# AN ERGODIC SZEMERÉDI THEOREM FOR COMMUTING TRANSFORMATIONS

*By*

H. FURSTENBERG AND Y. KATZNELSON

The classical Poincaré recurrence theorem asserts that under the action of a measure preserving transformation $T$ of a finite measure space $(X, \mathcal{B}, \mu)$, every set $A$ of positive measure recurs in the sense that for some $n > 0$, $\mu(T^{-n}A \cap A) > 0$. In [1] this was extended to *multiple recurrence*: the transformations $T, T^2, \cdots, T^k$ have a common power satisfying $\mu(A \cap T^{-n}A \cap \cdots \cap T^{-kn}A) > 0$ for a set $A$ of positive measure. We also showed that this result implies Szemerédi's theorem stating that any set of integers of positive upper density contains arbitrarily long arithmetic progressions. In [2] a topological analogue of this is proved: if $T$ is a homeomorphism of a compact metric space $X$, for any $\varepsilon > 0$ and $k = 1, 2, 3, \cdots$, there is a point $x \in X$ and a common power of $T, T^2, \cdots, T^k$ such that $d(x, T^n x) < \varepsilon$, $d(x, T^{2n}x) < \varepsilon, \cdots, d(x, T^{kn}x) < \varepsilon$. This (weaker) result, in turn, implies van der Waerden's theorem on arithmetic progressions for partitions of the integers. Now in this case a virtually identical argument shows that the topological result is true for any $k$ commuting transformations. This would lead one to expect that the measure theoretic result is also true for arbitrary commuting transformations. (It is easy to give a counterexample with noncommuting transformations.) We prove this in what follows.

**Theorem A.** *Let $(X, \mathcal{B}, \mu)$ be a measure space with $\mu(X) < \infty$, let $T_1, T_2, \cdots, T_k$ be commuting measure preserving transformations of $X$ and let $A \in B$ with $\mu(A) > 0$. Then*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{1}^{N} \mu(T_1^{-n}A \cap T_2^{-n}A \cap \cdots \cap T_k^{-n}A) > 0.$$

A corollary is the multidimensional extension of Szemerédi's theorem:

**Theorem B.** *Let $S \subset \mathbf{Z}^r$ be a subset with positive upper density and let $F \subset \mathbf{Z}^r$ be any finite configuration. Then there exists an integer $d$ and a vector $n \in \mathbf{Z}^r$ such that $n + dF \subset S$.*



JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. 34 (1978)

Here the upper density is taken with respect to any sequence of cubes

$$[a_n^{(1)}, b_n^{(1)}] \times [a_n^{(2)}, b_n^{(2)}] \times \cdots \times [a_n^{(r)}, b_n^{(r)}] \quad \textit{with } b_n^{(j)} - a_n^{(j)} \to \infty.$$

The proof of Theorem B on the basis of Theorem A is carried out as in the one-dimensional case ([1], [4]).

The ideas developed in this paper owe a good deal to conversations with D. Ornstein and J.-P. Thouvenot, to whom the authors would like to express their gratitude.

## 1. Relative ergodicity and weak mixing

Throughout the discussion we shall consider measure spaces on which a fixed group $\Gamma$ which is countable and commutative acts by measure preserving transformations. We say $(Y, \mathscr{D}, \nu)$ is a $\Gamma$-invariant *factor* of $(X, \mathscr{B}, \mu)$ if we have a map $\pi : X \to Y$ with $\pi^{-1}\mathscr{D} \subset \mathscr{B}$, $\pi\mu = \nu$ and for each $T \in \Gamma$, $T\pi(x) = \pi T(x)$. A factor of $(X, \mathscr{B}, \mu)$ is determined by a $\Gamma$-invariant closed subalgebra of $L^\infty(X, \mathscr{B}, \mu)$. $(X, \mathscr{B}, \mu)$ *is an extension* of $(Y, \mathscr{D}, \nu)$. We assume, as we may, that $(X, \mathscr{B}, \mu)$ is a "regular measure space" ([1], §4). Then we can associate to the factor $(Y, \mathscr{D}, \nu)$ a family of measures $\{\mu_y \mid y \in Y\}$ on $(X, \mathscr{B})$ such that for each $f \in L^1(X, \mathscr{B}, \mu)$, $f \in L^1(X, \mathscr{B}, \mu_y)$ for a.e. $y \in Y$, and $\int f d\mu_y$ is measurable and integrable in $(Y, \mathscr{D}, \nu)$ with

$$\int \left\{ \int f(x) d\mu_y(x) \right\} d\nu(y) = \int f(x) d\mu(x).$$

We write $\mu = \int \mu_y d\nu(y)$ and speak of this decomposition as the disintegration of $\mu$ with respect to the factor $(Y, \mathscr{D}, \nu)$. The $\mu_y$ are well defined up to sets of measure 0 in $Y$. The fact that $T \in \Gamma$ is measure preserving on $(X, \mathscr{B}, \mu)$ translates into $T\mu_y = \mu_{Ty}$ where, for any measure $\theta$, $T\theta$ is defined by $T\theta(A) = \theta(T^{-1}(A))$, or by

$$\int f(x) dT\theta(x) = \int f(Tx) d\theta(x).$$

We say that $(X, \mathscr{B}, \mu)$ is a *relatively ergodic extension* of $(Y, \mathscr{D}, \nu)$ for an element $T \in \Gamma$ if every $T$-invariant function on $X$ is (a.e.) a function on $Y$. Given two extensions $(X, \mathscr{B}, \mu)$ and $(X', \mathscr{B}', \mu')$ of $(Y, \mathscr{D}, \nu)$ we may form the fibre product $(\tilde{X}, \tilde{\mathscr{B}}, \tilde{\mu})$ where

$$\tilde{X} = X \times_Y X = \{(x, x') \in X \times X' : \pi(x) = \pi'(x')\},$$

$\tilde{\mathcal{B}}$ is the restriction of $\mathcal{B} \times \mathcal{B}'$ and $\tilde{\mu}$ is defined by the disintegration

$$\tilde{\mu}_y = \mu_y \times \mu'_y$$

where $\mu = \int \mu_y d\nu(y)$ and $\mu' = \int \mu'_y d\nu(y)$ are the disintegrations of $\mu$ and $\mu'$ respectively. We then say that $(X, \mathcal{B}, \mu)$ is a *relatively weak mixing* extension of $(Y, \mathcal{D}, \nu)$ for $T \in \Gamma$ if $(X \times_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$ is a relatively ergodic extension of $(Y, \mathcal{D}, \nu)$ for $T$.

**Lemma 1.1.** *Let $F: X \to M$ be a measurable map from a measure space $(X, \mathcal{B}, \mu)$ to a separable metric space $M$ and assume that the function $d(F(x), F(x'))$ is a.e. constant on $X \times X$. Then $F(x)$ is a.e. constant.*

**Proposition 1.2.** *If $(X, \mathcal{B}, \mu)$ is a relatively weak mixing extension of $(Y, \mathcal{D}, \nu)$ for $T \in \Gamma$ and $(X', \mathcal{B}', \mu')$ is a relatively ergodic extension of $(Y, \mathcal{D}, \nu)$ for $T$, then $(X \times_Y X', \tilde{\mathcal{B}}, \tilde{\mu})$ is a relatively ergodic extension of $(Y, \mathcal{D}, \nu)$ for $T$.*

**Proof.** Let $\pi: X \to Y$, $\pi': X' \to Y$ be the associated maps and assume that $H(x, x')$ is a $T$-invariant function on $X \times_Y X'$. Form the function $E(x_1, x_2)$ on $X \times_Y X$ given by

$$E(x_1, x_2) = \int |H(x_1, x') - H(x_2, x')| \, d\mu'_{\pi(x_1)}(x')$$

where $\mu' = \int \mu'_y d\nu(y)$. One sees that $E$ is $T$-invariant and so is a function of $\pi(x_1) = \pi(x_2)$. We apply Lemma 1.1 to the map $x \to H(x, \cdot)$ on $(X, \mathcal{B}, \mu_y)$ and conclude that it depends only on $\pi(x)$. Hence $H(x, x')$ is a function of $x'$ and by relative ergodicity of the extension $\pi': X' \to Y$, we see that $H$ depends only on $\pi(x) = \pi'(x')$. This proves the proposition.

If $(X, \mathcal{B}, \mu)$ is an extension of $(Y, \mathcal{D}, \nu)$ we denote by $E(f \mid Y)$ the conditional expectation which maps $L^p(X, \mathcal{B}, \mu)$ to $L^p(Y, \mathcal{D}, \nu)$ and is defined by $E(f \mid Y)(y) = \int f d\mu_y$ a.e. We shall frequently use the identity

$$\int_Y E(f \mid Y)^2 d\nu = \iiint f(x_1) f(x_2) d\mu_y(x_1) d\mu_y(x_2) d\nu(y) \tag{1.1}$$

$$= \int_{X \times_Y X} f(x_1) f(x_2) d\tilde{\mu}(x_1, x_2).$$

**Lemma 1.3.**[†] *Let $(X, \mathcal{B}, \mu)$ be a relatively weak mixing extension of $(Y, \mathcal{D}, \mu)$ for $T$ and let $\varphi, \psi \in L^2(X, \mathcal{B}, \mu)$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int [E(\psi T^n \varphi \mid Y) - E(\psi \mid Y) T^n E(\varphi \mid Y)]^2 d\nu(y) = 0.$$

**Proof.** We can assume $E(\psi \mid Y) = 0$. So we wish to evaluate

$$(1.2)\quad \frac{1}{N} \sum_{n=1}^{N} \int E(\psi T^n \varphi \mid Y)^2 d\nu(y) = \frac{1}{N} \sum_{n=1}^{N} \int \psi(x_1)\psi(x_2)\varphi(T^n x_1)\varphi(T^n x_2) d\tilde{\mu}(x_1, x_2)$$

by (1.1). Now a weakly convergent subsequence of $(1/N)\Sigma_{n=1}^{N} \varphi(T^n x_1)\varphi(T^n x_2)$ will converge to a $T$-invariant function on $X \times_Y X$, which, by hypothesis, is a function of $\pi(x_1) = \pi(x_2)$. The limit of (1.2) is then expressed in terms of $E(\psi \mid Y)^2 = 0$, and this being the case for any convergent subsequence we obtain the lemma.

We generalize the foregoing in the next theorem.

**Theorem 1.4.** *Let $(X, \mathcal{B}, \mu)$ be a relative weak mixing extension of $(Y, \mathcal{D}, \nu)$ for every $T \neq 1$, $T \in \Gamma$. Then if $f_1, \cdots, f_l \in L^\infty(X, \mathcal{B}, \mu)$ and $T_1, \cdots, T_l$ are distinct elements of $\Gamma$,*

$$(1.3)\qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \left[ E\left( \prod_{j=1}^{l} T_j^n f_j \mid Y \right) - \prod_{j=1}^{l} T_j^n E(f_j \mid Y) \right]^2 d\nu(y) = 0.$$

**Proof.** Write $g_j = T_j^n f_j$. If we express $E(\Pi g_j \mid Y) - \Pi E(g_j \mid Y)$ as

$$\sum E(g_1 g_2 \cdots g_{i-1}(g_i - E(g_i \mid Y) E(g_{i+1} \mid Y) \cdots E(g_l \mid Y) \mid Y)$$

we see that we can reduce (1.3) to the case where some $E(f_i \mid Y) = 0$. So we assume $E(f_l \mid Y) = 0$. Using (1.1) the problem is to prove

$$(1.4)\qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int_{X \times_Y X} \prod_{j=1}^{l} T_j^n f_j(x_1) T_j^n f_j(x_2) d\tilde{\mu}(x_1, x_2) = 0$$

given that $E(f_l \mid Y) = 0$. Since by Proposition 1.2, $X \times_Y X$ is a relatively weak mixing extension of $Y$ whenever $X$ is, (1.4) will follow if we prove that

[†] Here and elsewhere the operator $T$ is defined for $T \in \Gamma$ by $Tf(x) = f(Tx)$. Note that $TE(f \mid Y) = E(Tf \mid Y)$.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int_X \prod_{j=1}^{l} T_j^n f_j d\mu(x) = 0 \tag{1.5}$$

whenever $E(f_i \mid Y) = 0$.

When $l = 1$ the result is clear since $T$ is measure preserving so we proceed by induction and assume the theorem is valid for $l-1$. Set $S_i = T_i T_l^{-1}$, $i = 1, \cdots, l-1$. The $S_i$ are all distinct and also different from 1, and we assume that (1.3) holds with $l$ replaced by $l-1$ and the $T_i$ by the $S_i$.

Let $\mu_\Delta^{l-1}$ denote the diagonal measure on $X^{l-1}$ (see [1] for details on "standard measures" in product spaces) and let $\nu_\Delta^{l-1} = \pi\mu_\Delta^{l-1}$ be the diagonal measure on $Y^{l-1}$. Set

$$\mu_*^{l-1} = \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} (S_1 \times \cdots \times S_{l-1})^n \mu_\Delta^{l-1},$$

$$\nu_*^{l-1} = \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} (S_1 \cdots S_{l-1})^n \nu_\Delta^{l-1},$$

where the limits in question refer to convergence with respect to integration against functions of the form $g_1 \otimes \cdots \otimes g_{l-1}(x_1, \cdots, x_{l-1}) = g_1(x_1) \cdots g_{l-1}(x_{l-1})$, and $N_k$ is a subsequence for which these limits exist. We find

$$\begin{aligned}\int g_1 \otimes \cdots \otimes g_{l-1} d\mu_*^{l-1} &= \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} \int S_1^n g_1(x) \cdots S_{l-1}^n g_{l-1}(x) d\mu(x) \\ &= \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} \int E(S_1^n g_1 \cdots S_{l-1}^n g_{l-1} \mid Y) d\nu(y) \\ &= \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} \int S_1^n E(g_1 \mid Y) \cdots S_{l-1}^n E(g_{l-1} \mid Y) d\nu(y)\end{aligned}$$

by (1.3), and finally,

$$\int g_1 \otimes \cdots \otimes g_{l-1} d\mu_*^{l-1} = \int E(g_1 \mid Y) \otimes \cdots \otimes E(g_{l-1} \mid Y) d\nu_*^{l-1}. \tag{1.6}$$

We say that a measure on $X^{l-1}$ is a *conditional product measure* if it is related to its projection on $Y^{l-1}$ as in (1.6). (See [1] for details.) Equivalently, a measure on $X^{l-1}$ is a conditional product measure if it takes the same value at $g_1 \otimes \cdots \otimes g_{l-1}$ as it does at $E(g_1 \mid Y) \otimes \cdots \otimes E(g_{l-1} \mid Y)$.

Consider any measure of the form $d\theta = \psi \otimes \cdots \otimes \psi_{l-1} d\mu_*^{l-1}$ and form

$$\theta_* = \lim \frac{1}{N_k} \sum_1^{N_k} (S_1 \times \cdots \times S_{l-1})^n$$

passing to a subsequence if necessary. We shall show that $\theta_*$ is a conditional product measure. Namely

$$(1.7)\quad \int g_1 \otimes \cdots \otimes g_{l-1} d\theta_* = \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} \int \psi_1 S_1^n g_1 \cdots \psi_{l-1} S_{l-1}^n g_{l-1} d\mu_*^{l-1}$$
$$= \lim_{k\to\infty} \frac{1}{N_k} \sum_1^{N_k} \int E(\psi_1 S_1^n g_1 \mid Y) \cdots E(\psi_{l-1} S_{l-1}^n g_{l-1} \mid Y) d\nu_*^{l-1}.$$

But by Lemma 1.3 we can replace $E(\psi S^n g \mid Y)$ by $E(\psi \mid Y) S^n E(g \mid Y)$ "on the average". From this we readily see that

$$\int g_1 \otimes \cdots \otimes g_{l-1} d\theta_* = \int E(g_1 \mid Y) \otimes \cdots \otimes E(g_{l-1} \mid Y) d\theta_*$$

so that $\theta_*$ is a conditional product measure. Since linear combinations of $\psi_1 \otimes \cdots \otimes \psi_{l-1}$ are dense in $L^1(\mu_*^{l-1})$ the same result is true for any $\theta$ absolutely continuous with respect to $\mu_*^{l-1}$. In particular if $\theta$ is absolutely continuous with respect to $\mu_*^{l-1}$ and $S_1 \times \cdots \times S_{l-1}$-invariant it must be a conditional product measure.

Now let $f' \in L^\infty(X, \mathcal{B}, \mu)$ with $E(f' \mid Y) = 1$ and define the measure $\bar{\mu}_\Delta^{l-1}$ by setting

$$\int g_1 \otimes \cdots \otimes g_{l-1} d\bar{\mu}_\Delta^{l-1} = \int g_1(x) g_2(x) \cdots g_{l-1}(x) f'(x) d\mu(x).$$

$\bar{\mu}_\Delta^{l-1}$ is absolutely continuous with respect to $\mu_\Delta^{l-1}$ and if we form the limit

$$(1.8)\quad \bar{\mu}_*^{l-1} = \lim \frac{1}{N_k} \sum_1^{N_k} (S_1 \times \cdots \times S_{l-1})^n \bar{\mu}_\Delta^{l-1}$$

we will obtain a measure that is $S_1 \times \cdots \times S_{l-1}$-invariant and absolutely continuous with respect to $\mu_*^{l-1}$. Hence $\bar{\mu}_*^{l-1}$ is a conditional product measure. It is therefore determined by its image in $Y^{l-1}$. But the image of $\bar{\mu}_*^{l-1}$ on $Y^{l-1}$ is $\nu_\Delta^{l-1}$ since $E(f' \mid Y) = 1$. It follows that $\bar{\mu}_*^{l-1} = \mu_*^{l-1}$.

Finally take $f_l \in L^\infty(X, \mathcal{B}, \mu)$ with $E(f_l \mid Y) = 0$ and set $f' = f_l + 1$. Comparing (1.8) with the definition of $\mu_*^{l-1}$ we obtain

$$\lim \frac{1}{N_k} \sum_1^{N_k} (S_1 \times \cdots \times S_{l-1})^n (\bar{\mu}_\Delta^{l-1} - \mu_\Delta^{l-1}) = 0$$

or

$$\lim \frac{1}{N_k} \sum_1^{N_k} \int S_1^n f_1(x) \cdots S_{l-1}^n f_{l-1}(x) \cdot f_l(x) d\mu(x) = 0.$$

Replace $x$ by $T_l^n x$ and recall that $S_i T_l = T_i$:

$$\lim \frac{1}{N_k} \sum_1^{N_k} \int T_1^n f_1(x) \cdots T_{l-1}^n f_{l-1}(x) T_l^n f_l(x) d\mu(x) = 0. \tag{1.9}$$

But this gives (1.5) inasmuch as (1.9) is valid for some subsequence of any sequence. This completes the proof.

## 2. Compact extensions

In this section we shall describe what we will speak of as the *compactness* of an extension $(X, \mathscr{B}, \mu)$ of a $\Gamma$-invariant factor $(Y, \mathscr{D}, \nu)$ for the action of some $T \in \Gamma$. It will be convenient to extend this to the action of a subgroup of $\Gamma$, so suppose that $\Lambda$ is a finitely generated subgroup of $\Gamma$. Fix an epimorphism $Z^r \to \Lambda$ by writing $n \to T^{(n)}$, $n \in Z^r$. Let $\|n\| = \max |n_i|$ where $n = (n_1, \cdots, n_r)$. The ergodic theorem for $Z^r$-actions states that if $f \in L^1(X, \mathscr{B}, \mu)$ then

$$\lim_{N\to\infty} \frac{1}{(2N+1)^r} \sum_{\|n\| \leq N} f(T^{(n)}x) \tag{2.1}$$

exists for almost all $x \in X$ and defines a $\Lambda$-invariant function. We shall use the much more elementary fact that the limit in (2.1) exists weakly in $L^2(X, \mathscr{B}, \mu)$ for $f$ in this space.

Let $\mu = \int \mu_y d\nu$ be the disintegration of $\mu$ with respect to the factor $(Y, \mathscr{D}, \nu)$ of $(X, \mathscr{B}, \mu)$ and let $\pi : X \to Y$ be the map defining the factor. We shall denote the Hilbert-space $L^2(X, \mathscr{B}, \mu)$ by $\mathfrak{H}$ and $L^2(X, \mathscr{B}, \mu_y)$ by $\mathfrak{H}_y$. We have

$$\|f\|_{\mathfrak{H}}^2 = \int \|f\|_{\mathfrak{H}_y}^2 d\nu(y).$$

Also note that each $T \in \Gamma$ defines an isometry $f \to Tf$ of $\mathfrak{H}_{Ty}$ onto $\mathfrak{H}_y$ so that

$$\|Tf\|_{\mathfrak{H}_y} = \|f\|_{\mathfrak{H}_{Ty}}.$$

Let $H \in L^2(X \times_Y X, \tilde{\mathscr{B}}, \tilde{\mu})$ and $f \in L^2(X, \mathscr{B}, \mu)$. We define the convolution (relative to $(Y, \mathscr{D}, \nu)$) of $H$ and $f$

$$H * f(x) = \int H(x, x')f(x')d\mu_y(x')$$

where $y = \pi(x)$. We have

$$\| H * \varphi \|_{\mathfrak{H}_y} \leqq \| H \|_{\mathfrak{H}_y \otimes \mathfrak{H}_y} \| \varphi \|_{\mathfrak{H}_y},$$

and, in particular, if $\| H \|_{\mathfrak{H}_y \otimes \mathfrak{H}_y}$ is bounded, the operator $\varphi \to H * \varphi$ is a bounded operator on $\mathfrak{H}$. We shall say that $\varphi \in L^2(X, \mathcal{B}, \mu)$ is *fibrewise bounded* if $\| \varphi \|_{\mathfrak{H}_y}$ is bounded and similarly for $H \in L^2(X \times_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$.

Consider now the following properties of our extension $(X, \mathcal{B}, \mu)$ of $(Y, \mathcal{D}, \nu)$ with respect to the subgroup $\Lambda \subset \Gamma$:

$C_1$. The functions $\{H * \varphi\}$ span a dense subset of $L^2(X, \mathcal{B}, \mu)$ as $H$ ranges over fibrewise bounded $\Lambda$-invariant functions on $X \times_Y X$ and $\varphi \in L^2(X, \mathcal{B}, \mu)$.

$C_2$. There exists a dense subset $\mathcal{D} \subset L^2(X, \mathcal{B}, \mu)$ with the following property. If $f \in \mathcal{D}$ and $\delta > 0$, there exists a finite set of functions $g_1, \cdots, g_k \in L^2(X, \mathcal{B}, \mu)$ such that for each $T \in \Lambda$, $\min_{1 \leqq j \leqq k} \| Tf - g_j \|_{\mathfrak{H}_y} < \delta$ for a.e. $y \in Y$.

$C_3$. For each $f \in L^2(X, \mathcal{B}, \mu)$ the following holds. If $\varepsilon$, $\delta > 0$ are given, there exists a finite set of functions $g_1, \cdots, g_k \in L^2(X, \mathcal{B}, \mu)$ such that for each $T \in \Lambda$, $\min_{1 \leqq j \leqq k} \| Tf - g_j \|_{\mathfrak{H}_y} < \delta$ but for a set of $y$ of measure $< \varepsilon$.

$C_4$. For each $f \in L^2(X, \mathcal{B}, \mu)$ form the limit function

$$\tilde{P}(f(x, x')) = \lim_{N \to \infty} \frac{1}{(2N+1)^r} \sum_{\|n\| \leqq N} f(T^{(n)}x) \overline{f(T^{(n)}x')}$$

in $L^2(X \times_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$, then $\tilde{P}f$ does not vanish a.e. unless $f$ vanishes a.e.

**Theorem 2.1.** *The four properties* $C_1$–$C_4$ *of an extension* $(X, \mathcal{B}, \mu)$ *of* $(Y, \mathcal{D}, \nu)$ *with respect to a finitely generated subgroup* $\Lambda \subset \Gamma$ *are equivalent.*

**Proof.** $C_1 \Rightarrow C_2$. Let us say that $f \in L^2(X, \mathcal{B}, \mu)$ is AP (almost periodic) if for each $\delta > 0$, there exist $g_1, \cdots, g_k \in L^2(X, \mathcal{B}, \mu)$ with $\min_{1 \leqq j \leqq k} \| Tf - g_j \|_{\mathfrak{H}_y} < \delta$ for each $T \in \Lambda$ and a.e. $y \in Y$. Clearly any linear combination of AP functions is AP. To prove that $C_1 \Rightarrow C_2$ it will suffice to show that by an arbitrarily small modification of a function of the form $H * \varphi$, $H$ being $\Lambda$-invariant and fibrewise bounded, we obtain an AP function. Since $\varphi \to H * \varphi$ is bounded we can restrict to a dense subset of $\varphi$; in particular, we may assume that $\varphi$ is fibrewise bounded, say $\| \varphi \|_{\mathfrak{H}_y} \leqq M$.

Let $\eta > 0$ be given; we shall find an AP function $f \in L^2(X, \mathcal{B}, \mu)$ with $f = H * \varphi$ but for a set of $x \in X$ with measure $< \eta$ on which $f$ vanishes. In $L^2(X \times_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$, the functions of the form $\Sigma\, \psi_i(x)\psi'_i(x')$, $\psi_i, \psi'_i \in L^\infty(X, \mathcal{B}, \mu)$ are dense and so we can choose a sequence of such functions converging to $H$ in $L^2$. Passing to a subsequence we can assume that $H_n$ is a sequence of such functions with

$\|H - H_n\|^2_{\mathfrak{H}_y \otimes \mathfrak{H}_y} \to 0$ for almost all $y \in Y$. We can then find a subset $E_\eta \subset Y$ with $\nu(E_\eta) < \eta$ such that $\|H - H_n\|_{\mathfrak{H}_y \otimes \mathfrak{H}_y} \to 0$ uniformly for $y \notin E_\eta$. Let $F_\eta$ be the largest $\Lambda$-invariant set in $E_\eta : F_\eta = \bigcap_{T \in \Lambda} TE_\eta$. We shall show that the function

$$f(x) = \begin{cases} H * \varphi(x), & \pi(x) \notin F_\eta \\ 0, & \pi(x) \in F_\eta \end{cases} \tag{2.2}$$

is AP.

Let us say that a set of functions $g_1, \cdots, g_k$ is *$\delta$-spanning for $f$ on the set $B \subset Y$* if for each $y \in B$, and $T \in \Lambda$, $\min_j \|Tf - g_j\|_{\mathfrak{H}_y} < \delta$. The function 0 is $\delta$-spanning for $f$ in $F_\eta$ so it will suffice to find a $\delta$-spanning set in $Y \setminus F_\eta$. Note that if $g_1, \cdots, g_k$ is $\delta$-spanning in $B$ then by the isometry of $\mathfrak{H}_{Ty}$ with $\mathfrak{H}_y$, $Tg_1, \cdots, Tg_k$ is $\delta$-spanning in $TB$ if $T \in \Lambda$. Using this we can construct a $\delta$-spanning set in $\bigcup_{T \in \Lambda} TB$. Namely, enumerate the elements of $\Lambda : T_1, T_2, T_3, \cdots$ and for each $x \in \tilde{B} = \bigcup_{T \in \Lambda} TB$ let $T_x$ be the first $T_i$ with $T_i(x) \in B$. We then set $\tilde{g}_i(x) = g_i(T_x x)$ and so find that $\tilde{g}_1, \cdots, \tilde{g}_k$ is $\delta$-spanning in $\tilde{B}$.

In view of this we see that in order to prove that $f(x)$ given by (2.2) is AP it suffices to find a $\delta$-spanning set for $f$ in $Y \setminus E_\eta$.

Using the fact that $H$ is $\Lambda$-invariant we can simplify the study of $\{Tf : T \in \Lambda\} \subset \mathfrak{H}_y$ as follows. We have

$$T(H * \varphi)(x) = H * \varphi(Tx) = \int H(Tx, x')\varphi(x')d\mu_{Ty}(x')$$

$$= \int H(Tx.\, Tx')\varphi(Tx')d\mu_y(x') = H * T\varphi(x).$$

Since $\varphi \to T\varphi$ is an isometry of $\mathfrak{H}_{Ty} \to \mathfrak{H}_y$ we conclude that $\{T\varphi\colon T \in \Lambda\} \subset$ ball of radius $M$ in each $\mathfrak{H}_y$. Hence $g_1, \cdots, g_k$ will be $\delta$-spanning in $Y \setminus E_\eta$ for $H * \varphi$ with a fixed $\varphi$ satisfying $\|\varphi\|_y \leqq M$ for all $y$, if for all $\varphi$ satisfying $\|\varphi\|_y \leqq M$ we have $\min_{1 \leq j \leq k} \|H * \varphi - g_j\|_y < \delta$. To find this set of $g_j$, choose $n$ with $\|H - H_n\|_{\mathfrak{H}_y \otimes \mathfrak{H}_y} < \delta/2M$ for all $y \notin E_\eta$, and find $\{g_j\}$ with $\min_{1 \leq j \leq k} \|H_n * \varphi - g_j\|_{\mathfrak{H}_y} < \delta/2$ for all the $\varphi$ in question. Now if $H_n = \Sigma\, \psi_i(x)\psi_i'(x')$, $H_n * \varphi$ ranges over functions of the form $\Sigma\, \alpha_i\psi_i(x)$ with $|\alpha_i| \leqq M\|\psi_i'\|_{\mathfrak{H}_y}$ and since the $\psi_i$ are bounded, it is easy to produce a finite subset of these functions which can serve as $g_j$.

$C_2 \Rightarrow C_3$. If $f \in L^2(X, \mathscr{B}, \mu)$ is given and $f'$ is AP with $\|f - f'\| < \delta\sqrt{\varepsilon}$, then for each $T \in \Lambda$, $\|Tf - Tf'\| < \delta\sqrt{\varepsilon}$. If $g_1, \cdots, g_k$ is a $\delta$-spanning set for $f'$ on $Y$, then $\min\|Tf - g_j\|_{\mathfrak{H}_y} < 2\delta$ but for those $y$ on which $\|Tf - Tf'\|_{\mathfrak{H}_y} \geqq \delta$. But this set has measure $< \delta^2\varepsilon/\delta^2 = \varepsilon$.

$C_3 \Rightarrow C_4$. First let us reformulate $C_3$. Let us call $g_1, \cdots, g_k$ an $\varepsilon, \delta$-spanning set for $f$ if the condition of $C_3$ holds; i.e., if $\min\|Tf - g_j\|_{\mathfrak{H}_y} < \delta$ for $y$ outside of a set $E(T)$ with $\nu(E(T)) < \varepsilon$. For each $j = 1, \cdots, k$, let

$$F_j(T) = \{y : \| Tf - g_j \|_{\mathfrak{H}_y} < \delta\}$$

and let $\Omega \subset \Lambda$ be a finite subset large enough so that for each $j$,

$$\nu\Big(\bigcup_{T \in \Omega} F_j(T)\Big) > \nu\Big(\bigcup_{T \in \Lambda} F_j(T)\Big) - \varepsilon / k;$$

then $\min_{T' \in \Omega} \| Tf - T'f \|_{\mathfrak{H}_y} < 2\delta$ unless $y \in E(T)$ or

$$y \in \bigcup_j \Big\{ \bigcup_{T \in \Lambda} F_j(T) \setminus \bigcup_{T \in \Omega} F_j(T) \Big\}.$$

We see that the functions $\{T'f\colon T' \in \Omega\}$ form a $2\varepsilon, 2\delta$-spanning set.

Now assume that $\tilde{P}f = 0$. Evaluating $\int \overline{f(x)} f(x') \tilde{P}f(x, x') d\tilde{\mu}(x, x')$ we find that

$$\lim_{N \to \infty} \frac{1}{(2N+1)^r} \sum_{\|n\| \leq N} \left| \int \overline{f(x)} f(T^{(n)}x) d\mu_y(x) \right|^2 = 0 \tag{2.3}$$

in $L^2(Y, \mathfrak{D}, \nu)$.

Moreover $\tilde{P}f = 0$ implies $\tilde{P}Tf = 0$ for each $T \in \Lambda$ and we obtain from (2.3) that

$$\frac{1}{(2N+1)^r} \sum_{\|n\| \leq N} \Big\{ \sum_{T' \in \Omega} \left| \int \overline{T'f} T^{(n)} f d\mu_y \right|^2 \Big\} \to 0$$

in $L^2(Y, \mathfrak{D}, \nu)$. In particular for any $\varepsilon > 0$ there exists $T \in \Lambda$ with

$$\left| \int \overline{T'f} Tf d\mu_y \right| < \varepsilon \tag{2.4}$$

for all $T' \in \Omega$ and for all $y$ outside of a set of measure $< \varepsilon$. If we assume now $\Omega$ was chosen so that $\{T'f : T' \in \Omega\}$ is an $\varepsilon, \delta$-spanning set, then outside of a set of measure $< \varepsilon$,

$$\int | Tf - T'f |^2 d\mu_y < \delta^2 \tag{2.5}$$

for some $T'$ depending on $y$. But (2.4) and (2.5) give

$$\int | Tf |^2 d\mu_y < \delta^2 + 2\varepsilon$$

outside of a set of $y$ of measure $2\varepsilon$. Since $\varepsilon$, $\delta$ were arbitrary, we conclude that $f \equiv 0$.

$C_4 \Rightarrow C_1$. Suppose the functions of the form $H * \varphi$ were not dense as $H$ ranges over fibrewise bounded $\Lambda$-invariant functions on $X \times {}_Y X$, and $\varphi$ over $L^2(X, \mathcal{B}, \mu)$. Let $f \in L^2(X, \mathcal{B}, \mu)$ be orthogonal to all of these. Consider the function

$$H(x, x') = \lim \frac{1}{(2N+1)^r} \sum_{\|n\| \leq N} T^{(n)} f(x) \overline{T^{(n)} f(x')}.$$

This is $\Lambda$-invariant and belongs to $L^2(X \times {}_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$. In particular $\|H\|_{\mathfrak{H}_y \otimes \mathfrak{H}_y} < \infty$ for a.e. $y \in Y$. This norm is also $\Lambda$-invariant and we can find a $\Lambda$-invariant set $B \subset Y$ with $\nu(B)$ as close to 1 as we please on which $\|H\|_{\mathfrak{H}_y \otimes \mathfrak{H}_y}$ is bounded. Let $H_B = H \cdot 1_{\pi^{-1}(B)}$ and $f_B = f \cdot 1_{\pi^{-1}(B)}$; then,

$$H_B(x, x') = \lim_{N \to \infty} \frac{1}{(2N+1)^r} \sum_{\|n\| \leq N} T^{(n)} f_B(x) \overline{T^{(n)} f_B(x')}.$$

This function is fibrewise bounded and $f \perp H_B * f_B$ implies that $f_B \perp H_B * f_B$. But then

$$\int H_B(x, x') f_B(x') \overline{f_B(x)} \, d\tilde{\mu}(x, x') = 0, \tag{2.6}$$

or, $f_B(x)\overline{f_B(x')}$ is orthogonal to $H_B$ in $L^2(X \times {}_Y X, \tilde{\mathcal{B}}, \tilde{\mu})$. The same is then true of each $Tf_B(x)\overline{Tf_B(x')}$ and therefore also for any average of these functions. But then $H_B \perp H_B$ so that $H_B \equiv 0$. $C_4$ implies that $f_B \equiv 0$. Letting $B$ approximate $Y$ we conclude that $f \equiv 0$ and this proves $C_1$.

**Definition 3.1.** If $(Y, \mathcal{D}, \nu)$ is a $\Gamma$-invariant factor of $(X, \mathcal{B}, \mu)$ and $\Lambda$ is a finitely generated subgroup of $\Gamma$ for which one of the conditions $C_1$–$C_4$ holds, then we say that $(X, \mathcal{B}, \mu)$ is a *compact* extension of $(Y, \mathcal{D}, \nu)$ for the action of $\Lambda$.

Property $C_4$ of compact extension ensures a plentiful supply of $\Lambda$-invariant functions on $X \times {}_Y X$. If the extension is non-trivial these cannot all be functions on $Y$, since choosing $f$ with $E(f \mid Y) = 0$ implies $E(\tilde{P}f \mid Y) = 0$ and if $\tilde{P}f$ were a function on $Y$, this implies $\tilde{P}f = 0$. We see then that a compact extension is never relatively weak mixing for any $T \in \Lambda$. The converse is true in the following sense.

**Proposition 2.2.** *If $(Y, \mathcal{D}, \nu)$ is a $\Gamma$-invariant factor of $(X, \mathcal{B}, \mu)$ and for an element $T \in \Gamma$, the extension is not relatively weak mixing, then there exists a $\Gamma$-invariant factor $(X', \mathcal{B}', \mu')$ of $(X, \mathcal{B}, \mu)$ which is a non-trivial compact extension of $(Y, \mathcal{D}, \nu)$ for the action of the group generated by $T$.*

**Proof.** Let $H(x,x')$ be a bounded $T$-invariant function on $X \times_Y X$ which is not a function on $Y$. Replacing $H(x,x')$ by $H(x',x)$ if necessary we can assume that for some $\varphi \in L^\infty(X,\mathscr{B},\mu)$, $H * \varphi$ is not a function on $Y$. In the proof of Theorem 2.1 we showed that for each function $H * \varphi$ with $H$ and $\varphi$ fibrewise bounded, we could modify $H * \varphi$ on an arbitrarily small set to obtain an AP function. Hence, if $(X,\mathscr{B},\mu)$ is not a relatively weak mixing extension of $(Y,\mathscr{D},\nu)$ for $T \in \Gamma$, there exist AP functions on $(X,\mathscr{B},\mu)$ which are not functions on $(Y,\mathscr{D},\nu)$. Now it is clear that for any $\Lambda \subset \Gamma$, sums and products of bounded AP functions are AP functions. Moreover, functions in $L^\infty(Y,\mathscr{D},\nu)$ are AP. In addition, if $f$ is AP for $\Lambda$, and $S \in \Gamma$, then $Sf$ is again AP inasmuch as $\min \| Tf - g_j \|_{\mathfrak{H}_y} = \min \| TSf - Sg_j \|_{\mathfrak{H}_{y^{-1}y}}$. Thus if $\mathscr{B}'$ is the $\sigma$-algebra with respect to which all AP functions are measurable, then $\mathscr{B}'$ is $\Gamma$-invariant and $(X,\mathscr{B}',\mu)$ is a factor of $(X,\mathscr{B},\mu)$ which is a compact extension of $(Y,\mathscr{D},\nu)$ with respect to $\Lambda$. This proves the proposition.

Next we show that for a given $\Gamma$-invariant factor $(Y,\mathscr{D},\nu)$ of $(X,\mathscr{B},\mu)$, the set of $T$ such that $(X,\mathscr{B},\mu)$ is a compact extension of $(Y,\mathscr{D},\nu)$ for the group $\{T^n\}$ forms a subgroup of $\Gamma$. More precisely:

**Proposition 2.3.** *If $(X,\mathscr{B},\mu)$ is a compact extension of $(Y,\mathscr{D},\nu)$ for the actions of the subgroups $\Lambda_1,\Lambda_2 \subset \Gamma$, then it is compact for the action $\Lambda_1\Lambda_2$.*

**Proof.** We use the characterization $C_3$ of compactness. Let $f \in L^2(X,\mathscr{B},\mu)$ and $\varepsilon,\delta > 0$ be given. Choose $g_1,\cdots,g_k$ in $L^2(X,\mathscr{B},\mu)$ such that for each $T \in \Lambda_1$, $\min \| Tf - g_j \|_{\mathfrak{H}_y} < \delta/2$ but for $y \in E(T) \subset Y$, with $\nu(E(T)) < \varepsilon/2$. For each $g_j$, choose $h_{j1},\cdots,h_{jq_j} \in L^2(X,\mathscr{B},\mu)$ so that for each $S \in \Lambda_2$, $\min_{1 \le p \le q_j} \| Sg_j - h_{jp} \|_{\mathfrak{H}_y} < \delta/2k$ but for $y \in F_j(S)$, where $\nu(F_j(S)) < \varepsilon/2k$ then for $T \in \Lambda_1$, $S \in \Lambda_2$, and $y \notin S^{-1}E(T)$, $\min \| Tf - g_j \|_{\mathfrak{H}_{Sy}} < \delta/2$. Having chosen $j = j(y)$ to attain this minimum, we have $\| STf - Sg_j \|_{\mathfrak{H}_y} < \delta/2$. If, in addition, $y \notin F_j(S)$, then $\min_p \| Sg_j - h_{jp} \|_{\mathfrak{H}_y} \leqq \delta/2$. Thus outside of $S^{-1}E(T) \cup \bigcup_j F_j(S)$, $\min_{j,p} \| STf - h_{jp} \|_{\mathfrak{H}_y} < \delta$. Since $\nu(S^{-1}E(T) \cup \bigcup_j F_j(S)) < \varepsilon$, this proves the proposition.

Combining Propositions 2.2 and 2.3 we obtain the following "structure" theorem.

**Theorem 2.4.** *Assume $\Gamma$ is finitely generated and let $(Y,\mathscr{D},\nu)$ be a $\Gamma$-invariant factor of $(X,\mathscr{B},\mu)$. There exists a $\Gamma$-invariant proper extension $(X',\mathscr{B}',\mu')$ of $(Y,\mathscr{D},\nu)$ and a direct product decomposition $\Gamma = \Gamma_w \times \Gamma_c$ where $\Gamma_w$ and $\Gamma_c$ are two subgroups for which*

(i) *$(X',\mathscr{B}',\mu')$ is a relatively weak mixing extension of $(Y,\mathscr{D},\nu)$ for every $T \in \Gamma_w$, $T \neq I$.*

(ii) *$(X',\mathscr{B}',\mu')$ is a compact extension of $(Y,\mathscr{D},\nu)$ for the action of $\Gamma_c$.*

**Proof.** Let $\Gamma_c$ be a maximal subgroup of $\Gamma$ $(\cong \mathbf{Z}^m)$ for which there exists a non-trivial $\Gamma$-invariant compact extension of $(Y, \mathcal{D}, \nu)$ in $(X, \mathcal{B}, \mu)$, and denote by $(X', \mathcal{B}', \mu')$ the corresponding extension.

If $T \in \Gamma \backslash \Gamma_c$ then $(X', \mathcal{B}', \mu')$ is a relatively weak mixing extension of $(Y, \mathcal{D}, \nu)$. Otherwise, there would exist a $\Gamma$-invariant factor $(X'', \mathcal{B}'', \mu'')$ of $(X', \mathcal{B}', \mu')$ which is compact for $T$ (Proposition 2.2); and since $(X'', \mathcal{B}'', \mu'')$ is also compact for $\Gamma_c$, it would be compact for the group generated by $\Gamma_c$ together with $T$ in contradiction with the maximality of $\Gamma_c$. This also implies that if $T \notin \Gamma_c$ then $T^n \notin \Gamma_c$ for all $n \geqq 1$. $\Gamma/\Gamma_c$ is therefore torsion free and $\Gamma_c$ is a complemented subgroup of $\Gamma$. Take for $\Gamma_w$ any complement of $\Gamma_c$.

**Remark.** When one restricts $\Gamma$ to an invariant factor the representation need not be faithful, that is, some non-trivial elements of $\Gamma$ may act like the identity on the factor. In our decomposition above those elements which act trivially on $(X', \mathcal{B}', \mu')$ will clearly go to $\Gamma_c$.

We end this section with a modification of condition $C_2$ which will be the characterization of compact extensions which we will need in the next section.

**Proposition 2.5.** *Suppose $(X, \mathcal{B}, \mu)$ is a compact extension of $(Y, \mathcal{D}, \nu)$ for the action of a subgroup $\Lambda \subset \Gamma$. Then for each $f \in L^2(X, \mathcal{B}, \mu)$ and $\varepsilon, \delta > 0$, there exists a set $B \subset Y$ with $\nu(B) > 1 - \varepsilon$ and a set of functions $g_1, g_2, \cdots, g_k \in L^2(X, \mathcal{B}, \mu)$ such that if $f_B = f \cdot 1_{\pi^{-1}(B)}$, then for all $T \in \Lambda$ and a.e. $y \in Y$, $\min_{1 \leqq j \leqq k} \| Tf_B - g_j \|_{\mathfrak{H}_y} < \delta$.*

**Proof.** Let $f' \in L^2(X, \mathcal{B}, \mu)$ be an AP function with $\|f - f'\| < \delta\sqrt{\varepsilon}/2$ and let $g_1, \cdots, g_{k-1}$ be such that for $T \in \Lambda$ and a.e. $y \in Y$, $\min \| Tf' - g_j \|_{\mathfrak{H}_y} < \delta/2$. Let $g_k \equiv 0$ and let $B = \{y : \|f - f'\|_{\mathfrak{H}_y} < \delta/2\}$. Then $\nu(B) > 1 - \varepsilon$ and if $y \in T^{-1}B$, $\| Tf_B - Tf' \|_{\mathfrak{H}_y} = \| Tf - Tf' \|_{\mathfrak{H}_y} < \delta/2$, and so $\min_{1 \leqq j \leqq k-1} \| Tf_B - g_j \|_{\mathfrak{H}_y} < \delta$. If $y \notin T^{-1}B$, then $Tf_B = 0$ in $\mathfrak{H}_y$ and so $\| Tf_B - g_k \|_{\mathfrak{H}_y} < \delta$.

## 3. Proof of Theorem A

We denote by $\Gamma$ the group generated by the transformations $T_1, \cdots, T_k$ and since we do not assume that $\Gamma$ acts effectively we may assume $\Gamma \cong \mathbf{Z}^m$. We shall say that the action of a group $\Gamma$ on a probability measure space $(X, \mathcal{B}, \mu)$ is *SZ* if the statement of Theorem A is true whenever $T_1, \cdots, T_k$ belong to $\Gamma$. Thus, Theorem A states that every $Z^m$ action is *SZ*.

We prove Theorem A by "induction" on the $\Gamma$-invariant factors of $(X, \mathcal{B}, \mu)$. The action of $\Gamma$ on the trivial factor is trivially *SZ* and we show (a) that there exists a maximal factor for which the action of $\Gamma$ is *SZ*, and (b) that no proper factor of

$(X, \mathcal{B}, \mu)$ can be maximal for the property that the action on it is *SZ*. These two steps combined imply that the maximal factor must be $(X, \mathcal{B}, \mu)$ itself, and hence, that the action of $\Gamma$ on it is *SZ*.

**Lemma 3.1.** *Let $(Y, \mathcal{D}, \nu)$ be a $\Gamma$-invariant factor of $(X, \mathcal{B}, \mu)$. Let $A \in \mathcal{B}$, $A_0 \in \mathcal{D}$ and assume that for every $y \in A_0$, $\mu_y(A) \geqq 1 - \eta$. Then if $T_1, \cdots, T_k \in \Gamma$*

$$\mu\left(\bigcap_{j=1}^{k} T_j A\right) \geqq (1 - k\eta)\mu\left(\bigcap_{j=0}^{k} T_j A_0\right). \tag{3.1}$$

**Proof.** The intersection of $k$ sets of (probability) measures at least $1 - \eta$ each, has measure at least $1 - k\eta$. Thus for every $y \in \bigcap_{j=1}^{k} T_j A_0$ we have $\mu_y(\bigcap_{j=1}^{k} T_j A) \geqq 1 - k\eta$, and we obtain (3.1) by integrating on $\bigcap T_j A_0$.

The collection of all factors of $(X, \mathcal{B}, \mu)$ is partially ordered by inclusion (of the corresponding closed subalgebras of $L^\infty(X, \mathcal{B}, \mu)$). If $(Y_\alpha, \mathcal{D}_\alpha, \nu_\alpha)$ is a totally ordered family of factors we define its supremum, $(Y, \mathcal{D}, \nu) = \sup(Y_\alpha, \mathcal{D}_\alpha, \nu_\alpha)$, as the factor whose corresponding subalgebra is the closure of the union of the subalgebras corresponding to $(Y_\alpha, \mathcal{D}_\alpha, \nu_\alpha)$. In other words, a set $A \in \mathcal{B}$ belongs to $\mathcal{D}$ if for every $\varepsilon > 0$, there exists a set $A_0$ is some $\mathcal{D}_\alpha$ such that $\mu((A \setminus A_0) \cup (A_0 \setminus A)) < \varepsilon$. It is clear that if for every $\alpha$, $(Y_\alpha, \mathcal{D}_\alpha, \nu_\alpha)$ is $\Gamma$-invariant, so is $(Y, \mathcal{D}, \nu)$.

**Lemma 3.2.** *Let $(Y_\alpha, \mathcal{D}_\alpha, \mu_\alpha)$ be a totally ordered family of $\Gamma$-invariant factors. Assume that for each $\alpha$ the action of $\Gamma$ on $(Y_\alpha, \mathcal{D}_\alpha, \mu_\alpha)$ is SZ. Then the action of $\Gamma$ on $(Y, \mathcal{D}, \nu) = \sup(Y_\alpha, \mathcal{D}_\alpha, \mu_\alpha)$ is SZ.*

**Proof.** Let $T_1, \cdots, T_k \in \Gamma$ and let $A \in \mathcal{D}$, $\nu(A) > 0$. Take $\eta = (2k)^{-1}$ and $A_0' \in \mathcal{D}_{\alpha_0}$ such that

$$\mu((A \setminus A_0') \cup (A_0' \setminus A)) < \tfrac{1}{4}\eta\nu(A). \tag{3.2}$$

By (3.2), $\mu(A_0')\ (= \nu(A_0')) > \frac{3}{4}\mu(A) > 0$. Also the set of $y \in A_0'$ such that $\mu_y(A) < 1 - \eta$ has measure less than $\frac{1}{4}\mu(A)$, since otherwise $\mu(A_0' \setminus A) > \frac{1}{4}\eta\mu(A)$ which would contradict (3.2). If we denote by $A_0$ the subset of $A_0'$ of points $y$ for which $\mu_y(A) > 1 - \eta$, then $A_0 \in \mathcal{D}_{\alpha_0}$, $\mu(A_0) > \frac{1}{2}\mu(A)$, and since the action of $\Gamma$ on $\mathcal{D}_{\alpha_0}$ is *SZ* we have

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{1}^{N} \mu\left(\bigcap_{j=1}^{k} T_j^n A_0\right) = a > 0.$$

Applying Lemma 3.1 for $T_1^n, \cdots, T_k^n$, $n = 1, 2, \cdots$ we obtain

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{1}^{N} \nu\left(\bigcap_{j=1}^{k} T_j^n A\right) \geqq \frac{a}{2} > 0. \tag{3.3}$$

Since $A \in \mathscr{D}$ and $T_1, \cdots, T_k \in \Gamma$ were arbitrary, (3.3) is the statement that the action of $\Gamma$ on $\mathscr{D}$ is $SZ$.

**Proposition 3.3.** *The family of $\Gamma$-invariant factors on which the action of $\Gamma$ is $SZ$ has maximal elements (under inclusion).*

**Proof.** Zorn's lemma and Lemma 3.2.

We now turn to show that no proper $\Gamma$-invariant factor of $(X, \mathscr{B}, \mu)$ can be maximal for the property of $SZ$ action. In all that follows $(Y, \mathscr{D}, \mu)$ is a proper $\Gamma$-invariant factor and the action of $\Gamma$ on it is $SZ$.

**Lemma 3.4.** *Let $E_{j,l}$, $j = 1, \cdots, J$, $l = 1, \cdots, L$ be measurable sets and assume that for some $\delta > 0$ and every $j$ and $l$ we have $\mu(E_{j,1} \setminus E_{j,l}) \leqq \delta$. Then*

$$\mu\left(\bigcap_{j,l} E_{j,l}\right) \geqq \mu\left(\bigcap_{j} E_{j,1}\right) - JL\delta. \tag{3.4}$$

**Proof.** Replacing in $\bigcap E_{j,l}$ any term $E_{j,l}$ by $E_{j,1}$ may increase the measure of the intersection by at most $\delta$.

**Proposition 3.5.** *Assume that the action of $\Gamma$ on $(Y, \mathscr{D}, \nu)$ is $SZ$ and that $(X', \mathscr{B}', \mu')$ is a $\Gamma$-invariant extension of $(Y, \mathscr{D}, \nu)$ in $(X, \mathscr{B}, \mu)$ such that there exists a decomposition $\Gamma = \Gamma_w \times \Gamma_c$ as given by Theorem 2.4. Then the action of $\Gamma$ in $(X', \mathscr{B}', \mu')$ is $SZ$.*

**Proof.** Let $T_1, \cdots, T_k \in \Gamma$ and let $A \in \mathscr{B}'$ with $2a = \mu(A) > 0$. We have to show that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{1}^{N} \left(\bigcap_{j=1}^{k} T_j^n A\right) > 0.$$

We write $T_j = S'_j R'_j$ with $S'_j \in \Gamma_w$ and $R'_j \in \Gamma_c$ and then replace the set $\{T_j\}$ by the possibly larger set $\{S_j R_l\}$ where $\{S_j\}_{j=1}^{J}$ is the set of all the transformations $S'_j$ above renumbered so that possible repetitions are omitted, and similarly for $\{R_l\}_{l=1}^{L}$. There is no loss of generality in assuming that $R_1 =$ identity. We have enlarged the set of transformations and we are now going to (possibly) reduce $A$. We first look at $E(1_A \mid Y) = \mu_y(A)$ and take the intersection $A_1$ of $A$ with the set of fibers corresponding to points $y$ such that $\mu_y(A) > a$ $(= \frac{1}{2}\mu(A))$. Now, taking

$$\delta = (4JL)^{-1} a^J, \tag{3.5}$$

and using Proposition 2.5 for the action of $\Gamma_c$, we remove from $A_1$ a small set of fibers (that is, its intersection with a small set in $\mathcal{D}$) and obtain our final set $A_0$ such that $\mu(A_0) > 0$, $\mu_y(A_0) > a$ whenever $\mu_y(A_0) > 0$, and, denoting $f = 1_{A_0}$, there exist functions $\{g_j\}_{j=1}^K$ such that for every $y \in Y$ and $R \in \Gamma_c$

$$\min_{j=1,\cdots,K} \| Rf - g_j \|_{\mathfrak{G}_y} < \delta.$$

We now define the "coloring function" $c(R, y)$ on $\Gamma_c \times Y$ by setting $c(R, y) =$ the smallest integer $r$ such that $\| Rf - g_r \|_{\mathfrak{G}_y} = \min \| Rf - g_j \|_{\mathfrak{G}_y}$, and extend it to $\Gamma \times Y$ by $c(SR, y) = c(R, Sy)$. The "coloring function" assumes values in $\{1, \cdots, K\}$. Since $\Gamma \cong Z^m$ the set $G = \{S_j R_l\}$, $j = 1, \cdots, J$, $l = 1, \cdots, L$, can be viewed as a configuration in $Z^m$. By the multidimensional version of van der Waerden's theorem (see [3] for the proof of Grünwald or [2] for a simpler proof depending on the recurrence result in topological dynamics alluded to in our introduction) there exists a finite configuration $G_1$ (e.g. a large enough box) in $Z^m$ such that for any coloring of $G_1$ by $K$ colors one can find in $G_1$ a monochromatic translated homothetic copy of $G$. The constants of homothety are clearly bounded by some integer $H$ (e.g., the diameter of $G_1$). We denote by $\{T_\alpha\}$ a set in $\Gamma$ which corresponds, as above, to the configuration $G_1$. We have the following

**Fact.** *For every $y \in Y$ and $n \in Z$ there exists a $T \in \Gamma$ and an integer $h$, $1 \leqq h \leqq H$ such that*

$$\{S_j^{-nh} R_l^{-nh} Ty\}_{j,l} \subset \{T_\alpha^{-n} y\}_\alpha, \tag{3.6}$$

$$c(S_j^{-nh} R_l^{-nh}, Ty) = \text{const} \quad \textit{for } j = 1, \cdots, J,\ l = 1, \cdots, L. \tag{3.7}$$

Denote by $B_0$ the base of $A_0$ in $Y$, i.e., the set $\{y ; \mu_y(A_0) > a\}$ and apply the assumption that the action of $\Gamma$ in $(Y, \mathcal{D}, \nu)$ is $SZ$. There exists a positive number $b$ such that for all sufficiently large $N$, $\nu(\bigcap_\alpha T_\alpha^n B_0) > b$ for at least $bN$ values of $n$ in $[1, \cdots, N]$. Denote $B_n = \bigcap_\alpha T_\alpha^n B_0$. For $y \in B_n$ there exist $T$ and $h$ such that, by (3.6), $Ty \in \bigcap_{j,l} S_j^{nh} R_j^{nh} B_0$. We have pointed out before that $1 \leqq h \leqq H$ and it is equally clear that the number of possible $T$'s is bounded by the number of points in $G_1$. Thus we have a covering of $B_n$ by a finite number, say $H_1$, of subsets $B_n(T, h)$ containing the points of $B_n$ for which (3.6) and (3.7) are valid (for the specific choice of $T$ and $h$). It is clear that if $\nu(B_n) > b$, then, for some $(T, h)$, $\nu(B_n(T, h)) > b / H_1$.

If, for $y \in B_n(T, h)$, we look at the sets $S_j^{nh} R_l^{nh} A_0$ on the fibre of $Ty$, we obtain by (3.7) and Lemma 3.5 that

$$\mu_{Ty}\Big( \bigcap_{j,l} S_j^{nh} R_l^{nh} A_0 \Big) > \mu_{Ty}\Big( \bigcap_j S_j^{nh} A_0 \Big) - JL\delta \tag{3.8}$$

and by the choice of $\delta$, (3.5), any time that

$$\mu_{Ty}\left(\bigcap_j S_j^{nh}A_0\right) > \tfrac{3}{4}a^J \tag{3.9}$$

we have

$$\mu_{Ty}\left(\bigcap_j S_j^{nh}R_l^{nh}A_0\right) > \tfrac{1}{2}a^J. \tag{3.10}$$

Since $S_j \in \Gamma_w$, $j = 1, \cdots, J$ we obtain by Theorem 1.4 that for all sufficiently large $N$, (3.9) is valid for all the pairs $(y, n)$ such that $y \in B_n$ and $1 \leqq n \leqq N$, except for an arbitrarily small proportion of these.

Specifically, we obtain that for all sufficiently large $N$, there exists a subset $Q \subset [1, \cdots, N]$ such that $Q^* > \frac{1}{2}bN$ and such that for $n \in Q$ and an appropriate choice of $(T_n, h_n)$ we have (3.9) valid for all $y \in B'_n \subset B_n(T_n, h_n)$ such that

$$\nu(B'_n) > \frac{b}{2H_1}. \tag{3.11}$$

Integrating (3.10) on $B'_n$ we obtain that for $n \in Q$ and $h = h_n$

$$\mu\left(\bigcap_{j,l} S_j^{nh}R_l^{nh}A_0\right) > \tfrac{1}{4}H_1^{-1}ba^J = a_1. \tag{3.12}$$

Thus, for all large $N$, there exist at least $bN/2J$ integers $n$ in $[1, \cdots, JN]$ for which $\mu(\bigcap_{j,l} S_j^n R_l^n A_0) > a_1$ which clearly concludes the proof.

Theorem A follows immediately from Propositions 3.3 and 3.5.

## References

1. H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

2. H. Furstenberg and B. Weiss, *Topological dynamics and combinatorial number theory*, (this Vol.) J. Analyse Math. **34** (1978), 61–85.

3. R. Rado, *A note on combinatorial analysis*, Proc. London Math. Soc. V **48** (1943), 122–160.

4. J.-P. Thouvenot, *La démonstration de Furstenberg du théorème de Szemérédi sur les progressions arithmétiques*, Seminaire Bourbaki, No. 5/8, 1977/78.

DEPARTMENT OF MATHEMATICS
THE HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM, ISRAEL

(Received December 6, 1978)