ISRAEL JOURNAL OF MATHEMATICS 114 (1999), 271–288

AN APPLICATION OF ERGODIC THEORY TO A PROBLEM IN GEOMETRIC RAMSEY THEORY

BY

TAMAR ZIEGLER

Institute of Mathematics, The Hebrew University of Jerusalem
Givat Ram, Jerusalem 91904, Israel
e-mail: tamar@math.huji.ac.il

ABSTRACT

Let E be a measurable subset of $\mathbb{R}^k$, $k > 2$, with $\bar{D}(E) > 0$. Let $V = \{0, v_1, \ldots, v_{k+1}\} \in \mathbb{R}^k$, where $v_1, \ldots, v_{k+1}$ are affinely independent. We show that for r large enough, we can find an isometric copy of rV arbitrarily close to E. This is a generalization of a theorem of Furstenberg, Katznelson and Weiss [FKW] showing a similar property for $\mathbb{R}^2$, $V = \{0, v_1, v_2\}$.

1. Introduction

Let E be a measurable subset of $\mathbb{R}^k$, and let S range over all cubes in the space. We set

$$\bar{D}(E) = \limsup_{l(S) \to \infty} \frac{m(S \cap E)}{m(S)},$$

where $l(S)$ denotes the length of the side of S; $\bar{D}(E)$ is the upper density of E. We are interested in configurations which are necessarily contained in E. A theorem of Furstenberg, Katznelson and Weiss [FKW] states that if $E \subset \mathbb{R}^2$, with $\bar{D}(E) > 0$, all large distances in E are attained. More precisely:

THEOREM 1.1 ([FKW]): *If $E \subset \mathbb{R}^2$ with $\bar{D}(E) > 0$, there exists l_0 such that for any $l > l_0$ one can find a pair of points $x, y \in E$ with $\|x - y\| = l$.*

This result was also proved afterwards by Bourgain [Bo], using methods of harmonic analysis, and also by Falconer and Marstrand [FaMa]. Bourgain generalized this result:

Received November 27, 1997

THEOREM 1.2 (Bourgain): *Let $E \subset \mathbb{R}^k$ with $\bar{D}(E) > 0$ and $v_1, \ldots, v_{k-1}$ independent vectors in $\mathbb{R}^k$. Denote $V = \{0, v_1, \ldots, v_{k-1}\}$. Then there exists l_0 such that for any $l > l_0$, E contains an isometric copy of lV.*

It is natural to ask if the same is valid for larger configurations. Bourgain has shown by an example that this cannot be done in $\mathbb{R}^2$, and Graham [Gr] generalized this (using Rado's characterization of partition regular systems): We say that a set of points $S \subset \mathbb{R}^k$ is **spherical**, if S is contained on the surface of some sphere (with finite radius).

THEOREM 1.3 (Graham): *Let $V = v_1, \ldots, v_n \in \mathbb{R}^K$ be nonspherical. Then for any N there exists a set $E \subset \mathbb{R}^N$ with $\bar{D}(E) > 0$, and a set $T \subset \mathbb{R}$ with $\underline{D}(T) > 0$ such that E contains no congruent copy of tV for any $t \in T$ where*

$$\underline{D}(E) = \liminf_{l(S) \to \infty} \frac{m(S \cap E)}{m(S)}.$$

As some configurations may not be found in the set itself, it may be useful to weaken the condition, and try to find the configurations arbitrarily close to the set. A theorem by Furstenberg, Katznelson and Weiss [FKW] shows that with this weaker condition, one can find triangles in the plane:

THEOREM 1.4 ([FKW]): *Let $E \subset \mathbb{R}^2$ with $\bar{D}(E) > 0$, and let E_δ denote the points at distance $< \delta$ from E. Let $v, u \in \mathbb{R}^2$; then there exists l_0 such that for $l > l_0$ and any $\delta > 0$ there exists a triple $\{x, y, z\} \subset E_\delta$ forming a triangle congruent to $\{0, lu, lv\}$.*

The idea of the proof is to translate the geometric problem to a dynamical problem, where E corresponds to some measurable set E, with positive measure, in a measure preserving system $(X, \mathcal{B}, \mu, \mathbb{R}^2)$. The statement that E_δ contains a certain configuration corresponds to a recurrence condition on the set E. In the case of triangles (configurations formed by 2 vectors), the problem is reduced to the case where $(X, \mathcal{B}, \mu, \mathbb{R}^2)$ is a Kronecker action, and it can be shown that, in this case, the recurrence condition holds for configurations of all sizes (we give a short proof of this fact in the appendix). We generalize this result to higher dimensions.

MAIN THEOREM: *Let $E \subset \mathbb{R}^k$ (for $k > 2$) have positive upper density, and let E_δ denote the points of distance $< \delta$ from E. Let $u_1, \ldots, u_{k+1} \subset \mathbb{R}^k$ be $k+1$ points which are affinely independent. Then there exists l_0 such that for any $l > l_0$ there exist $\{x_1, \ldots, x_{k+2}\} \in E_\delta$ forming a configuration congruent to $\{0, lu_1, \ldots, lu_{k+1}\}$.*

One would expect, following Bourgain's generalization of the original problem, that increasing the dimension would allow us an 'extra vector', that is, configurations formed by k vectors in $\mathbb{R}^k$. However, it turns out that dimension 2 is not typical, and for all dimensions > 2 the reduction to the case of a Kronecker action can be carried out for configurations formed by $k+1$ vectors in $\mathbb{R}^k$.

ACKNOWLEDGEMENT: This paper is part of the author's master thesis, written under the supervision of Prof. H. Furstenberg. I thank Prof. Furstenberg for introducing me to Ergodic Theory, and for his encouragement and advice. I would also like to thank Prof. Bergelson for his valuable suggestions.

2. Translation of the geometric problem to a dynamical problem

We would like to solve the geometric problem using methods from ergodic theory, therefore we would like to translate it to a dynamical problem. The translation as shown here was done in [FKW].

Let $E \subset \mathbb{R}^k$ such that $\bar{D}(E) > 0$. Define

$$\varphi(u) = \min\{1, \operatorname{dist}(u, E)\}.$$

The functions $\varphi_v(u) = \varphi(u+v)$ form an equicontinuous, uniformly bounded family, and thus have compact closure in the topology of uniform convergence over bounded sets in $\mathbb{R}^k$. Denote this closure by X. $\mathbb{R}^k$ acts on X by $T_v\psi(u) = \psi(u+v)$ for $\psi \in X, u, v \in \mathbb{R}^k$. X is a compact metrizable space and we can identify Borel measures on X with functionals on $C(X)$. Since $\bar{D}(E) > 0$, there exists a sequence of cubes S_n such that

$$\frac{m(S_n \cap E)}{m(S_n)} \longrightarrow \bar{D}(E) > 0.$$

We wish to define a probability measure μ on X. We define the following probability measures: for $f \in C(X)$, let

$$\mu_n(f) = \frac{1}{m(S_n)} \int_{S_n} f(T_v\varphi)dm(v).$$

We have for some subsequence $\{n_k\}$

$$\mu_{n_k} \xrightarrow{w*} \mu$$

and μ is the desired probability measure. Set $f_0(\psi) = \psi(0)$; then f_0 is a continuous function on X. We define $\tilde{E} \subset X$ by

$$\psi \in \tilde{E} \iff f_0(\psi) = 0 \iff \psi(0) = 0;$$

$\tilde{E}$ is a closed subset of X and we have

$$\mu(\tilde{E}) = \lim_{l\to\infty} \int_X (1 - f_0(\psi))^l d\mu(\psi).$$

LEMMA 2.1: $\mu(\tilde{E}) > 0$.

Proof: It suffices to show that for any l,

$$\int_X (1 - f_0(\psi))^l d\mu(\psi) \geq \bar{D}(E),$$

but

$$\begin{aligned}\int_X (1 - f_0(\psi))^l d\mu(\psi) &= \lim_{k\to\infty} \frac{1}{m(S_{n_k})} \int_{S_{n_k}} (1 - f_0(T_v\varphi))^l dm(v) \\ &= \lim_{k\to\infty} \frac{1}{m(S_{n_k})} \int_{S_{n_k}} (1 - \varphi(v))^l dm(v) \\ &\geq \lim_{k\to\infty} \frac{m(S_{n_k} \cap E)}{m(S_{n_k})} = \bar{D}(E) > 0,\end{aligned}$$

since $\varphi(v) = 0$ for $v \in E$. ∎

The next proposition establishes the correspondence between E and $\tilde{E}$.

PROPOSITION 2.2: *Let $E \subset \mathbb{R}^k$ and $\tilde{E}$ as above. If for $u_1, \ldots, u_l \in \mathbb{R}^k$ we have*

$$\mu(\tilde{E} \cap T_{u_1}^{-1}\tilde{E} \cap \cdots \cap T_{u_l}^{-1}\tilde{E}) > 0, \tag{1}$$

then for all $\delta > 0$,

$$E_\delta \cap (E_\delta - u_1) \cap \cdots \cap (E_\delta - u_l) \neq \emptyset.$$

Proof: Define the function g on X by

$$g(\psi) = \begin{cases} \delta - f_0(\psi) & \text{if } f_0(\psi) < \delta, \\ 0 & \text{if } f_0(\psi) \geq \delta. \end{cases}$$

Since $g(\psi)$ is positive for $\psi \in \tilde{E}$, (1) implies

$$\int g(\psi) g(T_{u_1}\psi) \cdots g(T_{u_l}\psi) d\mu > 0.$$

In particular, for some $T_w\varphi$ the integrand is positive. As

$$g(T_w\varphi) > 0 \iff \varphi(w) < \delta \iff w \in E_\delta$$

we have

$$w \in E_\delta, w + u_1 \in E_\delta, \ldots, w + u_l \in E_\delta. \quad ∎$$

We can now forget the original set E, and the geometric problem takes the following dynamical form:

MAIN THEOREM (Dynamical Version): *Let $(X, B, \mu, (T_u)_{u\in\mathbb{R}^k})$ be an $\mathbb{R}^k$ action, $u_1, \ldots, u_{k+1} \in \mathbb{R}^k$ affinely independent, and $A \subset X$, $\mu(A) > 0$. There exists t_0 s.t. for all $t > t_0$, there exists $P \in \mathrm{SO}(k)$ s.t.*

$$\mu(A \cap T_{tPu_1}^{-1} A \cap \cdots \cap T_{tPu_{k+1}}^{-1} A) > 0,$$

where $\mathrm{SO}(k)$ is the special orthogonal group acting on $\mathbb{R}^k$.

3. Preliminaries

In the following section we give some of the measure theoretic and ergodic theory preliminaries required for understanding the proof of the theorem. The theorems are stated without proofs. For the proofs see [Fu1] and [Pe].

A **measure preserving system** (m.p.s.) is a system $(X, \mathcal{B}, \mu, G)$ where X is an arbitrary space, $\mathcal{B}$ is a σ-algebra of subsets of X, μ is a σ-additive probability measure on the sets of $\mathcal{B}$, and G is a locally compact group acting on X by measure preserving transformations. The action of G is **ergodic**, if $T^{-1}A = A$, $\forall A \in \mathcal{B}$, $T \in G$, implies $\mu(A) = 0$ or $\mu(A) = 1$. In this case we also say that μ is ergodic with respect to the action of G. Each T induces a natural operator on $L^2(X, \mathcal{B}, \mu)$ by $Tf = f \circ T$, and the ergodicity of the action of G is equivalent to the assertion that there are no non-constant G-invariant functions. We have:

THEOREM 3.1 (Mean Ergodic Theorem): *Let $(X, \mathcal{B}, \mu, T)$ be a m.p.s. and $f \in L^2(X)$. Then*

$$\frac{1}{N} \sum_{n=1}^{N} f \circ T^n \overset{L^2(X)}{\longrightarrow} Pf,$$

where Pf is the orthogonal projection of f on the subspace of the T invariant functions.

A **homomorphism** $\alpha : (X, \mathcal{B}, \mu, G) \to (Y, \mathcal{D}, \nu, G)$ of a m.p.s. is a homomorphism of measure spaces which commutes with the action of G. In this case we say that $(Y, \mathcal{D}, \nu, G)$ is a **factor** of $(X, \mathcal{B}, \mu, G)$, and $(X, \mathcal{B}, \mu, G)$ is an **extension** of $(Y, \mathcal{D}, \nu, G)$. The two measure preserving systems are **equivalent** if the homomorphism of one to the other is invertible.

A m.p.s. $(X, \mathcal{B}, \mu, G)$ is **regular** if X is a compact metric space, $\mathcal{B}$ the Borel algebra of X, μ a Borel measure. A m.p.s. is **separable** if $\mathcal{B}$ is generated by a countable subset. As every separable m.p.s. is equivalent to a regular m.p.s., we will confine our attention to a regular m.p.s.

Let α: $(X, \mathcal{B}, \mu, G) \to (Y, \mathcal{D}, \nu, G)$ be a homomorphism of a m.p.s. The map $f \to f^\alpha = f \circ \alpha$ identifies $L^2(Y, \mathcal{D}, \nu)$ with a closed subspace $L^2(Y, \mathcal{D}, \nu)^\alpha \subset$

$L^2(X,\mathcal{B},\mu)$. If P denotes the orthogonal projection of $L^2(X,\mathcal{B},\mu)$ to $L^2(Y,\mathcal{D},\nu)^\alpha$, then we define $E(f|Y)$ for $f \in L^2(X,\mathcal{B},\mu)$ by

$$E(f|Y) \in L^2(Y,\mathcal{D},\nu), \quad E(f|Y)^\alpha = Pf.$$

The operator $E(\cdot|Y)$ commutes with the action of G, i.e. for each $f \in L^1(X,\mathcal{B},\mu)$, and $T \in G$, $E(Tf|Y) = TE(f|Y)$.

3.1 Disintegration of Measure. Let $(X,\mathcal{B},\mu)$ be a regular measure space, and let α: $(X,\mathcal{B},\mu) \to (Y,\mathcal{D},\nu)$ be a homomorphism to another measure space (not necessarily regular). Suppose α is induced by a map φ: $X \to Y$. In this case the measure μ has a disintegration in terms of fiber measures μ_y, where μ_y is concentrated on the fiber $\varphi^{-1}(y) = X_y$. We denote by $\mathcal{M}(X)$ the compact metric space of probability measures on X.

Theorem 3.2: *There exists a measurable map from Y to $\mathcal{M}(X)$, $y \to \mu_y$ which satisfies:*

(1) *For every $f \in L^1(X,\mathcal{B},\mu)$, $f \in L^1(X,\mathcal{B},\mu_y)$ for a.e. $y \in Y$, and $E(f|Y)(y) = \int f d\mu_y$ for a.e. $y \in Y$*

(2) *$\int\{\int f d\mu_y\} d\nu(y) = \int f d\mu$ for every $f \in L^1(X,\mathcal{B},\mu)$.*

The map $y \to \mu_y$ is characterized by condition (1). We shall write $\mu = \int \mu_y d\nu$ and refer to this as the disintegration of the measure μ with respect to the factor Y.

If $(X,\mathcal{B},\mu,G)$ is a m.p.s., $\mathcal{D}$ the algebra of all G-invariant sets, $\mu = \int \mu_x d\mu(x)$ the disintegration of μ with respect to $\mathcal{D}$, then μ_x is G-invariant and ergodic, for a.e. x.

3.2 The Kronecker Factor. An action of a locally compact abelian group G by measure preserving transformations T_g on a measure space $(X,\mathcal{B},\mu)$ is a **Kronecker action** if X is a compact abelian group, μ the Haar measure on X, and there is a homomorphism τ, τ: $G \to X$ with $\tau(G)$ a dense subgroup of X, and

$$T_g(x) = x + \tau(g).$$

The system $(X,\mathcal{B},\mu,G)$ is called a **Kronecker system** (or an **almost periodic system**). Equivalently, $(X,\mathcal{B},\mu,G)$ is Kronecker if the eigenfunctions of G span $L^2(X)$. Every ergodic system has a maximal almost periodic factor:

Theorem 3.3: *Let $(X,\mathcal{B},\mu,T_g)$ be an ergodic measure preserving action of an abelian group G; then there is a map π: $X \to Z$ where Z is a compact abelian* *group, and a Kronecker action T_g on Z such that $T_g\pi(x) = \pi(T_g(x))$ for a.e. $x \in X$. For every character χ on Z the function $\chi'(x) = \chi(\pi(x))$ satisfies*

$$\chi'(T_gx) = \chi(\tau(g) + \pi(x)) = \chi(\tau(g))\chi'(x)$$

and so is an eigenvector of the G-action, and, moreover, every eigenfunction of the G-action comes about this way.

The factor system $(Z, \mathcal{D}, m, G)$, where $\mathcal{D}$ is the algebra of Borel sets, and m the Haar measure, is unique up to isomorphism and is called the Kronecker factor of $(X, \mathcal{B}, \mu, G)$. For the proof see [Fu2].

4. The Main Theorem

Throughout this section $(X, \mathcal{B}, \mu, (T_u)_{u\in\mathbb{R}^k})$ is an ergodic action of $\mathbb{R}^k$, $(Z, \mathcal{D}, m, (T_u)_{u\in\mathbb{R}^k})$ the corresponding Kronecker factor, τ: $\mathbb{R}^k \longrightarrow Z$ the homomorphism inducing the $\mathbb{R}^k$ action on Z. We have a map $\pi : X \to Z$ which defines a 'disintegration' of the measure μ to measures μ_z, $z \in Z$, with μ_z supported by $\pi^{-1}(z)$ for a.e. z. Let $\hat{f}(z)$ be the projection of $f \in L^2(X)$ to $L^2(Z)$, i.e.

$$\hat{f}(z) = \int f d\mu_z.$$

We denote by $\hat{f}^\pi(x)$ the lifting of $\hat{f}(z)$ to $L^2(X)$, i.e.

$$\hat{f}^\pi(x) = \hat{f} \circ \pi(x).$$

4.1 Reduction to the Kronecker Factor.

Definition 4.1: Let Z be a compact abelian group, τ: $\mathbb{R}^k \longrightarrow Z$ a homomorphism. We say that $u_1, \ldots, u_l$ are **τ-independent** if, given $\{\chi_i\}_{i=1}^l \in \hat{Z}$, $\prod_i \chi_i(\tau(u_i)) \neq 1$ unless the χ_i are all trivial.

Lemma 4.2: *If $u_1, \ldots, u_l$ are τ-independent, then for any $f_1, \ldots, f_l \in L^\infty$*

$$\frac{1}{N}\sum_{n=1}^{N} T_{nu_1+a_1}f_1 \ldots T_{nu_l+a_l}f_l \xrightarrow{L^2(Z)} \int f_1 \cdots \int f_l$$

uniformly in $a_1, \ldots, a_l$.

Proof: It is enough to show this for f_i characters on Z:

$$\int \left| \frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{l} T_{nu_i+a_i}\chi_i(z)\right|^2 dm(z)$$

$$= \int \Big| \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{l} \chi_i(z + \tau(nu_i + a_i)) \Big|^2 dm(z)$$
$$= \int \Big| \prod_{i=1}^{l} \chi_i(\tau(a_i))\chi_i(z) \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{l} (\chi_i(\tau(u_i)))^n \Big|^2 dm(z)$$
$$= \Big| \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{l} (\chi_i(\tau(u_i)))^n \Big|^2.$$

By the Weyl Theorem we have

$$\Big| \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{l} (\chi_i(\tau(u_i)))^n \Big|^2 \longrightarrow 0,$$

if the χ_i are not all trivial. ∎

LEMMA 4.3: *Let H be a Hilbert space, $\xi \in \Xi$ some index set, and let $u_n(\xi) \in H$ be uniformly bounded in n, ξ. Assume that for each m the limit*

$$\gamma_m(\xi) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle u_n(\xi), u_{n+m}(\xi) \rangle$$

exists uniformly and

$$\lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \gamma_m(\xi) = 0 \tag{2}$$

uniformly. Then

$$\frac{1}{N} \sum_{n=1}^{N} u_n(\xi) \xrightarrow{H} 0$$

uniformly in ξ.

Proof: Let M be large enough so that the expression in (2) is small uniformly in ξ. Let N be large enough with respect to M so that the two expressions

$$\frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} u_{n+m}(\xi), \quad \frac{1}{N} \sum_{n=1}^{N} u_n(\xi)$$

are close uniformly in ξ. We have

$$\begin{aligned}\|\frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M}u_{n+m}(\xi)\|^2 \leq & \frac{1}{N}\sum_{n=1}^{N}\|\frac{1}{M}\sum_{m=1}^{M}u_{n+m}(\xi)\|^2 \\ = & \frac{1}{NM^2}\sum_{n=1}^{N}\sum_{m_1,m_2=1}^{M}\langle u_{n+m_1}(\xi),u_{n+m_2}(\xi)\rangle \\ \overset{N\to\infty}{\longrightarrow} & \frac{1}{M^2}\sum_{m_1,m_2=1}^{M}\gamma_{m_2-m_1}(\xi),\end{aligned}$$

which is small, uniformly in ξ. ∎

LEMMA 4.4: *Let $(X,\mathcal{B},\mu,(T_u)_{u\in\mathbb{R}^k})$ be an ergodic action of $\mathbb{R}^k$. Suppose for some $v\in\mathbb{R}^k$, T_v acts ergodically on $(X,\mathcal{B},\mu)$. Then every eigenvector of T_v is an eigenvector of the $\mathbb{R}^k$ action.*

Proof: [FKW] The next proposition will enable us to evaluate averages of functions on X by evaluating the averages of the projections of the functions on the Kronecker factor.

PROPOSITION 4.5: *Let $(X,\mathcal{B},\mu,(T_v)_{v\in\mathbb{R}^k})$ be an ergodic action of $\mathbb{R}^k$, and let $u_1,\dots,u_l\in\mathbb{R}^k, l\leq k+1$ be s.t. for all $i\leq l$, $\{u_j-u_i\}_{j\neq i,j=1}^{l}$ are τ-independent, and $\{T_{u_j-u_i}\}_{j\neq i,j=1}^{l}$ act ergodically, and assume that $\{T_{u_i}\}_{i=1}^{l}$ also act ergodically. Let $f_1,\dots,f_l$ be bounded measurable functions on X. Then*

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{l}T_{nu_i+a_i}f_i(x)-\frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{l}T_{nu_i+a_i}\hat{f_i}^{\pi}(x)\overset{L^2(X)}{\longrightarrow}0$$

uniformly in $a_1,\dots,a_l$

Proof: The proof is by induction. For $l=1$

$$\frac{1}{N}\sum_{n=1}^{N}T_{nu+a}f(x)\longrightarrow\int f d\mu$$

uniformly in a, by the Mean Ergodic Theorem. Assume it is true for $l-1<k+1$, and suppose $u_1,\dots,u_l$ satisfy the conditions above. We apply Lemma 4.3 with $\xi=(a_1,\dots,a_l)$, $H=L_2(X,B,\mu)$ and

$$v_n=\prod_{i=1}^{l}T_{nu_i+a_i}f_i(x).$$

We have

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\langle v_n, v_{n+m}\rangle\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\prod_{i=1}^{l}T_{nu_i+a_i}f_i(x)T_{(n+m)u_i+a_i}\bar{f}_i(x)d\mu\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int f_1(x)T_{mu_1}\bar{f}_1(x)\prod_{i=2}^{l}T_{n(u_i-u_1)+(a_i-a_1)}(f_iT_{mu_i}\bar{f}_i)(x)d\mu\\
&=\int f_1(x)T_{mu_1}\bar{f}_1(x)\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\prod_{i=2}^{l}T_{n(u_i-u_1)+(a_i-a_1)}(f_iT_{mu_i}\bar{f}_i)(x)d\mu\\
&=\int f_1(x)T_{mu_1}\bar{f}_1(x)\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\prod_{i=2}^{l}T_{n(u_i-u_1)+(a_i-a_1)}\widehat{(f_iT_{mu_i}\bar{f}_i)}^{\pi}(x)d\mu\\
&=\int f_1(x)T_{mu_1}\bar{f}_1(x)d\mu\prod_{i=2}^{l}\int f_i(x)T_{mu_i}\bar{f}_i(x)d\mu\\
&=\prod_{i=1}^{l}\int f_i(x)T_{mu_i}\bar{f}_i(x)d\mu=\gamma_m
\end{aligned}$$

uniformly in $(a_1,\ldots,a_l)$, using Lemma 4.2 and the induction hypothesis, as $\{(u_j-u_1)-(u_i-u_1)\}_{j\neq i,j=1}^{l+1}=\{u_j-u_i\}_{j\neq i,j=1}^{l+1}$. Finally

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\gamma_m=\int\cdots\int f_1(x_1)\cdots f_l(x_l)F(x_1,\ldots,x_l)d\mu(x_1)\cdots d\mu(x_l)$$

where

$$F(x_1,\ldots,x_l)=\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\prod_i\bar{f}_i(T_{v_i}^{m}x_i),$$

which is well defined by the Mean Ergodic Theorem. Now suppose f_j is orthogonal to all eigenfunctions of the $\mathbb{R}^k$-action; then by Lemma 4.4 it is orthogonal to all eigenfunctions of T_{v_j}. Clearly

$$F(T_{v_1}x_1,\ldots,T_{v_l}x_l)=F(x_1,\ldots,x_l)$$

so we conclude that

$$\int f_j(x_j)F(x_1,\ldots,x_l)d\mu(x_j)=0,$$

and hence

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\gamma_m=0.$$

Hence in this case we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{i}^{l} T_{nu_i+a_i} f_i = 0$$

in $L_2(X)$ uniformly in $(a_1, \dots, a_l)$. Now set

$$f_j = \hat{f}_j^{\,\pi} + (\hat{f}_j^{\,\pi})^{\perp}$$

(in $L_2(X)$) and the result follows. ∎

4.2 Proof of the Theorem Subject to Condition (*). Without loss of generality, we may assume by disintegration of μ that the action of $\mathbb{R}^k$ is ergodic. Let $f = 1_A$ be the characteristic function of the set A, and $\hat{f}$ the projection of f on $L^2(Z)$, the Kronecker factor of X. Since $f \geq 0$, we have $\hat{f} \geq 0$, and $\mu(A) > 0$ implies $\int \hat{f} > 0$. Let W be a neighborhood of unity in Z such that, if $w_1, \dots, w_{k+1} \in W$, then

$$\int_Z \hat{f}(z)\hat{f}(z+w_1)\cdots\hat{f}(z+w_{k+1})dm(z) > a > 0,$$

for some a. Let φ be the homomorphism

$$\begin{aligned} \varphi\colon M_k(\mathbb{R}) &\longrightarrow Z^{k+1} \\ M &\longmapsto (\tau(Mu_1), \dots, \tau(Mu_{k+1})), \end{aligned}$$

where $M_k(\mathbb{R})$ is the set of $k \times k$ matrices over $\mathbb{R}$, and let Ω be the closure of the image of $M_k(\mathbb{R})$ in Z^{k+1}. We say that $M \in M_k(\mathbb{R}), P \in \mathrm{SO}(k)$ satisfy **condition** (*) if $Mv_1, \dots, Mv_{k+1}$ satisfy the conditions of Proposition 4.5, the image of $\{nM\}_{n=1}^{\infty}$ is dense in Ω, and M^tP is an antisymmetric matrix. Suppose we find such matrices; then since Ω is compact, there exists an $L \in \mathbb{N}$, such that for each $U \in M_k(\mathbb{R})$ and $L_0 \geq 0$, there exists an $L_0 \leq n \leq L + L_0$ s.t.

$$\varphi(nM+U) = (\tau(nMu_1 + Uu_1), \dots, \tau(nMu_{k+1} + Uu_{k+1})) \in W^{k+1}.$$

Hence for all $U \in M_k(\mathbb{R})$

$$\begin{aligned} &\frac{1}{N}\sum_{n=1}^{N} \int_Z \hat{f}(z)\hat{f}(z+\tau(nMu_1+Uu_1))\cdots\hat{f}(z+\tau(nMu_{k+1}+Uu_{k+1}))dm(z) \\ &> \frac{a}{2L} \end{aligned}$$

for all N greater than L. From Proposition 4.5 there exists an N_0 such that for all $N > N_0$

$$\frac{1}{N}\sum_{n=1}^{N}\int_X f(x)T_{nMu_1+Uu_1}f(x)\cdots T_{nMu_{k+1}+Uu_{k+1}}f(x)d\mu > \frac{a}{4L}$$

for all U. Hence

$$\frac{1}{N}\sum_{n=1}^{N}\mu(A\cap T_{(nM+U)u_1}A\cap\cdots\cap T_{(nM+U)u_{k+1}}A) > \frac{a}{4L}$$

for all $N > N_0$, for all U. Now take $U = tP$. For each t, there exists an $n < N_0$ s.t.

$$\mu(A\cap T_{(nM+tP)u_1}A\cap\cdots\cap T_{(nM+tP)u_{k+1}}A) > \frac{a}{4L}.$$

Since M^tP is antisymmetric, $M \in T_P(\mathrm{SO}(k))$, the tangent space of $\mathrm{SO}(k)$ at P. We have

$$P' := P\exp(\varepsilon nP^{-1}M) = P(I+\varepsilon nP^{-1}M+o(\varepsilon)) = P+\varepsilon nM+o(\varepsilon);$$

hence

$$\frac{1}{\varepsilon}(P+nM)-\frac{1}{\varepsilon}P' = o(1).$$

Now the T_u satisfy the following continuity condition:

$$\forall\varepsilon\exists\delta : \|u-u'\|\le\delta \Rightarrow |\mu(A\cap T_uA)-\mu(A\cap T_{u'}A)|\le\varepsilon.$$

Therefore, for $t = 1/\varepsilon$ large enough, P' as above,

$$\mu(A\cap T_{tP'u_1}A\cap\cdots\cap T_{tP'u_{k+1}}A) > \frac{a}{8L}.$$

We have thus proved the theorem. ∎

4.3 The Existence of the Matrices M, P Satisfying Condition $(*)$. We first show that if $\mathbb{R}^k$ acts ergodically, then the set of points which do not act ergodically is very small.

Proposition 4.6: *If (X, B, μ, T_u) is an ergodic action of $\mathbb{R}$, then but for a countable set of $u \in \mathbb{R}$, T_u acts ergodically. If $(X, \mathcal{B}, \mu, T_u)$ is an ergodic action of $\mathbb{R}^l$, then but for a countable set of $l-1$ dimensional hyperplanes, all $l-1$ dimensional hyperplanes through the origin act ergodically.*

Proof: [PS]. We now prove the main proposition. The idea is that while for $k = 2$ the orthogonal group is contained in a proper hyperplane in $\mathbb{R}^4$, for $k > 2$ the orthogonal group is large enough, and intersects each proper hyperplane in $\mathbb{R}^{k^2}$ in a subvariety of lower dimension.

PROPOSITION 4.7: *Let φ, Ω be as in the previous section. There exist matrices $M \in M_k(\mathbb{R}), P \in \mathrm{SO}(k)$ satisfying the following conditions:*

1. $\{\varphi(nM)\}_{n=1}^{\infty}$ *is dense in* Ω.
2. $\{M(u_j - u_i)\}_{j=1,(j\neq i)}^{k+1}$ *are τ-independent, for* $i = 1, \dots, k+1$.
3. $\{Mu_i\}_{i=1}^{k+1}, \{M(u_j - u_i)\}_{i,j=1,(j\neq i)}^{k+1}$ *act ergodically.*
4. $M^t P$ *is an antisymmetric matrix.*

For the proof of the proposition we will need the following lemmas:

LEMMA 4.8: *The group* $\mathrm{SO}(k)$ $(k \geq 3)$ *linearly spans* $M_k(\mathbb{R})$.

Proof: We separately prove the cases when k is odd and when k is even. For k odd, we denote

$$(E_{ij})_{mn} = \begin{cases} 1 & m = i,\, n = j, \\ 0 & \text{otherwise.} \end{cases}$$

One can get the E_{ij} using sums of matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 & 0 & & & \\ 0 & 0 & -1 & 0 & & & \\ 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ 0 & 0 & 0 & 0 & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & & \dots & & & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & -1 & 0 & 0 & & & \\ 0 & 0 & 0 & -1 & & & \\ 0 & 0 & 0 & 0 & -1 & & \\ \vdots & & & & & \ddots & \\ 0 & & \dots & & & 0 & -1 \end{pmatrix}.$$

For k even we can get, using similar matrices, matrices of the following form: for $1 \leq i < k$, $1 \leq j < k$

$$(E_{ij})_{mn} = \begin{cases} 1 & m = i,\, n = j \text{ or} \\ & m = i+1,\, n = j+1, \\ 0 & \text{otherwise.} \end{cases} \quad (F_{ij})_{mn} = \begin{cases} 1 & m = i,\, n = j, \\ -1 & m = i+1,\, n = j+2, \\ 0 & \text{otherwise.} \end{cases}$$

for $j = k$, $1 \leq i < k$

$$(E_{ik})_{mn} = \begin{cases} 1 & m = i,\, n = k, \\ -1 & m = i+1,\, n = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (F_{ik})_{mn} = \begin{cases} 1 & m = i,\, n = k \text{ or} \\ & m = i+1,\, n = 2, \\ 0 & \text{otherwise.} \end{cases}$$

and for $i = j = k$ we have

$$(E_{kk})_{mn} = \begin{cases} 1 & m = 1,\, n = 1 \text{ or } m = k,\, n = k, \\ 0 & \text{otherwise.} \end{cases}$$

For example,

$$\begin{pmatrix} 0 & 1 & 0 & 0 & & \\ 0 & 0 & 1 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ \vdots & & & & \ddots & \\ 0 & & \dots & & & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 & & \\ 0 & 0 & -1 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ \vdots & & & & \ddots & \\ 0 & & \dots & & & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ -1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & & 0 \\ \vdots & & & \ddots & \\ 0 & & \dots & & 0 \end{pmatrix}, \quad \dots,$$

which add up to matrices of the form

$$(E_{ij})_{mn} = \begin{cases} 1 & m = i,\ n = j, j+1, \\ 0 & \text{otherwise}; \end{cases}$$

for example,

$$\begin{pmatrix} 1 & 1 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & & \\ \vdots & & & & \ddots & \\ 0 & & \dots & & & 0 \end{pmatrix}.$$

As vectors in $\mathbb{R}^{k^2}$ they form the following $k^2 \times k^2$ matrix:

$$\begin{pmatrix} 1 & 1 & 0 & & & & & & \\ 0 & 1 & 1 & 0 & & & & & \\ & & \vdots & & & & & & \\ & & 0 & 1 & 1 & 0 & & & \\ & & & 0 & 1 & -1 & 0 & & \\ & & & & 0 & 1 & 1 & 0 & \\ & & & & & & \vdots & & \\ & & & & & & 0 & 1 & 1 \\ 1 & & & & & & & & 1 \end{pmatrix},$$

$$(E_{ij})_{mn} = \begin{cases} 1 & m = n \text{ or } m = n-1 \text{ for } m \neq lk\ (1 \leq l < k) \text{ or } m = k^2,\ n = 1, \\ -1 & m = lk,\ n = lk+1\ (1 \leq l < k), \\ 0 & \text{otherwise.} \end{cases}$$

The number of -1's in the secondary diagonal is $k-1$ (one for every k-th row but the last one), and is therefore an odd number, therefore the determinant is $\neq 0$, hence the vectors are independent. ∎

LEMMA 4.9: *Let $\{s_i\}_{i=1}^k$, $\{v_i\}_{i=1}^k \subset \mathbb{R}^k$, $f(M) = \sum_{i=1}^k \langle s_i, Mv_i \rangle$, and suppose $f \not\equiv 0$. Let $\mathcal{D}$ be the set of matrices in* $\mathrm{SO}(k)$ *which satisfy $f(D) = c$ for some constant $c \in \mathbb{R}$. Then* dim $\mathcal{D}$ *(as an algebraic variety)* $< \dim \mathrm{SO}(k)$.

Proof: Suppose $f(P) = c$ for all $P \in \mathrm{SO}(k)$. Let O be some matrix in $\mathrm{SO}(k)$; then $f(M - O) \not\equiv 0$, but $f(P - O) = 0$ for all $P \in \mathrm{SO}(k)$, in contradiction to

Lemma 4.8. Therefore there is a point $P \in \mathrm{SO}(k)$ which is not in $\mathcal{D}$. Hence $\mathcal{D}$ is a proper subvariety of $\mathrm{SO}(k)$. Since $\mathrm{SO}(k)$ is irreducible, $\dim \mathcal{D} < \dim \mathrm{SO}(k)$. ∎

LEMMA 4.10: *Let φ be the homomorphism*

$$\begin{aligned}\varphi\colon M_k(\mathbb{R}) &\longrightarrow Z^{k+1}\\ M &\longmapsto (\tau(Mu_1),\dots,\tau(Mu_{k+1})).\end{aligned}$$

Let Ω be the closure of the image of $M_k(\mathbb{R})$ in Z^{k+1}. Then for all matrices but a countable number of hyperplanes in $M_k(\mathbb{R}) = \mathbb{R}^{k^2}$, the image of $\{(nM)\}_{n=1}^{\infty}$ is dense in Ω.

Proof: Since Z^{k+1} is a compact abelian group, so is Ω. The image of $\{(nM)\}_{n=1}^{\infty}$ is dense in Ω if no character $\chi \neq 1$ in $\hat{\Omega}$ satisfies $\chi(\varphi(M)) = 1$. As $\chi \circ \varphi$ is a character on $\mathbb{R}^{k^2}$ it is of the form $\chi \circ \varphi(M) = e^{2\pi i \langle N, M\rangle}$ for some $N \in \mathbb{R}^{k^2}$. Since Ω is a compact metrizable abelian group, $\hat{\Omega}$ is countable. ∎

LEMMA 4.11: *For each $j = 1,\dots\infty$, let $\{v_{ij}\}_{i=1}^{k} \subset \mathbb{R}^k$ be a linearly independent set, and $\{s_{ij}\}_{i=1}^{k} \subset \mathbb{R}$, such that $(s_{j1},\dots,s_{jk}) \neq \vec{0}$. There exists an antisymmetric matrix $B \subset M_k(\mathbb{R})$ s.t.*

$$\forall j\colon f_{j,B}(M) \stackrel{\text{def}}{=} \sum_{i=1}^{k} \langle s_{ij}, MBv_{ij}\rangle \not\equiv 0.$$

Proof: Let $\mathcal{B}$ be the subspace of antisymmetric matrices. Since $f_{j,B}(M)$ is linear in M, we have $f_{j,B}(M) \equiv 0 \iff B$ satisfies the k^2 linear equations given by the standard basis for $\mathbb{R}^{k^2}$. Hence for each j, the 'bad' B form a linear subspace of $\mathcal{B}$. Since we have only a countable number of inequalities, it suffices to show that this linear subspace is a proper subspace of $\mathcal{B}$. So without loss of generality, we have only one inequality. Assume

$$\forall B \in \mathcal{B}\colon \sum_{i=1}^{k} \langle Ms_i, Bv_i\rangle \equiv 0 \quad (\text{w.l.o.g. } s_{11} \neq 0).$$

Taking (with the notation of Lemma 4.8) $M = E_{11}$ and $M = E_{21}$, we get

$$\sum_{l=1}^{k} b_{1l}\Big(\sum_{j=1}^{k} s_{j1}v_{jl}\Big) = s_{11}\Big(\sum_{j=1}^{k} b_{1j}v_{1j}\Big) + \cdots + s_{k1}\Big(\sum_{j=1}^{k} b_{1j}v_{kj}\Big) = 0,$$

$$\sum_{l=1}^{k} b_{2l}\Big(\sum_{j=1}^{k} s_{j1}v_{jl}\Big) = s_{11}\Big(\sum_{j=1}^{k} b_{2j}v_{1j}\Big) + \cdots + s_{k1}\Big(\sum_{j=1}^{k} b_{2j}v_{kj}\Big) = 0.$$

As this is true for all $b_{12}, \dots, b_{1k}, b_{23}, \dots, b_{2k} \in \mathbb{R}$,

$$s_{11}v_{1j} + \cdots + s_{k1}v_{kj} = 0$$

for $1 \le j \le k$. Hence

$$s_{11}v_1 + \cdots + s_{k1}v_k = 0,$$

which is a contradiction to the linear independence of the v_i. ∎

Proof of Proposition 4.7: We first wish to fulfill the τ-independence requirement. Suppose $w_1, \dots, w_k \in \mathbb{R}^k$; then $w_1, \dots, w_k$ are independent with respect to τ, if for all $\{\chi_i\}_{i=1}^k$ in $\hat{Z}$

$$\chi_1 \circ \tau(w_1) \cdots \chi_k \circ \tau(w_k) \neq 1.$$

As $\chi \circ \tau$ is a character on $\mathbb{R}^k$ it is of the form $\chi \circ \tau(w) = e^{2\pi i \langle s, w \rangle}$, so the condition above is equivalent to a countable number of inequalities of the form

$$\sum_{j=1}^{k} \langle s_{jm}, w_j \rangle \neq 1.$$

Hence the M we are seeking must satisfy the inequalities

$$g_m(M) = \sum_{(j \neq i), j=1}^{k+1} \langle s_{jm}, M(u_j - u_i) \rangle \neq 1 \quad \text{for } i = 1, \dots, k+1. \tag{3}$$

Since $u_1, \dots, u_{k+1}$ are affinely independent, for each $i = 1, \dots, k+1$ the $\{(u_j - u_i)\}_{j=1,(j \neq i)}^{k+1}$ are linearly independent, so $g_m \not\equiv 0$. Our M will also have to satisfy the inequalities arising from the requirement (2) (Lemma 4.10):

$$f_m(M) = \langle N_m, M \rangle = \sum_{i=1}^{k} \langle N_m e_i, M e_i \rangle \neq r_m. \tag{4}$$

Let B be the antisymmetric matrix of Lemma 4.11, for the inequalities $f_m(M)$, $g_m(M) \not\equiv 0$. From Lemma 4.9, for almost all $P \in \mathrm{SO}(k)$ (with respect to the Haar measure on $\mathrm{SO}(K)$), the matrix $M = PB$ satisfies the inequalities (3) and (4). We still have the ergodicity requirement, but by Proposition 4.6, it is satisfied by almost all $M = PB$. We have found the desired M and P. ∎

5. Appendix

If $(Z, \mathcal{B}, \mu, \mathbb{R}^k)$ is a Kronecker action, then the Main Theorem (dynamical version) holds for all configurations $u_1, \ldots, u_l \in \mathbb{R}^k$. There are several ways of proving this fact. One is implicit in the proof of the Main Theorem; we provide another one based on Theorem 1.1, for the case $k = 2$.

PROPOSITION 5.1: *Let $(Z, \mathcal{B}, \mu, \mathbb{R}^2)$ be a Kronecker action, $A \in \mathcal{B}$ a set of positive measure. Let $u_1, \ldots, u_l \in \mathbb{R}^2$. Then there exists $R \in \mathbb{R}$ such that for all $r > R$ there exists $P \in \mathrm{SO}(2)$ such that*

$$\mu(A \cap T_{rPv_1}^{-1} A \cap \cdots \cap T_{rPv_l}^{-1} A) > 0.$$

Proof: If we think of $u_1, \ldots, u_l$ as points in $\mathbb{C}$, then we must show that for $r > R$ there exists $|w| = r$ such that

$$\mu(A \cap T_{wu_1}^{-1} A \cap \cdots \cap T_{wu_l}^{-1} A) > 0.$$

Let $f = 1_A$, and let V be a neighborhood of the identity in Z such that $v_1, \ldots, v_l \in V$; then

$$\int f(z) f(z + v_1) \cdots f(z + v_l) > a > 0$$

for some a. It suffices to show that $\exists R$ such that for all $r > R$ there exists $|z| = r$ such that

$$\tau(zu_1), \ldots, \tau(zu_l) \in V.$$

Since Z is compact, $Z = \bigcup_{i=1}^n (V + z_i)$. Assign to each $z \in \mathbb{C}$ the l-tuple $(a_1(z), \ldots, a_l(z))$, where $a_j(z) = i$, if $\tau(zu_j) \in V + z_i$ (if there is more than one i corresponding to a certain j, just pick one of them arbitrarily). As $a_j(z) \leq n$ for all j, z, we have only a finite set of such l-tuples. Therefore there is a measurable set $E \subset \mathbb{C}$, with $\bar{D}(E) > 0$, so that if $z, z' \in E$, then $a_j(z) = a_j(z')$ for all j. It follows, by the definition of the a_j, that $z, z' \in E$ implies $\tau((z - z')u_j) \in V$ for $1 \leq j \leq l$. By Theorem 1.1 one can find all large distances in E, and hence one can find all large distances from the origin in $E - E$. ∎

References

[Bo] J. Bourgain, *A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$*, Israel Journal of Mathematics **54** (1986), 307–316.

[FM] K.J. Falconer and J.M. Marstrand, *Plane sets with positive density at infinity contain all large distances*, The Bulletin of the London Mathematical Society **18** (1986), 471–474.

[Fu1] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, 1981.

[Fu2] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, Journal d'Analyse Mathématique **31** (1977), 204–256.

[FKW] H. Furstenberg, Y. Katznelson and B. Weiss, *Ergodic theory and configurations in sets of positive density*, in *Mathematics of Ramsey Theory* (J. Nešetřil and V. Rődl, eds.), Springer-Verlag, Berlin, 1990.

[Gr] R. L. Graham, *Recent trends in Euclidean Ramsey theory*, Discrete Mathematics **136** (1994), 119–127.

[Pe] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1983.

[PS] C. Pugh and M. Shub, *Ergodic Elements of Ergodic Actions*, Compositio Mathematica **23** (1971), 115–122.