 3-joining of this action. Suppose that, for some sequences $\{g_j\}$ and $\{k_j\}$ of group elements where $g_j k_j^{-1} \to \infty$, we have that*

$$\eta_{g_j,k_j} \geq c_j \mu^3$$

for each $j \in \mathbb{N}$, where η_{g_j,k_j} are the induced joinings defined by (12). Then $\nu = \mu^3$.

Proof. Let $\{J_x\}_{x\in X}$ be the family of operators defined by ν. Note that, if f, g, h are bounded measurable functions of $(X, \mathcal{B}, \mu)$, then

$$\mu^3(f \otimes g \otimes h) = \int f \left\langle \int g, h \right\rangle d\mu.$$

To show that $\nu = \mu^3$, it is then sufficient to show that $J_x = \int d\mu$ for μ-almost all x.

Let $\mathcal{F}$ be a countable basis of (real-valued) bounded functions for $\mathcal{L}_2^0(X, \mu)$ such that $\|f\|_2 = 1$ for each $f \in \mathcal{F}$. Fix an f in $\mathcal{F}$; we shall show that $J_x f = 0$ for μ-almost all $x \in X$. For $\epsilon > 0$ and $g \in \mathcal{L}_\infty(X)$, define

$$E_{\epsilon,g} := \left\{ x \in X : \|J_x f - g\|_2 < \frac{\epsilon}{1 + \|f\|_\infty + \|g\|_\infty} \right\}.$$

We claim that, if $\mu(E_{\epsilon,g}) > 0$, then $\|g\|_2 < \sqrt{\epsilon}$. By assumption on η_{g_j,k_j},

$$\eta_{g_j,k_j} - c_j \mu^3 = (1 - c_j)\rho,$$

where $\rho \in M(1, 3)$. Hence for each $h \in \mathcal{L}_\infty(X)$

$$|\eta_{k_j,g_j}(f \otimes f \otimes h) - c_j \mu^3(f \otimes f \otimes h)| \leq (1 - c_j)\|f\|_\infty^2 \|h\|_\infty.$$

Therefore

$$\left| \int h(x) \langle J_{g_j x}^* J_{k_j x} f, f \rangle \, d\mu(x) \right| = \left| \int h(x) \left(\langle J_{g_j x}^* J_{k_j x} f, f \rangle - c_j \left\langle \int f, f \right\rangle \right) d\mu(x) \right|$$
$$\leq (1 - c_j)\|f\|_\infty^2 \|h\|_\infty.$$

By taking now $h = \chi_{A_j^\pm(\delta)}$, where

$$A_j^\pm(\delta) := \{x : \pm\langle J_{g_j x}^* J_{k_j x} f, f \rangle \geq \delta\},$$

we obtain that for $\delta > 0$

$$\mu(\{x : |\langle J_{g_j x}^* J_{k_j x} f, f \rangle| > \delta\}) < \frac{2(1 - c_j)\|f\|_\infty^2}{\delta}.$$

Let $\delta = 1 + \|f\|_\infty + \|g\|_\infty$. Assuming that $\mu(E_{\epsilon,g}) > 0$, and that $(X, \mathcal{B}, \mu, G)$ is mixing, we can find a j large enough such that $\mu(g_j^{-1} E_{\epsilon,g} \cap k_j^{-1} E_{\epsilon,g}) = \mu(k_j g_j^{-1} E_{\epsilon,g} \cap E_{\epsilon,g}) > 0$, and also so that

$$\{x : |\langle J_{g_j x}^* J_{k_j x} f, f \rangle| < \delta\} \cap g_j^{-1} E_{\epsilon,g} \cap k_j^{-1} E_{\epsilon,g} \neq \emptyset.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Thus for an x in this non-empty intersection, and for large j,

$$\begin{aligned}\langle g, g\rangle &= \langle g, g - J_{g_j x} f\rangle + \langle g - J_{k_j x} f, J_{g_j x} f\rangle + \langle J_{k_j x} f, J_{g_j x} f\rangle \\ &\leq \|g\|_\infty \|g - J_{k_j x} f\|_2 + \|J_{k_j x} f\|_2 \|g - J_{g_j x} f\|_2 + |\langle J_{k_j x} f, J_{g_j x} f\rangle| \\ &\leq (\|g\|_\infty + \|f\|_\infty + 1)\frac{\epsilon}{(1 + \|f\|_\infty + \|g\|_\infty)} = \epsilon,\end{aligned}$$

or $\|g\|_2 < \sqrt{\epsilon}$. This proves our claim. Now let $\mathcal{G}$ be a countable subset of bounded functions dense in $\{g \in \mathcal{L}_2(X, \mu) : \|g\|_2 > \sqrt{\epsilon}\}$. What we have proved above implies that, if $g \in \mathcal{G}$, then $\mu(E_{\epsilon,g}) = 0$. Let

$$F_\epsilon = \{x : \|J_x f\|_2 > \sqrt{\epsilon}\}.$$

By density of $\mathcal{G}$, there exists $g \in \mathcal{G}$ such that $\|J_x f - g\|_2 < \epsilon/(1 + \|f\|_\infty + \|g\|_\infty)$, so that

$$F_\epsilon \subset \bigcup_{g\in\mathcal{G}} E_{\epsilon,g},$$

which means that $\mu(F_\epsilon) \leq \sum_{g\in\mathcal{G}} \mu(E_{\epsilon,g}) = 0$. □

4. *D-approximations*

If $(X, \mathcal{B}, \mu, G)$ is a rank one mixing group action, then any level of the nth tower X^n cannot get mapped 'heavily' into itself by increasing elements of G. In other words, the situation

$$\overline{\lim}_n \frac{\mu(h_n X^n_{i_n} \cap X^n_{i_n})}{\mu(X^n_{i_n})} > 0$$

is not possible, for any sequence $i_n \in F^n$, and $h_n \to \infty$. This is not necessarily true for rank r actions; what is true though is that one can find two non-trivial (in measure) 'sub'-towers, Y^n and $Y^{n,2}$, such that the levels of Y^n do not get mapped heavily into the levels of $Y^{n,2}$. This is one of the things that we show in this section.

We first generalize Ryzhikov's definition of a *D-approximation* (see [**13**]). We say that the group action $(X, \mathcal{B}, \mu, G)$ has a D-approximation if there exist two Følner sequences $\{F^n\}$ and $\{G^n\}$ which generate the approximating towers $X^{n,1} := \bigcup_{f\in F^n} X^{n,1}_f$ and $X^{n,2} := \bigcup_{g\in G^n} X^{n,2}_g$ respectively, a sequence of group elements $h_n \to \infty$, and sequences of subsets of group elements $\{A^n\}, \{B^n\}$, satisfying:

(1) $\{A^n\}$ is Følner, $A^n \subset F^n$, $B^n \subset F^n$ and $|A^n| \geq \alpha_1|F^n|$, $|B^n| \geq \alpha_2|F^n|$, for some positive constants α_1, α_2;

(2) $h_n A^n \subset F^n$ and $h_n B^n \cap F^n = \emptyset$;

(3) for some $\delta > 0$, for each $f \in B^n$ there exists $Y^n_f \subset X^{n,1}_f$ such that $\mu(Y^n_f) \geq \delta\mu(X^{n,1}_f)$, $h_n\big(\bigcup_{f\in B^n} Y^n_f\big) \subset X^{n,2}$, and

$$\underline{\lim}_n \mu(h_n Y^n_{f_n} | X^{n,2}_{g_n}) = 0$$

whenever $f_n \in B^n$ and $g_n \in G^n$.

(Note that the two Følner sequences $\{F^n\}$ and $\{G^n\}$ need not necessarily be distinct.)

If $(X, \mathcal{B}, \mu, G)$ is a local rank ergodic action (so that the following remark also applies for finite rank actions) with local tower $X^n = \bigcup_{f\in F^n} X^n_f$ and (passing to a subsequence if

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

necessary) $\lim_n \mu(X^n) = c > 0$ then, as X^n is approximately invariant, by Lemma 1, we have, for any $A \in \mathcal{B}$,

$$\mu(A|X^n) \to_n \mu(A).$$

We use this in the following.

LEMMA 4. *Let $(X, \mathcal{B}, \mu, G)$ be a local rank mixing G-action, with associated Følner sequence $\{F^n\}$. Then*

$$\lim_{n\to\infty} \sup_{g\neq e} \frac{\mu(gX^n_{f_n} \cap X^n_{f_n})}{\mu(X^n_{f_n})} = 0 \tag{13}$$

whenever $f_n \in F^n$.

Proof. We will assume that $e \in F^n$ for each n and show that (13) is true for $f_n = e$; the case for an arbitrary sequence $\{f_n\}$ is similar. Suppose that the lemma is false, i.e. there exist sequences $\{n_k\} \subset \mathbb{N}$ and $\{g_k\} \subset G$ such that $g_k \neq e$ and

$$\frac{\mu(g_k X^{n_k}_e \cap X^{n_k}_e)}{\mu(X^{n_k}_e)} \to c \neq 0.$$

Passing to a subsequence if necessary, we can assume that $g_k \to \infty$. (For otherwise we can pass to a subsequence along which $g_k = g$ with $g \neq 0$. But since $\{F^{n_k}\}$ is Følner, for large k there exist distinct f and f', both in F^{n_k}, such that $fg = f'$. Therefore, $\mu(gX^{n_k}_e \cap X^{n_k}_e) = \mu(f^{-1}f'X^{n_k}_e \cap X^{n_k}_e) = \mu(X^{n_k}_{f'} \cap X^{n_k}_f) = 0$, a contradiction.)

Let $\beta = \lim_n \mu(X^n)$. Choose a set A with $0 < \mu(A) < c\beta$, and for each k choose a set $F^{n_k}(A) \subset F^{n_k}$ so that the sets

$$A^{n_k} = \bigcup_{f\in F^{n_k}(A)} X^{n_k}_f$$

satisfy $\mu((A \cap X^{n_k}) \triangle A^{n_k}) \to 0$.

We have

$$\mu(g_k A^{n_k} \cap A^{n_k}) \geq |F^{n_k}(A)|\mu(g_k X^{n_k}_e \cap X^{n_k}_e) \tag{14}$$

$$= \mu(X^{n_k}_e)|F^{n_k}(A)|\frac{\mu(g_k X^{n_k}_e \cap X^{n_k}_e)}{\mu(X^{n_k}_e)} \tag{15}$$

$$= \mu(A^{n_k})\frac{\mu(g_k X^{n_k}_e \cap X^{n_k}_e)}{\mu(X^{n_k}_e)} \to \beta\mu(A)c. \tag{16}$$

On the other hand,

$$\lim_{k\to\infty} \mu(g_k A^{n_k} \cap A^{n_k}) \leq \lim_{k\to\infty} \mu(g_k A \cap A) = (\mu(A))^2$$

which together with (16) leads to a contradiction. □

We use Lemma 4 to prove Proposition 1 below. We remark first that, given $\epsilon > 0$, we can always find a sequence of group elements $h_n \to \infty$ such that

$$\epsilon \leq \frac{|h_n^{-1}F^n \cap F^n|}{|F^n|} \leq 2\epsilon.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

For, if l is an element of infinite order, choose N so that

$$\frac{|l^{-1}F^n \cap F^n|}{|F^n|} = \frac{|lF^n \cap F^n|}{|F^n|} > 1 - \epsilon/2$$

for $n > N$. Letting

$$a_k := \frac{|l^k F^n \cap F^n|}{|F^n|},$$

we have, for $n > N$:

$$a_0 = 1, |a_{k+1} - a_k| < \epsilon, \text{ and } a_k \neq 0 \text{ iff } l^k \in F^n(F^n)^{-1}, \text{ a finite set.} \tag{17}$$

Choosing h_n to be the appropriate sequence of powers of l establishes the remark.

PROPOSITION 1. *If $(X, \mathcal{B}, \mu, G)$ is β-local with $\beta > 1/2$, then $(X, \mathcal{B}, \mu, G)$ has a D-approximation.*

Proof. Let $\alpha > 0$ be chosen so that $\beta(1 - 2\alpha) > 1/2$. By the remark above we can find a sequence $\{h_n^{-1}\}$ of group elements such that $h_n^{-1} \to \infty$ and

$$\alpha \leq \frac{|h_n^{-1}F^n \cap F^n|}{|F^n|} \leq 2\alpha.$$

In accord with the notation in the definition of a D-approximation, let $G_n = F_n$ for all n, and define $A^n := h_n^{-1}F^n \cap F^n$, $B^n := F^n \backslash A^n$. Then A^n, B^n are Følner sequences, $h_n A^n \subset F^n$, and $h_n B^n \cap F^n = \emptyset$. For $g \in B^n$, let $Y_g^n := X_g^n$. Recalling that a_n is the measure of any level of the local tower, we have:

$$\underline{\lim}_n \mu\left(\bigcup_{g \in B^n} X_g^n\right) = \underline{\lim}_n a_n|B^n| > \underline{\lim}_n a_n|F^n|(1 - 2\alpha) = \beta(1 - 2\alpha) > \frac{1}{2}$$

by choice of α. It follows that

$$\underline{\lim}_n \mu\left(h_n\left(\bigcup_{g \in B^n} X_g^n\right) \cap \left(\bigcup_{f \in F^n} X_f^n\right)\right) > 0.$$

Suppose now that there exist sequences $\{f_n\}$ and $\{g_n\}$ in B^n and F^n respectively such that

$$\overline{\lim}_n \mu(h_n X_{f_n}^n | X_{g_n}^n) = c > 0. \tag{18}$$

Then

$$\overline{\lim}_n \frac{\mu(f_n g_n^{-1} h_n X_{f_n}^n \cap X_{f_n}^n)}{\mu(X_{f_n})} = c > 0$$

and, since $f_n \in B^n = F^n \backslash h_n^{-1}F^n$ and $g_n \in F^n$, $f_n g_n^{-1} h_n \neq e$. This would contradict Lemma 4. Thus (18) is false, and Property 3 of the definition of D-approximation is also satisfied. □

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Let $(X, \mathcal{B}, \mu, G)$ be a rank r group action, generated by $\{F^{n,i}\}$, for $i = 1, 2, \ldots, r$. Let h be an element of infinite order in G, and let $\alpha > 0$ be small and fixed. Choose sequences of powers of h, say $h_n^{(i)} := h^{k_n(i)}$, for $i = 1, 2, \ldots, r$ such that

$$\frac{\alpha}{2^{r-1}} < \frac{|(h_n^{(i)})^{-1} F^{n,i} \cap F^{n,i}|}{|F^{n,i}|} < \frac{2\alpha}{2^{r-1}}. \tag{19}$$

If $f \in F^{n,i}$, define

$$X_f^{n,i\to j} = \{x \in X_f^{n,i} : h_n^{(i)} x \in X^{n,j}\}.$$

Also define

$$X^{n,i\to j} = \bigcup_{\{f \in F^{n,i} \setminus (h_n^{(i)})^{-1} F^{n,i}\}} X_f^{n,i\to j}.$$

If

$$\overline{\lim}_n \mu(X^{n,i\to i}) > 0$$

for some $i \in \{1, 2, \ldots, r\}$, then a D-approximation exists much in the same way that it existed for local rank actions: namely, for this i, we let $A^n := (h_n^{(i)})^{-1} F^{n,i} \cap F^{n,i}$, $B^n := F^{n,i} \setminus A^n$. These sequences will do the trick (letting $F^n = G^n = F^{n,i}$).

Thus we will make the assumption that

$$\overline{\lim}_n \mu(X^{n,i\to i}) = 0$$

for $i = 1, 2, \ldots, r$, since this is the only case that remains. In this case we will need some extra restrictions on our choices of $h_n^{(i)}$. We can choose the elements $h_n^{(i)}$ to satisfy both (19) and also

$$\frac{\left|\left(h_n^{(i)} \prod_{j\in J} h_n^{(j)}\right)^{-1} F^{n,i} \cap F^{n,i}\right|}{|F^{n,i}|} < \frac{2\alpha}{2^{r-1}} \tag{20}$$

whenever $J \subset \{1, \ldots, r\}$ and $i \notin J$. To see this, when choosing the $h_n^{(i)}$'s, simply let them be the highest power of l (the element of infinite order) such that

$$\frac{\alpha}{2^{r-1}} \leq \frac{|(h_n^{(i)})^{-1} F^{n,i} \cap F^{n,i}|}{|F^{n,i}|} \leq \frac{2\alpha}{2^{r-1}}.$$

Then, using the properties listed in (17), with $\epsilon = \alpha/2^{r-1}$, it can be seen that (20) will hold. We will also have

$$|F_0^{n,i}| < 2\alpha, \tag{21}$$

where

$$F_0^{n,i} = F^{n,i} \cap \bigcup_{\substack{J\subset\{1\ldots r\}\\ i\notin J}} \left(h_n^{(i)} \prod_{j\in J} h_n^{(j)}\right)^{-1} F^{n,i}.$$

For $f \in F^{n,i}$, $g \in F^{n,j}$, define

$$X_{f\to g}^{n,i\to j} := \{x \in X_f^{n,i} : h_n^{(i)} x \in X_g^{n,j}\};$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

so that $X^{n,i\to j}_{f\to g} \subset X^{n,i}_f$ for all $g \in F^{n,j}$. We say that $X^{n,i}_f$ goes ϵ-*heavily* into $X^{n,j}_g$ under $h^{(i)}_n$ if

$$\mu(X^{n,i\to j}_{f\to g}|X^{n,i}_f) \geq \epsilon$$

and, if $\mu(X^{n,i\to j}_{f\to g}|X^{n,i}_f) < \epsilon$, then we say that $X^{n,i}_f$ goes ϵ-*lightly* into $X^{n,j}_g$ under $h^{(i)}_n$. For $\epsilon > 0$ and $f \in F^{n,i}$, define

$$X^{n,i\to j}_f(\epsilon) := \bigcup_{\{g\in F^{n,j}:\mu(X^{n,i\to j}_{f\to g}|X^{n,i}_f)<\epsilon\}} X^{n,i\to j}_{f\to g}$$

for $j = 1, \ldots, r$. Thus $X^{n,i\to j}_f(\epsilon)$ is the part of $X^{n,i}_f$ which moves lightly into $X^{n,j}$.

PROPOSITION 2. *Let $(X, \mathcal{B}, \mu, G)$ be a rank r mixing action, generated by the Følner sequences $\{F^{n,i}\}$, $i = 1, 2, \ldots, r$, and let $h^{(i)}_n$ be as in (21). If*

$$\overline{\lim}_n \mu(X^{n,i\to i}) = 0$$

for $i = 1, 2, \ldots, r$, then there exist $i_0 \neq j_0 \in \{1, \ldots, r\}$, and there exists $\delta > 0$ such that for all $\epsilon > 0$, there exists $n = n(\epsilon)$ such that, for $n \geq n(\epsilon)$, there exists a set $B_n \subset F^{n,i_0}\backslash F^{n,i_0}_0$ where $|B^n| \geq \delta|F^{n,i_0}|$, and such that for all $f \in B_n$ we have

$$\mu(X^{n,i_0\to j_0}_f(\epsilon)|X^{n,i_0}_f) \geq \delta. \tag{22}$$

Setting $A^n := (h^{(i_0)}_n)^{-1}F^{n,i_0} \cap F^{n,i_0}$, and B^n as above, we obtain a D-approximation.

Proof. Suppose not. Then, for all $i \neq j$ and all $\delta > 0$, there exist some ϵ and a sequence $n_k \to \infty$ such that (22) with $n = n_k$ is true for less than $\delta|F^{n_k,i}|$ values of f in $F^{n_k,i}$.

In Figure 1, in the leftmost tower, region C represents the levels which are indexed by elements in $F^{n,1}_0$. Region C is fixed but small (it depends on the choice of α in (19)). Region B represents the levels in $X^{n_k,1}$ which do satisfy inequality (22), for some i_0, and region A represents the part of $X^{n_k,1}$ which moves ϵ-lightly into other towers. Both regions A and B can be made small by choosing δ small; and regions A, B and C are present and small in all towers.

We choose an $f \in F^{n_k,1}$ such that:

- f is not indexing a level in region B or C;
- $\exists j_1, \exists f_1 \in F^{n_k,j_1}$ with f_1 not indexing a level in region B or C and such that $\mu(X^{n_k,1\to j_1}_{f\to f_1}|X^{n,1}_f) \geq \epsilon$;
- $h^{(1)}_{n_k}X^{n_k,1}_f \cap X^{n_k,j_1}_{f_1}$ is in the part of $X^{n_k,j_1}_{f_1}$ which moves ϵ-heavily out of X^{n_k,j_1} under $h^{(j_1)}_{n_k}$, into $X^{n_k,j_2}_{f_2}$ where $f_2 \in F^{n_k,j_2}\backslash F^{n_k,j_2}_0$, so that

$$\mu(h^{(j_1)}_{n_k}h^{(1)}_{n_k}(X^{n_k,1}_f) \cap X^{n_k,j_2}_{f_2}|X^{n_k,1}_f) \geq \epsilon^2;$$

- repeating this process at most r times, until we see the same tower twice in our picture, we will find a $j_p \in \{1, \ldots, r\}$, $f_p \in F^{n_k,j_p}\backslash F^{n_k,j_p}_0$ and $f^* \in F^{n,j_p}$ such that

$$\mu(h^{(j_q)}_{n_k} \cdots h^{(j_p)}_{n_k}X^{n_k,j_p}_{f_p} \cap X^{n_k,j_p}_{f^*}|X^{n_k,j_p}_{f_p}) \geq \epsilon^r.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

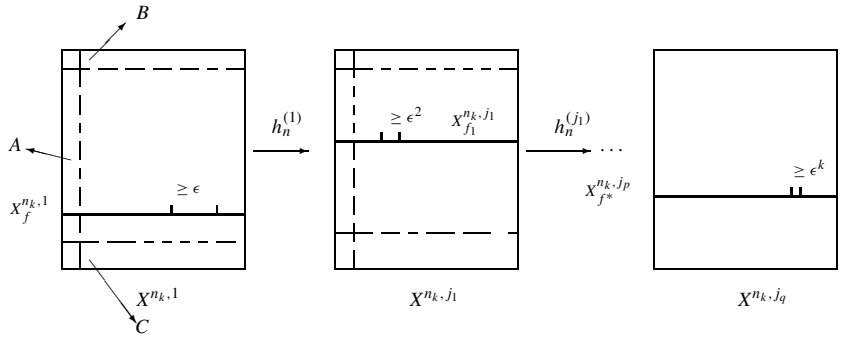


FIGURE 1. How one level moves heavily into following towers.

Since $f_p \in F^{n_k,j_p} \setminus F_0^{n_k,j_p}$, then

$$h_{n_k}^{(j_q)} \cdots h_{n_k}^{(j_p)} (f^*)^{-1} f_p \neq e.$$

Now it is possible to obtain a contradiction to mixing, using Lemma 4.

We now construct the D-approximation. Let $\epsilon_n \to 0$ and, by passing to a subsequence if necessary, find a sequence $\{B^n\}$ such that elements in B^n satisfy (22) for $\epsilon = \epsilon_n$. Suppose without loss of generality that we found $i_0 = 1$, $j_0 = 2$ in Proposition 2. Then let $h_n := (h_n^{(1)})^{-1}$, $A^n := h_n^{(1)} F^{n,1} \cap F^{n,1}$. Furthermore, for $g \in B^n$, let

$$Y_g^n := X_g^{n,1\to 2}(\epsilon_n) = \bigcup_{\{f \in F^{n,2} : \mu(X_{g\to f}^{n,1\to 2} | X_g^{n,1}) < \epsilon\}} X_{g\to f}^{n,1\to 2}.$$

Then

$$\underline{\lim}_n \mu\Big(\bigcup_{g\in B^n} Y_g^n \Big) \geq \underline{\lim}_n \delta |F^{n,1}| \delta a^{n,1} = \delta^2/2.$$

Also,

$$h_n\Big(\bigcup_{g\in B^n} Y_g^n \Big) \subset \bigcup_{f\in F^{n,2}} X_f^{n,2},$$

so that

$$\underline{\lim}_n \mu\Big(h_n\Big(\bigcup_{g\in B^n} Y_g^n \Big) \bigcap \Big(\bigcup_{f\in F^{n,2}} X_f^{n,2} \Big) \Big) = \delta^2/2.$$

Finally, note that, by construction, for any $g \in B^n$, $f \in F^{n,2}$, we have

$$\mu(h_n Y_g^n | X_f^{n,2}) < \epsilon_n \to_n 0,$$

i.e. condition (3) of the definition of D-approximation is also satisfied. □

We have therefore shown that a *D*-approximation exists for all rank r mixing group actions (X, G) where G is countable, Abelian and with an element of infinite order.

Henceforth we will assume that $A^n, B^n \subset F^{n,1}$ and that

$$h_n\Big(\bigcup_{g\in B^n} Y_g^n \Big) \subset X^{n,2}$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

(the other cases are similar). If $(k, l, m) \in A^n \times A^n \times B^n$, where A^n, B^n are the sets coming out of the D-approximation, then we will denote the cube (or rather cuboid) $X_k^{n,1} \times X_l^{n,1} \times Y_m^n$ by $Y_{k,l,m}^n$. We can then form the sets

$$R^n := \bigcup_{(k,l,m)\in A^n\times A^n\times B^n} Y_{k,l,m}^n \tag{23}$$

and

$$S^n := \bigcup_{(k,l,m)\in h_n A^n\times h_n A^n\times F^{n,2}} X_{k,l,m}^n. \tag{24}$$

If a pairwise independent ergodic measure ν gave R^n and S^n non-trivial mass, then we would be ready to prove Theorem 2. It is possible that $\liminf \nu(R^n) = 0$. However, Proposition 3 shows that we can find induced joinings $\eta_{g,k}$ which give the towers R^n, and therefore S^n, mass almost as large as their product mass. Furthermore, $\eta_{g,k}$ will also give the set $\left(\bigcup_{(k,l)\in A^n\times A^n} X_{k,l}^n\right) \times X$ measure comparable to its product mass; this will ensure that the induced joinings have a large pairwise independent component, as we shall see in Lemma 6.

In what follows, the sequence of sets $\{U_n\}$ is the sequence $\left\{\bigcup_{f\in A^n} X_f^{n,1}\right\}$ and $\{U_n'\}$ is the sequence $\left\{\bigcup_{g\in B^n} Y_g^n\right\}$.

PROPOSITION 3. *Let $(X, \mathcal{B}, G, \mu)$ be a group action, ν be pairwise independent, and $\{\eta_{g,k}\}_{g,k\in G}$ the family of induced measures generated by ν. Let $c_j \uparrow 1$. If $\{U_n\}, \{U_n'\}$ are sequences of measurable sets such that*

$$\lim_n \mu(U_n) = d > 0, \quad \lim_n \mu(U_n') = d' > 0,$$

then there exist sequences $\{g_j\}, \{k_j\}$ of group elements of infinite order such that

$$\lim_n g_j k_j^{-1} = \infty,$$

and

$$\overline{\lim}_n \eta_{g_j,k_j}(U_n \times U_n \times X) \geq c_j d^2 \tag{25}$$

and

$$\overline{\lim}_n \eta_{g_j,k_j}(U_n \times U_n \times U_n') \geq c_j d^2 d' \tag{26}$$

for each $j \in \mathbb{N}$.

Proof. Fix $n \in \mathbb{N}$ for the time being. Choose a natural number l large enough so that

$$\frac{\mu(U_n)^2 - l^{-1}\mu(U_n)}{1 - l^{-1}} > \mu(U_n)^2 c_1.$$

Note that l can be chosen independently of large n. Let

$$H := \{h \in G : h \text{ is of infinite order}\}.$$

Choose and fix a set $A \subset H$ such that $|A| = r(l)$, where $r(l)$ is given by Ramsey's Theorem. Let $g \in G$. Define $F_g : X^2 \to [0, 1]$ as

$$F_g(x, y) = \nu_{(gx,y)}(U_n).$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Then if $g, k \in G$, and $\mu_n := \mu_{U_n'}$, we have

$$\eta_{g,k}(U_n \times U_n \times U_n') = \mu(U_n') \int_X \int_X F_g(x, y) F_k(x, y)\, d\mu(y)\, d\mu_n(x)$$
$$= \mu(U_n')\langle F_g, F_k\rangle_{\mu\times\mu_n}.$$

Assumption. Suppose that there is no pair of group elements g, k in A such that both

$$\mu(U_n')\langle F_g, F_k\rangle_{\mu\times\mu_n} = \eta_{g,k}(U_n \times U_n \times U_n') \geq c_1\mu(U_n)^2\mu(U_n') \tag{27}$$

and

$$\langle F_g, F_k\rangle = \eta_{g,k}(U_n \times U_n \times X) \geq c_1(\mu(U_n))^2 \tag{28}$$

hold. Define the 2-colouring $*$ ($*$ depends on n) on A as: $g * k$ iff (28) does hold for the pair g, k. If $g * k$, then because of our assumption, (27) does not hold for the pair g, k.

By Ramsey's Theorem, A contains a set B of cardinality l where either $g * k$ for all $g, k \in B$ or $\neg(g * k)$ for all $g, k \in B$. (B is really B_n and $* = *_n$.) If $g * k$ for all $g, k \in B$, then (27) does not hold for all $g, k \in B$, i.e.

$$\langle F_g, F_k\rangle_{\mu\times\mu_n} < c_1(\mu(U_n))^2.$$

Similarly if $\neg(g * k)$ for $g, k \in B$, then

$$\langle F_g, F_k\rangle < c_1\ (\mu(U_n))^2$$

for $g, k \in B$. Lemma 2 tells us that neither possibility can happen. So our initial assumption was false, i.e. there exist elements $g, k \in A$ such that both (27) and (28) hold. Repeating this argument for all large n (with the same set A) there exists a sequence of pairs of elements $g_1(n), k_1(n)$ in A such that both (27) and (28) hold. As A is finite, then there exists a pair g_1, k_1 which work for infinitely many n's, i.e. such that both (25) and (26) hold for $j = 1$. The proof can be repeated for any $j \in \mathbb{N}$, except that we vary the set $A = A(j)$, taking it to consist of sufficiently spaced integers so that $g_j k_j^{-1} \to \infty$; and only working with the subsequence of n's that one obtained at stage $j - 1$. Taking a diagonal subsequence, we have (25) and (26) for all $j \in \mathbb{N}$. □

5. *Long orbits and almost pairwise independence*

In this section $(\mathcal{X}, G)$ is a rank r or local rank ($\beta > 1/2$) mixing G-action as in Theorem 2 and $\eta \in M(1, 3)$. For $(f, g, h) \in F^{n,i} \times F^{n,j} \times F^{n,k}$, where $1 \leq i, j, k \leq r$, we denote $X_f^{n,i} \times X_g^{n,j} \times X_h^{n,k}$ by $X_{f,g,h}^{n,i,j,k}$ and call this object a cube. For $1 \leq i, j, k \leq r$, we let

$$X^{n,i,j,k} := X^{n,i} \times X^{n,j} \times X^{n,k};$$

i.e. $X^{n,i,j,k}$ stands for the product of the towers i, j, and k at the nth stage. By an *orbit* in G^3 we mean any

$$G_{g_0,g_1}^3 := \{(l, lg_0, lg_1) : l \in G\},$$

that is, an orbit of the diagonal action of G on G^3 by translation. By an *n-orbit* in $X^{n,i,j,k}$ we mean

$$\mathcal{O}_{g_0,g_1}^{n,i,j,k} := \bigcup_{\{l \in F^{n,i} \cap F^{n,j} g_0^{-1} \cap F^{n,k} g_1^{-1}\}} X_{(l,lg_0,lg_1)}^{n,i,j,k}.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Thus each $X^{n,i,j,k}$ is partitioned into n-orbits and each n-orbit is a union of cubes of equal η-measure. The *length* of an n-orbit is the number of cubes in it. For $(f, g, h) \in F^{n,i} \times F^{n,j} \times F^{n,k}$, $\mathcal{O}(X^{n,i,j,k}_{f,g,h})$ will denote the n-orbit containing the cube $X^{n,i,j,k}_{f,g,h}$. We say that an n-orbit in $X^{n,i,j,k}$ is δ-*long* if its length is at least $\delta \min_{1\le i\le r} |F^{n,i}|$ (otherwise we call it δ-short), and let $\mathcal{O}^n(\delta^c)$ denote the union of the δ-short orbits, and $\mathcal{O}^{n,i,j,k}(\delta^c)$ the union of the δ-short orbits in $X^{n,i,j,k}$. Similarly, we call $F^{n,i} \cap gF^{n,j} \cap g^*F^{n,k}$ a δ-short intersection if $|F^{n,i} \cap gF^{n,j} \cap g^*F^{n,k}| \le \delta \min_{1\le i\le r} |F^{n,i}|$. In the next lemma we will show that the η-mass of $\mathcal{O}^{n,1,1,1}(\delta^c)$ is small in the limit. The argument can be copied to give the smallness of $\eta(\mathcal{O}^{n,i,j,k}(\delta^c))$ for any $1 \le i, j, k \le r$. Summing over these r^3 possibilities will give that $\underline{\lim}_n \eta(\mathcal{O}^n(\delta^c)) \le r^3 p^2 \delta$.

LEMMA 5. *If p is an intersection bound for $\{F^{n,1}, F^{n,2}, \ldots, F^{n,r}\}$ then*

$$\overline{\lim}_n \eta(\mathcal{O}^{n,1,1,1}(\delta^c)) \le p^2\delta.$$

Proof. We fix n; to ease notation, we assume that $e \in F^{n,1}$, so that $X^{n,1}_e$ will be considered as the 'base' of the tower $X^{n,1}$. We will consider all finite subsets $\gamma = \{g_1, g_2, \ldots, g_q\}$ of G with the properties that the $\{g_k F^{n,1}\}_{k=1}^q$ are pairwise disjoint and all intersect $F^{n,1}$ non-trivially, so that $|\gamma| \le p$. We call such a γ a *configuration* and let Γ denote the space of all configurations. Clearly Γ is finite. If $\gamma = \{g_1, g_2, \ldots, g_q\}$ is a configuration and $k \in G$, then $k\gamma$ is the set $\{kg_1, kg_2, \ldots, kg_q\}$.

For $x \in X$ we let $R^{n,1}(x) = \{g \in G : gx \in X^{n,1}_e\}$, the set of 'return times' to the base of the nth tower $X^{n,1}$. By an *n-block* in x we mean any $gF^{n,1}$ with $g \in R^{n,1}(x)$, and g is called the *base time* of this n-block. The n-blocks in x are disjoint translates of $F^{n,1}$. For $x \in X^{n,1}$ we denote by $B^n(x)$ the n-block in x containing $e \in G$, namely $B^n(x) = k^{-1}F^{n,1}$ if $x \in X^{n,1}_k$.

For $(x, y) \in X^{n,1}_k \times X$ we create a configuration $\gamma(x, y)$ by letting $g_1, \ldots, g_q$ denote the base times of the n-blocks in y which intersect $B^n(x)$ and defining

$$\gamma(x, y) = k\{g_1, g_2, \ldots, g_q\} = \{kg_1, \ldots, kg_q\}. \tag{29}$$

For $(x, y, z) \in X^{n,1}_k \times X^2$, if

$$\gamma(x, z) = k\{h_1, h_2, \ldots, h_s\}, \tag{30}$$

then let

$$\gamma_1(x, y, z) := \gamma(x, y) \quad \text{and} \quad \gamma_2(x, y, z) := \gamma(x, z).$$

We view the map $Q : (x, y, z) \to (\gamma_1(x, y, z), \gamma_2(x, y, z))$ as a partition of $X^{n,1} \times X^2$ indexed by Γ^2, so we will write $Q_{(\gamma_1,\gamma_2)} = Q^{-1}((\gamma_1, \gamma_2))$. Thus we are partitioning $X^{n,1} \times X^2$ according to the pattern, up to a shift, formed by $B^n(x)$ and the $F^{n,1}$-blocks in y and z which intersect $B^n(x)$.

For $(\gamma_1, \gamma_2) \in \Gamma^2$ and $k \in F^{n,1}$ let

$$Q_{(\gamma_1,\gamma_2),k} := \{(x, y, z) \in Q_{(\gamma_1,\gamma_2)} : x \in X^{n,1}_k\} \subset X^{n,1}_k \times X^2.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

We call $Q_{(\gamma_1,\gamma_2),k}$ a *cell* of $Q_{(\gamma_1,\gamma_2)}$. We claim that

$$\eta(Q_{(\gamma_1,\gamma_2),k}) = \eta(Q_{(\gamma_1,\gamma_2),e})$$

for each $k \in F^{n,1}$. This is true as $Q_{(\gamma_1,\gamma_2),e}$ is mapped to $Q_{(\gamma_1,\gamma_2),k}$ by the transformation k.

Define

$$S((\gamma_1,\gamma_2),1,1,1) := \bigcup_{\{(g,h):g\in\gamma_1,h\in\gamma_2\}} F^{n,1} \cap gF^{n,1} \cap hF^{n,1}.$$

It can be verified that $k^* \in S((\gamma_1,\gamma_2),1,1,1)$ iff $Q_{(\gamma_1,\gamma_2),k^*} \subset X^{n,1,1,1}$.

Now suppose that $k \in F^{n,1} \cap gF^{n,1} \cap hF^{n,1}$ and suppose that $\mathcal{O}(X^{n,1,1,1}_{k,l,m}) \cap Q_{(\gamma_1,\gamma_2),k}$ is non-trivial. Then $\mathcal{O}(X^{n,1,1,1}_{k,l,m})$ intersects $Q_{(\gamma_1,\gamma_2),k^*}$ non-trivially, whenever $k^* \in F^{n,1} \cap gF^{n,1} \cap hF^{n,1}$. Thus δ-short orbits can only intersect cells which belong to a δ-short intersection. There are at most p^2 intersections. Noting that the cells of $Q_{(\gamma_1,\gamma_2)}$ all have the same η-mass, we have

$$\eta(\mathcal{O}^{n,1,1,1}(\delta^c)) \le \sum_{\gamma_1,\gamma_2} p^2\delta|F^{n,1}|\eta(Q_{(\gamma_1,\gamma_2),e}) \le p^2\delta|F^{n,1}|\mu(X^{n,1}_e),$$

since the $Q_{(\gamma_1,\gamma_2),k}$ form a partition of $X^{n,1}_k \times X^2$ and $\eta \in M(1,3)$. Letting $n \to \infty$, we get the result. □

Note that, as long as we have the bounded intersection property for Følner sequences, a slightly modified proof of Lemma 5 shows that the δ-short orbits in $X^n \times X^n \times X^n$ are of small η-mass, whenever $\eta \in M(1,3)$ is a joining of a local rank action. The analogous result is also true for any 2-self-joining $\eta \in M(1,2)$ (ergodic or not) which preserves the action of G. This is used in the following lemma.

LEMMA 6. *Suppose $(X,\mathcal{B},\mu,G,\{A^n\})$ is a β-local mixing action. Let η be a (possibly non-ergodic) 2-self-joining of $(X,\mathcal{B},\mu,G)$ and suppose that*

$$\lim_n \eta\Bigg(\bigcup_{(k,l)\in A^n\times A^n} X^n_{k,l}\Bigg) = c > 0. \tag{31}$$

Then η either has an off-diagonal measure as a (non-trivial) component, or μ^2 as a non-trivial component.

If $\eta = \pi_{1,2}\eta_{g,k}$ is the projection of a joining induced by $\nu \in M(2,3)$, then $\eta_{g,k} \ge c\beta^{-2}\mu^2$, where

$$\beta = \lim_n \mu\Bigg(\bigcup_{k\in A^n} X^n_k\Bigg).$$

Proof. Let

$$\mathcal{O}^n(\delta) := \bigcup_{\{(k,l)\in A^n\times A^n:|A^nk^{-1}\cap A^nl^{-1}|\ge\delta|A^n|\}} \mathcal{O}(X^n_{k,l}).$$

Fix some small δ, such that

$$\underline{\lim}_n\, \eta(\mathcal{O}^n(\delta)) > 0.$$

By passing to a subsequence if necessary, assume that this liminf is a limit. Define

$$\eta^n = \eta_{\mathcal{O}^n(\delta)}.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

As in the proof of Lemma 1, a limit point of the measures η^n will be a non-trivial component of η. We can write η^n as

$$\eta^n = \sum_{\{(k,l):|A^n k^{-1} \cap A^n l^{-1}| \geq \delta |A^n|\}} a^n_{k,l} \eta^n_{k,l},$$

where $\eta^n_{k,l} := \eta_{\mathcal{O}(X^n_{k,l})}$ and $a^n_{k,l} := \eta(\mathcal{O}(X^n_{k,l}))/\eta(\mathcal{O}^n(\delta))$. Define

$$\lambda^n := \sum_{\{(k,l):|A^n k^{-1} \cap A^n l^{-1}| \geq \delta |A^n|\}} a^n_{k,l} \Delta^n_{k,l},$$

where $\Delta^n_{k,l}$ is the measure $\Delta_{k,l}$ conditioned on the set $\mathcal{O}(X^n_{k,l})$. (Recall that $\Delta_{k,l}$ is the graph joining: $\Delta_{k,l}(A \times B) := \mu(lk^{-1}A \cap B)$.) Since $(X, G, \{A^n\})$ is β-local, then $\{\lambda^n\}$ and $\{\eta^n\}$ have the same limit points. To see this, note that $\lambda^n(A_1 \times A_2) = \eta^n(A_1 \times A_2)$ whenever A_1, A_2 are unions of levels in X^j, for $n \geq j$. If A_1, A_2 are arbitrary measurable sets, find sequences $\{A^j_i\}_{j \in \mathbb{N}}$ for $i = 1, 2$ such that A^j_i is a union of levels in X^j and

$$\mu((A_i \cap X^j) \triangle A^j_i) \to_j 0.$$

Since

$$\lambda^n\left((A_1 \times A_2 \cap X^j \times X^j) \triangle (A^j_1 \times A^j_2) \leq \sum_{i=1}^{2} \lambda^n(((A_i \cap X^j) \triangle A^j_i) \times X)\right) \tag{32}$$

then, as $j \to \infty$, this quantity tends to 0 for each n. Inequality (32) also holds for η^n so, for a fixed j, the quantity on the right-hand side of inequality (32) will also be small. Thus λ^n and η^n have the same limit points.

There may exist a pair of group elements (k, l), and a subsequence of n's such that $\lim_n a^n_{(k,l)} > 0$, in which case by ergodicity of μ, and therefore of $\Delta_{k,l}$,

$$\begin{aligned}\eta &\geq \underline{\lim}_n \eta(\mathcal{O}^n(\delta)) \lim_n \eta^n = \underline{\lim}_n \eta(\mathcal{O}^n(\delta)) \lim_n \lambda^n \\ &\geq \underline{\lim}_n \eta(\mathcal{O}^n(\delta)) \lim_n a^n_{(k,l)} \lim_n \Delta^n_{k,l} = \underline{\lim}_n \eta(\mathcal{O}^n(\delta)) \lim_n a^n_{(k,l)} \Delta_{k,l},\end{aligned}$$

so that η has an off-diagonal measure as a component. If no such (k, l) exists then $\lim_n a^n_{k,l} = 0$ for each pair (k, l), in which case, by mixing,

$$\sum_{\{(k,l):|A^n k^{-1} \cap A^n l^{-1}| \geq \delta |A^n|\}} a^n_{k,l} \Delta_{k,l} \to_n \mu^2$$

and, by the ergodicity of μ^2 and the fact that λ^n is a non-trivial component of the above measure, it follows that the only limit point of λ^n is μ^2. Thus μ^2 is a component of η.

We now look at what happens when $\eta = \pi_{1,2}\eta_{g,k}$ is the projection of a joining induced by $\nu \in M(2, 3)$. In this case we discuss why the first possibility, that $\pi_{1,2}\eta_{g,k}$ has a diagonal measure as a non-trivial component, leads to a contradiction. Suppose that

$$\pi_{1,2}\eta_{g,k} \geq a\Delta_{(p,q)}$$

for some p, q in G. Thus if A, B are measurable sets, we have

$$\int_X \langle J^*_{gx} J_{kx} \chi_{q^{-1}pA}, \chi_B \rangle \, d\mu(x) \geq a\Delta(A \times B),$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

where Δ is diagonal measure. Let $D(X) := \{(x, x) : x \in X\}$. Expanding the measure on the left-hand side, we have

$$\int_X \int_X \nu_{kx,y} \circ q^{-1}p \times \nu_{gx,y}\, d\mu(y)\, d\mu(x) \geq a\Delta,$$

so that

$$\mu^2(\{(x, y) : \nu_{kx,y} \circ q^{-1}p \times \nu_{gx,y}(D(X)) > 0\}) > 0.$$

Now the only way that the measure $\nu_{kx,y} \circ q^{-1}p \times \nu_{gx,y}$ can give positive mass to $D(X)$ is if $\nu_{kx,y} \circ q^{-1}p$ and $\nu_{gx,y}$ have common discrete components: say

$$\nu_{gx,y} \geq a_{(x,y)}\delta_{t(x,y)}$$

and

$$\nu_{kx,y} \circ q^{-1}p \geq a_{(x,y)}\delta_{t(x,y)},$$

where $\delta_{t(x,y)}$ is the point mass at $t(x, y) \in X$ and $0 < a_{(x,y)} \leq 1$. Thus, if $\nu' := \nu \circ (q^{-1}p \times e \times k^{-1})$ and $\nu'' := \nu \circ (e \times e \times g^{-1})$, i.e.

$$\nu''(A \times B \times C) = \int_{g^{-1}C} \nu_x(A \times B)\, d\mu(x) = \int_C \int_B \nu_{gx,y}(A)\, d\mu(y)\, d\mu(x)$$

and

$$\nu'(A \times B \times C) = \int_C \int_B \nu_{kx,y} \circ q^{-1}p(A)\, d\mu(y)\, d\mu(x),$$

then ν' and ν'' have a common non-trivial component. By assumption, ν is ergodic, and hence so are ν', ν'', both being isomorphic to $(X^3, \mathcal{B} \times \mathcal{B} \times \mathcal{B}, G, \nu)$. Hence $\nu' = \nu''$. Thus ν' is a (ergodic) joining of the ergodic system $(X^2, \mathcal{B} \times \mathcal{B}, q^{-1}p \times k^{-1}g, \mu^2)$ with $(X, \mathcal{B}, e, \mu)$. Recall that $k^{-1}g$ was chosen to have infinite order and, if $q^{-1}p$ has infinite order, then $(X^2, \mathcal{B} \times \mathcal{B}, q^{-1}p \times k^{-1}g, \mu^2)$ is ergodic and so ν' must be product measure, which is only possible if ν itself were. If $q^{-1}p$ has finite order, then ν' joins a mixing system $((X, \mathcal{B}, \mu, k^{-1}g))$ with a periodic one, and this is also not possible (since then ν' joins the power of a mixing system, which is ergodic, with the identity) unless ν' (and hence ν) were product measure.

It remains to be shown that the 'size' of this component is $c\beta^{-2}$ when $\eta = \pi_{1,2}\eta_{g,k}$. Let

$$\eta = k\, d\mu^2 + \theta$$

be the Radon–Nikodym decomposition of η with respect to μ^2, where θ is a measure which is singular with respect to μ^2 and $k > 0$. Now

$$\overline{\lim}_n\, \theta\Bigg(\bigcup_{(p,q)\in A^n\times A^n} X^n_{p,q}\Bigg) = 0;$$

or else we could repeat the whole argument above to conclude that θ has μ^2 as a component (by the above argument, η, and therefore θ, cannot have an off-diagonal as a component), and this contradicts the fact that $\theta \perp \mu^2$. Hence

$$c = \lim_n \eta\Bigg(\bigcup_{(p,q)\in A^n\times A^n} X^n_{p,q}\Bigg) = k \lim_n \mu^2\Bigg(\bigcup_{(p,q)\in A^n\times A^n} X^n_{p,q}\Bigg) = k\beta^2$$

or $k = c\beta^{-2}$. □

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

We have shown therefore that, if $\eta_{g,k}$ is an induced joining satisfying (31), then

$$\pi_{1,2}\eta_{g,k} = \kappa\mu^2 + \theta, \tag{33}$$

where $\kappa \geq c\beta^{-2}$ and $\theta \perp \mu^2$. Note that (33) implies that there exists $\eta^*_{g,k}$ pairwise independent with

$$\eta^*_{g,k} \geq \kappa\eta_{g,k} \geq c\beta^{-2}\eta_{g,k}.$$

6. *Proof of the main theorem*

Recall ((23) and (24)) that

$$R^n := \bigcup_{(k,l,m)\in A^n\times A^n\times B^n} Y^n_{k,l,m}$$

and

$$S^n := \bigcup_{(k,l,m)\in h_nA^n\times h_nA^n\times F^{n,2}} X^n_{k,l,m}$$

were the products of towers constructed using the sets A^n and B^n coming from the D-approximation. R^n and S^n were constructed so that the third component of R^n moved μ-lightly into the third component of S^n. In the next lemma we move this up to the level of joinings. Namely, we shall show that if η is pairwise independent such that

$$\overline{\lim}_n\, \eta(h_nR^n \cap S^n) > 0,$$

then 'one of R^n or S^n will have substantially many η-light cubes'. This is made rigorous in Lemma 7.

Let $c_j \uparrow 1$. Proposition 3 gives us a sequence of induced joinings η_{g_j,k_j} which satisfy

$$\overline{\lim}_n\, \eta_{g_j,k_j}(R^n) \geq c_jd^2d'$$

and

$$\overline{\lim}_n\, \eta_{g_j,k_j}\left(\left(\bigcup_{k,l\in A^n\times A^n} X^n_{k,l}\right)\times X\right) \geq c_jd^2,$$

where $\lim_n \mu\big(\bigcup_{k\in A^n} X^n_k\big) = d > 0$ and $\lim_n \mu\big(\bigcup_{k\in B^n} Y^n_k\big) = d' > 0$.

Lemma 6 tells us that

$$\eta_{g_j,k_j} \geq c_j\eta^*_{g_j,k_j},$$

where $\eta^*_{g_j,k_j}$ is pairwise independent. Thus, for a large enough j,

$$\overline{\lim}_n\, \eta^*_{g_j,k_j}(R^n) > c^*_jd^2d',$$

where $c^*_j \to_j 1$. We will fix such a large enough j, and for ease of notation write this measure $\eta^*_{g_j,k_j} = \eta^j$. We define a cube $Y^n_{k,l,m} \in R^n$ to be ϵ-light (for η) if $\eta(Y^n_{k,l,m}) \leq \epsilon(a^{n,1})^2$; similarly for a cube in S_n. Let

$$L^n_\epsilon = \bigcup\{Y^n_{k,l,m} : (k,l,m) \in A^n \times A^n \times B^n \text{ and } Y^n_{k,l,m} \text{ is } \epsilon\text{-light}\}$$

and

$$\underline{L}^n_\epsilon = \bigcup\{X^n_{k,l,m} : (k,l,m) \in h_nA^n \times h_nA^n \times F^{n,2} \text{ and } X^n_{k,l,m} \text{ is } \epsilon\text{-light}\}.$$

Let $Di(\eta) = \lim_{\epsilon\to 0} \underline{\lim}_n\, \eta(L^n_\epsilon)$ and $\underline{Di}(\eta) = \lim_{\epsilon\to 0} \underline{\lim}_n\, \eta(\underline{L}^n_\epsilon)$.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

LEMMA 7. *Let $\eta \in M(2,3)$, where $(X, \mathcal{B}, \mu, G)$ is finite rank or β-local rank, $\beta > 1/2$, and G is countable, Abelian, and has an element of infinite order. If*

$$\overline{\lim}_n \eta(R^n) > 0,$$

then $Di(\eta) > 0$ or $\underline{Di}(\eta) > 0$.

Proof. Suppose that both $Di(\eta) = 0$ and $\underline{Di}(\eta) = 0$. Given any ι small, there exists a (fixed) ϵ such that $\underline{\lim}_n \eta(L^n_\epsilon) < \iota$ and $\underline{\lim}_n \eta(\underline{L}^n_\epsilon) < \iota$. If $C^n_{k,l} := \bigcup_{m \in B^n} Y^n_{k,l,m}$, then $h_n C^n_{k,l} \subset \underline{C}^n_{h_n k, h_n l} := \bigcup_{m \in F^{n,2}} X^n_{h_n k, h_n l, m}$. Since η is pairwise independent, then

$$\eta(C^n_{k,l}) \le \eta(X^n_{k,l} \times X) = (a^{n,1})^2;$$

so there exist at most $1/\epsilon$ ϵ-heavy cubes in $C^n_{k,l}$, and similarly for $\underline{C}^n_{h_n k, h_n l}$. If H^n_ϵ, $\underline{H}^n_\epsilon$ are the complements of L^n_ϵ, $\underline{L}^n_\epsilon$ in R^n, S^n respectively then, if we have chosen ι small enough, we can state that

$$\overline{\lim}_n \eta(h_n H^n_\epsilon \cap \underline{H}^n_\epsilon) > 0. \tag{34}$$

Since η is pairwise independent then for any $m \in B^n$, $m' \in F^{n,2}$,

$$\eta(h_n(Y^n_{k,l,m}) \cap X^n_{h_n k, h_n l, m'}) \le a^{n,1} a^{n,2} \mu(h_n Y^n_m | X^n_{m'}).$$

Now the image of an ϵ-heavy cube in $C^n_{k,l}$ under h_n can intersect at most $1/\epsilon$ heavy cubes in $\underline{C}^n_{h_n k, h_n l}$. Thus

$$\begin{aligned}
\mu(h_n H^n_\epsilon \cap \underline{H}^n_\epsilon) &= \sum_{k,l \in A^n} \mu(h_n(C^n_{k,l} \cap H^n_\epsilon) \cap (\underline{C}^n_{h_n k, h_n l} \cap \underline{H}^n_\epsilon)) \\
&\le \sum_{k,l \in A^n} \frac{1}{\epsilon^2} a^{n,1} a^{n,2} \sup_{\{m \in B^n, m' \in F^{n,2}\}} \mu(h_n Y^n_m | X^n_{m'}) \\
&\le |F^{n,1}|^2 \frac{1}{\epsilon^2} a^{n,1} a^{n,2} \sup_{\{m \in B^n, m' \in F^{n,2}\}} \mu(h_n Y^n_m | X^n_{m'}).
\end{aligned}$$

The sets R^n and S^n were built in such a way (Property 3 in the definition of D-approximation) that

$$\overline{\lim}_n \sup_{\{m \in B^n, m' \in F^{n,2}\}} \mu(h_n Y^n_m | X^n_{m'}) = 0.$$

Since ϵ is fixed, and we are assuming that $0 < \underline{\lim} |F^{n,1}|/|F^{n,2}| < \infty$, then

$$\mu(h_n H^n_\epsilon \cap \underline{H}^n_\epsilon) \to 0.$$

This contradicts (34). □

Now we can proceed to show the following.

THEOREM 4. *If $\eta \in M(2,3)$ is ergodic and $Di(\eta) > 0$ or $\underline{Di}(\eta) > 0$, then $\eta = \mu^3$.*

Proof. Assume that $Di(\eta) > 0$ (the case where $\underline{Di}(\eta) > 0$ is similar). Let $\epsilon_n \to 0$; by passing to a subsequence if necessary, we can assume that $\lim_n \eta(L^n_{\epsilon_n}) > 0$. Using Lemma 5, we choose a δ small enough so that, if

$$L^n := L^n_{\epsilon_n} \cap \mathcal{O}^n(\delta),$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

then $\overline{\lim}\ \eta(L^n) > 0$. We will now start slicing L^n into '2-dimensional' slices and then '1-dimensional' fibres. If for $m \in B^n$ we let

$$L^n(m) := \bigcup_{\{(k,l)\in A^n\times A^n : Y^n_{k,l,m}\subset L^n\}} Y^n_{k,l,m}$$

and

$$L^n(m,\delta) := \bigcup_{\{(k,l)\in A^n\times A^n : Y^n_{k,l,m}\subset L^n\}} \mathcal{O}(Y^n_{k,l,m}),$$

then

$$\eta(L^n) = \sum_{m\in B^n} \eta(L^n(m)) \le \frac{1}{\delta|F^{n,1}|}\sum_{m\in B^n}\eta(L^n(m,\delta))$$

and, since $|B^n| \ge \delta|F^{n,1}|$, then, taking limits, there exists a sequence $\{m_n\}$ such that (passing to a subsequence if necessary)

$$\lim_n \eta(L^n(m_n,\delta)) > 0.$$

Now

$$L^n(m_n,\delta) = \bigcup_{l\in A^n} L^n(l,m_n,\delta);$$

here $L^n(l,m_n,\delta)$ is the fibre $\bigcup_{\{k:Y^n_{k,l,m_n}\in L^n(m_n,\delta)\}}\mathcal{O}(Y^n_{k,l,m_n})$. Thus we can find a d small enough so that we can pass down to a subset of these fibres indexed by $A^{n,*}\subset A^n$ each satisfying $\eta(L^n(l,m_n,\delta)) > da^{n,1}$ and also so that

$$\lim_n \eta\left(\bigcup_{l\in A^{n,*}} L^n(l,m_n,\delta)\right) > 0.$$

Let $L^n(m_n,\delta) := \bigcup_{l\in A^{n,*}} L^n(l,m_n,\delta)$ be the union of these suitable fibres. This will ensure later on that the Blum–Hanson coefficients are uniformly small.

Note that $L^n(m_n,\delta)$ is approximately invariant: if $g\in G$ and $K\in\mathbb{N}$ are given, choose n so large that $|g^{-1}F^{n,1}\cap F^{n,1}| > (1-\delta/K)|F^{n,1}|$. If $(k,l,m_n)\in L^n(m_n)$, then $\mathcal{O}(Y^n_{k,l,m_n})$ is at least δ-long, so at least $\delta(1-3/K)|F^{n,1}|$ of the cubes in $\mathcal{O}(Y^n_{k,l,m_n})$ stay in $\mathcal{O}(Y^n_{k,l,m_n})$ after g acts on them. This is true for the orbit of any $Y^n_{k,l,m}$ in $L^n(m_n)$. By letting $K\to\infty$, $L^n(m_n,\delta)$ becomes increasingly invariant. Thus, by Lemma 1,

$$\eta_{L^n(m_n,\delta)} \to_n \eta.$$

We can write

$$\eta_{L^n(m_n,\delta)} = \sum_{l\in A^{n,*}} a^n_l \eta^n_{l,m_n},$$

where

$$\eta^n_{l,m_n} := \eta_{L^n(l,m_n,\delta)}$$

and

$$a^n_l := \frac{\eta(L^n(l,m_n,\delta))}{\eta(L^n(m_n,\delta))}.$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

We can write η^n_{l,m_n} as $\sum_{k\in A^n} b^n_k \eta^n_{k,l,m_n}$, where

$$\eta^n_{k,l,m_n} := \eta_{\mathcal{O}(Y^n_{k,l,m_n})}$$

and $b^n_k := \eta(\mathcal{O}(Y^n_{k,l,m_n}))/\eta(L^n(l,m_n,\delta))$. Note that

$$b^n_k \le \frac{\epsilon_n (a^{n,1})^2 \delta |F^{n,1}|}{d a^{n,1}} < \text{constant} \cdot \epsilon_n \to 0$$

uniformly for k as $n \to \infty$. This will justify the use of the Blum–Hanson Theorem shortly.

Now define the measures $\tau^n_{k,l,m} := (\mu^3)_{\mathcal{O}(Y^n_{k,l,m})}$, and also

$$\lambda^n := \sum_{l\in A^{n,*}} a^n_l \sum_{k\in A^n} b^n_k \tau^n_{k,l_n,m_n}.$$

We can show, in a way similar to what we did in Lemma 6 (by comparing λ^n to $\eta_{L^n(m_n,\delta)}$ on $A\times B\times C$ where A, B, C are unions of levels of the towers, and then using approximation arguments), that

$$\lambda^n \to \eta.$$

Let

$$\tilde{\mathcal{O}}(X^n_{k,l,m}) := \bigcup_{\{g\in F^{n,1}k^{-1}\cap F^{n,1}l^{-1}\}} g X^n_{k,l,m} = \bigcup_{\{g\in F^{n,1}k^{-1}\cap F^{n,1}l^{-1}\}} X^n_{gk} \times X^n_{gl} \times g X^n_m$$

and

$$\theta^n_{k,l,m} := \mu^3_{\tilde{\mathcal{O}}(X^n_{k,l,m})}.$$

Note that $\theta^n_{k,l,m} \circ (k^{-1}e_1 \times i \times i) = \theta^n_{e_1,l,m}$ and

$$\frac{\mu^3(\mathcal{O}(X^n_{e_1,l,m}))}{\mu^3(\mathcal{O}(X^n_{k,l,m}))} \le \frac{|F^n|}{\delta|F^n|} = \frac{1}{\delta};$$

hence $f^n_m := d\tau^n_{k,l,m}/d\theta^n_{l,l,m} \le 1/\delta$. We have

$$\begin{aligned}
&\tau^n(A\times B\times C)\\
&\quad = \sum_l a^n_l \sum_m b^n_{l,m} \left(\int \chi_{A\times B\times C}(x,y,z)\, d\tau^n_{e,l,m}(x,y,z)\right)\\
&\quad = \sum_l a^n_l \sum_m b^n_{l,m} \left(\int \chi_{A\times B\times C}(x,y,z)\, f_{m,n}(x,y,z)\, d\theta^n_{e,l,m}(x,y,z)\right)\\
&\quad \le \frac{1}{\delta} \sum_l a^n_l \sum_m b^n_{l,m} \left(\int \chi_{A\times B\times C}(x,y,z)\, d\theta^n_{e,l,m}(x,y,z)\right)\\
&\quad = \frac{1}{\delta} \sum_l a^n_l \sum_m b^n_{l,m} \int \chi_{A\times B}(x,y)(\chi_C(z)-\mu(C))\, d\theta^n_{e,l,m}(x,y,z)\\
&\qquad + \frac{1}{\delta} \sum_l a^n_l \sum_m b^n_{l,m} \int \mu(C)\chi_{A\times B}(x,y)\, d\theta^n_{e,l,m}(x,y,z)\\
&\quad \le \frac{1}{\delta} \sum_l a^n_l \left| \sum_m b^n_{l,m} \int \chi_{A\times B}(x,y)(\chi_C(z)-\mu(C))\, d\theta^n_{e,l,m}(x,y,z)\right|\\
&\qquad + \frac{1}{\delta} \sum_l a^n_l \sum_m b^n_{l,m} \int \mu(C)\chi_{A\times B}(x,y)\, d\theta^n_{e,l,m}(x,y,z).
\end{aligned}$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

Now the second term

$$\frac{1}{\delta}\sum_l a_l^n \sum_m b_{l,m}^n \int \mu(C)\chi_{A\times B}(x,y)\,d\theta_{e,l,m}^n(x,y,z)$$

is just $(1/\delta)\mu(C)\sum_l a_l^n \sum_m b_{l,m}^n \theta_{e,l,m}^n(A\times B\times X) \to_n (1/\delta)\mu(C)\mu^2(A\times B)$. To see this, note that the probability measures are given by

$$\begin{aligned}\theta^n(A\times B) &:= \sum_l a_l^n \sum_m b_{l,m}^n \theta_{e,l,m}^n(A\times B\times X)\\ &= \sum_l a_l^n \mu^2_{\left(\bigcup_{g\in F^{n,1}\cap F^{n,1}l^{-1}} X_{ge,gl}^n\right)}(A\times B)\\ &= \sum_l \frac{a_l^n}{|F^n a\cap F^n a l^{-1}|f_n^2}\mu^2\left(A\times B\bigcap\left(\bigcup_{g\in F^{n,1}\cap F^{n,1}l^{-1}} X_{ge,gl}^n\right)\right)\end{aligned}$$

and

$$\frac{a_l^n}{|F^{n,1}\cap F^{n,1}l^{-1}|f_n^2} \le \frac{a_l^n}{\delta|F^{n,1}|f_n^2} \le \frac{|F^{n,1}|f_n^2}{\nu(\mathcal{O}^{n,*}(e,\delta))\delta|F^{n,1}|f_n^2} \le K$$

for all n, l. Thus a weak-* limit of the measures θ^n has to be absolutely continuous with respect to μ^2, as well as being invariant. By ergodicity of μ^2, the limit in fact is μ^2.

As for the first term, we have

$$\begin{aligned}&\frac{1}{\delta}\sum_l a_l^n \left|\sum_m b_{l,m}^n \int \chi_{A\times B}(x,y)(\chi_C(z)-\mu(C))\,d\theta_{e,l,m}^n(x,y,z)\right|\\ &\quad= \frac{1}{\delta}\sum_l a_l^n \left|\sum_m b_{l,m}^n \int \chi_{A\times B}(x,y)(\chi_C(m^{-1}z)-\mu(C))\,d\theta_{e,l,e}^n(x,y,z)\right|\\ &\quad\le \frac{1}{\delta}\sum_l a_l^n \int \left|\sum_m b_{l,m}^n(\chi_{mC}(z)-\mu(C))\chi_{A\times B}(x,y)\right| d\theta_{e,l,e}^n(x,y,z)\\ &\quad\le \frac{1}{\delta}\sum_l a_l^n \int \left|\sum_m b_{l,m}^n(\chi_{mC}(z)-\mu(C))\,d\pi_3\theta_{e,l,e}^n(x,y,z)\right|\\ &\quad\le \frac{1}{\delta^2}\sum_l a_l^n \int \left|\sum_m b_{l,m}^n\chi_{mC}(z)-\mu(C)\right| d\mu\\ &\quad\le \frac{1}{\delta^2}\sum_l a_l^n \left\|\sum_m b_{l,m}^n\chi_{mC}-\mu(C)\right\|_{2,\mu}\end{aligned}$$

and, by the Blum–Hanson Theorem, this last term tends to 0 as $n\to\infty$. This completes the proof when η is ergodic. □

We apply Theorem 4 to the pairwise independent components η'_{g_j,k_j} (which we agreed to call η^j) of the chosen induced joinings η_{g_j,k_j}. If η^j is not necessarily ergodic, we only have a component of η^j being product measure. For, supposing that $Di(\eta^j) > 0$, we proceed as in the proof of Theorem 4, but this time we can only say that

$$\eta^j \ge \eta^j(L^n)(\eta^j)_{L^n}.$$

Theorem 4 shows that $\lim_n(\eta^j)_{L^n} = \mu^3$. But now suppose that

$$\eta^j = k_j\mu^3 + \theta^j$$

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

is the Radon–Nikodym decomposition of η^j with respect to product measure. As we argued in Lemma 6, θ^j satisfies

$$\overline{\lim}_n \theta^j(L^n) = 0,$$

or else we could repeat the procedure above to show that $Di(\theta_j) > 0$ and then that θ^j has a component which is product measure, a contradiction, since $\mu^3 \perp \theta^j$. Hence

$$k_j = \lim_n \frac{\eta^j(L^n)}{\mu^3(L^n)} \geq c'_j,$$

since the induced joinings were chosen so that $\lim_n \eta^j(L^n) \geq c'_j \lim_n \mu^3(L^n)$; and so $\eta^j \geq c'_j \mu^3$. Thus

$$\eta_{g_j,k_j} \geq c_j \eta'_{g_j,k_j} \geq c_j c'_j \mu^3.$$

Now we can apply Lemma 3 to get the long-awaited fact that $\nu = \mu^3$.

Proof of Theorem 1. We give the argument for $k = 3$. If $\mathcal{X}$ is not 3-mixing then there exist measurable sets A, B, C and sequences $\{k_j\}$ and $\{l_j\}$ of group elements, both tending to infinity, and $\epsilon > 0$ such that $k_j^{-1} l_j \to_j \infty$, satisfying

$$|\mu(A \cap l_j B \cap k_j C) - \mu(A)\mu(B)\mu(C)| \geq \epsilon$$

for all j. Consider the joining $\Delta_{k_j,l_j}(E_1 \times E_2 \times E_3) := \mu(E_1 \cap l_j E_2 \cap k_j E_3)$. If Δ^* is a limit point of the sequence $\{\Delta_{k_j,l_j}\}$, then $\Delta^* \neq \mu^3$. On the other hand, 2-fold mixing implies that $\Delta^* \in M(2, 3)$. This contradicts Theorem 2. □

Acknowledgements. The author would like to thank J. Choksi and A. del Junco, and especially I. Klemes, for several discussions.

References

[1] J. R. Blum and D. L. Hanson. On the mean ergodic theorem for subsequences. *Bull. Amer. Math. Soc.* **66** (1960), 308–311.

[2] S. Ferenczi. Systèmes de rang un gauche. *Ann. Inst. Henri Poincaré* **21**(2) (1985), 177–186.

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

[4] B. Host. Mixing of all orders and pairwise independent joinings of systems with singular spectrum. *Israel J. Math.* **76** (1991), 289–298.

[5] A. del Junco. A simple map with no prime factors. *Israel J. Math.* **104** (1998), 301–320.

[6] A. del Junco and D. Rudolph. On ergodic actions whose self-joinings are graphs. *Ergod. Th. & Dynam. Sys.* **7** (1987), 531–557.

[7] A. del Junco and R. Yassawi. Multiple mixing and rank one group actions. *Can. J. Math.* **52**(2) (2000), 332–347.

[8] S. A. Kalikow. Twofold mixing implies threefold mixing for rank one transformations. *Ergod. Th. & Dynam. Sys.* **4** (1984), 237–259.

[9] F. Ledrappier. Un champ markovien peut être d'entropie nulle et mélangeant. *C. R. Acad. Sci. Paris Ser. A* **287** (1978), 561–562.

[10] J. J. van Lint and R. M. Wilson. *A Course in Combinatorics.* Cambridge University Press, 1992.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press

[11] D. S. Ornstein, D. J. Rudolph and B. Weiss. Equivalence of measure preserving transformations. *Mem. Amer. Math. Soc.* **262** (1982).
[12] V. V. Ryzhikov. Mixing, rank, and minimal self-joining of actions with an invariant measure. *Russ. Acad. Sci. Sb. Math.* **75**(2) (1993), 405–427.
[13] V. V. Ryzhikov. Joinings and multiple mixing of actions of finite rank. Translation in *Funct. Anal. Appl.* **27**(2) (1992), 128–140.
[14] A. A. Tempel'man. Ergodic theorems for general dynamical systems. English translation in *Sov. Math. Dokl.* **8**(5) (1967), 1213–1216.

https://doi.org/10.1017/S0143385702001657 Published online by Cambridge University Press