Canad. J. Math. Vol. **52** (2), 2000 pp. 332–347

Multiple Mixing and Rank One Group Actions

Andrés del Junco and Reem Yassawi

Abstract. Suppose G is a countable, Abelian group with an element of infinite order and let $\mathcal{X}$ be a mixing rank one action of G on a probability space. Suppose further that the Følner sequence $\{F_n\}$ indexing the towers of $\mathcal{X}$ satisfies a "bounded intersection property": there is a constant p such that each $\{F_n\}$ can intersect no more than p disjoint translates of $\{F_n\}$. Then $\mathcal{X}$ is mixing of all orders. When $G = \mathbf{Z}$, this extends the results of Kalikow and Ryzhikov to a large class of "funny" rank one transformations. 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. This method generalizes Ryzhikov's technique.

1 Introduction

In this paper we discuss the question of whether mixing implies multiple mixing for certain rank one group actions. Rohlin [9] first asked this question in the case of measure-preserving transformations, that is when the group is $\mathbf{Z}$. Kalikow [7] showed that rank one mixing transformations were 3-mixing, and Host [3] proved that mixing transformations with singular spectrum are mixing of all orders. Ryzhikov [10] shows that rank one mixing transformations are mixing of all orders by showing that pairwise independent self-joinings of the rank one system are necessarily product measure. Here we generalize Ryzhikov's result to certain rank one group actions.

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 $T_g \colon X \to X$; however we will mostly use g to denote both the element of the group, and the measure-preserving transformation it represents. A sequence $\{F_n\}_{n=1}^{\infty}$ of finite subsets of G is (left) *Følner* if $\forall g \in G$,

$$\lim_{n\to\infty} \frac{|gF_n \Delta F_n|}{|F_n|} = 0.$$

We will say that $\{F_n\}$ satisfies the *bounded intersection property* if there exists a number p such that whenever $g_1F_n, \ldots, g_kF_n$ are disjoint translates of F_n with $g_iF_n \cap F_n \neq \varnothing$ then $k \leq p$. We will call such a p an *intersection bound* for $\{F_n\}$. Examples of G's which have natural Følner sequences satisfying this condition are $\mathbf{Z}^n$, countable direct sums of finite cyclic groups and direct sums of the these two cases.

The Følner sequence $\{F_n\}$ satisfies the *Tempel'man condition* if $F_n \subset F_{n+1}$ for each n and there exists $K \in \mathbf{N}$ such that $|F_nF_n^{-1}| \leq K|F_n|$ for all $n \in \mathbf{N}$. We mention the Tempel'man condition [12] 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.

Received by the editors January 1998; revised March 17, 1999 and September 2, 1999.
AMS subject classification: 28D15.

$\mathcal{X}$ is a *rank one* group action if there exists a Følner sequence $\{F_n\}$ and a sequence of measurable partitions

$$\mathcal{P}_n := \{\{X_g^n\}_{g\in F_n}, X \setminus X_n\}$$

of X, where $X_n := \bigcup_{g\in F_n} X_g^n$, such that

1. $\mu(X \setminus X_n) \to 0$,
2. $hX_g^n = X_{hg}^n$ whenever $g \in F_n \cap h^{-1}F_n$,
3. For each measurable set A, there exists a sequence $\{A_n\}$ where A_n is a union of elements of $\{X_g^n\}_{g\in F_n}$ and $\mu(A\Delta A_n) \to 0$.

We will use f_n to denote the common value of $\mu(X_g^n)$. 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 $\{F_n\}$ the Følner sequence associated with the rank one group action $\mathcal{X}$ and write $(\mathcal{X}, \{F_n\})$ for $\mathcal{X}$ when we wish to specify $\{F_n\}$. We will say that $(\mathcal{X}, \{F_n\})$ has the bounded intersection property if $\{F_n\}$ has the bounded intersection property. For a general scheme for constructing rank one group actions, see [4]. If $(\mathcal{X}, \{F_n\})$ is a rank one action, then without loss of generality we can assume that the identity element e of G belongs to F_n for each n. When $G = \mathbf{Z}$, what we are calling a rank one action is in fact called a *"funny rank one"* action and is more general than the classical notion, which requires that the F_n be intervals in $\mathbf{Z}$. Ferenczi [2] constructs a funny rank one transformation which is not loosely Bernoulli, and therefore not of finite rank. His example in fact has the bounded intersection property with $p = 4$.

If $\{g_n\} \subset G$ we write $g_n \to \infty$, if whenever $V \subset G$ is finite then there exists an N such that $g_n \in G \setminus V$ for $n \geq N$. A group action $\mathcal{X}$ is (2)-*mixing* if

$$\lim_{n\to\infty} \mu(A_1 \cap g_n A_2) = \mu(A_1)\mu(A_2)$$

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

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

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

Our main theorem is

Theorem 1 *Suppose G has an element of infinite order and $\mathcal{X}$ is a mixing rank one G-action with the bounded intersection property. Then $\mathcal{X}$ is mixing of all orders.*

This result applies, for example, when G is the direct sum of $\mathbf{Z}^n$ with finitely or countably many finite cyclic groups. We remark that it is possible to modify Ferenczi's construction to yield mixing funny rank one $\mathbf{Z}$-actions with the bounded intersection property which are not loosely Bernoulli. Thus even in the case $G = \mathbf{Z}$ our result applies in some situations where Ryzhikov's result for finite rank [11] does not.

To prove Theorem 1, we use the method of joinings, generalizing [10].

The measure λ is a *2-joining* of $\mathcal{X}$ and $\mathcal{Y}$ if $(X \times Y, \mathcal{B} \otimes \mathcal{F}, \lambda, G)$ is a measure-preserving group action with the additional condition that

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

and

$$\lambda(X \times B) = \nu(B)$$

for $A, B \in \mathcal{B}, \mathcal{F}$ respectively. For a detailed account of joinings, see [6]. If $\{(X_i \mathcal{B}_i, \nu_i)\}_{i=1}^n$ are n probability spaces and λ is a measure on $(\Pi_{i=1}^n X_i, \bigotimes_{i=1}^n \mathcal{B}_i)$, then we define $\pi_{i_1,i_2,\ldots,i_k}\lambda$ to be the projection of λ on $(\Pi_{j=1}^k X_{i_j}, \bigotimes_{j=1}^k \mathcal{B}_{i_j})$. λ is an *n-joining* of $\{(X_i, \mathcal{B}_i, \nu_i, G)\}_{i=1}^n$ if $(\Pi_{i=1}^n X_i, \bigotimes_{i=1}^n \mathcal{B}_i, G, \lambda)$ is a measure-preserving system so that $\pi_k \lambda = \nu_k$ for $k = 1, \ldots, n$. For $n > 2$ it is natural to impose stronger conditions on λ: In particular, if λ is an n-joining of $\{X_i\}_{i=1}^n$, then we can require that

$$\pi_{i_1,i_2,\ldots,i_k}\lambda = \Pi_{j=1}^k \nu_{i_j}$$

whenever $i_1, i_2, \ldots, i_k$ are k distinct elements of $\{1, 2, \ldots, n\}$. In this case we write $\lambda \in M(k, n)$ and say λ is k-fold independent. Note that $M(k, n) \subset M(1, n) \,\forall k \geq 1$, and that $M(1, n)$ (and $M(k, n)$) are subsets of $M(X^n)$, the set of all probability measures on X^n. We define a topology on $M(1, n)$ where $\lambda_j \to_j \lambda$ if and only if $\lambda_j(A_1 \times A_2 \times \cdots \times A_n) \to_j \lambda(A)$. This turns $M(1, n)$ into a compact metric space if $(X, \mathcal{B}, \mu)$ is regular, *i.e.*, if X is a compact metric space and $\mathcal{B}$ is the Borel σ-algebra. This topology in fact coincides with the weak-star topology on $M(1, n)$ as a subset of the Borel measures on the compact metric space X^n.

Theorem 1 follows from the following theorem:

Theorem 2 *Suppose $\mathcal{X}$ is a G-action on a regular measure space satisfying the hypotheses of Theorem 1. Then $M(2, k) = \{\mu^k\}$ for $k > 2$.*

Note that regularity is required in the statement of Theorem 2, but not Theorem 1. The proof of Theorem 1 clarifies this.

Proof of Theorem 1 The rank one 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 point action. This shows that there is no harm in assuming that $\mathcal{X}$ itself is regular. (Like most dynamical notions, the concepts of mixing and rank one are obviously invariant under isomorphism of actions at the level of measure algebras. The concept of a joining however is not.)

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

$$|\mu(A \cap j_n B \cap k_n C) - \mu(A)\mu(B)\mu(C)| \geq \epsilon$$

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

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. Throughout the remainder of the paper we will assume that the underlying measure space is regular and convergence of measures will always mean weak-star convergence.

The second author thanks Jal Choksi and Ivo Klemes for several helpful discussions. This work forms part of the second author's PhD thesis.

2 Preliminaries

We have the following lemma for rank one mixing group actions:

Lemma 1 *Let $\mathcal{X}$ be a rank one mixing G-action. Then*

$$\lim_{n\to\infty} \max_{g\neq e} \frac{\mu(gX_e^n \cap X_e^n)}{\mu(X_e^n)} = 0.$$

Proof Suppose the lemma is false, *i.e.*, there exist sequences $\{n_k\} \subset \mathbf{N}$ and $\{g_k\} \subset G$ such that $g_k \neq e$ and

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

We note that $g_k \to \infty$, because if the g_k do not diverge, we can pass to a subsequence along which g_k is constantly $g \neq e$. But since $\{F_{n_k}\}$ is Følner, for some k there exists $f \neq f'$, both in F_{n_k} such that $fg = f'$. Therefore, $\mu(gX_e^{n_k} \cap X_e^{n_k}) = \mu(f^{-1}f'X_e^{n_k} \cap X_e^{n_k}) = \mu(X_{f'}^{n_k} \cap X_f^{n_k}) = 0$, a contradiction.

Choose a set A with $0 < \mu(A) < c$, and for each k choose a set $F_{n_k}(A) \subset F_{n_k}$ so that the sets

$$A_{n_k} = \bigcup_{f \in F_{n_k}(A)} X_f^{n_k}$$

satisfy $\mu(A \triangle A_{n_k}) \to 0$.

We have

$$\mu(g_k A \cap A) \to \big(\mu(A)\big)^2$$

and

$$|\mu(g_k A \cap A) - \mu(g_k A_{n_k} \cap A_{n_k})| \to 0,$$

so

$$\mu(g_k A_{n_k} \cap A_{n_k}) \to \big(\mu(A)\big)^2.$$

But also

$$\begin{aligned}\mu(g_k A_{n_k} \cap A_{n_k}) &\geq |F_{n_k}(A)| \mu(g_k X_e^{n_k} \cap X_e^{n_k}) \\ &= \mu(X_e^{n_k}) |F_{n_k}(A)| \frac{\mu(g X_e^{n_k} \cap X_e^{n_k})}{\mu(X_e^{n_k})} \\ &= \mu(A_{n_k}) \frac{\mu(g X_e^{n_k} \cap X_e^{n_k})}{\mu(X_e^{n_k})} \to c\mu(A),\end{aligned}$$

a contradiction. ■

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

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

The next simple lemma is used repeatedly in [10].

Lemma 2 *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) $\lim_n \nu(gI_n \Delta I_n) = 0$ *for each* $g \in G$.

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

Proof

$$\nu \geq \nu(I_n)\nu_{I_n}$$

so if λ is any limit point of the sequence $\{\nu_{I_n}\}$, then λ is G-invariant by (2), and $\nu \geq c\lambda$ by (1). By the ergodicity of ν, $\lambda = \nu$. This proves the lemma. ■

We will also need a version of the Blum-Hanson Theorem for group actions [1]. If $\phi \in \mathcal{L}_2(\mu)$, and $g \in G$, then we write $\phi \circ g$ to denote the function $\phi \circ T_g$.

Theorem 3 *Let $\mathcal{X}$ be a mixing action of a countable Abelian group G. Suppose that $\{a^n\}_{n\in\mathbf{N}}$ is a sequence of non-negative valued functions $a^n \colon G \to [0, \infty)$ satisfying*

(1) $\sum_{g\in G} a^n(g) = 1 \, \forall n \in \mathbf{N}$
(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/10.4153/CJM-2000-015-0 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 set $\mathcal{O}_\epsilon$ such that $|\langle \phi \circ g, \phi \rangle| < \epsilon/2$ whenever g is in the complement of $\mathcal{O}_\epsilon$. 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 $\langle \sum_{g\in G} a^n(g)\phi\circ g, \sum_{g\in G} a^n(g)\phi\circ g\rangle$ 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

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

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

3 Proof of Theorem 2

In this section $(\mathcal{X}, \{F_n\})$ is a rank one mixing G-action as in Theorem 2 and $\nu \in M(2,3)$. Our aim is to show that ν must be product measure μ^3.

For $(k, l, m) \in F_n^3$ we denote $X_k^n \times X_l^n \times X_m^n$ by $X_{k,l,m}^n$ and call this object a cube. By an orbit in G^3 we mean any

$$G^3_{g_0,g_1} := \{(k, kg_0, kg_1) : k \in G\},$$

that is, an orbit of the diagonal action of G on G^3 by translation. By an *n-orbit* in X_n^3 we mean

$$\mathcal{O}^n_{g_0,g_1} := \bigcup \{X^n_{\mathbf{g}} : \mathbf{g} \in G^3_{g_0,g_1} \cap F_n^3\} = \bigcup \{X^n_{(k,kg_0,kg_1)} : k \in F_n \cap F_n g_0^{-1} \cap F_n g_1^{-1}\}.$$

Thus X_n^3 is partitioned into n-orbits and each n-orbit is a union of cubes of equal ν-measure. The length of an n-orbit is the number of cubes in it. By a strand of $\mathcal{O}^n_{g_0,g_1}$ we mean its intersection with the G-orbit of any $(x, y, z) \in \mathcal{O}^n_{g_0,g_1}$ and by a slice of $\mathcal{O}^n_{g_0,g_1}$ we mean any measurable union of strands of $\mathcal{O}^n_{g_0,g_1}$. For $(k, l, m) \in F_n^3$, $\mathcal{O}(X^n_{k,l,m})$ will denote the n-orbit containing the cube $X^n_{k,l,m}$.

Let us say an n-orbit is δ-long if its length is at least $\delta|F_n|$ and let $\mathcal{O}_n(\delta)$ denote the union of the δ-long n-orbits.

Lemma 3 *If p is an intersection bound for $\{F_n\}$ then*

$$\liminf_n \nu\big(\mathcal{O}_n(\delta)\big) \geq 1 - p^2\delta.$$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

Proof Fixing n, we will consider all finite subsets $\gamma = \{g_1, \ldots, g_r\}$ of G with the properties that the $\{F_n g_i\}$ are pairwise disjoint and all intersect F_n non-trivially, so $|\gamma| \leq p$. We call such a γ a *configuration* and let Γ denote the space of all configurations. Clearly Γ is finite.

For $x \in X$ we let $R_n(x) = \{g \in G : gx \in X_e^n\}$, the set of "return times" to the base of the n-th tower. By an *n-block* in x we mean any gF_n with $g \in R_n(x)$, and g is called the *base time* of this n-block. The n-blocks in x are disjoint translates of F_n. For $x \in X_n$ we denote by $B_n(x)$ the n-block in x containing $e \in G$, namely $B_n(x) = k^{-1}F_n$ if $x \in X_k^n$.

For $(x, y) \in X_n \times X$ we create a configuration $\gamma(x, y)$ by letting $g_1, \ldots, g_r$ denote the base times of the n-blocks in y which intersect $B_n(x)$ and defining $\gamma(x, y) = \{g^{-1}g_1, \ldots, g^{-1}g_r\}$, where g is the base time of $B_n(x)$. For $(x, y, z) \in X_n \times X^2$ 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 $Q\colon (x, y, z) \mapsto \big(\gamma_1(x, y, z), \gamma_2(x, y, z)\big)$ as a partition of $X_n \times X^2$ indexed by Γ^2, so we will write $Q_{(\gamma_1,\gamma_2)} = Q^{-1}\big((\gamma_1, \gamma_2)\big)$. Thus we are partitioning $X_n \times X^2$ according to the pattern, up to a shift, formed by $B_n(x)$ and the n-blocks in y and z which intersect $B_n(x)$.

For $(\gamma_1, \gamma_2) \in \Gamma^2$ let

$$S(\gamma_1, \gamma_2) = \bigcup \{F_n \cap g_1 F_n \cap g_2 F_n : g_1 \in \gamma_1, g_2 \in \gamma_2\}$$

and for $k \in S(\gamma_1, \gamma_2)$ let

$$Q_{(\gamma_1,\gamma_2),k} = \{(x, y, z) \in Q_{(\gamma_1,\gamma_2)} : x \in X_k^n\}.$$

We will refer to the $Q_{(\gamma_1,\gamma_2),k}$ as the cells of $Q_{(\gamma_1,\gamma_2)}$. The cells of a given $Q_{(\gamma_1,\gamma_2)}$ all have the same ν-measure as they are mapped to each other by the action of G. Moreover for $g_1 \in \gamma_1$, $g_2 \in \gamma_2$

$$Q_{(\gamma_1,\gamma_2)} \cap \mathcal{O}^n_{g_1,g_2} = \bigcup \{Q_{(\gamma_1,\gamma_2),k} : k \in F_n \cap {g_1}^{-1} F_n \cap {g_2}^{-1} F_n\}$$

is a slice of $\mathcal{O}^n_{g_1,g_2}$ and $Q_{(\gamma_1,\gamma_2)} \cap X_n^3$ is a disjoint union of such slices. The conditional measure $\nu_{Q_{(\gamma_1,\gamma_2)}}\big(X_n^3 \setminus \mathcal{O}_n(\delta)\big)$ of the short orbit slices in $Q_{(\gamma_1,\gamma_2)}$ is the number of cells in $Q_{(\gamma_1,\gamma_2)}$ belonging to short orbit slices divided by the total number of cells in $Q_{(\gamma_1,\gamma_2)}$. Thus

$$\nu_{Q_{(\gamma_1,\gamma_2)}}\big(X_n^3 \setminus \mathcal{O}_n(\delta)\big) \leq \frac{|\gamma_1|\,|\gamma_2|\delta|F_n|}{|F_n|} \leq p^2\delta. \tag{1}$$

Since $X_n^3 \setminus \mathcal{O}_n(\delta)$ is partitioned by the sets $Q_{(\gamma_1,\gamma_2)}$, we get

$$\nu\big(X_n^3 \setminus \mathcal{O}_n(\delta)\big) \leq p^2\delta.$$

Since ν is a joining,

$$\nu(X_n^3) \geq 1 - 3\mu(X \setminus X_n) \to 1 \quad \text{as} \quad n \to \infty,$$

so

$$\nu(X^3 \setminus X_n^3) \to 0 \quad \text{as} \quad n \to \infty.$$

Thus

$$\begin{aligned}\limsup_{n\to\infty} \nu\big(X^3 \setminus \mathcal{O}_n(\delta)\big) &= \limsup_{n\to\infty}\big[\nu(X^3 \setminus X_n^3) + \nu\big(X_n^3 \setminus \mathcal{O}_n(\delta)\big)\big] \\ &= \limsup_{n\to\infty}\big[\nu\big(X_n^3 \setminus \mathcal{O}_n(\delta)\big)\big] \\ &\le p^2\delta,\end{aligned}$$

and so

$$\liminf_{n\to\infty} \nu\big(\mathcal{O}_n(\delta)\big) \ge 1 - p^2\delta. \qquad \blacksquare$$

3.1 Lightness

Let $\mathcal{X}$ be rank one mixing with $\{F_j\}$ the associated Følner sequence. Recall that $\mu(X_g^j) = f_j$ for $g \in F_j$. If $\epsilon > 0$, we say that the cube $X_{k,l,m}^j$ is ϵ-light if

$$\nu(X_{k,l,m}^j) \le \epsilon f_j^2.$$

We also define

$$L_\epsilon^j = \bigcup \{X_{k,l,m}^j : X_{k,l,m}^j \epsilon - \text{light}\}.$$

We will need to know that $\nu(L_\epsilon^j)$ is substantial for large j and small ϵ. To this end, define

$$\mathrm{Di}(\nu) := \lim_{\epsilon\to 0} \liminf_{j\to\infty} \nu(L_\epsilon^j).$$

To prove that $\mathrm{Di}(\nu) > 0$, we will need the following lemma.

Lemma 4 *Suppose G has an element of infinite order, and let $\{F_j\}$ be a Følner sequence in G. If $0 < \delta < 1/4$, then there exist a sequence k_n of group elements such that $k_n \to \infty$ and for all sufficiently large n,*

$$1 - 2\delta \le \frac{|k_n F_n \cap F_n|}{|F_n|} \le 1 - \delta.$$

Proof Let $l \in G$ be an element of infinite order. Given $\epsilon > 0$, choose N so that for all $n \ge N$,

$$\frac{|F_n \cap lF_n|}{|F_n|} > 1 - \epsilon/2.$$

If $n \ge N$ and

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

then $a_0 = 1$ and $|a_{k+1} - a_k| < \epsilon$. Note also that $a_k \ne 0$ if and only if $l^k \in F_n {F_n}^{-1}$, a finite set. Thus k can be chosen large enough so that $a_k = 0$. The lemma follows immediately. $\blacksquare$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

Lemma 5 *If* $(\mathcal{X}, \{F_n\})$ *is rank one mixing and* $\nu \in M(2,3)$ *is ergodic, then* $\mathrm{Di}(\nu) > 0$.

Proof We show this by contradiction. If $\mathrm{Di}(\nu) = 0$, then letting H_ϵ^j stand for the union of the ϵ-heavy cubes (ϵ-heavy cubes are cubes which are not ϵ-light), given $\eta > 0$, we can find $\epsilon > 0$ such that

$$\lim_{j\to\infty} \nu(H_\epsilon^j) > 1 - \eta.$$

(We have dropped to a subsequence but without loss of generality we maintain the same indexing.)

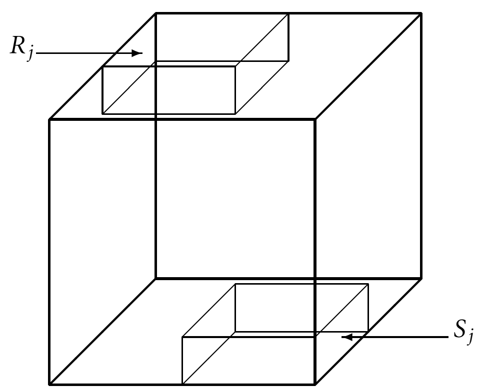

Figure 1

Fix $0 < \delta < 1/8$. If

$$A_j := F_j \cap k_j^{-1} F_j,$$

(the set obtained in Lemma 4) and $B_j := F_j \setminus A_j$, then $\bigcup_{k\in B_j} X_k^j$ is always of nontrivial μ-measure, greater than $\delta\mu(X_j)$. Hence for each j we can find $B_j^* \subset F_j$ such that

$$\liminf_j \mu\Big(k_j\Big(\bigcup_{k\in B_j} X_k^j\Big) \cap \Big(\bigcup_{k\in B_j^*} X_k^j\Big)\Big) \ge \delta^2$$

where $\mu(\bigcup_{k\in B_j^*} X_k^j) \ge \delta$. If we form the sets

$$R_j := \bigcup_{(k,l,m)\in A_j\times A_j\times B_j} X_{k,l,m}^j$$

and

$$S_j := \bigcup_{(k,l,m)\in k_jA_j\times k_jA_j\times B_j^*} X_{k,l,m}^j$$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

we have that

$$\nu(R_j) \geq \nu(X \times X \times \pi_3 R_j) - \nu\bigl(X \times (X \setminus \pi_2 R_j) \times \pi_3 R_j\bigr) - \nu\bigl((X \setminus \pi_1 R_j) \times X \times \pi_3 R_j\bigr)$$

and using the fact that $\nu \in M(2,3)$, we have

$$\liminf_j \nu(R_j) \geq \delta - 2 \cdot 2\delta \cdot 2\delta = \delta(1 - 8\delta) > 0.$$

Figure 1 illustrates the case when $G = \mathbf{Z}$ and $F_j = [0, h_j)$. Similarly,

$$\begin{aligned} \liminf_j \nu(k_j R_j \cap S_j) &= \liminf_j \nu\bigl(k_j \pi_1 R_j \times k_j \pi_2 R_j \times (k_j \pi_3 R_j \cap \pi_3 S_j)\bigr) \\ &\geq \lim_j \mu(X_j)\delta^2(1 - 4\delta) > 0. \end{aligned}$$

Therefore if η is chosen small enough, then

$$\liminf_{j \to \infty} \nu\bigl(k_j(R_j \cap H^j_\epsilon) \cap (S_j \cap H^j_\epsilon)\bigr) > 0.$$

If $k, l \in A_j$, we define the sets

$$C^j_{k,l} := \bigcup_{m \in B_j} X^j_{k,l,m},$$

and

$$\underline{C}^j_{k_j k, k_j l} := \bigcup_{m \in B^*_j} X^j_{k_j k, k_j l, m}.$$

Note that

$$k_j(C^j_{k,l}) \cap S_j \subset \underline{C}^j_{k_j k, k_j l}.$$

Since $\nu \in M(2,3)$, then $\nu(C^j_{k,l}) \leq f_j^2$; similarly for the lower columns $\underline{C}^j_{k,l}$. This means that any such column has at most $1/\epsilon$ heavy cubes, and that there are at most $1/\epsilon^2$ intersections of heavy cubes in the expression

$$\nu\bigl(k_j(C^j_{k,l} \cap H^j_\epsilon) \cap \underline{C}^j_{k_j k, k_j l} \cap H^j_\epsilon\bigr).$$

Hence

$$\begin{aligned} \nu\bigl(k_j(R_j \cap H^j_\epsilon) \cap (S_j \cap H^j_\epsilon)\bigr) &= \sum_{k,l \in A_j} \nu\bigl(k_j(C^j_{k,l} \cap H^j_\epsilon) \cap \underline{C}^j_{k_j k, k_j l} \cap H^j_\epsilon\bigr) \\ &\leq \frac{|A_j|^2}{\epsilon^2} \sup_{k,l \in A_j, m \in B_j, p \in B^*_j} \nu(k_j X^j_{k,l,m} \cap X^j_{k_j k, k_j l, p}). \end{aligned}$$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

We now use the fact that $\nu \in M(2,3)$ to state that

$$\begin{aligned}\nu(k_j X^j_{k,l,m} \cap X^j_{k_j k, k_j l, p}) &= \nu\big(X^j_{k_j k} \times X^j_{k_j l} \times (k_j X^j_m \cap X^j_p)\big)\\ &\leq \nu\big(X \times X^j_{k_j l} \times (k_j X^j_m \cap X^j_p)\big)\\ &= f_j \mu(k_j X^j_m \cap X^j_p)\\ &= f_j^2 \frac{\mu(p^{-1} k_j m X^j_e \cap X^j_e)}{\mu(X^j_e)}.\end{aligned}$$

Note that by choice of the sequence k_j in Lemma 4, $p^{-1}k_j m \neq e$ whenever $m \in B_j$ and $p \in F_j$. Using Lemma 1, we can conclude that

$$\nu\big(k_j(R_j \cap H^j_\epsilon) \cap (S_j \cap H^j_\epsilon)\big) \leq \frac{|A_j|^2}{\epsilon^2} f_j^2 \sup_{m,p \in B_j} \frac{\mu(p^{-1}k_j m X^j_e \cap X^j_e)}{\mu(X^j_e)} \to_j 0.$$

This gives the desired contradiction. We can now prove Theorem 2.

Proof of Theorem 2 Suppose G satisfies the conditions in the statement of Theorem 2 and $\nu \in M(2,3)$ is given so that $(X^3, \mathcal{B} \otimes \mathcal{B} \otimes \mathcal{B}, G, \nu)$ is ergodic (we will show that any ergodic $\nu \in M(2,3)$ has to be product measure. This implies, using the ergodic decomposition theorem, that any pairwise independent ν is also product measure). Lemma 5 tells us that $\mathrm{Di}(\nu) > 0$, and using Lemma 3, we choose δ so that $\liminf_n \nu\big(\mathcal{O}_n(\delta)\big) > 1 - \frac{\mathrm{Di}(\nu)}{2}$. Define

$$\mathcal{O}^*_n(\delta, \epsilon) := \mathcal{O}_n(\delta) \cap L^n_\epsilon,$$

that is $\mathcal{O}^*_n(\delta,\epsilon)$ is the union of all ϵ-light cubes whose n-orbits are of length at least $\delta|F_n|$. Given any sequence $\{\epsilon_n\} \downarrow 0$, there exists a sequence j_n such that

$$\lim_{n\to\infty} \nu\big(\mathcal{O}^*_{j_n}(\delta, \epsilon_n)\big) = d > 0.$$

After all, given $\epsilon > 0$, once j_n is sufficiently large, we get

$$\nu(L^{j_n}_\epsilon) > \frac{3}{4}\,\mathrm{Di}(\nu)$$

and

$$\nu\big(\mathcal{O}_{j_n}(\delta,\epsilon)\big) > 1 - \frac{\mathrm{Di}(\nu)}{2},$$

so

$$\nu\big(\mathcal{O}^*_{j_n}(\delta,\epsilon)\big) > \frac{\mathrm{Di}(\nu)}{4}.$$

We will assume that $j_n = n$. Let

$$\mathcal{O}_n(k,\delta) := \bigcup_{\{(l,m):X^n_{k,l,m}\subset \mathcal{O}^*_n(\delta,\epsilon_n)\}} \mathcal{O}(X^n_{(k,l,m)});$$

$\mathcal{O}_n(k,\delta)$ is the "slab" of the n-orbits of ϵ_n-light δ-long cubes at $k \in F_n$. Note that although the notation for the k-th slab makes no mention of ϵ_n (in an attempt to stop the notation from burgeoning any more), it is implicitly contained in the definition. Since

$$\lim_{n\to\infty} \nu\big(\mathcal{O}^*_n(\delta,\epsilon_n)\big) > 0$$

and

$$\nu\big(\mathcal{O}^*_n(\delta,\epsilon_n)\big) = \sum_{k\in F_n} \nu\big(\mathcal{O}^*_n(\delta,\epsilon_n) \cap (X^n_k \times X_n \times X_n)\big) \le \frac{1}{\delta|F_n|} \sum_{k\in F_n} \nu\big(\mathcal{O}_n(k,\delta)\big),$$

we can find some sequence $\{k_n\}$ such that

$$\liminf_{n\to\infty} \nu\big(\mathcal{O}_n(k_n,\delta)\big) > 0.$$

For ease of notation we can assume that $k_n = e$ for each $n \in \mathbf{N}$. Henceforth we will only be interested in $\mathcal{O}_n(e,\delta)$. Similarly define the "fibre" over $(e,l) \in F_n^2$ as

$$\mathcal{O}_n(e,l,\delta) := \bigcup_{\{m:X^n_{(e,l,m)}\in \mathcal{O}^*_n(\delta,\epsilon_n)\}} \mathcal{O}(X^n_{(e,l,m)})$$

splitting up $\mathcal{O}_n(e,\delta) = \bigcup_{l\in F_n} \mathcal{O}_n(e,l,\delta)$ as a disjoint union.

Furthermore, we can find $d^* > 0$ so that, setting

$$D_n = \{l \in F_n : \nu\big(\mathcal{O}_n(e,l,\delta)\big) \ge d^*|F_n|^{-1}\},$$

when n is sufficiently large,

$$\nu\Big(\bigcup_{l\in D_n} \mathcal{O}_n(e,l,\delta)\Big) \ge d^*.$$

For otherwise, for all $d^* > 0$ and arbitrarily large n we would get

$$\begin{aligned}\nu\big(\mathcal{O}_n(e,\delta)\big) &\le \nu\Big(\bigcup_{l\in D_n} \mathcal{O}_n(e,l,\delta)\Big) + \nu\Big(\bigcup_{l\in F_n\setminus D_n} \mathcal{O}_n(e,l,\delta)\Big)\\ &< d^* + d^*,\end{aligned}$$

contradicting the fact that

$$\liminf_{n\to\infty} \nu\big(\mathcal{O}_n(e,\delta)\big) > 0.$$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

Thus letting $\mathcal{O}_n^*(e,\delta)$ denote $\bigcup_{l\in D_n}\mathcal{O}_n(e,l,\delta)$, we have

$$\liminf_{n\to\infty}\nu\big(\mathcal{O}_n^*(e,\delta)\big)>0.$$

In parallel with this notation, we shall split up $\nu^n(\delta):=\nu_{\mathcal{O}_n^*(e,\delta)}$ into a convex combination of probability measures: letting $\nu_{l,m}^n:=\nu_{\mathcal{O}(X^n_{(e,l,m)})}$, we define

$$\nu_l^n(\delta):=\sum_{\{m:X^n_{(e,l,m)}\in\mathcal{O}_n^*(\delta,\epsilon_n)\}}b_{l,m}^n\nu_{l,m}^n$$

where $b_{l,m}^n:=\frac{\nu(\mathcal{O}(X^n_{e,l,m}))}{\nu(\mathcal{O}_n(e,l,\delta))}$. Note that $\nu_l^n(\delta)=\nu_{\mathcal{O}_n(e,l,\delta)}$ and by lightness

$$b_{l,m}^n\le\frac{\epsilon_n|F_n|^{-1}}{d^*|F_n|^{-1}}=\frac{\epsilon_n}{d^*}\to_n 0$$

for all such l, m. Similarly, we have

$$\nu^n(\delta):=\nu_{\mathcal{O}_n^*(e,\delta)}=\sum_{\{l\in D_n\}}a_l^n\nu_l^n(\delta)$$

where $a_l^n:=\frac{\nu(\mathcal{O}_n(e,l,\delta))}{\nu\left(\mathcal{O}_n^*(e,\delta)\right)}$.

Next, we define sums of restrictions of μ^3, beginning with the measures

$$\tau_{e,l,m}^n:=\mu^3_{\mathcal{O}(X^n_{(e,l,m)})}$$

but then averaging using the same weights $b_{l,m}^n$ and a_l^n defined using ν:

$$\tau_l^n:=\sum_{\{m:X^n_{(e,l,m)}\in\mathcal{O}_n^*(\delta,\epsilon_n)\}}b_{l,m}^n\tau_{e,l,m}^n$$

and

$$\tau^n:=\sum_{\{l\in D_n\}}a_l^n\tau_l^n.$$

Since $\liminf_{n\to\infty}\nu\big(\mathcal{O}_n^*(e,\delta)\big)>0$ and $\{\mathcal{O}_n^*(e,\delta)\}$ is approximately invariant ($\{\mathcal{O}_n^*(e,\delta)\}$ is the union of δ-long n-orbits, which are approximately invariant), then by Lemma 2, $\lim_{n\to\infty}\nu^n(\delta)=\nu$. Note that $\tau_{e,l,m}^n(A)=\nu_{e,l,m}^n(A)$ whenever A is a union of cubes in X_n^3. It follows that $\tau^n(A)=\nu^n(\delta)(A)$ for all such A, and since $\mathcal{X}$ is rank one, that

$$\tau^n\to_n\nu.$$

If $m\in F_n$, let

$$\tilde{\mathcal{O}}(X^n_{(e,l,m)}):=\bigcup_{\{g\in F_n\cap F_nl^{-1}\}}gX^n_{(e,l,m)}=\bigcup_{\{g\in F_n\cap F_nl^{-1}\}}X_g^n\times X_{gl}^n\times gX_m^n$$

— $\tilde{\mathcal{O}}(X^n_{(e,l,m)})$ is a longer orbit than $\mathcal{O}(X^n_{(e,l,m)})$ —and define

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

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

$$\frac{\mu^3\big(\tilde{\mathcal{O}}(X^n_{(e,l,m)})\big)}{\mu^3\big(\mathcal{O}(X^n_{(e,l,m)})\big)} \le \frac{|F_n|}{\delta|F_n|} = \frac{1}{\delta};$$

hence $f^n_{l,m} := \frac{d\tau^n_{(e,l,m)}}{d\theta^n_{(e,l,m)}} \le 1/\delta$. We have

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

Now the second term $\frac{1}{\delta}\sum_l a^n_l \sum_m b^n_{l,m} \int \mu(C)\chi_{A\times B}(x,y)\, d\theta^n_{e,l,m}(x,y,z)$ is just $\frac{1}{\delta}\mu(C)\sum_l a^n_l \sum_m b^n_{l,m}\theta^n_{e,l,m}(A \times B \times X) \to_n \frac{1}{\delta}\mu(C)\mu^2(A \times B)$. To see this, note that the probability measures

$$\begin{aligned}
\theta_n(A \times B) &:= \sum_l a^n_l \sum_m b^n_{l,m}\theta^n_{e,l,m}(A \times B \times X) \\
&= \sum_l a^n_l \mu^2_{(\bigcup_{g\in F_n\cap F_n l^{-1}} X^n_{ge,gl})}(A \times B) \\
&= \sum_l \frac{a^n_l}{|F_n \cap F_n l^{-1}| f^2_n} \mu^2\Big(A \times B \cap \Big(\bigcup_{g\in F_n\cap F_n l^{-1}} X^n_{ge,gl}\Big)\Big)
\end{aligned}$$

https://doi.org/10.4153/CJM-2000-015-0 Published online by Cambridge University Press

and

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

for all n, l. Thus a weak star 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

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

Note that by the proof of the Blum-Hanson theorem, the convergence of

$$\Big\|\sum_m b_{l,m}^n \chi_{mC}-\mu(C)\Big\|_{2,\mu}$$

depends only on C and

$$b^n := \max_m b_{l,m} \le \frac{\epsilon_n}{d^*}$$

and this bound is independent of $l \in F_n$. Hence

$$\sum_l a_l^n \Big\|\sum_m b_{l,m}^n \chi_{mC}-\mu(C)\Big\|_{2,\mu} \to_n 0.$$

Thus we have shown that $\nu \le \frac{1}{\delta}\mu^3$. By ergodicity, $\nu = \mu^3$. This completes the proof.

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 (2) **21**(1985), 177–186.
[3] B. Host, *Mixing of all orders and pairwise independent joinings of systems with singular spectrum.* Israel J. Math **76**(1991), 289–298.
[4] A. del Junco, *A weakly mixing simple map with no prime factors.* Israel J. Math., to appear.
[5] ———, *A Transformation with Simple Spectrum which is not Rank One.* Canad. J. Math. (3) **29**(1977), 655–663.
[6] A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs.* Ergodic Theory Dynamical Systems **7**(1987), 531–557.
[7] S. A. Kalikow, *Twofold mixing implies threefold mixing for rank one transformations.* Ergodic Theory Dynamical Systems **4**(1984), 237–259.
[8] D. Ornstein and B. Weiss, *The Shannon-McMillan-Breiman theorem for amenable groups.* Israel J. Math. (1) **44**(1983), 53–60.
[9] V. A. Rokhlin, *On ergodic compact Abelian groups.* Izv. Akad. Nauk. SSSR Ser. Mat. **13**(1949), 323–340 (Russian).
[10] V. V. Ryzhikov, *Mixing, Rank, and Minimal Self-Joining of Actions with an Invariant Measure.* Russian Acad. Sci. Sb. Math. (2) **75**(1993), 405–427.
[11] ———, *Joinings and multiple mixing of the actions of finite rank.* Funct. Anal. Appl. **27**(1993), 128–140.
[12] A. A. Tempel'man, *Ergodic Theorems for General Dynamical Systems.* English transl.: Soviet Math. Doklady (5) **8**(1967), 1213–1216.

Department of Mathematics
University of Toronto
Toronto, ON
M5S 1A1

Department of Mathematics and Statistics
McGill University
805 Sherbrooke Street West
Montreal, QC
H3A 2K6