NILFACTORS OF $\mathbb{R}^m$-ACTIONS AND CONFIGURATIONS
IN SETS OF POSITIVE UPPER DENSITY IN $\mathbb{R}^m$

T. ZIEGLER

ABSTRACT. We use ergodic theoretic tools to solve a classical problem in geometric Ramsey theory. Let $E$ be a measurable subset of $\mathbb{R}^m$, with $\bar{D}(E) > 0$. Let $V = \{0, v_1, \ldots, v_k\} \subset \mathbb{R}^m$. 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 [FuKaW] showing a similar property for $m = k = 2$.

1. Introduction

Let $E$ be a measurable subset of $\mathbb{R}^m$. We set
\[ \bar{D}(E) := \limsup_{l(S) \to \infty} \frac{m(S \cap E)}{m(S)}, \]
where $S$ ranges over all cubes in $\mathbb{R}^m$, and $l(S)$ denotes the length of a side of $S$. $\bar{D}(E)$ is the upper density of $E$. We are interested in configurations which are necessarily contained in $E$. Furstenberg, Katznelson, and Weiss [FuKaW] showed, using methods from ergodic theory, that if $E \subset \mathbb{R}^2$, with $\bar{D}(E) > 0$, all large distances in $E$ are attained. More precisely:

1.1. Theorem (FuKaW). 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, using different methods, by Bourgain [Bo], and by Falconer and Marstrand [FM]. It is natural to ask if the same is valid for larger configurations. Bourgain has shown by an example that this can not be done [Bo].

As some configurations may not be found in the set itself, we try to find the configurations arbitrarily close to the set. In the same paper Furstenberg, Katznelson, and Weiss [FuKaW] show that with this weaker condition, one can find triangles in the plane:

1.2. Theorem (FuKaW). Let $E \subset \mathbb{R}^2$ with $\bar{D}(E) > 0$, and let $E_\delta$ 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 $(x, y, z) \subset E_\delta^3$ 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 $\hat{E}$, with positive measure, in a measure preserving system $(X^0, \mathcal{B}, \mu, \mathbb{R}^2)$. The statement that
$E_\delta$ contains a certain configuration, corresponds to a recurrence condition on the set $\tilde{E}$. In the case of triangles (configurations formed by 2 vectors), the recurrence phenomenon in question is reduced to the case where $(X^0, \mathcal{B}, \mu, \mathbb{R}^2)$ is a Kronecker action. The problem for a general configuration reduces to the study of pro-nilsystems (defined later). We prove the following theorem:

1.3. **Theorem.** Let $E \subset \mathbb{R}^m$ have positive upper density, and let $E_\delta$ denote the points of distance $< \delta$ from $E$. Let $(u_1, \ldots, u_k) \subset (\mathbb{R}^m)^k$. Then there exists $l_0$ such that for any $l > l_0$, and any $\delta > 0$ there exists $\{x_1, x_2, \ldots, x_{k+1}\} \in E_\delta^{k+1}$ forming a configuration congruent to $\{0, lu_1, \ldots, lu_k\}$.

**Acknowledgment** I thank Sasha Leibman and Hillel Furstenberg for helpful comments.

2. **Translation of the Geometric Problem to a Dynamical Problem.**

We start by translating the geometric problem to a dynamical problem. The translation as shown here was done in [FuKaW]. We bring it here for the sake of completeness.

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

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

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 $\mathbb{R}^m$. Denote this closure by $X^0$. $\mathbb{R}^m$ acts on $X^0$ by $T_v \psi(u) = \psi(u + v)$ for $\psi \in X^0, u, v \in \mathbb{R}^m$. $X^0$ is a compact metrizable space and we can identify Borel measures on $X^0$ with functionals on $C(X^0)$. Since $\bar{D}(E) > 0$, there exists a sequence of cubes $S_n$ such that

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

We define a probability measure $\mu$ on $X^0$ as follows. We define the following probability measures: for $f \in C(X^0)$, let

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

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

$$\mu_{n_k} \overset{w*}{\to} \mu.$$ 

Set $f_0(\psi) = \psi(0)$, then $f_0$ is a continuous function on $X^0$. We define $\tilde{E} \subset X^0$ by

$$\psi \in \tilde{E} \iff f_0(\psi) = 0 \iff \psi(0) = 0.$$
$\tilde{E}$ is a closed subset of $X^0$ and we have:

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

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

Proof. It suffices to show that for any $l$,

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

Indeed

$$\int_X (1 - f_0(\psi))^l d\mu(\psi) = \lim_{k \to \infty} \frac{1}{m(S_{n_k})} \int_{S_{n_k}} (1 - f_0(T_v \varphi))^l dm(v)$$

$$= \lim_{k \to \infty} \frac{1}{m(S_{n_k})} \int_{S_{n_k}} (1 - \varphi(v))^l dm(v)$$

$$\geq \lim_{k \to \infty} \frac{m(S_{n_k} \cap E)}{m(S_{n_k})} = \bar{D}(E) > 0,$$

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

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

2.2. Proposition. Let $E \subset \mathbb{R}^m$ and $\tilde{E}$ be as above. If for $(u_1, \ldots, u_l) \in (\mathbb{R}^m)^l$ we have

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

then for all $\delta > 0$,

$$E_{\delta} \cap (E_{\delta} - u_1) \cap \ldots \cap (E_{\delta} - u_l) \neq \emptyset.$$

Proof. Define the function $g$ on $X^0$ 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 $\psi \in \tilde{E}$, equation (1) implies that

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

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

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

we have

$$w \in E_{\delta}, w + u_1 \in E_{\delta}, \ldots, w + u_l \in E_{\delta}.$$ 

We now forget the original set $E$, and the geometric problem takes the following dynamical form:
2.3. Theorem (Dynamical Version). Let \((X, \mathcal{B}, \mu, \mathbb{R}^m)\) be a \(\mathbb{R}^m\) action, and let \(T_u\) denote the action of \(u \in \mathbb{R}^m\). Let \((u_1, \ldots, u_k) \in (\mathbb{R}^m)^k\), and let \(A \subset X\), with \(\mu(A) > 0\). There exists \(t_0 \in \mathbb{R}^+\) s.t. for all \(t > t_0\), there exists a rotation \(P \in SO(m)\) such that
\[
\mu(A \cap T_{tP_{u_1}}^{-1} A \cap \ldots \cap T_{tP_{u_k}}^{-1} A) > 0.
\]
(Here \(SO(m)\) is the special orthogonal group acting on \(\mathbb{R}^m\)).

3. Preliminaries.

In the following section we give some measure theoretic and ergodic theory preliminaries. The theorems are stated without proofs. For the proofs see [Fu1], [Pe].

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

3.1. Theorem (Mean Ergodic Theorem). Let \(X = (X^0, \mathcal{B}, \mu, T)\) be a m.p.s., and \(f \in L^2(X)\). Then
\[
\frac{1}{N} \sum_{n=1}^{N} f \circ T^n \xrightarrow{L^2(X)} \mathbb{P} f,
\]
where \(\mathbb{P} f\) is the orthogonal projection of \(f\) on the subspace of the \(T\)-invariant functions.

Let \(X = (X^0, \mathcal{B}, \mu, G)\) be a measure preserving system (m.p.s). Let \(Y = (Y^0, \mathcal{D}, \nu, G)\) be a homomorphic image of \(X\); i.e., we have a map \(\pi : X^0 \to Y^0\) with \(\pi^{-1} \mathcal{D}_Y \subset \mathcal{B}_X\), \(\pi \mu_X = \mu_Y\) and \(\pi\) commutes with the \(G\) action. Then \(Y\) is a factor of \(X\), \(X\) is an extension of \(Y\), and abusing the notation we write \(\pi : X \to Y\) for the factor map. A factor of \(X\) is determined by a \(G\)-invariant subalgebra of \(L^\infty(X)\). The map \(\pi\) induces two natural maps \(\pi^* : L^2(Y) \to L^2(X)\) given by \(\pi^* f = f \circ \pi\), and \(\pi_* : L^2(X) \to L^2(\mu_Y)\) given by \(\pi_* f = E(f|\mathcal{D}_Y)\) (the orthogonal projection of \(f\) on \(\pi^* L^2(Y)\)). The two measure preserving systems are equivalent if the homomorphism of one to the other is invertible. We shall simplify the notation writing \(E(f|\mathcal{D}_Y)\) for \(E(f|\mathcal{B}_Y)\).

A m.p.s. \(X\) is regular if \(X^0\) is a compact metric space, \(\mathcal{B}\) the Borel algebra of \(X^0\), \(\mu\) a measure on \(\mathcal{B}\). 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 regular m.p.s.

3.1. **Disintegration of Measure.** Let \((X^0, \mathcal{B}, \mu)\) be a regular measure space, and let \(\alpha : (X^0, \mathcal{B}, \mu) \to (Y^0, \mathcal{D}, \nu)\) be a homomorphism to another measure space (not necessarily regular). Suppose \(\alpha\) is induced by a map \(\varphi : X^0 \to Y^0\). In this case the measure \(\mu\) has a disintegration in terms of fiber measures \(\mu_y\), where \(\mu_y\) is concentrated on the fiber \(\varphi^{-1}(y) = X_y\). We denote by \(\mathcal{M}(X)\) the compact metric space of probability measures on \(X^0\).

3.2. **Theorem.** There exists a measurable map from \(Y^0\) to \(\mathcal{M}(X^0)\), \(y \to \mu_y\) which satisfies:

1. For every \(f \in L^1(X^0, \mathcal{B}, \mu)\), \(f \in L^1(X^0, \mathcal{B}, \mu_y)\) for a.e. \(y \in Y^0\), and \(E(f|Y)(y) = \int f \, d\mu_y\) for a.e. \(y \in Y^0\).
2. \(\int f \, d\mu_y \, d\nu(y) = \int f \, d\mu\) for every \(f \in L^1(X^0, \mathcal{B}, \mu)\).

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 \(\mu\) with respect to \(\mathcal{D}\).

If \((X^0, \mathcal{B}, \mu, G)\) is a m.p.s., \(\mathcal{D}\) the algebra of all \(G\)-invariant sets, \(\mu = \int \mu_x \, d\mu(x)\) the disintegration of \(\mu\) with respect to \(\mathcal{D}\), then \(\mu_x\) is \(G\)-invariant and ergodic for a.e. \(x\).

3.3. **Nilsystems and Characteristic Factors** A \(k\)-step nilflow is a system \(X = (N/\Gamma, \mathcal{B}, m, G)\) where \(N\) is a \(k\)-step nilpotent Lie group, \(\Gamma\) a cocompact lattice, \(\mathcal{B}\) the (completed) Borel algebra, \(m\) the Haar measure, and the action of \(G\) is by translation by elements of \(N\): \(T_g n\Gamma = a_g n\Gamma\) where \(g \to a_g\) is a homomorphism of \(G\) to \(N\). We will sometimes denote this system by \((N/\Gamma, G)\), or \((N/\Gamma, a)\) if \(G = \mathbb{Z}\) and \(1 \to a\). If \(G\) is connected and \((N/\Gamma, G)\) is an ergodic nilflow, then we may assume that \(N\) is connected so that \(X^0 = N/\Gamma\) is connected and is a homogeneous space of the identity component of \(N\). A \(k\)-step pro-nilflow is an inverse limit of \(k\)-step nilflows.

3.4. **Theorem** (Cf. [Pa1]). Let \(X = (N/\Gamma, a)\) be an ergodic nilflow, then \(X\) is uniquely ergodic. Let \(f\) be a continuous function on \(N/\Gamma\). Then the averages \(\frac{1}{N} \sum_{n=1}^{N} f(a^n x)\) converge uniformly to \(\int f(x) \, dm\).

3.5. Let \(N\) be a connected simply connected nilpotent Lie group, \(\Gamma\) a cocompact lattice in \(N\), and \(X^0 = N/\Gamma\). Let \(\pi : N \to X^0\) be the natural projection, and let \(M\) be a closed connected subgroup of \(N\) such that \(\pi(M)\) is a closed submanifold of \(X^0\). Let \(G = \mathbb{R}^k\) and let \(\varphi : G \to N\) be a homomorphism. For \(x \in X^0\) let \(O(x) = \overline{Gx}\), and for \(x \in X^0\), \(y \in G\) let \(O_y(x) = \{\varphi(ny)x\}_{n \in \mathbb{Z}}\); these are subnilmanifolds of \(X^0\) (see for example [Le]).

Introducing Malcev coordinates on \(N\) and \(M\) (Ma) we can identify these groups topologically with, say, \(\mathbb{R}^l\) and \(\mathbb{R}^m\), \(l \geq m\). Call a proper subspace of \(\mathbb{R}^d\) countably linear if it is contained in a countable union of proper linear
Call a subset of $\mathbb{R}^d$ polynomial if it is the set of zeroes of some nonzero polynomial in $\mathbb{R}^d$ (i.e. an algebraic variety of co-dimension 1), and countably polynomial if it is contained in a countable union of proper polynomial subsets. The following proposition is due to Sasha Leibman:

**3.6. Proposition.** There exists a connected subnilmanifold $V$ of $X^0$ such that $O(\pi(a)) \subseteq aV$ for all $a \in M$, and there exists a countably linear set $B \subset G$ such that for every $g \in G \setminus B$ there is a countably polynomial set $A_g \subset M$ such that $O_g(\pi(a)) = aV$ for all $a \in M \setminus A_g$.

**Proof.** Define a mapping $\eta: G \times M \to N$ by $\eta(g, a) = a^{-1}\varphi(g)a$. In Malcev coordinates on $M$ and $N$, $\eta$ is a polynomial mapping $\mathbb{R}^{k+m} \to \mathbb{R}'$ (see [Mal]). Moreover for each $a \in M$, $\eta(\cdot, a)$ is a homomorphism $G \to N$. Let $H$ be the closure of the subgroup generated by $\eta(G \times M)$. Let $V$ be the closure of $\pi(H)$ in $N/\Gamma$. Then $V$ is a subnilmanifold $V = \pi(K)$ for some closed subgroup $K$ of $N$ (see [Sh]) $(\pi(H)$ itself is not necessarily closed). We then have $a^{-1}\varphi(g)\pi(a) = \pi(a^{-1}\varphi(g)a) \in V$, thus $\varphi(g)\pi(a) \in aV$ for any $a \in M$, and $g \in G$. So, $O(\pi(a)) \subseteq aV$ for all $a \in M$.

Let $\tilde{L}$ be the set $\{l \in N : lV = V\}$. Then $\tilde{L}$ is a group, $\eta(G \times M)$ and $K$ are subsets of $\tilde{L}$, and $V = \pi(\tilde{L})$. Let $L$ be the identity component of $\tilde{L}$, then $\eta(G \times M) \subseteq L$, and therefore $H \subseteq L$. $V$ is connected, and therefore a homogeneous subspace of $L$; $V = L/L \cap \Gamma$. Let $W$ be the maximal torus factor of $V$, $W = L/([L, L](L \cap \Gamma))$, and let $p : V \to W$ be the natural projection. Let $\hat{W}$ be the group of characters of $W$, and let $\chi \in \hat{W}$. The character $\chi$ can be lifted to a homomorphism $\zeta_{\chi} : L \to \mathbb{R}$. For each $\chi \in \hat{W}$, let $\psi_{\chi} := \zeta_{\chi} \circ \eta$. Then $\psi_{\chi}$ are polynomials on $G \times M$, which for each $a \in M$ are linear with respect to $G$. Moreover, each $\psi_{\chi}$ is a nonzero polynomial; otherwise $\eta(G \times M)$ would be contained in the kernel of the corresponding homomorphism $\chi \circ p \circ \pi : N \to S^1$. This is a closed subgroup of $N$ containing $\eta(G \times M)$, thus contains the subgroup $H$. Therefore $\chi \circ p \circ \pi(H) = 1$, but this implies that $\chi \circ p(V) = 1$, i.e. $\chi$ is the trivial character.

Let $C_{\chi} \subset G \times M$ be the set of zeroes of $\psi_{\chi}$, and let $C = \bigcup_{\chi \in \hat{W}} C_{\chi}$. Then $C$ is a countably polynomial subset of $G \times M$.

For any $(g, a) \notin C$ one has $\chi \circ p(a^{-1}\varphi(g)\pi(a)) = \chi \circ p(\pi(a^{-1}\varphi(g)a)) \neq 0$ for all $\chi \in \hat{W}$, so the projection of $a^{-1}\varphi(g)\pi(a) \in V$ to $W$ is not contained in any proper subtorus of $W$. Consider the following $\mathbb{Z}$ action on $V$: for $n \in \mathbb{Z}$, $v \to a^{-1}\varphi(n\varphi)av$. Since the projection of $a^{-1}O_g(\pi(a)) \subseteq V$ to $W$ is a closed subgroup of $W$, i.e. a subtorus of $W$, it is equal to $W$. By Parry ([Par1]) this implies the $\mathbb{Z}$ action is minimal and therefore $a^{-1}O_g(\pi(a)) = V$, and so $O_g(\pi(a)) = aV$.

Now let

$$B = \{g \in G : \{g\} \times M \subseteq C\}, \quad M_{\chi}(g) = \{a \in M : \psi_{\chi}(g,a) = 0\}.$$ 

If $(g) \times M \subseteq C$, then $M = \bigcup_{\chi \in \hat{W}} M_{\chi}(g)$. As $M$ is connected, if $M_{\chi}(g)$ has non-empty interior, then $M_{\chi}(g) = M$. By the Baire category theorem $M_{\chi}$
is non-empty for some \( \chi \in \hat{W} \). Therefore

\[
B = \bigcup_{\chi \in \hat{W}} \{ g \in G : \psi_{\chi}(g,a) = 0 \text{ for all } a \in M \}.
\]

Then \( B \) is a countably linear subset of \( G \), and for each \( g \in G \setminus B \),

\[
A_g = C \cap (\{ g \} \times M) = \bigcup_{\chi \in \hat{W}} \{ a \in M : \psi_{\chi}(g,a) = 0 \}
\]

is a countably polynomial subset of \( M \).

3.7. **Theorem.** Let \( X = (X^0, \mathcal{B}, \mu, \mathbb{R}^m) \) be an ergodic \( \mathbb{R}^m \) action. We can associate with \( X \) an inverse sequence of factors \( \cdots \to Y_k(X) \to Y_{k-1}(X) \to \cdots \to Y_1(X) \), where \( Y_k(X) \) is a \( k-1 \)-step pro-nilflow such that the following holds: If \( u_1, \ldots, u_k \in \mathbb{R}^m \) are such that the actions of \( T_{u_i} \) and \( T_{u_i-u_j} \) for \( i \neq j \) are ergodic, then for any bounded measurable functions \( f_1, \ldots, f_k \) the limits in \( L^2(X) \)

\[
(2) \lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nu_j}f_j(x), \quad \lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} \pi^* \prod_{j=1}^{k} T_{nu_j} E(f_j|Y_k)(x)
\]

exist and are equal. The factor \( Y_k(X) \) is called the \( k \)-universal characteristic factor (\( k \)-u.c.f) of \( X \). Let

\[
\tau_{\hat{u}}(T) := T_{u_1} \times \cdots \times T_{u_k},
\]

let \( \triangle_k(\mu) \) be the diagonal measure on \( X^k \), then

\[
\bar{\triangle}_{\hat{u}}(\mu) := \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \tau_{\hat{u}}^* \triangle_k(\mu)
\]

is well defined. If \( F \) is a function invariant under \( \tau_{\hat{u}} \) with respect to the measure \( \bar{\triangle}_{\hat{u}}(\mu) \) and if \( E(f_j|B_{Y_k}) = 0 \) for some \( 1 \leq j \leq k \), then

\[
\int f_1(x_1) \cdots f_k(x_k) F(x_1, \ldots, x_k) d\bar{\triangle}_{\hat{u}} = 0.
\]

The factors \( Y_k(X) \) were constructed for an ergodic m.p.s \( X = (X^0, \mathcal{B}, \mu, T) \) by Host and Kra [HK2] and independently by Ziegler [Z]. Frantzikinakis and Kra [FKK2] showed that if \( X_i = (X^0, \mathcal{B}, \mu, T_i) \) are ergodic measure preserving systems on the same space \( X^0 \), where \( T_i \) commute, then \( B_{Y_{\hat{u}}(X_i)} = B_{Y_{\hat{u}}(X_j)} = B_{Y_k} \) for any \( i, j \), and if the action of \( T_i^{-1}T_j \) is ergodic for all \( i \neq j \), then equation (2) holds (replacing \( T_{nu_i} \) with \( T_{nu_i} = T_i^{nu_i} \)). Thus if we have a \( \mathbb{R}^m \) action then the systems \( X_u = (X^0, \mathcal{B}, \mu, T_u) \), \( u \in \mathbb{R}^m \) for which the action of \( T_u \) is ergodic, share the same sequence of factors \( Y_k(X) = Y_k(X_u) \). The fact that \( Y_k(X) \) is a factor of the \( \mathbb{R}^m \) action follows from [Z] corollary 2.4. We will show that the action of \( \mathbb{R}^m \) on \( Y_k(X) \) preserves the pro-nil structure:
3.8. Definition. Let \( Y = (Y^0, \mathcal{B}_Y, \mu_Y, \mathbb{R}^m) \) be a \( j \)-step pronilflow; \( Y = \lim_{\rightarrow} N_i/\Gamma_i \). We say that the action of \( \mathbb{R}^m \) on \( Y \) preserves the pro-nil structure if the action of \( \mathbb{R}^m \) on \( Y \) induces a \( \mathbb{R}^m \) action on \( N_i/\Gamma_i \) by group rotations.

3.9. Proposition. Let \( Y = (Y^0, \mathcal{B}, \mu, T) \) be a \( j \)-step ergodic pronilflow; \( Y = \lim_{\rightarrow} N_i/\Gamma_i \). Let \( \{T_c\}_{c \in \mathbb{R}^m} \) be a \( \mathbb{R}^m \) action on \( (Y^0, \mathcal{B}, \mu) \) that commutes with the action of \( T \). Then the action of \( \mathbb{R}^m \) on \( Y \) preserves the pro-nil structure.

Proof. For \( j = 1 \), \( Y \) is a Kronecker action, and any factor of \( Y \) is a Kronecker action. Thus it is enough to check that eigenfunctions of the \( T \) action are also eigenfunctions of the \( \mathbb{R}^m \) action. If \( \psi \) is an eigenfunction, \( T\psi(y) = \lambda \psi(y) \), then as \( T \) and \( T_c \) commute \( \psi(TT_c y) = \lambda \psi(T_c y) \). Combining the two equations we get

\[
T \left( \frac{\psi(T_c y)}{\psi(y)} \right) = 1.
\]

By ergodicity of \( T \) we get \( \psi(T_c y) = \delta_c \psi(y) \).

We proceed by induction on \( j \). Let \( Y \) be a \( j \)-step ergodic pronilflow; \( Y = \lim_{\rightarrow} M_i/\Lambda_i \). We first show that the \( \mathbb{R}^m \) action on \( Y \) induces a \( \mathbb{R}^m \) action on \( M_i/\Lambda_i \). Let \( \pi : Y \to M_i/\Lambda_i \) be the projection. Let \( p : Y \to Y_j(M_i/\Lambda_i) \) be the projection onto the \( j \) u.c.f of \( M_i/\Lambda_i \). \( Y_j(M_i/\Lambda_i) \) is a \( j-1 \)-step nilflow, we denote it \( N_i/\Gamma_j \). The space \( L^2(M_i/\Lambda_i) \circ \pi \subset L^2(Y) \) is spanned by functions \( f \) satisfying the following condition:

\[
Tf(y) = g(y)f(y)
\]

where \( g = \tilde{g} \circ p \) with \( \tilde{g} \) of type \( j \) (see \([Z]\) theorem 6.1). As \( T, T_c \) commute for any \( c \in \mathbb{R}^m \)

\[
TT_c f(y) = T_c Tf(y) = T_c g(y) T_c f(y).
\]

Thus

\[
T \left( \frac{f(T_c y)}{f(y)} \right) = \frac{T_c g(y) f(T_c y)}{g(y) f(y)}.
\]

By the induction hypothesis the action of \( \mathbb{R}^m \) on \( Y \) induces an action on \( Y_j(M_i/\Lambda_i) = N_i/\Gamma_i \), and this action is given by rotation by an element \( a_i(c) \in N_i \). By proposition 6.37 in \([Z]\), as \( \mathbb{R}^m \) commutes with the action of \( T \) on \( N_i/\Gamma_i \) given by rotation by \( a \in N_i \), there exists a family of measurable functions \( \{f_c : N_i/\Gamma_i \to \mathcal{S}^1\}_{c \in \mathbb{R}^m} \) and a family of constants \( \{\lambda_c\}_{c \in \mathbb{R}^m} \) such that

\[
\frac{T_c g(p(y))}{g(p(y))} = \lambda_c \frac{T_c f_c(p(y))}{f_c(p(y))}.
\]

We get

\[
T \left( \frac{f(T_c y)}{f(y)f_c(p(y))} \right) = \lambda_c \frac{f(T_c y)}{f(y)f_c(p(y))}.
\]
This implies that \( \lambda_c \) is an eigenvalue of \( T \), but as it is multiplicative in \( c \in \mathbb{R}^m \), \( \lambda_c \equiv 1 \). Therefore by ergodicity of the \( T \) action

\[
\frac{f(T_c y)}{f(y)f_c(p(y))} = \delta_c'
\]
or

\[
f(T_c y) = \delta_c f(y)f_c(p(y)) \in L^2(M_i/\Lambda_i) \circ \pi.
\]

This shows that the \( \mathbb{R}^m \) action on \( Y \) induces an \( \mathbb{R}^m \) action on \( M_i/\Lambda_i \). The fact that this action is given by group rotations was shown by Parry \([Pa2]\) in the case where \( M_i \) is connected. Alternatively, \( M_i/\Lambda_i \) can be presented as a torus extension of \( Y_j(M_i/\Lambda_i) = N_i/\Gamma_i \) with \( g : N_i/\Gamma_i \to \mathbb{T}^n \) a cocycle of type \( j \). Without loss of generality we can assume \( n = 1 \). Now the tuples \((a, g), (a_i(c), f_c)\) belong to the group \( G \) defined in \([Z]\) proposition 6.37.

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

The following is a version of the van der Corput Lemma (see \([FuKaW]\)).

**3.11. Lemma.** Let \( H \) be a Hilbert space, \( \xi \in \Xi \) some index set, and let \( u_n(\xi) \in H \) for \( n \in \mathbb{N} \) be uniformly bounded in \( n, \xi \). Assume that for each \( r \) the limit

\[
\gamma_r(\xi) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle u_n(\xi), u_{n+r}(\xi) \rangle
\]

exists uniformly and

\[
\lim_{R \to \infty} \frac{1}{R} \sum_{r=1}^{R} \gamma_r(\xi) = 0
\]

uniformly. Then

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

uniformly in \( \xi \).

**3.12. Multidimensional Szemerédi.** The following generalization of Szemerédi’s theorem was proved by Furstenberg and Katznelson \([FuKa]\).

**3.13. Theorem.** Let \( X = (X^0, \mathcal{B}, \mu, \mathbb{Z}^k) \) be a m.p.s., and let \( T_1, \ldots, T_k \) be the generators of the \( \mathbb{Z}^k \) action. Let \( f \geq 0 \) be a bounded measurable function on \( X \) with \( \int f \, d\mu > 0 \). Then

\[
\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int f(x)T_n^0 f(x) \cdots T_n^0 f(x) \, d\mu(x) > 0.
\]
4. The Main Theorem

Denote \( M_m(\mathbb{R}) \) the \( m \times m \) matrices over \( \mathbb{R} \), and \( SO(m) \) the special orthogonal group. Recall that if \((N/\Gamma,G)\) is a nilflow the action of \( T_g \) for \( g \in G \) is given by \( T_g n \Gamma = a_g n \Gamma \) where \( a_g \in N \).

4.1. Lemma. Let \((N/\Gamma, \mathbb{R}^m)\) be an ergodic measure preserving action of \( \mathbb{R}^m \) on a nilmanifold \( N/\Gamma \), where \( N \) is connected. Let \( f_j \) be continuous functions on \( N/\Gamma \). Let \((u_1, \ldots, u_l) \in (\mathbb{R}^m)^l \). Then there exists a countably linear set \( S \subset M_m(\mathbb{R}) \) such that for any \( F \in M_m(\mathbb{R}) \setminus S \) the function

\[
g_{F,L}(x) := \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{l} T_{(nF+L)u_j} f_j(x)
\]

is independent of \( L \in M_m(\mathbb{R}) \) for a.e. \( x \in N/\Gamma \). Furthermore for any such \( F \) the convergence is uniform in \( L \).

Proof. Let \( M \) be the diagonal of \( N^l \) and let \( G = M_m(\mathbb{R}) = \mathbb{R}^{m^2} \) (thought of as an additive group). Let \( \varphi : M_m(\mathbb{R}) \to N^l \) be given by

\[
\varphi(F) = (a_{Fu_1}, \ldots, a_{Fu_l}).
\]

By proposition 3.6 there exists a submanifold \( V \) of \((N/\Gamma)^l\), and there exists a countably linear set \( S \subset M_m(\mathbb{R}) \) such that for every \( F \in M_m(\mathbb{R}) \setminus S \) there is a countably polynomial set \( A_F \subset M \) such that for \( (a, \ldots, a) \notin A_F \),

\[
\{\varphi(nF) \pi(a, \ldots, a)^l \}_{n \in \mathbb{Z}} = (a, \ldots, a)V,
\]

and

\[
G \pi(a, \ldots, a) \subseteq (a, \ldots, a)V \quad \text{(therefore } (a, \ldots, a)V).\]

For any \( F \in M_m(\mathbb{R}) \setminus S \), and \( (a, \ldots, a) \notin A_F \) we have

\[
T_{Lu_1} \times \ldots \times T_{Lu_l} \pi(a, \ldots, a) \in (a, \ldots, a)V.
\]

The action of \( \varphi(F) \) on \((a, \ldots, a)V\) is ergodic, and by theorem 3.4 it is uniquely ergodic. The point \( (T_{Lu_1}a \Gamma, \ldots, T_{Lu_l}a \Gamma) \in (a, \ldots, a)V \). By theorem 3.3 the convergence in equation (4) is uniform in \( L \), and \( g_{F,L}(a \Gamma) \) is independent of \( L \).

\[
\square
\]

4.2. Corollary. Let \( Y = (Y^0, \mathcal{B}, \mu, \mathbb{R}^m) \) be an ergodic pro-nilflow. Let \( f_j \) be bounded measurable functions on \( Y^0 \). Let \((u_1, \ldots, u_l) \in (\mathbb{R}^m)^l \). Then there exists a countably linear set \( S \subset M_m(\mathbb{R}) \) such that for any \( F \in M_m(\mathbb{R}) \setminus S \), and all \( L \in M_m(\mathbb{R}) \) the function

\[
g_{F,L}(y) := \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{l} T_{(Fn+L)u_j} f_j(y)
\]

where the limit is in \( L^2(Y) \), is a constant function of \( L \in M_m(\mathbb{R}) \) and the convergence is uniform in \( L \).

Proof. If \( Y^0 = \lim_{n} N_j/\Gamma_j \), the continuous functions on \( N_j/\Gamma_j \) lifted to \( Y^0 \), for all \( j \), are dense in \( C(Y^0) \).

\[
\square
\]
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 factor \( Y_k(X) \) described in 3.7.

4.3. Proposition. Let \( X = (X^0, \mathcal{B}, \mu, \mathbb{R}^m) \) be an ergodic action of \( \mathbb{R}^m \), and let \((u_1, \ldots, u_k) \in (\mathbb{R}^m)^k \). Let \( Y_k \) be the factor described in theorem [3.7], and let \( \pi : X \to Y_k \) be the factor map. Let \( f_1 \ldots f_k \) be bounded measurable functions on \( X \). Then there exists a countably linear subset \( \mathcal{S} \subset M_m(\mathbb{R}) \) such that for any \( M \in M_m(\mathbb{R}) \setminus \mathcal{S} \), satisfying \( T_M u_i, T_{M(u_i-u_j)} \) for \( i, j = 1, \ldots, k, i \neq j \) are ergodic, and for all \( P \in M_m(\mathbb{R}) \) we have

\[
\frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{(nM+P)u_j} f_j(x) - \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{(nM+P)u_j} \pi^* E(f_j|Y_k)(x) \xrightarrow{L^2(X)} 0
\]

uniformly in \( P \).

Proof. We prove this inductively. For \( k = 1 \), let \( u \in \mathbb{R}^k \), \( u \neq 0 \). If \( T_M u \) is ergodic then

\[
\frac{1}{N} \sum_{n=1}^{N} T_{nMu+Pu} f(x) = T_{Pu} \left( \frac{1}{N} \sum_{n=1}^{N} T_{nMu} f(x) \right) \xrightarrow{\text{ergodic}} \int f(x) d\mu
\]

uniformly in \( P \), by the Mean Ergodic Theorem. Assume the statement holds for \( k \): i.e., for \( M \) outside a countably linear set satisfying \( T_M u_i, T_{M(u_i-u_j)} \) for \( i, j = 1, \ldots, k, i \neq j \) are ergodic, and all \( P \in M_m(\mathbb{R}) \) we have

\[
\frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{(nM+P)u_j} f_j(x) - \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{(nM+P)u_j} \pi^* E(f_j|Y_k)(x) \xrightarrow{L^2(X)} 0
\]

uniformly in \( P \). We show this for \( k + 1 \). Let \( \mathcal{S} \subset M_m(\mathbb{R}) \) be the set from corollary [4.2] corresponding to \( Y_k \) and \( u_1, \ldots, u_{k+1} \). For \( M \in M_m(\mathbb{R}) \setminus \mathcal{S} \) the \( L^2 \) limit

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k+1} T_{(nM+P)u_j} E(f_j|Y_k)(y)
\]

is independent of \( P \), and the convergence to the limit in uniform in \( P \). Let \( M \in M_m(\mathbb{R}) \setminus \mathcal{S} \) satisfy \( T_M u_i, T_{M(u_i-u_j)} \) are ergodic for \( i, j = 1, \ldots, k + 1, i \neq j \). It is enough to show that if for some \( 1 \leq j \leq k + 1 \), \( E(f_j|Y_{k+1}) = 0 \) then

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k+1} T_{(nM+P)u_j} f_j(x) = 0
\]

uniformly in \( P \). We use the Van der Corput Lemma (lemma [5.11]). Let \( v_n(M, P) := \prod_{j=1}^{k+1} T_{nMu_j+Pu_j} f_j(x) \). Then

\[
\langle v_n(M, P), v_{n+r}(M, P) \rangle = \int \prod_{j=1}^{k+1} T_{nMu_j+Pu_j} f_j(x) T_{(n+r)Mu_j+Pu_j} \tilde{f}_j(x) \, d\mu,
\]
and
\[
\gamma_r(M,P) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \langle v_n(M,P), v_{n+r}(P,M) \rangle
\]
\[
= \lim_{N \to \infty} \int f(x)T_{rMu_1}f(x) \frac{1}{N} \sum_{n=1}^{N} \prod_{j=2}^{k+1} T_{nM(u_j-u_1)+P(u_j-u_1)}(f_j(x)T_{rMu_j}f_j(x)) \ d\mu.
\]
By the induction hypothesis this limit is equal (uniformly in \(P\)) to the following limit
\[
\lim_{N \to \infty} \int f(x)T_{rMu_1}f(x) \frac{1}{N} \sum_{n=1}^{N} \prod_{j=2}^{k+1} T_{nM(u_j-u_1)+P(u_j-u_1)}\pi^*E(f_jT_{rMu_j}f_j|Y_k)(x) \ d\mu,
\]
which equals
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k+1} \int T_{nM(u_j-u_1)+P(u_j-u_1)}\pi^*E(f_jT_{rMu_j}f_j|Y_k)(x) \ d\mu.
\]
The limit in equation (6) is a limit on a \(k - 1\) step pronilflow, as \(M \in M_m(\mathbb{R}) \ \backslash \ S\) it is the same for all \(P\), and the convergence is uniform in \(P\). By 3.7 the limit in equation (6) is equal to
\[
\int \prod_{j=1}^{k+1} T_{P(u_j-f_j(x_j))T_{rMu_j+P(u_j-f_j(x_j))}d\bar{\Delta}_{M\mathbb{U}}(\mu)(x_1,\ldots,x_{k+1}),
\]
where \(\bar{\Delta}_{M\mathbb{U}}(\mu)\) is a measure on \(X^{k+1}\). Now
\[
\lim_{R \to \infty} \frac{1}{R} \sum_{r=1}^{R} \prod_{j=1}^{k+1} T_{rMu_j+P(u_j-f_j(x_j))}
\]
converges uniformly in \(P\) to a function \(F\) in \(L^2(\bar{\Delta}_{M\mathbb{U}}(\mu))\) which is invariant under \(T_{Mu_1} \times \ldots \times T_{Mu_{k+1}}\) (by the Mean Ergodic Theorem, as in the case \(k = 1\)). Finally by 3.7 if \(E(f_j|Y_{k+1}) = 0\) for some \(1 \leq j \leq k + 1\), then
\[
\lim_{R \to \infty} \frac{1}{R} \sum_{r=1}^{R} \gamma_r(M,P) = 0.
\]
(uniformly in \(P\)).

4.4. Remark. Corollary 4.2 and proposition 4.3 remain valid if we replace \(M_m(\mathbb{R})\) by a linear subspace of \(M_m(\mathbb{R})\). We apply this for the case \(m = 2\), replacing \(M_2(\mathbb{R})\) by the embedding \(\mathbb{C} \hookrightarrow M_2(\mathbb{R})\). Then, thinking of \(u_1,\ldots,u_k\) as points in \(\mathbb{C}\) we can replace the matrices \(F,L \in M_2(\mathbb{R})\) with \(c,d \in \mathbb{C}\) where \(c\) is outside countably many lines in \(\mathbb{C}\).
4.5. Lemma. For each \( r = 1, \ldots, \infty \), let \( \{s_i^r\}_{i=1}^m \subset \mathbb{R}^m \), such that for each \( r \), \( s_i^r \neq 0 \) for some \( 1 \leq l \leq m \). Let \( \{e_i\}_{i=1}^m \) be the standard basis for \( \mathbb{R}^m \). There exists an antisymmetric matrix \( B \subset M_m(\mathbb{R}) \) s.t. \( Bu_i, B(u_i - u_j) \neq \vec{0} \) for \( 1 \leq i, j \leq k, i \neq j \), and

\[
\forall r : f_{r,B}(M) \overset{\text{def}}{=} \sum_{l=1}^m \langle s_i^r, MBe_l \rangle \neq 0.
\]

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

\[
\forall B \in B : \sum_{l=1}^m \langle Ms_l, Be_l \rangle \equiv 0
\]

Without loss off generality \( s_{11} \neq 0 \). Let \( E_{21} \) be an \( m \times m \) matrix with 1 at the index 21, and 0 elsewhere. Then

\[
\sum_{l=1}^m \langle E_{21}s_l, Be_l \rangle = s_{11}b_{21} + s_{21}b_{22} + \ldots + s_{m1}b_{2m}
\]

\[
= -s_{11}b_{12} + s_{31}b_{23} + \ldots + s_{m1}b_{2m} = 0
\]

As \( s_{11} \neq 0 \) this is a non trivial linear condition on antisymmetric matrices. Finally, the conditions \( Bu_i = \vec{0} \) or \( B(u_i - u_j) = \vec{0} \) are non trivial linear conditions on antisymmetric matrices. \( \square \)

4.6. Lemma. Let \( S \) be a countably linear set in \( M_m(\mathbb{R}) \), \( m \geq 3 \). Let \( u_1, \ldots, u_k \in \mathbb{R}^m \). There exist matrices \( M \in M_m(\mathbb{R}) \setminus S \), and \( P \in SO(m) \) such that \( M^tP \) is antisymmetric, and \( T_{Mu_i}, T_{Mu_i-uj} \) for \( i, j = 1, \ldots, k \), \( i \neq j \) are ergodic.

Proof. The set \( S \) is countably linear therefore it is a countable union of sets of the form

\[
S_r = \{N \in M_m(\mathbb{R}) : \sum_{l=1}^m \langle s_l^r, Ne_l \rangle = c_r \},
\]

Where \( e_l \) is the standard basis for \( \mathbb{R}^m \), \( s_l^r \in \mathbb{R}^m \), \( c_r \in \mathbb{R} \). By lemma 4.5 there exists an antisymmetric matrix \( B \), such that \( Bu_i, B(u_i - u_j) \neq \vec{0} \) for \( 1 \leq i, j \leq k, i \neq j \), and for all \( r \)

\[
f_r(M) = \sum_{l=1}^m \langle s_l^r, MBe_l \rangle \neq 0.
\]

For each \( r \), the set of \( M \) with \( f_r(M) = c_r \) is a hyperplane in \( M_m(\mathbb{R}) \). This subspace intersects \( SO(m) \) in a proper algebraic subvariety of \( SO(m) \). Therefore for a.e. \( P \in SO(m) \) (with respect to the Haar measure on
SO(m)), \( M = PB \) will avoid the bad set \( \tilde{S} \). Clearly, if \( M = PB \) avoids \( S \) then \( tM = tPB \) avoids \( S \) for any \( t > 0 \). By proposition 3.10 for a.e. \( P \in SO(m) \) and a.e. \( t \in \mathbb{R} \) \( T_{tPBu} \) act ergodically.

4.7. Proof of theorem 2.3 Without loss of generality, we may assume by disintegration of \( \mu \), that the action of \( \mathbb{R}^m \) is ergodic. Let \( f = 1_A \) be the characteristic function of the set \( A \), and let \( \mu(A) = \lambda \). Let \( Y_k \) be the factor described in theorem 3.7 and let \( E(f|Y_k) \) be the projection of \( f \) on \( L^2(Y_k) \).

We first prove the theorem for \( m > 2 \). By corollary 4.2, proposition 4.3 and lemma 4.6, there exist matrices \( M \in M_m(\mathbb{R}) \), \( P \in SO(m) \) such that \( M'P \) is antisymmetric, and for all \( t \in \mathbb{R} \) we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} E(f|Y_k)(y) = g(y)
\]

in \( L^2(Y_k) \), and

\[
\frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} f(x) - \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} \pi^* g(x) \to 0
\]

in \( L^2(X) \), where the convergence is uniform in \( t \). Then

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} f(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{nM_{uj}} f(x) = \pi^* g(x),
\]

and the convergence is uniform in \( t \). By theorem 3.13

\[
\int f(x) \pi^* g(x) d\mu_X > C > 0.
\]

Uniform convergence implies that there exists \( N_0 \), such that for all \( t \)

\[
\left| \frac{1}{N_0} \sum_{n=1}^{N_0} \int f(x) \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} f(x) d\mu_X - \int f(x) \pi^* g(x) d\mu_X \right| < \frac{C}{2}.
\]

Therefore for \( N_0 \), and for all \( t \in \mathbb{R} \)

\[
\frac{1}{N_0} \sum_{n=1}^{N_0} \int f(x) \prod_{j=1}^{k} T_{nM_{uj} + tP_{uj}} f(x) d\mu_X > \frac{C}{2}.
\]

This implies that for all \( t \in \mathbb{R} \) there exists \( n \leq N_0 \) with

\[
\mu(A \cap T_{(nM+tP)u_1} A \cap \ldots \cap T_{(nM+tP)u_k} A) = \int f(x) \prod_{j=1}^{k} T_{(nM+tP)u_j} f(x) d\mu_X > \frac{C}{2}.
\]

Now the \( T_u \) satisfy the following continuity condition:

(7) \( \forall \varepsilon \exists \delta : ||u - u'|| \leq \delta \Rightarrow |\mu(A \cap T_u A) - \mu(A \cap T_{u'} A)| \leq \varepsilon. \)
As $M^t P$ is antisymmetric, $M \in T_P(SO(m))$ - the tangent space of $SO(m)$ at $P$. Thus

$$P' := P \exp(\epsilon n P^{-1} M) = P(I + \epsilon n P^{-1} M + o(\epsilon)) = P + \epsilon n M + o(\epsilon)$$

belongs to $SO(m)$. But

$$\left( \frac{1}{\epsilon} P + n M \right) - \frac{1}{\epsilon} P' = o(1),$$

and if $t = \frac{1}{\epsilon}$ is large enough, then by equation (7)

$$\mu(A \cap T_{1P'u_1} \cap \ldots \cap T_{1P'u_k}) > \frac{C}{4}.$$  

For $m = 2$ the proof is similar. By remark 4.4 there exists $c \in \mathbb{C}$ such that for all $t \in \mathbb{R}$ we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{ncu_j + it cu_j} E(f|Y_k)(y) = g(y)$$

in $L^2(Y)$, and

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{ncu_j + it cu_j} f(x) - \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{k} T_{ncu_j + it cu_j} \pi^* E(f|Y_k)(x) \to 0$$

in $L^2(X)$, where the convergence is uniform in $t$. As in the proof for $m > 2$, there exists $N_0$, such that for all $t$

$$\frac{1}{N_0} \sum_{n=1}^{N_0} \int f(x) \prod_{j=1}^{k} T_{ncu_j + it cu_j} f(x) d\mu_X > \frac{C}{2}.$$  

This implies that for all $t \in \mathbb{R}$ there exists $n \leq N_0$ with

$$\mu(A \cap T_{(n+it)cu_1} A \cap \ldots \cap T_{(n+it)cu_k} A) =$$

$$\int f(x) \prod_{j=1}^{k} T_{(n+it)cu_j} f(x) d\mu_X > \frac{C}{2}.$$  

If $t$ is large enough, then

$$(n + it)cu_j \sim \frac{t}{|n + it|} (n + it)cu_j,$$

and $\left| \frac{t}{|n + it|} (n + it) \right| = t$.

REFERENCES

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

[FM] Falconer, K.J., Marstrand, J.M. Plane sets with positive density at infinity contain all large distances. Bull. Lond. Math. Soc. 18, 471 – 474 (1986).

[Ful] Furstenberg, H. Recurrence in ergodic theory and combinatorial number theory. Princeton University Press, Princeton, NJ, USA (1981).
[Fu2] Furstenberg, H. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. d'Analyse Math 31, 204 – 256 (1977).

[FrKr] Frantzikinakis, N., Kra, B. Convergence of multiple ergodic averages for some commuting transformations. To appear in Ergodic Theory Dynam. Systems.

[FuKa] Furstenberg, H., Katznelson, Y. An ergodic Szemerédi theorem for commuting transformations. J. Analyse Math 34, 275-291 (1978).

[FuKaW] Furstenberg, H., Katznelson, Y., Weiss, B. Ergodic theory and configurations in sets of positive density. Mathematics of Ramsey Theory, Springer Verlag 1990, Nešetřil, Rödl, eds.

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

[HKr] Host, B., Kra, B. Non-conventional ergodic averages and nilmanifolds. To appear in the Ann. of Math.

[Le] Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. Ergodic Theory Dynam. Systems 25 no. 1, 201 – 213 (2005).

[Ma] Malcev, A. On a class of homogeneous spaces. Izvestiya Akad. Nauk SSSR, Ser. Mat 13, 9-32 (1949).

[Pa1] Parry, W. Ergodic properties of affine transformations and flows on nilmanifolds. Amer. J. Math 91 757-771 (1969).

[Pa2] Parry, W. Metric classification of ergodic nilflows and unipotent affines Amer. J. of Math, 93, pp. 819-828 (1969).

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

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

[Sh] Shah, N. Invariant measures and orbit closures on homogeneous spaces for actions of subgroups. Lie groups and ergodic theory (Mumbai, 1996), 229-271, Tata Inst. Fund. Res. Stud. Math., 14, Tata Inst. Fund. Res., Bombay, 1998.

[Z] Ziegler, T. Universal Characteristic Factors and Furstenberg Averages preprint available on-line at [math.DS/0403212](http://math.DS/0403212)