A Roth Theorem for Amenable Groups

BY

VITALY BERGELSON,

RANDALL McCUTCHEON,

AND

QING ZHANG

November 30, 2005

Abstract: We prove the following mean ergodic theorem: for any two commuting measure preserving actions $\{T_g\}$ and $\{S_g\}$ of a countable amenable group G on a probability space $(X, \mathcal{A}, \mu)$, and any φ, $\psi \in L^2(X, \mathcal{A}, \mu)$, $\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x)$ exists in $L^1(X, \mathcal{A}, \mu)$, where $\{\Phi_n\}$ is any left Følner sequence for G. This generalizes Furstenberg's ergodic Roth theorem, which corresponds to the case $G = \mathbf{Z}$, $T_g = S_g$, as well as a more general result of Conze and Lesigne (which corresponds to the case $G = \mathbf{Z}$ with no restrictions on T_g and S_g). The limit is identified, and two combinatorial corollaries are obtained. The first of these states that in any subset $E \subset G \times G$ which is of positive upper density (with regard to any left Følner sequence in $G \times G$), we may find triangular configurations of the form $\{(a, b), (ga, b), (ga, gb)\}$. This result has as corollaries Roth's theorem on arithmetic progressions of length three and a theorem of Brown and Buhler guaranteeing solutions to the equation $x + y = 2z$ in any sufficiently big subset of an abelian group of odd order. The second corollary states that if $G \times G \times G$ is partitioned into finitely many cells, one of these cells contains configurations of the form $\{(a, b, c), (ga, b, c,), (ga, gb, c), (ga, gb, gc)\}$.

1 Introduction

A far reaching extension of the Poincaré recurrence theorem was given in 1977 by H. Furstenberg ([7]), who showed that if T is a measure preserving transformation on a probability space $(X, \mathcal{A}, \mu)$, one has multiple recurrence: for any positive integer k, and any $A \in \mathcal{A}$, $\mu(A) > 0$, there exists an integer $n > 0$ such that

$$\mu(A \cap T^{-n}A \cap \cdots \cap T^{-(k-1)n}A) > 0.$$

Besides being of intrinsic interest to ergodic theory, Furstenberg's multiple recurrence theorem may be used to reprove an important combinatorial result - Szemerédi's theorem ([16]), which states that any set $E \subset \mathbf{N}$ of positive upper density

$$\overline{d}(E) = \limsup_{N\to\infty} \frac{|E \cap \{1, \cdots, N\}|}{N}$$

contains arbitrarily long arithmetic progressions.

As a matter of fact, Furstenberg and Katznelson showed ([9]) that for any commuting measure preserving transformations $T_1, \cdots, T_k$ on a probability space $(X, \mathcal{A}, \mu)$, and any $A \in \mathcal{A}$, with $\mu(A) > 0$, one has

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

This leads naturally to the investigation of limits of expressions of the form

$$F_N = \frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) \cdots f_k(T_k^n x),$$

where $\{f_i : 1 \leq i \leq k\}$ is a set of bounded measurable functions and $\{T_i : 1 \leq i \leq k\}$ is a family of commuting measure preserving transformations. Although it is believed that, in the general case, $\lim_{N\to\infty} F_N$ exists in norm, or even a.e., the instances for which this has been proved are few. Even in the case of powers of the same transformation, $T_i = T^i$, $i = 1, \cdots, k$, existence of the limit in norm is known only for $k = 2$ and $k = 3$. In the case $k = 2$, this was proved in Furstenberg's pioneering work [7] (see also [8], section 4.4). Furstenberg calls this result the *ergodic Roth theorem*, since it gives as a corollary the first non-trivial case of Szemerédi's theorem, Roth's theorem ([15]), which states that any set of positive upper density in $\mathbf{N}$ contains arithmetic progressions of length 3. For the case $k = 3$, see Conze and Lesigne [4], [5], or Furstenberg and Weiss [12]. In the more general case of commuting transformations, the best result to date is due to Conze and Lesigne, who established the existence of the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) f_2(T_2^n x) \tag{1}$$

in $L^1(X, \mathcal{A}, \mu)$ for $f_1,\ f_2 \in L^2(X, \mathcal{A}, \mu)$.

One might ask whether the Furstenberg and Katznelson multiple recurrence theorem generalizes to measure preserving actions of more general groups. Thus, if $\{T_g^{(i)} : g \in G\}_{1\leq i\leq k}$ are commuting (i.e. $T_h^{(i)} T_g^{(j)} = T_g^{(j)} T_h^{(i)}$ for all $1 \leq i \neq j \leq k$ and $h, g \in G$) measure preserving G-actions on a probability space $(X, \mathcal{A}, \mu)$, does there exist, for any $A \in \mathcal{A}$, $\mu(A) > 0$, some $g \in G$, $g \neq e$, such that $\mu(A \cap T_g^{(1)} A \cap T_g^{(2)} A \cap \cdots \cap T_g^{(k)} A) > 0$? Evidence presented in [1], where a topological version of this question is dealt with, namely $T_g^{(i)}$ are homeomorphisms of a compact space and A is an open set, suggests that the answer in general may be no. However, if G is amenable and one writes instead

$$\mu(A \cap T_g^{(1)} A \cap T_g^{(1)} T_g^{(2)} A \cap \cdots \cap T_g^{(1)} \cdots T_g^{(k)} A) > 0, \tag{2}$$

there are some reasons for hope (see [1, 2]). Indeed, in this paper we will prove this for $k = 2$.

We will first show that if G is any countable amenable group (equivalently, G is countable and has a Følner sequence), and $\{\Phi_n\}$ is any left Følner sequence for G (that is, $\Phi_n \subset G$, with $|\Phi_n| < \infty$, $n \in \mathbf{N}$, and $|\Phi_n| \to \infty$, $\frac{|\Phi_n \cap g\Phi_n|}{|\Phi_n|} \to 1$ for all $g \in G$), then for any commuting measure preserving G-actions $\{T_g\}$ and $\{S_g\}$ on a probability space $(X, \mathcal{A}, \mu)$, the limit

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x)$$

exists in $L^1(X, \mathcal{A}, \mu)$ for every φ, $\psi \in L^2(X, \mathcal{A}, \mu)$. Following Furstenberg, we call this an *ergodic Roth theorem for amenable groups.* This result includes as a special case the existence of the limit (1). Furthermore, we will identify the limit by means of an explicit formula (§4). Then we will show, in (§5), that if $\mu(A) > 0$, the set

$$\{g: \ \mu(A \cap T_g^{-1}A \cap (T_g S_g)^{-1}A) > 0\}$$

is both left and right syndetic. (A subset $B \subset G$ is said to be left (right) syndetic if $G = \bigcup_{i=1}^n g_i B$ $(G = \bigcup_{i=1}^n B g_i)$ for some $g_1, \cdots, g_n \in G$.) As an application of this result, a combinatorial corollary will be proved (Theorem 6.1) which may be viewed as a generalization of Roth's combinatorial result about arithmetic triples. In the abelian case, this will also give as a consequence (Corollary 6.4) a theorem of Brown and Buhler ([3], [6]). In order to formulate this corollary now, we indicate the following definition from (§5). If G is an amenable group with a left Følner sequence $\{\Phi_n\}$, and $A \subset G$, the upper density of A with respect to $\{\Phi_n\}$ is the number $\overline{d}(A) = \limsup_{n\to\infty} \frac{|\Phi_n \cap A|}{|\Phi_n|}$. We also make note of the fact that if a group G is amenable then $G \times G$ is as well.

Combinatorial Result 1. (Theorem 6.1) *Suppose G is a countable amenable group and that $E \subset G \times G$ has positive upper density with respect to a left Følner sequence $\{\Phi_n\}$ for $G \times G$. Then the set*

$$\big\{g \in G: \ \textit{there exists } (a, b) \in G \times G \textit{ such that } \{(a, b), (ga, b), (ga, gb)\} \subset E\big\}$$

is both left and right syndetic in G.

Finally, in (§7), we will prove, as a further application of the results of (§5), a topological multiple recurrence theorem for three commuting actions of a countable amenable group G as homeomorphisms of a compact metric space. This theorem will have the following consequence, which, in light of some counter-examples (in free groups G) given in [1], is somewhat surprising:

Combinatorial Result 2. (Corollary 7.2) *Suppose that G is a countable amenable group, $r \in \mathbf{N}$, and $G \times G \times G = \bigcup_{i=1}^r C_i$. Then the set*

$$\big\{g \in G: \ \textit{there exists } i, 1 \le i \le r, \ \textit{and } (a, b, c) \in G \times G \times G \textit{ such that}$$

$$\{(a, b, c), (ga, b, c), (ga, gb, c), (ga, gb, gc)\} \subset C_i\big\}$$

is both left and right syndetic in G. **Acknowledgement:** We thank H. Furstenberg for inspiring the present work and for many helpful suggestions.

2 Preliminaries

In this section we give a brief account of the machinery that we will be using. For further information, the interested reader may consult ([7], [11], [17], and [18]). We note, however, that our treatment differs slightly from that of Zimmer's, in that we consider left actions $\{T_g\}$ (that is, $T_{gh}x = T_g T_h x$), of our space X, whereas he considers right actions. Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are probability measure spaces and that $\pi : \ X \to Y$ is measure preserving. Then there exists a family of probability measures $\{\mu_y : y \in Y\}$ on $(X, \mathcal{A})$, having the following properties:

(i) $\mu_y(\pi^{-1}(y)) = 1$ for a.e. $y \in Y$.

(ii) For every $f \in L^1(X, \mathcal{A}, \mu)$, the function $y \to \int f d\mu_y$ is $\mathcal{B}$-measurable, and $\int f d\mu = \int \{\int f d\mu_y\} d\nu$.

This "decomposition" $\{\mu_y : y \in Y\}$ of the measure μ is essentially unique; that is, if $\{\mu_y'\}$ is another family with the above properties, then $\mu_y = \mu_y'$ for a.e. $y \in Y$. Conditional expectation is given by

$$\mathbf{E}(f \mid \pi^{-1}(\mathcal{B}))(x) = \int f(t) d\mu_{\pi(x)}(t)$$

for $f \in L^1(X, \mathcal{A}, \mu)$. $\langle \cdot, \cdot \rangle_y$ and $|| \cdot ||_y$ will be used to denote the inner product and norm with respect to the measure μ_y.

In a small abuse of terminology, we will often regard $\mathcal{B}$ itself as being a σ-algebra on X, namely the σ-algebra $\pi^{-1}(\mathcal{B})$. Also, since any sub-σ-algebra of $\mathcal{A}$ may be realized as $\pi^{-1}(\pi(\mathcal{A}))$ for some factor map π onto some space Y, we will often speak of such σ-algebras as being themselves "factors". Suppose that $\psi(x)$ is a $\mathcal{B}$-measurable (that is, measurable with respect to $\pi^{-1}(\mathcal{B})$) function on X. Upon removal of a set of measure zero from X, $\psi(y) = \psi(\pi^{-1}(y))$ will then be well-defined. Thus ψ may be thought of as a function on Y.

A closed subspace $\mathcal{M} \subset L^2(X, \mathcal{A}, \mu)$ will be called a *$\mathcal{B}$-module* if, for every $\mathcal{B}$-measurable function ψ and $\varphi \in \mathcal{M}$ satisfying $\psi\varphi \in L^2(X, \mathcal{A}, \mu)$, we have $\psi\varphi \in \mathcal{M}$. A (finite or infinite) subset $S \subset L^2(X, \mathcal{A}, \mu)$ spans a $\mathcal{B}$-module $\mathcal{M}$ if the set

$$\{\psi_1\varphi_1 + \cdots + \psi_k\varphi_k : \ \psi_1, \ldots, \psi_k \in L^\infty(X, \mathcal{B}, \mu) \text{ and } \varphi_1, \ldots, \varphi_k \in S\ \}$$

is dense in $\mathcal{M}$.

If $\mathcal{M}$ is a $\mathcal{B}$-module and $y \in Y$, we will denote by $\mathcal{M}_y$ the image of $\mathcal{M}$ under the restriction map $f \to f_y = f\mid_{\pi^{-1}(y)}$. $\mathcal{M}_y$ is a subspace of $L^2(X, \mathcal{A}, \mu_y)$, and for any $f \in L^2(X, \mathcal{A}, \mu)$, the set $\{y : f_y \in \mathcal{M}_y\}$ is $\mathcal{B}$-measurable. If there exists $r < \infty$ such that for a.e. $y \in Y$, $\dim \mathcal{M}_y \leq r$, then $\mathcal{M}$ will be said to be *finite dimensional.* If $\dim \mathcal{M}_y$ is constant a.e., then we will say that $\mathcal{M}$ is of *uniform dimension.* Conversely, if for a.e. $y \in Y$, $\mathcal{M}_y$

is a subspace of $L^2(X, \mathcal{A}, \mu_y)$, and if for any $f \in L^2(X, \mathcal{A}, \mu)$, the set $\{y : f_y \in \mathcal{M}_y\}$ is measurable, then the subspace $\{f : f_y \in \mathcal{M}_y \ a.e.\} \subset L^2(X, \mathcal{B}, \mu)$ will be a $\mathcal{B}$-module which we denote by $\mathcal{M} = \int_{\oplus} \mathcal{M}_y \, d\nu(y)$.

For any $\mathcal{B}$-module $\mathcal{M}$, a (finite or infinite) sequence $\{\psi_n\} \subset \mathcal{M}$ is called a *global orthonormal sequence* if $\langle \psi_i, \psi_j \rangle_y = 0$ for $i \neq j$ and $\|\psi\|_y = 1$ a.e. If $\{\psi_j\}$ also spans $\mathcal{M}$, then it will be called a *global orthonormal basis* for $\mathcal{M}$. It is clear that if $\{\psi_j\}$ is a global orthonormal basis for $\mathcal{M}$, then for a.e. $y \in Y$, the non-zero elements in $\{\psi_{j,y}\}$ form an orthonormal basis for $\mathcal{M}_y$. We remark that, given a $\mathcal{B}$-module $\mathcal{M}$ of uniform dimension, one can derive from any spanning sequence (using a modified Gram-Schmidt procedure), a global orthonormal basis.

Suppose G is a countable amenable group, and that $\{T_g : g \in G\}$ and $\{S_g : g \in G\}$ (sometimes denoted by simply $\{T_g\}$ and $\{S_g\}$), are measure preserving G-actions on X and Y, respectively, and that $\pi T_g = S_g \pi$ for all $g \in G$. Then $(X, \mathcal{A}, \mu, \{T_g\})$ will be called an *extension* of $(Y, \mathcal{B}, \nu, \{S_g\})$, and $(Y, \mathcal{B}, \nu, \{S_g\})$ will be called a factor of $(X, \mathcal{A}, \mu, \{T_g\})$. If there exists a $\{T_g\}$-invariant measurable set $X' \subset X$ and an $\{S_g\}$-invariant measurable set $Y' \subset Y$, with $\mu(X') = \nu(Y') = 1$, such that π is a bimeasurable bijection from X' to Y', then $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Y, \mathcal{B}, \nu, \{S_g\})$ will be said to be *isomorphic*.

In general, if $(Y, \mathcal{B}, \nu, \{S_g\})$ is a factor of $(X, \mathcal{A}, \mu, \{T_g\})$, then $\pi^{-1}(\mathcal{B})$ is a $\{T_g\}$-invariant sub-σ-algebra of $\mathcal{A}$. As before, we will sometimes write $\mathcal{B}$ instead of $\pi^{-1}(\mathcal{B})$, and speak of the factor $(X, \mathcal{B}, \mu, \{T_g\})$, or just the factor $\mathcal{B}$.

For $f \in L^2(X, \mathcal{A}, \mu)$, write $T_g f(x) = f(T_g x)$. Then T_g is a unitary operator and $\{T_g\}_{g \in G}$ is a unitary anti-action of G (i.e., $T_{gh} f = T_h T_g f$). More generally we have, for $g \in G$ and $y \in Y$, operators

$$T_{g,y} : L^2(X, \mathcal{A}, \mu_{T_g y}) \to L^2(X, \mathcal{A}, \mu_y), \ \ T_{g,y} f(x) = f(T_g x).$$

For a.e. $y \in Y$, $T_{g,y}$ is unitary for all $g \in G$. A $\mathcal{B}$-module $\mathcal{M}$ is said to be $\{T_g\}$-invariant if $T_g \mathcal{M} \subset \mathcal{M}$ for all $g \in G$. Equivalently, for a.e. $y \in Y$, $T_{g,y} \mathcal{M}_{T_g y} = \mathcal{M}_y$ for all $g \in G$.

We denote by $K(\mathcal{B}, \{T_g\})$ the smallest closed sub-space of $L^2(X, \mathcal{A}, \mu)$ containing all finite dimensional $\{T_g\}$-invariant $\mathcal{B}$-modules. One can show that there exists a $\{T_g\}$-invariant sub-σ-algebra $\mathcal{B}_T$ such that $L^2(X, \mathcal{B}_T, \mu) = K(\mathcal{B}, \{T_g\})$. If $\mathcal{B}_T = \mathcal{A}$, $(X, \mathcal{A}, \mu, \{T_g\})$ will be called a *compact* extension of $(Y, \mathcal{B}, \nu, \{T_g\})$.

A *cocycle representation* on a $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$ is a family of unitary maps

$$\alpha(g, y) : \mathcal{M}_{T_g y} \to \mathcal{M}_y; \ g \in G, \ y \in Y$$

satisfying $\alpha(gh, y) = \alpha(h, y)\alpha(g, T_h y)$, such that for every φ, $\psi \in \mathcal{M}$ and $g \in G$, the pruduct $\langle \varphi_y, \alpha(g, y)\psi_{T_g y} \rangle_y$ is a measurable function of y. It is clear that

$$T_{g,y} : \ \mathcal{M}_{T_g y} \to \mathcal{M}_y$$

defines a cocycle representation $\{T_{g,y} : g \in G, y \in Y\}$ on $\mathcal{M}$, called the *natural cocycle representation* induced by $\{T_g\}$. We will denote this cocycle by $\{T_g \mid_{\mathcal{M}}\}$. Suppose next that H is a compact group. We call a measurable map $\alpha : G \times Y \to H$ an *anti-cocycle* if for a.e. $y \in Y$,

$$\alpha(gh, y) = \alpha(g, T_h y)\alpha(h, y).$$

If $(Y, \mathcal{B}, \nu, \{T_g\})$ is ergodic, then every finite dimensional, $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$ will be of uniform dimension, and therefore will have a global orthonormal basis $\varphi_1, \cdots, \varphi_n$. Then we will have $\mathcal{M} = \{f \in L^2(X, \mathcal{A}, \mu) : f(x) = \lambda_1(x)\varphi_1(x) + \cdots + \lambda_n(x)\varphi_n(x) : \lambda_i$ is $\mathcal{B}$-measurable, $1 \le i \le n.\}$ By $\{T_g\}$-invariance of $\mathcal{M}$, we may write

$$T_g\varphi_i(x) = \sum_{j=1}^{n} \lambda_{ij}(g, \pi(x))\varphi_j(x),\ 1 \le i \le n.$$

Since $T_{g,y} : \mathcal{M}_{T_g y} \to \mathcal{M}_y$ is unitary, $\{T_g\varphi_i : 1 \le i \le n\}$ will also be a basis and for a.e. $y \in Y$, $(\lambda_{ij}(g, y))_{1 \le i,j \le n}$ will be a unitary matrix $M(g, y)$. $\mathbf{U}(n)$, the group of $n \times n$ unitary matrices over $\mathbf{C}$, is a compact group, and $\{M(g, y)\}$ is an anti-cocycle, called the anti-cocycle induced by $(\mathcal{M}, \{T_g\})$ with respect to $\varphi_1, \cdots, \varphi_n$. It is characterized by:

$$\begin{pmatrix} T_g\varphi_1(x) \\ \vdots \\ T_g\varphi_n(x) \end{pmatrix} = M(g, \pi(x)) \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix}.$$

If we fix some $y \in Y$ and $x \in \pi^{-1}(y)$, we can write the above equation as

$$\begin{pmatrix} T_{g,y}\varphi_{1,T_g y}(x) \\ \vdots \\ T_{g,y}\varphi_{n,T_g y}(x) \end{pmatrix} = M(g, y) \begin{pmatrix} \varphi_{1,y}(x) \\ \vdots \\ \varphi_{n,y}(x) \end{pmatrix}.$$

Hence the matrix $M(g, y)$ induces the operator $T_{g,y} : \mathcal{M}_{T_g y} \to \mathcal{M}_y$ relative to the global basis (φ_i).

Suppose that two systems $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, and that $\mathcal{M}$ and $\mathcal{N}$ are finite dimensional $\{T_g\}$ and $\{S_g\}$-invariant $\mathcal{B}$-modules in $L^2(X, \mathcal{A}, \mu)$ and $L^2(Z, \mathcal{C}, \omega)$ respectively. Let $\pi : X \to Y$ and $\tau : Z \to Y$ be the factor maps. Two cocycle representations α on $\mathcal{M}$ and β on $\mathcal{N}$ will be said to be *cohomologous*, or *equivalent*, if there exists a family of unitary maps $V_y : \mathcal{M}_y \to \mathcal{N}_y$, $y \in Y$, such that for any $\varphi \in \mathcal{N}$, $\psi \in \mathcal{M}$, $\langle \varphi_y, V_y\psi_y \rangle_y$ is measurable in y, and

$$V_y \alpha(g, y) V_{T_g y}^{-1} = \beta(g, y)$$

a.e.

Proposition 2.1 *Suppose $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, and that $\mathcal{M}$ and $\mathcal{N}$ are, respectively, $\{T_g\}$ and $\{S_g\}$-invariant finite dimensional $\mathcal{B}$-modules in $L^2(X, \mathcal{A}, \mu)$ and $L^2(Z, \mathcal{C}, \omega)$. Then $\{T_g \mid_{\mathcal{M}}\}$ is equivalent to $\{S_g \mid_{\mathcal{N}}\}$ if and only if for any global orthonormal basis $\varphi_1, \cdots, \varphi_n$ for $\mathcal{M}$, there exists a global orthonormal basis $\psi_1, \cdots, \psi_n$ of $\mathcal{N}$ such that the anti-cocycle $\{M(g, y)\}$ induced by $(\mathcal{M}, \{T_g\})$ with respect to $\varphi_1, \cdots, \varphi_n$ is equal a. e. to the anti-cocycle $\{N(g, y)\}$ induced by $(\mathcal{N}, \{S_g\})$ with respect to $\psi_1, \cdots, \psi_n$.*

Proof. Suppose that $\{T_g \mid_{\mathcal{M}}\}$ is equivalent to $\{S_g \mid_{\mathcal{N}}\}$ via a family $\{V_y : \mathcal{M}_y \to \mathcal{N}_y\}$ of unitary operators satisfying

$$S_{g,y} = V_y T_{g,y} V_{T_g y}^{-1}.$$

Suppose that $\{\varphi_1, \cdots, \varphi_n\}$ is any global orthonormal basis for $\mathcal{M}$. Let $\psi_i \in \mathcal{N}_y$ be defined fiberwise by $\psi_{i,y} = V_y \varphi_{i,y}$, $1 \le i \le n$, $y \in Y$. Then the anti-cocycle $\{M(g,y)\}$ induced by $(\mathcal{M}, \{T_g\})$ with respect to $\varphi_1, \cdots, \varphi_n$ is equal to the anti-cocycle $\{N(g,y)\}$ induced by $(\mathcal{N}, \{S_g\})$ with respect to $\psi_1, \cdots, \psi_n$.

For the converse, simply define the family V_y by $V_y : \varphi_{i,y} \to \psi_{i,y}$, $1 \le i \le n$, $y \in Y$. □

Suppose that systems $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$. We define $\mathcal{A} \otimes \mathcal{C}$ to be the σ-algebra on $X \times Z$ generated by $\{A \times C : A \in \mathcal{A}, C \in \mathcal{C}\}$. We will denote by $\mu \times_{\mathcal{B}} \omega$ the measure on $(X \times Z, \mathcal{A} \otimes \mathcal{C})$ defined by

$$\int F \, d\mu \times_{\mathcal{B}} \omega = \int \left(\int F \, d\mu_y \times \omega_y \right) d\nu(y)$$

for any $\mathcal{A} \otimes \mathcal{C}$-measurable function F. $\mu \times_{\mathcal{B}} \omega$ is characterized by its effect on functions of the form $\varphi \otimes \psi$, where $\varphi \otimes \psi(x,z) = \varphi(x)\psi(z)$. One easily checks that $\mu \times_{\mathcal{B}} \omega$ is $\{T_g \times S_g\}$-invariant. Hence $(X \times Z, \mathcal{A} \otimes \mathcal{C}, \mu \times_{\mathcal{B}} \omega, \{T_g \times S_g\})$ is a measure preserving system having $(Y, \mathcal{B}, \nu, \{T_g\})$ as a factor. The support of $\mu \times_{\mathcal{B}} \omega$ is seen to be $\{(x,z) \in X \times Z : \pi(x) = \tau(z)\}$. It follows that the factor maps from $X \times Z$ onto Y, $\pi'(x,z) = \pi(x)$ and $\tau'(x,z) = \tau(z)$, are equal a.e. (with respect to $\mu \times_{\mathcal{B}} \omega$). The following definition appears in [11].

Definition 2.1 *A positive definite symmetric kernel (or PDS kernel) for $\mathcal{B}$ is a function $H(x,x') \in L^\infty(X \times X, \mathcal{A} \otimes \mathcal{A}, \mu \times_{\mathcal{B}} \mu)$ which satisfies the following for a.e. (with respect to $\mu \times_{\mathcal{B}} \mu$) $(x,x') \in X \times X$:*

(i) $H(x,x') = \overline{H(x',x)}$;

(ii) $\int H(x,x')\psi(x)\overline{\psi(x')} \, d\mu \times_{\mathcal{B}} \mu \ge 0$ *for any* $\psi \in L^\infty(X, \mathcal{A}, \mu)$.

Every $H \in L^2(X \times X, \mathcal{A} \otimes \mathcal{A}, \mu \times_{\mathcal{B}} \mu)$ induces an operator $\mathbf{H}$ on $L^2(X, \mathcal{A}, \mu)$, given by

$$(\mathbf{H}\psi)(x) = \int H(x,x')\psi(x') \, d\mu_{\pi(x)}(x').$$

If H is a PDS kernel, $\mathbf{H}$ will be self-adjoint and positive (semi-)definite. We also have a fiberwise decomposition $\mathbf{H} = \int_\oplus \mathbf{H}_y \, d\nu(y)$, where

$$\mathbf{H}_y : L^2(X, \mathcal{A}, \mu_y) \to L^2(X, \mathcal{A}, \mu_y)$$

is given by

$$(\mathbf{H}_y\psi)(x) = \int H(x,x')\psi(x')d\mu_y(x').$$

For a.e. $y \in Y$, $\mathbf{H}_y$ is a Hilbert-Schmidt operator.

According to the spectral theorem for compact operators, for a.e. $y \in Y$ the spectrum of $\mathbf{H}_y$ consists of a bounded sequence of positive numbers and possibly zero, with zero the only possible limit point of this set. Each non-zero element of the spectrum is an eigenvalue with finite multiplicity. The eigenspace corresponding to $\lambda = 0$ (the nullspace of $\mathbf{H}_y$) may be empty, of finite dimension, or of infinite dimension. One can list the non-zero eigenvalues of $\mathbf{H}_y$, repeating each according to its multiplicity, as $\lambda_1(y) \geq \lambda_2(y) \geq \cdots$. $\{\lambda_1(y) \geq \lambda_2(y) \geq \cdots\}$ will be called a complete set of eigenvalues. We now have the following:

Theorem 2.2 *The functions $\lambda_j(y)$ are $\mathcal{B}$-measurable. Moreover, there exists a sequence of functions $\{\theta_n\} \subset L^2(X, \mathcal{A}, \mu)$ such that for all $m, n \in \mathbf{N}$, and a.e. $y \in Y$,*

$$\langle \theta_{n,y}, \theta_{m,y} \rangle_y = \delta_{nm} \text{ and } \mathbf{H}_y \theta_{n,y} = \lambda_n(y) \theta_{n,y}.$$

A proof of this theorem as well as the following corollary can be found in [11, 3.7].

Corollary 2.3 *Suppose that H is a PDS kernel. Let $\{\theta_n\}$ be a sequence of functions such that $\langle \theta_{n,y}, \theta_{m,y} \rangle_y = \delta_{nm}$ and $\mathbf{H}_y \theta_{n,y} = \lambda_n(y) \theta_{n,y}$. Then*

$$H(x, x') = \sum_n \lambda_n(y) \theta_n(x) \overline{\theta_n(x')}$$

and

$$\lambda_n(y) = \int H(x, x') \overline{\theta_n(x)} \theta_n(x') \, d\mu_y(x) d\omega_y(x') \quad a.e.$$

Suppose H is a $\{T_g \times T_g\}$-invariant PDS kernel. Then

$$\begin{aligned} \mathbf{H}(T_g \psi)(x) &= \int H(x, x') \psi(T_g x') \, d\mu_y(x') \\ &= \int H(x, T_g^{-1} x') \psi(x') \, d\mu_{T_g y}(x') \\ &= \int H(T_g x, x') \psi(x') \, d\mu_{T_g y}(x') \\ &= \mathbf{H}\psi(T_g x) = T_g(\mathbf{H}\psi)(x). \end{aligned}$$

Thus $\mathbf{H}T_g = T_g \mathbf{H}$. Therefore $\mathbf{H}_{T_g y} = T_{g,y}^{-1} \mathbf{H}_y T_{g,y}$, which implies that $\lambda_n(T_g y) = \lambda_n(y)$. Since we are assuming that $(Y, \mathcal{B}, \nu, \{T_g\})$ is ergodic, $\lambda_1(y), \lambda_2(y), \cdots$ must be constant a.e., and we may merely write $\lambda_1, \lambda_2, \cdots$. The numbers λ_i will be called the *$\mathcal{B}$–eigenvalues* of $\mathbf{H}$. We also call $\theta_1, \theta_2, \cdots$ (as in Theorem 2.2) the *$\mathcal{B}$–eigenfunctions* corresponding to $\lambda_1, \lambda_2, \cdots$. The following proposition follows immediately from the fact that $\mathbf{H}T_g = T_g \mathbf{H}$.

Proposition 2.4 *Suppose that H is a PDS kernel with a complete set of eigenvalues*

$$\{\lambda_1(y) \geq \lambda_2(y) \geq \cdots\}.$$

For $\lambda \in \{\lambda_1, \lambda_2, \cdots\}$, let $\mathcal{M}_{\lambda,y} \subset L^2(X, \mathcal{A}, \mu_y)$ be the finite dimensional subspace spanned by $\{\theta_{n,y} : \lambda_n = \lambda\}$ and let $\mathcal{M}_\lambda = \int_\oplus \mathcal{M}_{\lambda,y} \, d\nu(y)$. Then $\mathcal{M}_\lambda \subset K(\mathcal{B}, \{T_g\})$.

Corollary 2.5 *If H is a $\{T_g \times T_g\}$-invariant PDS kernel, then the image of $\mathbf{H}$ is contained in $K(\mathcal{B}, \{T_g\})$.*

Suppose again that $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, with factor maps $\pi : X \to Y$ and $\tau : Z \to Y$, respectively. Let $\mathcal{M}$, $\mathcal{N}$ be $\{T_g\}$- and $\{S_g\}$-invariant finite dimensional $\mathcal{B}$-modules in $L^2(X, \mathcal{A}, \mu)$ and in $L^2(Z, \mathcal{C}, \omega)$, respectively. If $\{T_g\,|_{\mathcal{M}}\}$ is equivalent to $\{S_g\,|_{\mathcal{N}}\}$, then by Proposition 2.1 there exist bases $\{\varphi_1, \cdots, \varphi_N\}$ for $\mathcal{M}$ and $\{\psi_1, \cdots, \psi_N\}$ for $\mathcal{N}$ such that the anti-cocycle $\{M(g, y)\}$ induced by $(\mathcal{M}, \{T_g\})$ with respect to $\{\varphi_1, \cdots, \varphi_N\}$ is equal a.e. to the anti-cocycle $\{N(g, y)\}$ induced by $(\mathcal{N}, \{S_g\})$ with respect to $\{\psi_1, \cdots, \psi_N\}$. In this case, let

$$\eta(x, z) = \sum_{n=1}^{N} \varphi_n(x)\overline{\psi_n(z)}. \tag{3}$$

Then η is $\{T_g \times S_g\}$-invariant with respect to $\mu \times_{\mathcal{B}} \omega$, since whenever $\pi(x) = y = \tau(z)$, we have

$$\begin{aligned} (T_g \times S_g)\eta(x, z) &= \sum_{n=1}^{N} \varphi_n(T_g x)\overline{\psi_n(S_g z)} \\ &= \Big(\varphi_1(x), \cdots, \varphi_N(x)\Big) M^{\tau}(g, y)\overline{N(g, y)} \begin{pmatrix} \overline{\psi_1(z)} \\ \vdots \\ \overline{\psi_N(z)} \end{pmatrix} \\ &= \sum_{n=1}^{N} \varphi_n(x)\overline{\psi_n(z)} = \eta(x, z). \end{aligned}$$

(Recall that $M(g, y)$ is unitary and equal to $N(g, y)$ a.e.)

Now suppose that $\theta(x, z)$ is any function which is $\{T_g \times S_g\}$-invariant with respect to $\mu \times_{\mathcal{B}} \omega$. We will show that $\theta(x, z)$ is a limit of functions of the form (3). Define, for $\pi(x) = \pi(x') = y = \tau(z) = \tau(z')$,

$$H_1(x, x') = \int \theta(x, z)\overline{\theta(x', z)}\, d\omega_y(z)$$

and

$$H_2(z, z') = \int \theta(x, z)\overline{\theta(x, z')}\, d\mu_y(x).$$

Then H_1, H_2 are $\{T_g \times T_g\}$- and $\{S_g \times S_g\}$-invariant PDS kernels on $X \times X$ and $Z \times Z$, respectively, inducing fiberwise Hilbert-Schmidt operators $\mathbf{H}_1$ on $L^2(X, \mathcal{A}, \mu)$ and $\mathbf{H}_2$ on $L^2(Z, \mathcal{C}, \omega)$ (see the discussion following Definition 2.1).

θ induces a family of operators $\Theta_y : L^2(X, \mathcal{A}, \mu_y) \to L^2(Z, \mathcal{C}, \omega_y)$ by

$$(\Theta_y \varphi)(z) = \int \overline{\theta(x, z)}\varphi(x)\, d\mu_y(x).$$

The dual of Θ_y, $\Theta_y^* : L^2(Z,\mathcal{C},\omega_y) \to (X,\mathcal{A},\mu_y)$, is given by

$$(\Theta_y^*\psi)(x) = \int \theta(x,z)\psi(z)\ d\omega_y(z).$$

Let $\Theta = \int_\oplus \Theta_y\ d\nu$. One may check that $\mathbf{H}_{1,y} = \Theta_y^*\Theta_y$, and $\mathbf{H}_{2,y} = \Theta_y\Theta_y^*$, hence one has $\mathbf{H}_1 = \Theta^*\Theta$ and $\mathbf{H}_2 = \Theta\Theta^*$.

Let $\{\lambda_1 \geq \lambda_2 \geq \cdots\}$ be the complete set of $\mathcal{B}$-eigenvalues for $\mathbf{H}_1$, and let $\{\varphi_1, \varphi_2, \cdots\}$ be the corresponding $\mathcal{B}$-eigenfunctions. Then $\Theta\varphi_j \neq 0$ for $j = 1, 2, \cdots$ (otherwise $\mathbf{H}_1\varphi_j = 0$), and

(i) $\mathbf{H}_2\Theta\varphi_j = \Theta\Theta^*\Theta\varphi_j = \Theta\mathbf{H}_1\varphi_j = \lambda_j\Theta\varphi_j$;

(ii) $\langle\Theta\varphi_j, \Theta\varphi_i\rangle = \langle\varphi_j, \mathbf{H}_1\varphi_j\rangle = \lambda_j(y)\delta_{ij}$.

Hence, $\{\lambda_1(y) \geq \lambda_2(y) \geq \cdots\}$ is a subset of a complete set of eigenvalues for $\mathbf{H}_2$, and $\Theta\varphi_1, \Theta\varphi_2, \cdots$ are corresponding eigenfunctions. We have therefore shown that the complete set of $\mathcal{B}$-eigenvalues for $\mathbf{H}_1$ is a subset of the complete set of $\mathcal{B}$-eigenvalues for $\mathbf{H}_2$. A similar argument shows reverse inclusion, hence the two sets are equal and we have the following:

Proposition 2.6 *If $\{\lambda_1 \geq \lambda_2 \geq \cdots\}$ is the complete set of $\mathcal{B}$-eigenvalues for $\mathbf{H}_1$, and $\varphi_1, \varphi_2, \cdots$ are corresponding $\mathcal{B}$-eigenfunctions, then $\{\lambda_1 \geq \lambda_2 \geq \cdots\}$ is also the complete set of $\mathcal{B}$-eigenvalues for $\mathbf{H}_2$, and $\Theta\varphi_1, \Theta\varphi_2, \cdots$ are corresponding $\mathcal{B}$-eigenfunctions.*

A consequence of Proposition 2.6 is that for all $\varphi \in L^2(X,\mathcal{A},\mu)$ with $\Theta\varphi \neq 0$, $\Theta\varphi$ lies in the span of the $\mathcal{B}$-eigenfunctions of $\mathbf{H}_2$, so we have $\mathbf{H}_2\Theta\varphi \neq 0$, which implies that $\mathbf{H}_1\varphi \neq 0$. Hence $\ker\Theta = \ker\mathbf{H}_2$. Now, for a.e. $z \in Z$, we have $\theta(\cdot, z) \in L^2(X,\mathcal{A},\mu_y)$, where $y = \tau(z)$, so that

$$\theta(x,z) = \sum_j \left(\int \theta(x',z)\overline{\varphi_j(x')}d\mu_y(x') \right)\varphi_j(x) = \sum_j \overline{\Theta_y\varphi_j(z)}\varphi_j(x)$$

for a.e. $x \in X$ (with respect to μ_y). We now have the following corollary:

Corollary 2.7 *If $\{\lambda_1 \geq \lambda_2 \geq \cdots\}$ is the complete set of $\mathcal{B}$-eigenvalues of $\mathbf{H}_1$, and $\varphi_1, \varphi_2, \cdots$ are the corresponding $\mathcal{B}$-eigenfunctions, then for a.e. (x,z), with respect to $\mu \times_\mathcal{B} \omega$,*

$$\theta(x,z) = \sum_j \varphi_j(x)\overline{(\Theta\varphi_j)(z)}.$$

For $0 \neq \lambda \in \{\lambda_1, \lambda_2, \cdots\}$, let $\mathcal{M}_{\lambda,y}$, $\mathcal{N}_{\lambda,y}$ be the finite dimensional subspaces spanned by $\{\varphi_{n,y} : \lambda_n = \lambda\}$ and $\{\Theta^*\varphi_{n,y} : \lambda_n = \lambda\}$ in $L^2(X,\mathcal{A},\mu_y)$ and in $L^2(Z,\mathcal{C},\omega_y)$, respectively. Let

$$\mathcal{M}_\lambda = \int_\oplus \mathcal{M}_{\lambda,y}\ d\nu(y) \text{ and } \mathcal{N}_\lambda = \int_\oplus \mathcal{N}_{\lambda,y}\ d\nu(y).$$

$\mathcal{M}_\lambda$ and $\mathcal{N}_\lambda$ are $\{T_g\}$- and $\{S_g\}$-invariant finite dimensional $\mathcal{B}$-modules, respectively. Let $\mathbf{A}_\lambda : \mathcal{M}_\lambda \to \mathcal{N}_\lambda$ be defined by

$$(\mathbf{A}_\lambda\varphi_n)(z) = \frac{\Theta\varphi_n(z)}{\lambda} \quad \text{whenever } \lambda_n = \lambda.$$

Then $\mathbf{A}_{\lambda,y} : \mathcal{M}_{\lambda,y} \to \mathcal{N}_{\lambda,y}$ is unitary a.e., and by $T_g \times S_g$-invariance of Θ, we have, for all $g \in G$,

$$(S_g \mid_{\mathcal{N}_\lambda})\mathbf{A}_\lambda = \mathbf{A}_\lambda(T_g \mid_{\mathcal{M}_\lambda}).$$

Moreover, the anti-cocycle induced by $(\mathcal{M}_\lambda, \{T_g\})$ with respect to $\{\varphi_n : \lambda_n = \lambda\}$ is equal a.e. to the anti-cocycle induced by $(\mathcal{N}_\lambda, \{S_g\})$ with respect to $\{\frac{\Theta\varphi_n}{\lambda} : \lambda_n = \lambda\}$. Therefore, we have the following theorems:

Theorem 2.8 *If $\theta(x,z)$ is a $\{T_g \times S_g\}$-invariant function, then θ is a limit of linear combinations of functions of the form* (3).

Theorem 2.9 *If $\theta(x,z)$ is a $\{T_g \times S_g\}$-invariant function bounded a.e. with respect to $\mu \times_{\mathcal{B}} \omega$, then for any $f \in L^2(Z, \mathcal{C}, \omega)$,*

$$\Theta^* f(x) = \int \theta(x,z) f(z)\, d\omega_y(z) \in K(\mathcal{B}, \{T_g\}).$$

Remark. A proof of Theorem 2.9 can also be found in [18, page 561].

3 Irreducible Finite Dimensional Modules

If $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ are measure preserving systems with a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, then by Theorem 2.8, the space of $\{T_g \times S_g\}$-invariant functions in $L^2(X \times Z, \mathcal{A} \otimes \mathcal{C}, \mu \times_{\mathcal{B}} \omega)$ is spanned by functions of the form (3). In this section, we will find an orthonormal basis of the space of $\{T_g \times S_g\}$-invariant functions.

Definition 3.1 *A $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$ is said to be irreducible if for any $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{N} \subset \mathcal{M}$, we have $\mathcal{N} = \mathcal{M}$ or $\mathcal{N} = \{0\}$.*

The following lemma has been indicated (without proof) by Conze and Lesigne in [4, page 149]. For the sake of the completeness, we give a proof.

Proposition 3.1 *Suppose that $\mathcal{M} = \int_\oplus \mathcal{M}_y \, d\nu(y)$ is an irreducible finite dimensional $\{T_g\}$-invariant module, and that $\{\varphi_1, \cdots, \varphi_m\}$ is a global orthonormal basis for $\mathcal{M}$. Let $\{M(g,y)\}$ be the anti-cocycle induced by $(\mathcal{M}, \{T_g\})$ with respect to $\{\varphi_1, \cdots, \varphi_m\}$. Then:*

(i) If $A(y)$ is a $\mathcal{B}$-measurable $m \times m$-matrix-valued function, and $\{N(g,y) : g \in G\}$ is a family of $\mathcal{B}$-measurable matrix-valued functions such that

$$A(T_g y) M(g,y) = N(g,y) A(y)$$

for a.e. $y \in Y$ and all $g \in G$, then either $A(y) = 0$ a.e. or $A(y)$ is an $m \times m$ invertible matrix a.e.

(ii) If $A(y)$ is a $\mathcal{B}$-measurable $m \times m$-matrix-valued function such that

$$A(T_g y) M(g,y) = M(g,y) A(y)$$

for a.e. $y \in Y$ and all $g \in G$, then $A(y) = \lambda I$ a.e. (Here λ is a constant and I is the $m \times m$ identity matrix.)

Proof. (i) Let

$$\mathcal{M}_A = \left\{ f \in \mathcal{M} :\ f(x) = \Big(b_1(y) \cdots b_m(y)\Big) A(y) \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_m(x) \end{pmatrix} \right\}.$$

(Where $y = \pi(x)$ and b_i is $\mathcal{B}$-measurable, $1 \le i \le m$.) The $\mathcal{B}$-module $\mathcal{M}_A$ is $\{T_g\}$-invariant, for if $f(x) \in \mathcal{M}_A$ has the form indicated above, then, for all $g \in G$ one has

$$\begin{aligned} T_g f(x) &= \Big(T_g b_1(y) \cdots T_g b_m(y)\Big) A(T_g y) \begin{pmatrix} T_g\phi_1(x) \\ \vdots \\ T_g\phi_m(x) \end{pmatrix} \\ &= \Big(T_g b_1(y) \cdots T_g b_m(y)\Big) A(T_g y) M(g,y) \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_m(x) \end{pmatrix} \\ &= \Big(T_g b_1(y) \cdots T_g b_m(y)\Big) N(g,y) A(y) \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_m(x) \end{pmatrix} \in \mathcal{M}_A. \end{aligned}$$

Since $\mathcal{M}$ is irreducible, and $A(y)$ is not zero, we have $\mathcal{M}_A = \mathcal{M}$, from which it is easy to conclude that $A(y)$ is invertible a.e.

(ii) By (i), we know that $A(y)$ is invertible a.e., and that

$$A(T_g y) = M(g,y) A(y) M(g,y)^{-1}.$$

This implies that the eigenvalues of $A(y)$ coincide with those of $A(T_g y)$. By ergodicity on Y, we may conclude that the eigenvalues of $A(y)$ are constant a.e. Let λ be one of these eigenvalues and put

$$\mathcal{M}_\lambda = \left\{ f(x) = \Big(b_1(y) \cdots b_m(y)\Big) \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_m(x) \end{pmatrix} : \right.$$
$$\left. \Big(b_1(y) \cdots b_m(y)\Big) A(y) = \lambda \Big(b_1(y) \cdots b_m(y)\Big) \text{ a.e.} \right\}.$$

(Again, $y = \pi(x)$ and b_i's are $\mathcal{B}$-measurable.) Then $\mathcal{M}_\lambda$ is a $\mathcal{B}$-module. Suppose that $f \in \mathcal{M}_\lambda$ has the form indicated. Then

$$T_g f(x) = \Big(b_1(T_g y) \cdots b_m(T_g y)\Big) M(g,y) \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_m(x) \end{pmatrix}.$$

We have

$$\begin{aligned}\Big(b_1(T_gy)\cdots b_m(T_gy)\Big)M(g,y)A(y) &= \Big(b_1(T_gy)\cdots b_m(T_gy)\Big)A(T_gy)M(g,y)\\ &= \lambda\Big(b_1(T_gy)\cdots b_m(T_gy)\Big)M(g,y).\end{aligned}$$

Hence $T_gf \in \mathcal{M}_\lambda$ and $\mathcal{M}_\lambda$ is T_g-invariant. Hence $\mathcal{M}_\lambda = \mathcal{M}$ and $A(y) = \lambda I$ a.e.

□

For any $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$, we write $\mathcal{M}^* = \{\overline{\varphi} :\ \varphi \in \mathcal{M}\}$. For any cocycle representation α on $\mathcal{M}$, we define a cocycle representation $\overline{\alpha}$ on $\mathcal{M}^*$ by $\overline{\alpha}(y,g)\overline{\varphi}_y = \overline{\alpha(y,g)\varphi_y}$, $\varphi \in \mathcal{M}$. One may show that if $\alpha = \{T_g\ |_{\mathcal{M}}\}$, then $\overline{\alpha} = \{T_g\ |_{\mathcal{M}^*}\}$. Assume $\mathcal{M}$ and $\mathcal{N}$ are $\{T_g\}$- and $\{S_g\}$-invariant finite dimensional $\mathcal{B}$-modules on X and Z, respectively, and that α, β are cocycle representations on $\mathcal{M}$ and $\mathcal{N}$. Let $\mathcal{M} \otimes \mathcal{N}$ be the smallest closed subspace of $L^2(X \times Z, \mathcal{A} \otimes \mathcal{C}, \mu \times_{\mathcal{B}} \omega)$ containing $\{\varphi \otimes \psi : \varphi \in \mathcal{M}$ and $\psi \in \mathcal{N}\}$. $\mathcal{M} \otimes \mathcal{N}$ is a $\{T_g \times S_g\}$-invariant $\mathcal{B}$-module. Let $\alpha \otimes \beta$ be the cocycle representation on $\mathcal{M} \otimes \mathcal{N}$ defined by

$$(\alpha \otimes \beta)(g,y)\varphi_y \otimes \psi_y = \alpha(g,y)\varphi_y \otimes \beta(g,y)\psi_y.$$

A measurable family of operators $\{\mathbf{A}_y : \mathcal{N}_y \to \mathcal{M}_y |\ y \in Y\}$ will be called an *intertwining field* for α and β if for a.e. $y \in Y$, $\alpha(g,y)A_{T_gy} = A_y\beta(g,y)$ for all $g \in G$. We will write $S(\alpha,\beta)$ for the set of intertwining fields. $S(\alpha,\beta)$ is a vector space, and $int(\alpha,\beta) = \dim(S(\alpha,\beta))$ will be called the *intertwining number* of α and β. Proof of the following proposition can be found in [17, page 386].

Proposition 3.2 *For any cocycle representations α on $\mathcal{M}$ and β on $\mathcal{N}$,*

$$int(\alpha,\beta) = int(I, \alpha \otimes \overline{\beta}).$$

(I is the identity cocycle representation on the one-dimensional $\mathcal{B}$-module $L^2(Y,\mathcal{B},\nu)$.)

We now have the following corollary:

Corollary 3.3 *The dimension of the $\mathcal{B}$-module spanned by the $\{T_g \times S_g\}$-invariant functions in $\mathcal{M} \otimes \mathcal{N}^*$ is equal to $int(\{T_g|_{\mathcal{M}}\}, \{S_g|_{\mathcal{N}}\})$.*

Proof. If we write $L^2(Y,\mathcal{B},\nu) = \int_\oplus \mathcal{E}_y\ d\nu(y)$, and $I = \int_\oplus I_y\ d\nu(y)$, then $\dim \mathcal{E}_y = 1$ a.e. and we may identify $\mathcal{E}_y$ with $\mathbf{C}$, so that $I_y : \mathcal{E}_{T_gy} \to \mathcal{E}_y$ is just the identity $\mathbf{C} \to \mathbf{C}$ and we may suppress it. Also this identification allows us to view $S(I, \{T_g \times S_g|_{\mathcal{M}\otimes\mathcal{N}^*}\})$ as the set of all measurable families $\{A_y\}$ of linear functionals $A_y : \mathcal{M}_y \otimes \mathcal{N}_y^* \to \mathbf{C}$ satisfying

$$A_{T_gy} = A_y(T_{g,y} \times S_{g,y}). \tag{4}$$

Such families of functionals are given by integration against a function $h \in \mathcal{M} \otimes \mathcal{N}^*$:

$$A_yf = \int f(x,z)h(x,z)\ d\mu_y \times \omega_y(x,z) \tag{5}$$

for $f \in \mathcal{M}_y \otimes \mathcal{N}_y$. Conversely, any $h \in \mathcal{M} \otimes \mathcal{N}^*$ gives rise to a measurable family of functionals given by (5). Suppose that $\{A_y\}$ is given by (5). By (4), we have

$$\begin{aligned}\int f(T_g x, S_g z) h(x,z)\, d\mu_y \times \omega_y(x,z) &= \int f(x,z) h(x,z)\, d\mu_{T_g y} \times \omega_{T_g y}(x,z) \\ &= \int f(T_g x, S_g z) h(T_g x, S_g z)\, d\mu_y \times \omega_y(x,z),\end{aligned}$$

hence $\{A_y\} \in S(I, \{T_g \otimes S_g|_{\mathcal{M}\otimes\mathcal{N}^*}\})$ if and only if h is $\{T_g \times S_g\}$-invariant.

□

Lemma 3.4 *Suppose that* $(Y, \mathcal{B}, \nu\{T_g\})$ *is a common ergodic factor of* $(X, \mathcal{A}, \mu, \{T_g\})$ *and* $(Z, \mathcal{C}, \omega, \{S_g\})$*,* $\mathcal{M}$ *and* $\mathcal{N}$ *are* $\{T_g\}$*- and* $\{S_g\}$*-invariant irreducible finite dimensional* $\mathcal{B}$*-modules, respectively, and that* $\{T_g\,|_{\mathcal{M}}\}$ *is equivalent to* $\{S_g\,|_{\mathcal{N}}\}$*. Then* $int(\{T_g\,|_{\mathcal{M}}\}, \{S_g\,|_{\mathcal{N}}\}) = 1$.

Proof. Let $\{\varphi_1, \cdots, \varphi_m\}$ be a global orthonormal basis for $\mathcal{M}$. By Proposition 2.1, we can choose a basis $\{\psi_1, \cdots, \psi_m\}$ for $\mathcal{N}$ so that the anti-cocycle $\{M(g,y)\}$ induced by $\{T_g\,|_{\mathcal{M}}\}$ is equal to the anti-cocycle $\{N(g,y)\}$ induced by $\{S_g\,|_{\mathcal{N}}\}$. Since $\{T_g\,|_{\mathcal{M}}\}$ is equivalent to $\{S_g\,|_{\mathcal{N}}\}$, we have an intertwining field $\{\mathbf{A}_y : \mathcal{N}_y \to \mathcal{M}_y\}$. We introduce a family of matrices $\{A(y)\}$ given by

$$\mathbf{A}_y : \begin{pmatrix} b_1(y) \cdots b_m(y) \end{pmatrix} \begin{pmatrix} \psi_{1,y} \\ \vdots \\ \psi_{m,y} \end{pmatrix} \to \begin{pmatrix} b_1(y) \cdots b_m(y) \end{pmatrix} A(y) \begin{pmatrix} \varphi_{1,y} \\ \vdots \\ \varphi_{m,y} \end{pmatrix}.$$

One may check that $A(T_g y) M(g,y) = M(g,y) A(y)$. By Proposition 3.1 (ii), there exists a constant δ such that $A(y) = \delta I$ a.e. Since this is true for the matrix function induced by any intertwining field, the result follows.

□

Lemma 3.5 *Same hypotheses as Lemma 3.4, except that* $\{T_g\,|_{\mathcal{M}}\}$ *is not equivalent to* $\{S_g\,|_{\mathcal{N}}\}$*. Then* $int(\{T_g\,|_{\mathcal{M}}\}, \{S_g\,|_{\mathcal{N}}\}) = 0$.

Proof. Suppose that $int(\{T_g\,|_{\mathcal{M}}\}, \{S_g\,|_{\mathcal{N}}\}) > 0$. Let

$$\{\mathbf{A}_y : \mathcal{N}_y \to \mathcal{M}_y\} \in S(\{T_g\,|_{\mathcal{M}}\}, \{S_g\,|_{\mathcal{N}}\})$$

and put $\mathbf{A} = \int_{\oplus} A_y\, d\nu(y)$, $\mathbf{A} : \mathcal{N} \to \mathcal{M}$. It is easy to see that $\mathbf{A}\mathcal{N}$ is a $\{T_g\}$-invariant sub-module of $\mathcal{M}$, and therefore is equal to $\mathcal{M}$, and that $\ker \mathbf{A}$ is an $\{S_g\}$-invariant sub-module of $\mathcal{N}$, hence is trivial. We may conclude that $\mathcal{M}$ and $\mathcal{N}$ have the same dimension n. Choose bases for $\mathcal{M}$ and $\mathcal{N}$ inducing anti-cocycles $M(g,y)$ and $N(g,y)$, respectively. With respect to these bases (see proof of Lemma 3.4), $\{A_y\}$ will give rise to a family of $n \times n$ matrices $\{A(y)\}$ such that

$$A(T_g y) M(g,y) = N(g,y) A(y). \tag{6}$$

By Proposition 3.1 (i), $A(y)$ is invertible a.e. ¿From (6) we also have

$$M^*(g,y)A^*(T_g y) = A^*(y)N^*(g,y). \tag{7}$$

(Here W^* is the conjugate transpose of a matrix W. We note that W is unitary if and only if $WW^* = I$.) Since $M(g,y)$ and $N(g,y)$ are unitary matrices, multiplying (6) by (7) gives

$$A(T_g y)A^*(T_g y)N(g,y) = N(g,y)A(y)A^*(y).$$

By Proposition 3.1 (ii), we have $A(y)A^*(y) = \delta I$ a.e. for some constant δ. If ξ is an n-row column vector of norm 1 we have $\delta = \langle A(y)A^*(y)\xi, \xi\rangle = \langle A^*(y)\xi, A^*(y)\xi\rangle > 0$. Hence $V(y) = \frac{1}{\sqrt{\delta}}A(y)^{-1}$ is unitary a.e. and gives rise, via the established bases, to a family of unitary maps $V_y : \mathcal{M}_y \to \mathcal{N}_y$ satisfying

$$V_y\, T_{g,y}\, V^{-1}_{T_g y} = S_{g,y},$$

contradicting the non-equivalence of $\{T_g\mid_{\mathcal{M}}\}$ and $\{S_g\mid_{\mathcal{N}}\}$.

□

The following two theorems are consequences of Proposition 3.2, Corollary 3.3, Lemma 3.4 and Lemma 3.5.

Theorem 3.6 *Suppose $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, and that $\mathcal{M}$ and $\mathcal{N}$ are $\{T_g\}$- and $\{S_g\}$-invariant irreducible finite dimensional $\mathcal{B}$-modules, respectively. Then the following are equivalent:*

- *(i) $\{T_g\mid_{\mathcal{M}}\}$ is equivalent to $\{S_g\mid_{\mathcal{N}}\}$.*
- *(ii) $int(\{T_g\mid_{\mathcal{M}}\}, \{S_g\mid_{\mathcal{N}}\}) = 1$.*
- *(iii) The dimension of the space spanned by $\{T_g \times S_g\}$-invariant functions in $\mathcal{M} \otimes \mathcal{N}^*$ is one.*

Theorem 3.7 *Same hypotheses as Theorem 3.6. The following are equivalent:*

- *(i) $\{T_g\mid_{\mathcal{M}}\}$ is not equivalent to $\{S_g\mid_{\mathcal{N}}\}$.*
- *(ii) $int(\{T_g\mid_{\mathcal{M}}\}, \{S_g\mid_{\mathcal{N}}\}) = 0$.*
- *(iii) the dimension of the space spanned by $\{T_g \times S_g\}$-invariant functions in $\mathcal{M} \otimes \mathcal{N}^*$ is zero.*

Lemma 3.8 *There exists a sequence $\{\mathcal{M}_i\}$ of irreducible, $\{T_g\}$-invariant finite dimensional $\mathcal{B}$-modules such that $K(\mathcal{B}, \{T_g\}) = \bigoplus_i \mathcal{M}_i$.*

Proof. Let $\Re$ denote the set of all finite dimensional $\{T_g\}$-invariant $\mathcal{B}$-modules. Let

$$\Im = \{\Lambda \subset \Re : \ \mathcal{M}, \mathcal{N} \in \Lambda \Rightarrow \mathcal{M} \perp \mathcal{N} \text{ or } \mathcal{M} = \mathcal{N}\}.$$

$\Im$ is partially ordered by $\subset$. . Suppose that $\{\Lambda_\lambda\}$ is a totally ordered chain in $\Im$. It is clear that $\cup_\lambda \Lambda_\lambda \in \Im$. By Zorn's lemma, there is an element $\Lambda_0 \in \Im$ such that if $\Lambda \in \Im$ and $\Lambda_0 \subset \Lambda$, then $\Lambda = \Lambda_0$. Since $L^2(X, \mathcal{A}, \mu)$ is a separable space, Λ_0 is a countable set $\{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \cdots\}$. We claim that $K(\mathcal{B}, \{T_g\}) = \bigoplus \mathcal{M}_i$. Suppose that this is not the case. Then there is a finite dimensional $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$ such that

$$span\{\mathcal{M}, \mathcal{M}_1, \mathcal{M}_2, \cdots\} \neq \bigoplus_i \mathcal{M}_i.$$

Let $\mathcal{M}_0 = \{\varphi \in span\{\mathcal{M}, \mathcal{M}_1, \mathcal{M}_2, \cdots\} : \varphi \perp \bigoplus_i \mathcal{M}_i\}$. It is clear that $\mathcal{M}_0$ is a finite dimensional $\{T_g\}$-invariant $\mathcal{B}$-module. Then $\{\mathcal{M}_0\} \cup \Lambda_0 \supset \Lambda_0$ but $\{\mathcal{M}_0\} \cup \Lambda_0 \neq \Lambda_0$.

□

Definition 3.2 *A sequence of finite dimensional, irreducible, $\{T_g\}$-invariant $\mathcal{B}$-modules $\{\mathcal{M}_i\}$ will be called a global orthogonal decomposition of $K(\mathcal{B}, \{T_g\})$ if $\mathcal{M}_i \perp \mathcal{M}_j$ for $i \neq j$ and $K(\mathcal{B}, \{T_g\}) = \bigoplus_i \mathcal{M}_i$.*

Suppose measure preserving systems $(X, \mathcal{A}, \mu, \{T_g\})$ and $(Z, \mathcal{C}, \omega, \{S_g\})$ have a common ergodic factor $(Y, \mathcal{B}, \nu, \{T_g\})$, and that $\{\mathcal{M}_i\}$, $\{\mathcal{N}_i\}$ are global orthogonal decompositions of $K(\mathcal{B}, \{T_g\})$ and $K(\mathcal{B}, \{S_g\})$, respectively. For every pair (i, j) satisfying

$$\int(\{T_g \mid_{\mathcal{M}_i}\}, \{S_g \mid_{\mathcal{N}_j}\}) = 1,$$

choose a $\{T_g \times S_g\}$-invariant function $F_{ij}(x, z)$ in $\mathcal{M}_i \otimes \mathcal{N}_j^*$ with $\|F_{ij}\| = 1$. Put $F_{ij} = 0$ otherwise. Since F_{ij} is $\{T_g \times S_g\}$-invariant, $\int |F_{ij}(x, z)|^2 d\mu_y \times \omega_y(x, z)$ is a $\{T_g\}$-invariant function of y. By ergodicity, $\|F_{ij}\|_y = 1$ for a.e. $y \in Y$.

Theorem 3.9 *In $L^2(X \times Z, \mathcal{A} \otimes \mathcal{C}, \mu \times_{\mathcal{B}} \omega)$,*

$$\{F_{ij} : \ int(\{T_g \mid_{\mathcal{M}_i}\}, \{S_g \mid_{\mathcal{N}_j}\}) = 1\}$$

is an orthonormal basis of the space spanned by $\{T_g \times S_g\}$-invariant functions.

Proof. Clearly the set is orthonormal. Let f be any $\{T_g \times S_g\}$-invariant function. By Theorem 2.8, $f \in K(\mathcal{B}, \{T_g\}) \otimes K(\mathcal{B}, \{S_g\})$. Therefore, we may write $f = \sum_{i,j} f_{ij}$, where $f_{ij} \in \mathcal{M}_i \otimes \mathcal{N}_j^*$. It is then easy to see that f_{ij} must be $\{T_g \times S_g\}$-invariant for all i, j, hence $f_{ij} = c_{ij} F_{ij}$ for constants c_{ij}.

□

4 Convergence Theorem

Throughout this section, G will be a countable amenable group. The following mean ergodic theorem for amenable groups is well known:

Theorem 4.1 *Suppose that $\{T_g\}$ is a measure preserving G-action of a probability space $(X, \mathcal{A}, \mu)$. If $\{\Phi_n\}$ is any left Følner sequence for G, then for every $f \in L^2(X, \mathcal{A}, \mu)$,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} f(T_g x) = Pf(x)$$

exists in $L^2(X, \mathcal{A}, \mu)$. Furthermore, P is the orthogonal projection in $L^2(X, \mathcal{A}, \mu)$ onto the space of $\{T_g\}$-invariant functions.

Note that the constant functions are $\{T_g\}$-invariant. Therefore, if $A \in \mathcal{A}$, we have $||P(1_A)|| \geq \mu(A)$. It follows that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \mu(A \cap T_g^{-1} A) &= \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \int 1_A(x)\, 1_A(T_g x)\, d\mu(x) \\ &= \int 1_A\, P(1_A)\, d\mu = ||P(1_A)||^2 \geq (\mu(A))^2. \end{aligned}$$

Hence, for any left Følner sequence $\{\Phi_n\}$, and any $\epsilon > 0$, there exists $n \in \mathbf{N}$ and $g \in \Phi_n$ such that $\mu(A \cap T_g^{-1} A) \geq \big(\mu(A)\big)^2 - \epsilon$. We will use this fact in the next section. The following Lemma is motivated by van der Corput's fundamental inequality.

Lemma 4.2 *Suppose that $\{u_g : g \in G\}$ is a bounded set in a Hilbert space H, and that $\{\Phi_n\}$ is a left Følner sequence for G. If*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|^2} \Big(\limsup_{m\to\infty} \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \sum_{h,k\in\Phi_n} \langle u_{hg}, u_{kg} \rangle \Big) = 0,$$

then $\lim_{n\to\infty} ||\frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} u_g|| = 0$.

Proof. Let $\epsilon > 0$ be arbitrary. Fix n. We have

$$\frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} u_g = \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \frac{1}{|\Phi_n|} \sum_{h\in\Phi_n} u_{hg} + \Psi'_m = \Psi_m + \Psi'_m,$$

where $||\Psi'_m|| \to 0$ as $m \to \infty$. (Since n is fixed, $|\bigcap_{h\in\Phi_n} h\Phi_m|/|\Phi_m| \to 1$ as $m \to \infty$, and for $g \in \bigcap_{h\in\Phi_n} h\Phi_m$, u_g is represented $|\Phi_n|$ times in the sum on the right, namely as $\{u_{h(h^{-1}g)} : h \in \Phi_n\}$.) By the Cauchy-Schwartz inequality,

$$||\Psi_m||^2 \leq \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \Big|\Big| \frac{1}{|\Phi_n|} \sum_{h\in\Phi_n} u_{hg} \Big|\Big|^2 = \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \frac{1}{|\Phi_n|^2} \sum_{h,k\in\Phi_n} \langle u_{hg}, u_{kg} \rangle.$$

By choosing n big enough at the outset, we may have this last quantity as small as desired for large m, so that $||\Psi_m|| < \epsilon$ for large m. Since $||\Psi'_m|| \to 0$ and ϵ was arbitrary, the proof is complete.

□

Now suppose that $\{T_g\}$ and $\{S_g\}$ are measure preserving G-actions on a probability space $(X, \mathcal{A}, \mu)$ with $T_g S_h = S_h T_g$ for all $g, h \in G$, and let $\mathcal{B} \subset \mathcal{A}$ be the σ-algebra of $\{S_g\}$-invariant sets. Let $(Y, \mathcal{B}, \nu)$ be the factor determined by $\mathcal{B}$, and let $\pi : X \to Y$ be the factor map. Note that for a.e. $x \in X$, $T_g(\pi(x)) = \pi(T_g x)$ is well-defined for all $g \in G$ and forms a measure-preserving G-action of $(Y, \mathcal{B}, \nu)$. Furthermore, $(Y, \mathcal{B}, \nu, \{T_g\})$ is a common factor of $(X, \mathcal{A}, \mu, \{T_g\})$ and $(X, \mathcal{A}, \mu, \{S_g T_g\})$. We let $\{\mathcal{M}_i\}$, $\{\mathcal{N}_i\}$ be global orthogonal decompositions of $K(\mathcal{B}, \{T_g\})$ and $K(\mathcal{B}, \{S_g T_g\})$, respectively. We also choose global orthonormal bases $\{\varphi_1^i, \cdots, \varphi_{m(i)}^i\}$ for $\mathcal{M}_i$ and $\{\psi_1^i, \cdots \psi_{n(i)}^i\}$ for $\mathcal{N}_i$, $i \in \mathbf{N}$, in such a way that whenever $\{T_g \mid_{\mathcal{M}_i}\}$ is equivalent to $\{S_g T_g \mid_{\mathcal{N}_j}\}$, the matrix anti-cocycles generated by these with respect to the given bases are equal.

Theorem 4.3 *If $\varphi \perp K(\mathcal{B}, \{T_g\})$ or $\psi \perp K(\mathcal{B}, \{S_g T_g\})$, then*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x) = 0$$

in $L^1(X, \mathcal{A}, \mu)$.

Proof. Assume that $\varphi \perp K(\mathcal{B}, \{T_g\})$ (the other case is similar). We first prove that the result holds when $\psi \in L^\infty(X, \mathcal{A}, \mu)$. Let $u_g(x) = \varphi(T_g x)\psi(S_g T_g x) \in L^2(X, \mathcal{A}, \mu)$. For fixed $h, k \in G$, we have

$$\begin{aligned}
& \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \langle u_{hg}, u_{kg} \rangle \\
= \quad & \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \int \varphi(T_{hg}x)\psi(S_{hg}T_{hg}x)\overline{\varphi(T_{kg}x)\psi(S_{kg}T_{kg}x)}\, d\mu \\
= \quad & \frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} \int \left(\varphi(T_h x)\overline{\varphi(T_k x)}\right)\left(\psi(S_{hg}T_h x)\overline{\psi(S_{kg}T_k x)}\right) d\mu \\
= \quad & \int (\varphi(T_h x)\overline{\varphi(T_k x)}) \left(\frac{1}{|\Phi_m|} \sum_{g\in\Phi_m} S_h T_h \psi(S_g x)\overline{S_k T_k \psi(S_g x)} \right) d\mu \\
\to \quad & \int (T_h\varphi T_k\overline{\varphi})\mathbf{E}(S_h T_h \psi S_k T_k \overline{\psi} \mid \mathcal{B})\, d\mu \\
= \quad & \int \mathbf{E}(T_h\varphi T_k\overline{\varphi} \mid \mathcal{B})\mathbf{E}(S_h T_h \psi S_k T_k \overline{\psi} \mid \mathcal{B})\, d\mu \\
= \quad & \int (T_h \times S_h T_h)(\varphi \otimes \psi)\ (T_k \times S_k T_k)(\overline{\varphi \otimes \psi})\, d\mu \times_{\mathcal{B}} \mu = a_{h,k}.
\end{aligned}$$

By Theorem 4.1, we now have, since $\varphi \otimes \psi$ is orthogonal to the $\{T_g \times S_g T_g\}$-invariant functions,

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|^2} \sum_{h,k\in\Phi_n} a_{h,k} = \lim_{n\to\infty} \left|\left| \frac{1}{|\Phi_n|} \sum_{l\in\Phi_n} T_l \times S_l T_l(\varphi \otimes \psi) \right|\right| = 0.$$

By Lemma 4.2 we have

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x) = 0,$$

this limit in $L^2(X,\mathcal{A},\mu)$. Since $\mu(X) = 1 < \infty$, convergence will be in $L^1(X,\mathcal{A},\mu)$ as well. We can extend our result to $\psi \in L^2(X,\mathcal{A},\mu)$, approximating ψ (in L^2) by functions $f \in L^\infty(X,\mathcal{A},\mu)$ satisfying $f \perp K(\mathcal{B},\{T_g\})$. (Recall that $K(\mathcal{B},\{T_g\}) = L^2(X,\mathcal{E},\mu)$ for some $\{T_g\}$-invariant σ-algebra $\mathcal{E}$. Thus, for any $f \in L^\infty(X,\mathcal{A},\mu)$ which is close to ψ in $L^2(X,\mathcal{A},\mu)$, $f - \mathbf{E}(f \mid \mathcal{E}) \in L^\infty(X,\mathcal{A},\mu)$ will be orthogonal all $\mathcal{E}$-measurable functions and be at least as close to ψ in $L^2(X,\mathcal{A},\mu)$.)

□

Let $\mathbf{S}^k$ denote the unit sphere in $\mathbf{C}^k$. If $\xi^1,\cdots,\xi^k \in \mathbf{S}^k$, we will denote by $[\xi^1\cdots\xi^k]$ the $k\times k$ matrix whose columns are the vectors ξ^i. We use $|\cdot|_k$ to denote the norm on $\mathbf{C}^k$ and λ_k to denote normalized Lebesgue measure on $\mathbf{S}^k$ ($\lambda_k(\mathbf{S}^k) = 1$). ξ_i will denote the ith coordinate of a vector $\xi \in \mathbf{S}^k$.

Lemma 4.4 *For any $\delta < 1$ and $k \in \mathbf{N}$, there exists $\epsilon = \epsilon(\delta,k) > 0$ such that for any $E \subset \mathbf{S}^k$ with $\lambda_k(E) < \epsilon$, there exist $\xi^1,\cdots,\xi^k \in \mathbf{S}^k\backslash E$ with the property that for any $\xi \in \mathbf{C}^k$, $\left|[\xi^1\cdots\xi^k]^\tau\xi\right|_k \geq \delta|\xi|_k$.*

Proof. Suppose $\delta < 1$ and $k \in \mathbf{N}$ are given. Let $\{e_i : 1 \leq i \leq k\}$ be the natural coordinate basis for $\mathbf{S}^k$ (so that $I_k = [e_1\cdots e_k]$ is the $k\times k$ identity matrix). For some $\gamma > 0$, we will have that whenever $|\xi_i - e_i| < \gamma$, $1 \leq i \leq k$, $\left|[\xi_1\cdots\xi_k]^\tau\xi\right|_k > \delta|\xi|_k$. Clearly there is some $\epsilon > 0$ such that for any $E \subset \mathbf{S}^k$ with $\lambda_k(E) < \epsilon$, there exist $\xi_i \in \mathbf{S}^k \setminus E$ with $|\xi_i - e_i| < \gamma$, $1 \leq i \leq k$.

□

Theorem 4.5 *Suppose that $M(g,y) = [m_{is}(g,y)]$ is an anti-cocycle on $(Y,\mathcal{B},\nu,\{T_g\})$, and that $N(g,y) = [n_{jt}(g,y)]$ is an anti-cocycle on $(Y,\mathcal{B},\nu,\{T_g\})$. For any $a(y) \in L^\infty(Y,\mathcal{B},\nu)$, $1 \leq i,s \leq m$, and $1 \leq t,j \leq n$, let*

$$\varphi^{st}_{ijd}(y) = \frac{1}{|\Phi_d|}\sum_{g\in\Phi_d} a(T_g y)m_{is}(g,y)n_{jt}(g,y).$$

Then $\lim_{d\to\infty}\varphi^{st}_{ijd}$ exists in $L^1(Y,\mathcal{B},\nu)$.

Proof. Let $\mathcal{B}^m$, $\mathcal{B}^n$ be the Borel σ-algebras on $\mathbf{S}^m$ and $\mathbf{S}^n$ respectively. Let

$$\overline{X} = Y\times\mathbf{S}^m\times\mathbf{S}^n,\ \overline{\mathcal{B}} = \mathcal{B}\otimes\mathcal{B}^m\otimes\mathcal{B}^n,\ \text{and}\ \overline{\mu} = \nu\times\lambda^m\times\lambda^n.$$

Define measure-preserving transformations $\{R_g : g \in G\}$ on $\overline{X}$ by

$$R_g(y,\xi,\eta) = (T_g y, M(g,y)\xi, N(g,y)\eta).$$

One may check that $R_{gh} = R_g R_h$, that is, $\{R_g : g \in G\}$ is a G-action. Suppose now that $a(y) \in L^\infty(Y, \mathcal{B}, \nu)$. For $1 \le i \le m$, $1 \le j \le n$, define $\xi_{ij} \in L^\infty(\overline{X}, \overline{\mathcal{B}}, \overline{\mu})$ by $\xi_{ij}(y, \xi, \eta) = a(y)\xi_i \eta_j$. Then

$$\begin{aligned}\xi_{ij}(R_g(y,\xi,\eta)) &= a(T_g y)[M(g,y)\xi]_i [N(g,y)\eta]_j \\ &= a(T_g y)\sum_{s,t} m_{is}(g,y) n_{jt}(g,y)\xi_s \eta_t.\end{aligned}$$

Hence, by Theorem 4.1,

$$\lim_{d\to\infty} \sum_{s,t} \varphi_{ijd}^{st}(y)\xi_s\eta_t = \lim_{d\to\infty} \frac{1}{|\Phi_d|} \sum_{g\in\Phi_d} \xi_{ij}(R_g(y,\xi,\eta))$$

exists in $L^1(\overline{X}, \overline{\mathcal{B}}, \overline{\mu})$. Let $\varepsilon = \varepsilon(\delta, m)$ (see Lemma 4.4). Let D_0 be so large that whenever $d, d' > D_0$, we have

$$\int \left| \sum_{s,t} (\varphi_{ijd}^{st}(y) - \varphi_{ijd'}^{st}(y))\xi_s\eta_t \right| \, d\overline{\mu}(y,\xi,\eta) < \frac{\varepsilon^2}{m}.$$

Let φ_{ijd} be the $m \times n$ matrix $[\varphi_{ijd}^{st}(y)]_{1\le s\le m,\ 1\le t\le n}$. We have

$$\int |\xi^\tau(\varphi_{ijd}(y) - \varphi_{ijd'}(y))\eta| \, d\overline{\mu}(y,\xi,\eta) < \frac{\varepsilon^2}{m},$$

which implies that

$$\lambda_m\left(\left\{\xi : \int\int |\xi^\tau(\varphi_{ijd}(y) - \varphi_{ijd'}(y))\eta| \, d\nu(y)d\lambda_n(\eta) > \frac{\varepsilon}{m}\right\}\right) < \varepsilon.$$

By choice of ϵ, there exist column vectors $\xi^1, \cdots, \xi^m \in \mathbf{S}^m$ such that

(i) $\int\int \sum_{r=1}^m |(\xi^r)^\tau(\varphi_{ijd}(y) - \varphi_{ijd'}(y))\eta| \, d\nu(y)d\lambda_n(\eta) < \frac{\varepsilon}{m}$; $1 \le r \le m$

(ii) $\left|[\xi^1 \cdots \xi^m]^\tau \xi\right|_m \ge \frac{1}{2}|\xi|_m$ for all $\xi \in \mathbf{C}^m$.

Since all norms on a finite dimensional space are equivalent, there exist numbers K_1 and K_2 such that

$$K_1 \left(\sum_{r=1}^m |a_i|^2\right)^{1/2} \ge \sum_{r=1}^m |a_i| \ge K_2 \left(\sum_{r=1}^m |a_i|^2\right)^{1/2}.$$

Therefore,

$$\begin{aligned}\varepsilon &> \int\int \sum_{r=1}^m |(\xi^r)^\tau(\varphi_{ijd}(y) - \varphi_{ijd'}(y))\eta| \, d\nu(y)d\lambda_n(\eta) \\ &\ge K_2 \int\int \left|[\xi^1 \cdots \xi^m]^\tau(\varphi_{ijd}(y) - \varphi_{ijd'}(y))\eta\right|_m d\nu(y)d\lambda_n(\eta)\end{aligned}$$

$$\begin{aligned}
&\geq \frac{K_2}{2}\int\int |(\varphi_{ijd}(y)-\varphi_{ijd'}(y))\eta|_m \; d\nu(y)d\lambda_n(\eta)\\
&\geq \frac{K_2}{2K_1}\int\int\sum_{s=1}^{m}|[(\varphi_{ijd}(y)-\varphi_{ijd'}(y))\eta]_s| \; d\nu(y)d\lambda_n(\eta)\\
&= \frac{K_2}{2K_1}\sum_{s=1}^{m}\int\int\left|\sum_{t=1}^{n}(\varphi_{ijd}^{st}(y)-\varphi_{ijd'}^{st}(y))\eta_t\right| \; d\nu(y)d\lambda_n(\eta),
\end{aligned}$$

which implies that $\sum_{t=1}^{n}\varphi_{ijd}^{st}(y)\eta_t$ converges in $L^1(Y\times \mathbf{S}^n, \mathcal{B}\otimes\mathcal{B}^n, \nu\times\lambda_n)$, $1\leq s\leq m$. Another application of Lemma 4.4, with $k=n$, gives by the same steps convergence of $\varphi_{ijd}^{st}(y)$ in $L^1(Y,\mathcal{B},\nu)$, $1\leq s\leq m$, $1\leq t\leq n$.

□

We now suppose $\mathcal{M}=\mathcal{M}_{n_1}\in\{\mathcal{M}_i\}$ and $\mathcal{N}=\mathcal{N}_{n_2}\in\{\mathcal{N}_j\}$ are elements of the global orthogonal decompositions having global orthonormal bases $\{\varphi_1,\cdots,\varphi_m\}$ and $\{\psi_1,\cdots,\psi_n\}$ (we have suppressed the superscripts for convenience). Let $[m_{is}(g,y)]$ and $[n_{jt}(g,y)]$ be the induced anti-cocycles. Recall that if $\{T_g\mid_{\mathcal{M}}\}$ is equivalent to $\{S_gT_g\mid_{\mathcal{M}}\}$, then $m=n$ and the induced anti-cocyles are equal. In this case, we let

$$F_{n_1,n_2}(x,z)=F_{\mathcal{M},\mathcal{N}}(x,z)=\sum_{k=1}^{n}(\varphi_k\otimes\overline{\psi_k})(x,z). \tag{8}$$

If $\{T_g\mid_{\mathcal{M}}\}$ is not equivalent to $\{S_gT_g\mid_{\mathcal{M}}\}$, we will let $F_{\mathcal{M},\mathcal{N}}=0$. We now have

Proposition 4.6 *If $(Y,\mathcal{B},\nu,\{T_g\})$ is ergodic, then for any $\varphi\in\mathcal{M}$ and any $\psi\in\mathcal{N}$,*

$$\lim_{d\to\infty}\frac{1}{|\Phi_d|}\sum_{g\in\Phi_d}\varphi(T_gx)\overline{\psi(S_gT_gx)}=\left(\int(\varphi\otimes\overline{\psi})\;\overline{F_{\mathcal{M},\mathcal{N}}}\;d\mu_y\times\mu_y\right)\;F_{\mathcal{M},\mathcal{N}}(x,x)$$

in $L^1(X,\mathcal{A},\mu)$. (Here $y=\pi(x)$.)

Proof. Suppose first that $\varphi(x)=a(x)\varphi_i(x)$, $\psi(x)=b(x)\psi_j(x)$, where $a(x)$, $b(x)\in L^\infty(X,\mathcal{B},\mu)$. Then

$$\varphi(T_gx)\overline{\psi(S_gT_gx)}=a(T_gx)\overline{b(S_gT_gx)}T_g\varphi_i(x)\overline{\psi_j(S_gT_gx)}.$$

Since b is $\{S_g\}$-invariant, we may write $a(T_gx)\overline{b(S_gT_gx)}=a\overline{b}(T_gx)$. Setting

$$\varphi_{ijd}^{st}(x)=\frac{1}{\Phi_d}\sum_{g\in\Phi_d}a\overline{b}(T_gx)m_{is}(g,x)n_{jt}(g,x),$$

$\{\varphi_{ijd}^{st}\;:\;d\in\mathbf{N}\}$ will be uniformly bounded in $L^\infty(X,\mathcal{B},\mu)$, and, for fixed i,j,s,t will converge, by Theorem 4.5, to a function φ_{ij}^{st} as $d\to\infty$, hence

$$\lim_{d\to\infty}\frac{1}{|\Phi_d|}\sum_{g\in\Phi_d}\varphi(T_gx)\overline{\psi(S_gT_gx)}$$

$$
\begin{aligned}
&= \lim_{d\to\infty} \sum_{s,t} \left(\frac{1}{|\Phi_d|} \sum_{g\in\Phi_d} a\overline{b}(T_g x) m_{is}(g,x) n_{jt}(g,x) \right) \varphi_s(x)\overline{\psi_t(x)} \\
&= \lim_{d\to\infty} \sum_{s,t} \varphi_{ijd}^{st}(x)\varphi_s(x)\overline{\psi_t(x)} = \sum_{s,t} \varphi_{ij}^{st}(x)\varphi_s(x)\overline{\psi_t(x)}.
\end{aligned}
$$

Also, the projection of $\varphi\otimes\overline{\psi}$ onto the $\{T_g \times S_g T_g\}$-invariant functions is just the projection onto the zero- or one-dimensional subspace generated by $F_{\mathcal{M},\mathcal{N}}$ and we have

$$
\begin{aligned}
&\left(\int (\varphi\otimes\overline{\psi})\ \overline{F_{\mathcal{M},\mathcal{N}}}\ d\mu \times_{\mathcal{B}} \mu \right) F_{\mathcal{M},\mathcal{N}} \\
&= \lim_{d\to\infty} \frac{1}{|\Phi_d|} \sum_{g\in\Phi_d} (T_g \times S_g T_g)\varphi\otimes\overline{\psi} \\
&= \lim_{d\to\infty} \sum_{s,t} \left(\frac{1}{|\Phi_d|} \sum_{g\in\Phi_d} a\overline{b}(T_g y) m_{is}(g,y) n_{jt}(g,y) \right) \varphi_s \otimes \overline{\psi_t} \\
&= \sum_{s,t} \varphi_{ij}^{st}\ \varphi_s \otimes \overline{\psi_t}.
\end{aligned}
$$

Hence, if $\{T_g\ |_{\mathcal{M}}\}$ is not equivalent to $\{S_g T_g\ |_{\mathcal{M}}\}$, then $F_{\mathcal{M},\mathcal{N}} = 0$ and $\varphi_{ij}^{st} = 0$ for all i,j,s,t. This completes the proof in this case. If $\{T_g\ |_{\mathcal{M}}\}$ is equivalent to $\{S_g T_g\ |_{\mathcal{M}}\}$, then the above gives $\varphi_{ij}^{st} = \left(\int a\overline{b}(y)\ d\nu(y)\right)\delta_{ij}\delta_{st}$, hence

$$
\begin{aligned}
\lim_{d\to\infty} \frac{1}{|\Phi_d|} \sum_{g\in\Phi_d} \varphi(T_g x)\overline{\psi(S_g T_g x)} &= \delta_{i,j} \sum_{k=1}^{n} \left(\int a\overline{b}(y)\ d\nu(y) \right) \varphi_k(x)\overline{\psi(x)} \\
&= \left(\int (\varphi\otimes\overline{\psi}) \overline{F_{\mathcal{M},\mathcal{N}}}\ d\mu_y \times \mu_y \right) F_{\mathcal{M},\mathcal{N}}(x,x).
\end{aligned}
$$

This completes the proof for the case

$$\varphi(x) = a(x)\varphi_i(x),\ \psi(x) = b(x)\psi_j(x),\ a,b \in L^\infty(X,\mathcal{B},\mu).$$

It is clear that the result will therefore hold for $\varphi = \sum_i a_i\varphi_i$, $\psi = \sum_j b_j\psi_j$, where $a_i, b_i \in L^\infty(X,\mathcal{B},\mu)$. This is sufficient for the general case, as such functions φ are dense in $\mathcal{M}$ (under the L^2 norm) and such functions ψ are dense in $\mathcal{N}$.

□

We now have the following.

Theorem 4.7 *Suppose that G is a countable amenable group with a left Følner sequence $\{\Phi_n\}$, and that $\{T_g\}$, $\{S_g\}$ are measure preserving G-actions on a measure space $(X,\mathcal{A},\mu)$ which commute in the sense that $T_g S_h = S_h T_g$ for all $g,\ h \in G$. Let $\mathcal{B} \subset \mathcal{A}$ be the σ-algebra of all $\{S_g\}$-invariant sets and assume that the natural factor $(Y,\mathcal{B},\nu,T_g)$ is ergodic. Let $\{\mathcal{M}_k\}$, $\{\mathcal{N}_k\}$ be global orthogonal decompositions of $K(\mathcal{B},\{T_g\})$ and $K(\mathcal{B},\{S_g T_g\})$, respectively, and* *choose orthonormal bases* $\{\varphi_1^i, \cdots, \varphi_{m(i)}^i\}$ *for* $\mathcal{M}_i$*, and* $\{\psi_1^i, \cdots, \psi_{n(i)}^i\}$ *for* $\mathcal{N}_i$*, such that the matrix anti-cocycles for* $\{T_g \mid_{\mathcal{M}_i}\}$ *and for* $\{S_g T_g \mid_{\mathcal{N}_j}\}$ *with respect to these bases are identical whenever* $\{T_g \mid_{\mathcal{M}_i}\}$ *is equivalent to* $\{S_g T_g \mid_{\mathcal{N}_j}\}$*. Let* $F_{ij} \in \mathcal{M}_i \otimes \mathcal{N}_j^*$ *be given by*

$$F_{ij}(x,z) = \sum_{k=1}^{m(i)} \varphi_k^i \otimes \overline{\psi_k^j}(x,z)$$

if $\{T_g \mid_{\mathcal{M}_i}\}$ *is equivalent to* $\{S_g T_g \mid_{\mathcal{N}_j}\}$*, and zero otherwise. Let* $\pi : X \to Y$ *be the natural projection. Then for any* φ*,* $\psi \in L^2(X, \mathcal{A}, \mu)$*,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x) = \sum_{i,j} \left(\int (\varphi \otimes \psi)\ \overline{F_{ij}}\ d\mu_y \times \mu_y \right) F_{ij}(x,x)$$

in $L^1(X, \mathcal{A}, \mu)$*. (Where* $y{=}\pi(x)$*.)*

Proof. We can approximate φ arbitrarily closely in $L^2(X, \mathcal{A}, \mu)$ by a function of the form $\varphi_0 + \varphi_1 + \cdots + \varphi_m$, where $\varphi_0 \perp K(\mathcal{B}, \{T_g\})$ and $\varphi_i \in \mathcal{M}_i$, $1 \le i \le m$, and a similar statement is true of ψ. Therefore the result follows from Proposition 4.6 and Theorem 4.3.

□

Remark. Since $\mathcal{B}$ consists of the $\{S_g\}$-invariant sets in $\mathcal{A}$, $(Y, \mathcal{B}, \nu, \{T_g\})$ will be ergodic if and only if there are no non-trivial measurable sets which are both $\{T_g\}$- and $\{S_g\}$-invariant, that is, when $\{T_g S_h : (g,h) \in G^2\}$ is an ergodic G^2-action on $(X, \mathcal{A}, \mu)$.

We now come to our main result.

Theorem 4.8 *Suppose that* G *is a countable amenable group and* $\{\Phi_n\}$ *is a left Følner sequence for* G*. Suppose that* $(X, \mathcal{A}, \mu)$ *is a probability space, and that* $\{T_g\}$ *and* $\{S_g\}$ *are measure preserving* G*-actions on* X *with* $T_g S_h = S_h T_g$ *for all* g*,* $h \in G$*. Then for any* φ*,* $\psi \in L^2(X, \mathcal{A}, \mu)$*,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x)$$

exists in $L^1(X, \mathcal{A}, \mu)$*.*

Proof. By Theorem 4.7 and the Remark above, the limit in question exists whenever $(\{T_g S_h : (g,h) \in G^2\})$ is an ergodic G^2-action on $(X, \mathcal{A}, \mu)$. For the general case, let $\mathcal{C} \subset \mathcal{A}$ be the σ-algebra of sets which are both $\{T_g\}$- and $\{S_g\}$-invariant and denote the factor associated with $\mathcal{C}$ by $(Z, \mathcal{C}, \gamma)$. Denote by $\pi : X \to Z$ the natural projection, and by $\{\mu_z : z \in Z\}$ the decomposition of μ with respect to Z. Then for almost every $z \in Z$, the measure-preserving G^2-action $\{T_g S_h\}$ on $(X, \mathcal{A}, \mu_z)$ is ergodic. Suppose φ, $\psi \in L^2(X, \mathcal{A}, \mu)$. Then $\varphi_z, \psi_z \in L^2(X, \mathcal{A}, \mu_z)$ a.e. We will assume that $||\varphi||_z$, $||\psi||_z$ are bounded by some number N. This assumption is without loss of generality since such functions are dense in $L^2(X, \mathcal{A}, \mu)$. For a.e. z we have existence of the limit

$$f_z(x) = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g x)$$

in $L^1(X, \mathcal{A}, \mu_z)$. Clearly $\{f_z : z \in Z\}$ is a measurable family, and $\int |f_z|\, d\mu_z \leq N^2$ a.e. Let $f = \int_{\oplus} f_z\, d\gamma(z)$. Then, by the dominated convergence theorem,

$$\lim_{n\to\infty} \int \left| \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi(S_g T_g) - f(x) \right| d\mu(x)$$

$$= \lim_{n\to\infty} \int\int \left| \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \varphi(T_g x)\psi_(S_g T_g x) - f_z(x) \right| d\mu_z d\gamma(z) \to 0.$$

□

5 Positivity of the Limit and Multiple Recurrence

As before, we assume that $\{T_g\}$ and $\{S_g\}$ are commuting measure preserving G-actions on $(X, \mathcal{A}, \mu)$, and we let $\mathcal{B} \subset \mathcal{A}$ be the sub-σ-algebra of all $\{S_g\}$-invariant sets. Then there are sub-σ-algebras $\mathcal{B}_T$ and $\mathcal{B}_{ST}$ such that

$$L^2(X, \mathcal{B}_T, \mu) = K(\mathcal{B}, \{T_g\}) \text{ and } L^2(X, \mathcal{B}_{ST}, \mu) = K(\mathcal{B}, \{S_g T_g\}).$$

We will let $\mathbf{P}_1(f) = \mathbf{E}(f \mid \mathcal{B}_T)$, $\mathbf{P}_2(f) = \mathbf{E}(f \mid \mathcal{B}_{ST})$ for $f \in L^2(X, \mathcal{A}, \mu)$.

Definition 5.1 *Suppose $\{\Phi_n\}$ is a left or right Følner sequence of G. For any $\Gamma \subset G$, define*

$$\overline{d}(\Gamma) = \limsup_{n\to\infty} \frac{|\Gamma \cap \Phi_n|}{|\Phi_n|} \text{ and } \underline{d}(\Gamma) = \liminf_{n\to\infty} \frac{|\Gamma \cap \Phi_n|}{|\Phi_n|}.$$

$\overline{d}(G)$ and $\underline{d}(G)$ will be called the upper and lower density of Γ with respect to $\{\Phi_n\}$.

Definition 5.2 *Suppose G is a discrete group. A subset $A \subset G$ will be called left (right) syndetic if there exists a finite set $K \subset G$ such that $\bigcup_{g\in K} gA = G$ ($\bigcup_{g\in K} Ag = G$).*

With respect to any left Følner sequence $\{\Phi_n\}$, $\underline{d}(A) > 0$ for any left syndetic set $A \subset G$. (Otherwise, passing to a sub-Følner sequence $\{\Phi_{n_k}\}$, we may assume that $\overline{d}(A) = 0$. It is clear from the definitions that $\overline{d}(C \cup B) \leq \overline{d}(C) + \overline{d}(B)$, and that $\overline{d}(gC) = \overline{d}(C)$ for all $B, C \subset G$ and $g \in G$. We have $G = g_1 A \cup \cdots \cup g_k A$, so that $\overline{d}(G) \leq k\overline{d}(A) = 0$, a contradiction.) Conversely, if $A \subset G$ has the property that for any left Følner sequence $\{\Phi_n\}$, there exists $n \in \mathbf{N}$ such that $A \cap \Phi_n \neq \phi$, then A is necessarily left syndetic. (Suppose A is not left syndetic. Let $\{\Phi_n\}$ be any left Følner sequence. For any sequence $(h_n)_{n=1}^{\infty} \subset G$, $\{\Phi_n h_n\}$ will be a left Følner sequence as well, since for any $g \in G$, $n \in \mathbf{N}$, we have $|g(\Phi_n h_n) \cap \Phi_n h_n| = |(g\Phi_n \cap \Phi_n)h_n| = |g\Phi_n \cap \Phi_n|$. Since A is not left syndetic, for every $n \in \mathbf{N}$ we have $G \neq \Phi_n^{-1} A$. Letting $h_n \in G \setminus \Phi_n^{-1} A$, we have $A \cap \Phi_n h_n = \phi$ for all $n \in \mathbf{N}$.)

The proof of the following lemma makes use of the fact that if $\{\Phi_n\}$ is a left Følner sequence and $g_1, \cdots, g_k \in G$, then setting $\Phi'_n = \Phi_n \cap g_1\Phi_n \cap \cdots \cap g_k\Phi_n$, $\{\Phi'_n\}$ will be a left Følner sequence as well. (Simply note that if $|g_i\Phi_n \cap \Phi_n|, |hg_i\Phi_n \cap \Phi_n| \geq (1-\epsilon)|\Phi_n|$,

$1 \le i \le k$, and $|h\Phi_n \cap \Phi_n| \ge (1-\epsilon)|\Phi_n|$, then $|h\Phi'_n \cap \Phi'_n| \ge (1-(2k+1)\epsilon)|\Phi_n$.) Also, in order to formulate the lemma in its full strength, we introduce the following notation. For any (finite or infinite) sequence $(g_i) \subset G$, we let

$$FP\big((g_i)\big) = \{g_{n_1} g_{n_2} \cdots g_{n_k} : n_1 < n_2 < \cdots < n_k\}.$$

Lemma 5.1 *Suppose that $\{T_g\}$ is a measure-preserving G-action of a countable amenable group G on a probability space $(X, \mathcal{A}, \mu)$, and that $\{\Phi_n\}$ is a left Følner sequence for G. For any set $A \in \mathcal{A}$, $\mu(A) > 0$, and any δ, $0 < \delta < 1$, there exists a sequence $(h_i)_{i=1}^{\infty} \subset G$ with the property that for every $h \in FP\big((h_i)_{i=1}^{\infty}\big)$, $h \in \Phi_n$ for some $n \in \mathbf{N}$, and such that for all $j \in \mathbf{N}$,*

$$\mu\left(\bigcap_{h \in \{e\} \cup FP\left((h_i)_{i=1}^{j}\right)} T_h^{-1} A\right) > \Big(\delta\mu(A)\Big)^{2^j}.$$

Proof. Choose $n_1 \in \mathbf{N}$ and $h_1 \in \Phi_{n_1}$ such that $\mu(A \cap T_{h_1}^{-1} A) > \delta\big(\mu(A)\big)^2$ (see the discussion after Theorem 4.1). Put $A_1 = (A \cap T_{h_1}^{-1} A)$. Choose $n_2 \in \mathbf{N}$ and $h_2 \in (\Phi_{n_2} \cap h_1^{-1}\Phi_{n_2})$ such that $\mu(A_1 \cap T_{h_2}^{-1} A_1) > \delta\big(\mu(A_1)\big)^2$. Put $A_2 = (A_1 \cap T_{h_2}^{-1} A_1)$. Then

$$\mu(A \cap T_{h_1}^{-1} A \cap T_{h_2}^{-1} A \cap T_{h_1 h_2}^{-1} A) > \delta^3\big(\mu(A)\big)^4.$$

Choose $n_3 \in \mathbf{N}$ and

$$h_3 \in (\Phi_{n_3} \cap h_1^{-1}\Phi_{n_3} \cap h_2^{-1}\Phi_{n_3} \cap (h_1 h_2)^{-1}\Phi_{n_3})$$

such that $\mu(A_2 \cap T_{h_3}^{-1} A_2) > \delta\big(\mu(A_2)\big)^2$. Continuing in this fashion, we get what we need. □

Theorem 5.2 *Suppose that G is a countable amenable group with a left Følner sequence $\{\Phi_n\}$. Suppose $\{S_g\}$, $\{T_g\}$ are measure preserving G-actions on a probability space $(X, \mathcal{A}, \mu)$ with $T_g S_h = S_h T_g$ for any g, $h \in G$. Then for any set $A \in \mathcal{A}$, $\mu(A) > 0$,*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} \mu(A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A) > 0.$$

Proof. Let $f = 1_A$, $f_1 = \mathbf{P}_1 f$, and $f_2 = \mathbf{P}_2 f$. (These projections were introduced at the beginning of this section.) By Theorems 4.3 and 4.8, we know that

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} \mu(A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A) = \lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} \int f\ T_g f_1\ S_g T_g f_2\ d\mu.$$

(In particular, this limit exists.) We will show that

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g \in \Phi_n} \int f\ T_g f_1\ S_g T_g f_2\ d\mu > 0.$$

Recall that $\mathbf{P}_1$ is an orthogonal projection. Since

$$\begin{aligned}\int f 1_{\{x:f_1(x)=0\}}\, d\mu &= \int \mathbf{P}_1\left(f 1_{\{x:f_1(x)=0\}}\right) d\mu \\ &= \int f_1 1_{\{x:f_1(x)=0\}}\, d\mu = 0,\end{aligned}$$

$f_1(x) \neq 0$ for a.e. $x \in A$. Similarly $f_2(x) > 0$ for a.e. $x \in A$. Therefore, there exists some $\alpha > 0$, and a set $A' \subset A$, $\mu(A') > 0$, such that $f_1(x) f_2(x) > \alpha$ for all $x \in A'$. Furthermore, there exists $\beta > 0$, and a set $B \subset Y$, $\nu(B) = 2\xi > 0$, such that for all $y \in B$, $\mu_y(A') > \beta$. It follows that $\int f f_1 f_2\, d\mu_y > \alpha\beta$ for all $y \in B$. Notice that ξ, α, and β do not depend on the Følner sequence $\{\Phi_n\}$.

Let $\epsilon = \frac{\alpha\beta}{16}$. We may approximate f_1 by a function ϕ_1 which is contained in a finite-dimensional $\{T_g\}$-invariant $\mathcal{B}$-module $\mathcal{M}$. Likewise, we may approximate f_2 by a function ϕ_2 which is contained in a finite-dimensional $\{T_g S_g\}$-invariant $\mathcal{B}$-module $\mathcal{N}$. We make these approximations so close that there exists a set $B' \subset B$, $\nu(B') > \xi$, such that for all $y \in B'$, $||f_1 - \phi_1||_y < \epsilon$ and $||f_2 - \phi_2||_y < \epsilon$. It is not difficult to see, since $\mathcal{M}$ and $\mathcal{N}$ are finite-dimensional, that there exists a finite family of functions $h_1, \cdots, h_l \in L^2(X, \mathcal{A}, \mu)$ having the property that for a.e. $y \in Y$ and all $g \in G$, there exist $k_1 = k_1(g, y)$ and $k_2 = k_2(g, y)$ such that $||T_g\phi_1 - h_{k_1}||_y < \epsilon$ and $||S_g T_g \phi_2 - h_{k_2}||_y < \epsilon$.

Let $M = l^2 + 1$. By Lemma 5.1, there exists $\gamma > 0$, depending only on $\mu(B')$ and M, and $g_1, \cdots, g_M \in G$, with $\mu(B' \cap T_{g_1}^{-1} B' \cap \cdots \cap T_{g_M}^{-1} B') > \gamma$, such that $g_j g_i^{-1} \in \Phi_n$ for some $n \in \mathbf{N}$ whenever $1 \leq i < j \leq M$. (Let $g_i = h_1 h_2 \cdots h_i$, where the h_i's are as in that lemma.) Fix $y \in B' \cap T_{g_1}^{-1} B' \cap \cdots \cap T_{g_M}^{-1} B'$. There exist numbers $i = i(y)$, $j = j(y)$, $1 \leq i < j \leq M$, such that $k_1(g_i, y) = k_1(g_j, y)$ and $k_2(g_i, y) = k_2(g_j, y)$, so that

$$||T_{g_i}\phi_1 - T_{g_j}\phi_1||_y < 2\epsilon \text{ and } ||S_{g_i} T_{g_i}\phi_2 - S_{g_j} T_{g_j}\phi_2||_y < 2\epsilon.$$

We now have $||\phi_1 - T_{g_j g_i^{-1}}\phi_1||_{T_{g_i} y} < 2\epsilon$. Also, since $T_{g_i} y \in B'$, we have $||\phi_1 - f_1||_{T_{g_i} y} < \epsilon$. On the other hand, since $T_{g_j} y \in B'$, we have

$$||T_{g_j g_i^{-1}} f_1 - T_{g_j g_i^{-1}}\phi_1||_{T_{g_i} y} = ||T_{g_j} f_1 - T_{g_j}\phi_1||_y = ||f_1 - \phi_1||_{T_{g_j} y} < \epsilon.$$

This, finally, gives $||f_1 - T_{g_j g_i^{-1}} f_1||_{T_{g_i} y} < 4\epsilon$. Similarly, $||f_2 - S_{g_j g_i^{-1}} T_{g_j g_i^{-1}} f_2||_{T_{g_i} y} < 4\epsilon$. Let $g = g(y) = g_{j(y)} g_{i(y)}^{-1}$. It follows that, for $y \in B' \cap T_{g_1}^{-1} \cap \cdots \cap T_{g_M}^{-1} B'$,

$$\int f\, T_g f_1\, S_g T_g f_2\, d\mu_y > \int f f_1 f_2\, d\mu_y - 8\epsilon > \frac{\alpha\beta}{2}.$$

Let $C \subset B' \cap T_{g_1}^{-1} B' \cap \cdots \cap T_{g_M}^{-1} B'$ satisfy $\nu(C) > \frac{\gamma}{M^2}$ and have the property that $g_0 = g(y)$ is constant on C. Then

$$\int f\, T_{g_0} f_1\, S_{g_0} T_{g_0} f_2\, d\mu > \frac{\alpha\beta\gamma}{2M^2} > 0.$$

Recall that $g_0 \in \Phi_n$ for some $n \in \mathbf{N}$. Let

$$\Gamma = \left\{ g : \int f\ T_g f_1\ S_g T_g f_2\ d\mu > \frac{\alpha\beta\gamma}{2M^2} \right\}.$$

Note that f_1 and f_2 do not depend on the arbitrary left Følner sequence $\{\Phi_n\}$, and we have found $n \in \mathbf{N}$ and $g_0 \in \Phi_n \cap \Gamma$. It follows that Γ is left syndetic, so that $\overline{d}(\Gamma) > 0$ with respect to every left Følner sequence (in particular, with respect to $\{\Phi_n\}$), and

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} \int f\ T_g f_1\ S_g T_g f_2\ d\mu \geq \overline{d}(\Gamma) \frac{\alpha\beta\gamma}{2M^2} > 0.$$

□

We now have the following corollaries.

Corollary 5.3 *Under the conditions of Theorem 5.2, for any $f \in L^2(X, \mathcal{A}, \mu)$ with $f \geq 0$ and $f \not\equiv 0$, we have*

$$\lim_{n\to\infty} \frac{1}{|\Phi_n|} \sum_{g\in\Phi_n} T_g f\ S_g T_g f \not\equiv 0.$$

Corollary 5.4 *Under the conditions of Theorem 5.2, for any set $A \in \mathcal{A}$ with $\mu(A) > 0$, there exists $\gamma > 0$ such that the set*

$$R_\gamma = \{g : \mu(A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A) > \gamma\}$$

is both left and right syndetic.

Proof. Suppose that for every $k \in \mathbf{N}$, $R_{1/k}$ is not left syndetic. Then there exist left Følner sequences $\{\Phi_n^{(k)}\}$ such that $R_{1/k} \cap \Phi_n^{(k)} = \phi$ for all $n, k \in \mathbf{N}$. For some sequence $n_k \to \infty$, $\{\Phi'_k\} = \{\Phi_{n_k}^{(k)}\}$ will be a left Følner sequence. For this sequence we will have

$$\lim_{k\to\infty} \frac{1}{|\Phi'_k|} \sum_{g\in\Phi'_k} \mu(A \cap T_g^{-1} A \cap (T_g S_g)^{-1} A) = 0,$$

contradicting Theorem 5.2. Therefore, R_γ is left syndetic for some $\gamma > 0$.

By the same token, we have some $\alpha > 0$ such that the set

$$S_\alpha = \{g :\ \mu(A \cap S_g^{-1} A \cap (S_g T_g)^{-1} A) > \alpha\}$$

is left syndetic. One may check that $R_\alpha = S_\alpha^{-1}$, so that R_α is right syndetic.

□

6 Applications to Combinatorics

In this section, we will use the recurrence theorem of §5 and an amenable group version of Furstenberg's correspondence principle to prove the following combinatorial theorem, which is a density version, for amenable groups, of a Ramsey-theoretic result for general groups proved in [1] (for Furstenberg's correspondence principle, see [8, page 73]).

Theorem 6.1 *Suppose that G is a countable amenable group and that $E \subset G \times G$ has positive upper density with respect to a left Følner sequence $\{\Phi_n\}$ for $G \times G$. Then the set*

$$\{g \in G : \text{ there exists } (a, b) \in G \times G \text{ such that } \{(a, b), (ga, b), (ga, gb)\} \subset E\}$$

is both left and right syndetic in G.

Let us start with an amenable group version of Furstenberg's correspondence principle. Let $\Omega = \{0, 1\}^{G\times G}$. With the product topology, Ω is a compact metrizable space. Elements of Ω are given by 1_E for subsets $E \subset G \times G$. Commuting G-actions $\{T_g\}$ and $\{S_g\}$ may be defined on Ω by $(S_g\xi)(g_1, g_2) = \xi(g_1, g_2g)$ and $(T_g\xi)(g_1, g_2) = \xi(g_1g, g_2)$. One then has the following:

Proposition 6.2 *Suppose that G is a countable amenable group and $\{\Phi_n\}$ is a left Følner sequence for $G \times G$. Suppose $E \subset G \times G$. Let $X = \overline{\{T_gS_h1_E : g, h \in G\}}$. If*

$$\overline{d}(E) = \limsup_{n\to\infty} \frac{|E \cap \Phi_n|}{|\Phi_n|} > 0,$$

then there exists a $\{T_g\}$- and $\{S_g\}$-invariant probability measure μ on X such that

$$\mu(\{\eta \in X : \eta(e, e) = 1\}) > 0.$$

Proof. Let $A = \{\eta \in X : \eta(e, e) = 1\}$. A is closed and open in X, so 1_A is continuous. There exists a sub-sequence $\{\Phi_{n_k}\}$ such that

$$\lim_{k\to\infty} \frac{|\Phi_{n_k} \cap E|}{|\Phi_{n_k}|} = \overline{d}(E)$$

exists. Then

$$\lim_{k\to\infty} \frac{1}{|\Phi_{n_k}|} \sum_{(g_1,g_2)\in\Phi_{n_k}} 1_A(S_{g_2}T_{g_1}1_E) = \overline{d}(E).$$

Since $C(X)$ is separable, one can choose a countable dense set $\Im \subset C(X)$ with $1_A \in \Im$ and a further subsequence (still written $\{\Phi_{n_k}\}$) such that the limit

$$\lim_{k\to\infty} \frac{1}{|\Phi_{n_k}|} \sum_{(g_1,g_2)\in\Phi_{n_k}} f(S_{g_1}T_{g_2}1_E) = I(f)$$

exists for all $f \in \Im$. One may then check that I is a bounded, positive linear functional and is therefore given by integration against a positive finite measure μ on X. It is clear that $\mu(A) = \overline{d}(E) > 0$ and that μ is $\{S_g\}$- and $\{T_g\}$-invariant.

□

Proof of Theorem 6.1. Let $\xi = 1_E$ and $X = \overline{\{T_g S_h \xi : \ g, h \in G\}}$. Let μ be an $\{S_g\}$- and $\{T_g\}$-invariant measure on X with $\mu(A) > 0$, where $A = \{\eta \in X : \eta(e,e) = 1\}$. By Corollary 5.4, the set

$$R = \{g \in G : \ \mu(A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A) > 0\}$$

is both left and right syndetic. For $g \in R$, choose $\xi' \in A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A$. Since A is open and $\xi' \in \overline{\{S_g T_h \xi : \ g, h \in G\}}$, there exist $a, b \in G$ such that $S_b T_a \xi \in \left(A \cap T_g^{-1} A \cap (S_g T_g)^{-1} A\right)$. Therefore

$$\xi(a,b) = \xi(ga,b) = \xi(ga,gb) = 1,$$

which means $\{(a,b),(ga,b),(ga,gb)\} \subset E$.

□

Corollary 6.3 *Suppose that G is a countable amenable group and that $F \subset G \times G$ has positive upper density with respect to a right Følner sequence $\{\Psi_n\}$ for $G \times G$. Then the set*

$$\{g \in G : \ \textit{there exists } (a,b) \in G \times G \textit{ such that } \{(a,b),(ag,b),(ag,bg)\} \subset F\}$$

is both left and right syndetic.

Proof. Let $E = F^{-1}$ and $\Phi_n = \Psi_n^{-1}$ and apply Theorem 6.1.

□

We remark that configurations of the type appearing in Corollary 6.3 cannot, in general, be found in a set E which is of positive upper density with respect to a *left* Følner sequence $\{\Phi_n\}$, even if $\overline{d}(E) = 1$. For example, let G be an amenable group having the property that for any finite subsets $E, F \subset G$, with $e \notin F$, there exists $h \in G$ such that $(h^{-1} E h \cap F) = \emptyset$ (the set of permutations of $\mathbf{N}$ which move finitely many elements is such a group). Let $\{\Phi_n\}$ be any left Følner sequence for $G \times G$. We will construct a sequence $(h_n)_{n=1}^{\infty} \subset G$ such that for all $n \in \mathbf{N}$, there exists no configuration of the form $\{(a,b),(ag,b),(ag,bg)\}$ in $\Phi_n(e,h_n)$. (Note that $\{\Phi_n(e,h_n)\}$ is again a left Følner sequence for $G \times G$.) Let

$$L_n = \{g \in G : \ \text{there exists } h \in G \text{ such that } (g,h) \in \Phi_n \text{ or } (h,g) \in \Phi_n\}.$$

L_n is finite, so there exists $h_n \in G$ such that $\left(h_n^{-1}(L_n^{-1} L_n) h_n \cap (L_n^{-1} L_n \setminus \{e\})\right) = \emptyset$. Suppose there exists $(a_1, a_2) \in G \times G$, and $e \neq g \in G$, such that

$$\{(a_1, a_2 h_n), (a_1 g, a_2 h_n), (a_1 g, a_2 h_n g)\} \subset \Phi_n(e, h_n).$$

Then $a_1, a_1 g \in L_n$, so that $g \in (L_n^{-1} L_n \setminus \{e\})$. Also $a_2 h_n, a_2 h_n g \in L_n h_n$, so that $a_2 \in L_n$ and $g \in h_n^{-1}(L_n^{-1} L_n) h_n$, a contradiction. One may check, incidentally, that this same construction gives a counterexample to existence of configurations of the type $\{(a,b),(ag,b),(a,bg)\}$.

We do not know, however, the answer to the following:

Question *Suppose that G is a countable amenable group and that $E \subset G \times G$ has positive upper density with respect to a left Følner sequence for $G \times G$. Does E necessarily contain a configuration of the form $\{(a,b),(ga,b),(a,gb)\}$?*

It is not difficult to see that the answer to this question is *yes* in the case of an abelian group G. In fact one can prove that if G is a countable abelian group and $E \subset G \times G$ has positive upper density with respect to a Følner sequence $\{\Phi_n\}$ for $G \times G$ then the set

$$\{g \in G : \text{ there exists}(a,b) \in G \times G \text{ such that } \{(a,b),(g+a,b),(a,g+b)\} \subset E\}$$

is syndetic in G. (One can in fact prove this via Corollary 5.4 and Proposition 6.2 by considering the commuting shifts $(T_g\xi)(g_1,g_2) = (g_1+g,g_2)$ and $(S_g\xi)(g_1,g_2) = (g_1-g,g_2+g)$.) We will now use this result for abelian groups to prove a configuration theorem for *finite* groups.

Corollary 6.4 *For every $\epsilon > 0$ there exists $k_0 = k_0(\epsilon)$ such that if G is any finite group of order $n > k_0$ and $B \subset G \times G$ satisfies $|B| > \epsilon n^2$ then B contains a configuration of the form $\{(a,b),(ga,b),(a,gb)\}$, $g \neq e$.*

Proof. Let $\epsilon > 0$ be given. Our first claim is that there exists m_0 such that for all $m > m_0$, if $B \subset \{1,2,\cdots,m\} \times \{1,2,\cdots,m\}$ and $|B| > \epsilon m^2$ then B contains a triangular configuration $\{(a,b),(g+a,b),(a,g+b)\}$, $g \neq 0$. (Suppose this is false. Then for some increasing sequence $(m_n)_{n=1}^\infty$ there exists, for all n, $B_n \subset \{1,2,\cdots,m_n\} \times \{1,2,\cdots,m_n\}$ containing no such triangle $\{(a,b),(g+a,b),(a,g+b)\}$ with $|B_n| > \epsilon m_n^2$. It follows that for any rapidly enough increasing sequence $(a_n)_{n=1}^\infty$, the set $E = \bigcup_{n=1}^\infty \big((a_n,a_n)+B_n\big)$ contains no such triangle. According to the discussion immediately preceding this corollary (applied to $G = \mathbf{Z}$), this contradicts the fact that $\overline{d}(E) \geq \epsilon$ with respect to the Følner sequence $\{\Phi_n\}$, where $\Phi_n = \{a_n+1,a_n+2,\cdots,a_n+m_n\}^2$.)

Our next claim is that if $q > 1$ then there exists $t_0 = t_0(q)$ such that for all $k > t_0$, if $B \subset C_q^k \times C_q^k$ (here C_q is a cyclic group of order q) and $|B| > \epsilon q^{2k}$ then B contains a configuration of the form $\{(a,b),(g+a,b),(a,g+b)\}$, $g \neq 0$. (Suppose now that this is false. Consider the group

$$G = \oplus_{i=1}^\infty C_q^{(i)},$$

where for all i $C_q^{(i)}$ is a cyclic group of order q. Let $S_n = C_q^{(1)} \times \cdots \times C_q^{(n)} \subset G$. For some increasing sequence $(k_n)_{n=1}^\infty$, there exists, for all n, $B_n \subset S_{k_n} \times S_{k_n}$ containing no configuration $\{(a,b),(g+a,b),(a,g+b)\}$, $g \neq 0$, with $|B_n| > \epsilon q^{2k_n}$. It follows that for any sparse enough sequence $(g_n)_{n=1}^\infty \subset G$, the set $E = \bigcup_{n=1}^\infty \big((g_n,g_n)+B_n\big)$ will contain no configuration $\{(a,b),(g+a,b),(a,g+b)\}$, $g \neq 0$. This contradicts the fact that $\overline{d}(E) \geq \epsilon$ with respect to the Følner sequence $\{\Phi_n\}$, where $\Phi_n = (g_n,g_n)+(S_{k_n} \times S_{k_n})$.)

Let $M = \prod_{q=2}^{m_0} q^{t_0(q)}$. Since any finite abelian group is isomorphic to a direct sum of cyclic groups, M has the property that any abelian group A of order $|A| > M$ has either a cyclic subgroup of order m, where $m > m_0$, or a subgroup isomorphic to C_q^k, where $2 \leq q \leq m_0$ and $k > t_0(q)$.

We need now the following group theoretic fact: for p prime and $n > 2$, any p-group of order $p^{\frac{1}{2}n(n-1)}$ has an abelian subgroup of order p^n (see, for example, [13], p. 120). In light of this it is easy to see that there exists a number k_0 having the property such that any finite group of order greater than k_0 has an abelian subgroup of order greater then M.

Suppose now that G is a finite group of order $n > k_0$ and $B \subset G \times G$ satisfies $|B| > \epsilon n^2$. G contains an abelian subgroup of order greater than M, and this abelian group contains therefore a group A which is either a cyclic group of order m, where $m > m_0$, or a subgroup isomorphic to C_q^k, where $2 \leq q \leq m_0$ and $k > k_0(q)$.

For some $(h_1, h_2) \in G \times G$, we will now have $|(A \times A)(h_1, h_2) \cap B| > \epsilon |A|^2$. Let

$$E = \big\{(a_1, a_2) \in A \times A : (a_1 h_1, a_2 h_2) \in B\big\}.$$

Then $|E| > \epsilon |A|^2$, so E contains a configuration $\{(x, y), (gx, y), (x, gy)\}$, $g \neq 0$. Letting $a = xh_1$ and $b = yh_2$ we have $\{(a, b), (ga, b), (a, gb)\} \subset B$.

□

One may notice that the configurations of Corollary 6.4 come from the existence of a large abelian subroup of G. We should therefore mention that a much stronger statement than that given in the corollary (in particular, an analogous statement in G^k for arbitrary $k \in \mathbf{N}$) may be obtained using a result of Furstenberg and Katznelson ([10], Theorem 9.3), which (among other things) guarantees the existence of all kinds of configurations in sizeable subsets of finite abelian groups. We will not concern ourselves here with the details. At any rate, the triangular configurations in $G \times G$ guaranteed by Corollary 6.4 yield the following configurations in G.

Corollary 6.5 *For every $\epsilon > 0$ there exists $k_0 = k_0(\epsilon)$ such that if G is any finite group of order $n > k_0$ and $B \subset G$ satisfies $|B| > \epsilon n$ then B contains a configuration of the form $\{h, gh, hg^{-1}\}$, $g \neq e$.*

Proof. Let $k = k_0(\epsilon)$ be as in Corollary 6.4 and suppose that G is a finite group with $|G| = n > k_0$ and $B \subset G$ with $|B| > \epsilon n$. Let

$$C = \big\{(a, b) \in G \times G : ab^{-1} \in B\big\}.$$

Then $|C| = n|B| > \epsilon n^2$, so C contains a configuration $\{(a, b), (ga, b), (a, gb)\}$, $g \neq e$. That is, $\{ab^{-1}, gab^{-1}, ab^{-1}g^{-1}\} \subset B$. Letting $h = ab^{-1}$ we are done.

□

We now indicate how one may in the same fashion get Roth's theorem on arithmetic progressions of length 3 ([15]).

Corollary 6.6 *For any $\epsilon > 0$ there exists $n_0 = n_0(\epsilon)$ such that for all $n > n_0$, if $B \subset \{1, 2, \cdots, n\}$ and $|B| > \epsilon n$ then B contains an arithmetic progression of length 3.*

Proof. Let $n_0(\epsilon) = m_0(\frac{\epsilon^2}{2})$, where m_0 is as in the first claim of the proof of Corollary 6.4. Suppose now that $n > n_0$ and $B \subset \{1, 2, \cdots, n\}$ satisfies $|B| > \epsilon n$. Let

$$C = \big\{(a, b) : 1 \leq b < a \leq n, a - b \in B\big\}.$$

Then one may show that $|C| > \frac{\epsilon^2}{2}$, so that C contains a configuration $\{(a,b),(g+a,b),(a,b+g)\}$. In other words, letting $x = a - b - g$, we have $\{x, x+g, x+2g\} \subset B$.

□

By restricting Corollary 6.5 to the case of G abelian and of odd order, we get the following theorem of Brown and Buhler ([3], [6]).

Corollary 6.7 *For every $\epsilon > 0$ there exists $k_0 = k_0(\epsilon)$ such that if A is any finite abelian group of odd order $n > k_0$ and $B \subset A$ satisfies $|B| > \epsilon n$ then B contains distinct elements x, y, z with $x + y = 2z$.*

Proof. Apply Corollary 6.5, letting $x = gh$, $z = h$, and $y = hg^{-1}$, noting that $x \neq y$ since A is of odd order.

□

7 Topological Recurrence Theorem

As an application of Theorem 6.1, we prove in this section a topological recurrence theorem for three commuting actions of a countable amenable group G. The proof is along the lines of some proofs of van der Waerden's theorem on arithmetic progressions, or its natural generalization in $\mathbf{Z}^n$, Grünwald's theorem (see, eg., [8]). The idea in such proofs is to inductively extend the validity of a topological multiple recurrence assertion from n commuting $\mathbf{Z}$-actions to $n+1$ commuting $\mathbf{Z}$-actions. However, when $\mathbf{Z}$ is replaced by a non-commutative group G, the induction already fails when $n = 2$. Indeed, in [1], counter-examples are given in some special cases, where G is a free group, to what might have looked to be possible non-abelian generalizations of van der Waerden's theorem. Specifically, there appears there a result, valid for two G-actions, the natural extension of which to three actions fails. It is therefore pleasing, and somewhat surprising, that for amenable groups, Theorem 6.1 fills the gap in the proof in going from two to three actions. Indeed, it will be easy to see that if one were to be given the analog to Theorem 6.1 for n dimensions, the methods of this section would give a topological result for $n+1$ actions, $n \in \mathbf{N}$. (Of course, we do not have such an analog for Theorem 6.1 at the present time, even for $n = 3$.)

Theorem 7.1 *Suppose that G is a countable amenable group and that (X, ρ) is a compact metric space. Suppose that $\{T_g\}$, $\{S_g\}$, and $\{R_g\}$ are G-actions by homeomorphisms on X such that $T_g S_h = S_h T_g$, $T_g R_h = R_h T_g$, and $R_g S_h = S_h R_g$ for all $g, h \in G$. Then for every $\epsilon > 0$, the set*

$$J_\epsilon = \{g \in G : \text{ there exists } x \in X \text{ such that } \rho(x, R_g x) < \epsilon,$$

$$\rho(x, R_g S_g x) < \epsilon,\ \rho(x, R_g S_g T_g x) < \epsilon\}$$

is both left and right syndetic in G.

Proof. We may, by passing to a closed subset of X if necessary, assume that X is minimal with respect to the G^3-action $\{R_f T_g S_h : f, g, h \in G\}$. We claim that for any open set

$U \subset X$, there exists a set H, which is both left and right syndetic in G, such that for all $g \in H$, there exists $z \in U$ with $\{z, S_g z, S_g T_g z\} \subset U$. We now prove the claim.

Let $U \subset X$ be open. Find $x \in U$ and $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. Let $Y \subset X$ be a closed set which is minimal with respect to the G^2-action $\{S_g T_h : \; g, h \in G\}$. One may check that $\overline{\bigcup_{g \in G} R_g Y}$ is $\{R_g\}$-, $\{S_g\}$-, and $\{T_g\}$-invariant, and is therefore equal to X. It follows that for some $g_0 \in G$, $R_{g_0}^{-1} B_{\epsilon/2}(x) \cap Y \neq \phi$. Let $\delta > 0$ be so small that if $y, y' \in Y$, with $\rho(y, y') < \delta$, then $\rho(R_{g_0} y, R_{g_0} y') < \frac{\epsilon}{2}$. Let $U' \subset R_{g_0}^{-1} B_{\frac{\epsilon}{2}}(x) \cap Y$ be an open set in Y of diameter less that δ. Let $y_0 \in Y$. Since the action $\{S_g T_h : \; g, h \in G\}$ is minimal on Y, the set

$$R = \{(g, h) : \; S_g T_h y_0 \in U'\}$$

is left syndetic in $G \times G$, and therefore of positive lower density with respect to every left Følner sequence of $G \times G$. It follows from Theorem 6.1 that the set

$$H = \{g \in G : \text{ there exists } (a, b) \in G \times G \text{ such that } \{(a, b), (ga, b), (ga, gb)\} \subset R\}$$

is both left and right syndetic in G. For $g \in H$, set $y = S_a T_b y_0 \in U'$, where

$$\{(a, b), (ga, b), (ga, gb)\} \subset R.$$

Then $S_g y \in U'$ and $S_g T_g y \in U'$, so that, letting $z = R_{g_0} y$, we have $z \in B_{\epsilon/2}(x)$, $\rho(z, S_g z) < \epsilon/2$, and $\rho(z, S_g T_g z) < \epsilon/2$. Therefore $\{z, S_g z, S_g T_g z\} \subset U$, establishing the claim.

Let $\epsilon > 0$, and let $\{\Phi_n\}$ be any left Følner sequence for G. Choose $x_0 \in X$ arbitrarily. In what follows, "small" will mean "of diameter less that $\epsilon/2$." Let U_0 be a small open set containing x_0. By the claim proved above, there exists a set $H_0 \subset G$, which is left syndetic, such that for every $g \in H_0$, there exists $z \in U_0$ with $\{z, S_g z, S_g T_g z\} \subset U_0$. Since any left syndetic set intersects a member of every left Følner sequence, we may find $n_0 \in \mathbf{N}$ and $h_0 \in \Phi_{n_0} \cap H_0$. There exists $y_0 \in U_0$ with $\{y_0, S_{h_0} y_0, S_{h_0} T_{h_0} y_0\} \subset U_0$. Put $x_1 = R_{h_0}^{-1} y_0$ and let U_1 be a small open set containing x_1 such that for every $x \in U_1$, we have $\{R_{h_0} x, R_{h_0} S_{h_0} x, R_{h_0} S_{h_0} T_{h_0} x\} \subset U_0$.

Suppose now that we have chosen small open sets $U_0, U_1, \cdots, U_t$ containing points $x_0, x_1, \cdots, x_t$, and $h_0, h_1, \cdots, h_{t-1} \in G$, $n_0, n_1, \cdots, n_{t-1} \in \mathbf{N}$ such that whenever $0 \leq i < j \leq t$ we have

(a) $h_{i,j} = h_i h_{i+1} \cdots h_{j-1} \in \Phi_{n_{j-1}}$, and
(b) $\{R_{h_{i,j}} x, R_{h_{i,j}} S_{h_{i,j}} x, R_{h_{i,j}} S_{h_{i,j}} T_{h_{i,j}} x\} \subset U_i$ for all $x \in U_j$.

We show now that we can continue this choosing. By our claim there exists a set H_t, left syndetic in G, such that for all $g \in H_t$, there exists $z \in X$ such that $\{z, S_g z, S_g T_g z\} \subset U_t$. For every $n \in \mathbf{N}$, let

$$\Phi'_n = \Phi_n \cap h_{t-1}^{-1} \Phi_n \cap (h_{t-2} h_{t-1})^{-1} \Phi_n \cap \cdots \cap (h_0 h_1 \cdots h_{t-1})^{-1} \Phi_n.$$

Then $\{\Phi'_n\}$ is a left Følner sequence for G, and we may, therefore, find $n_t \in \mathbf{N}$ and $h_t \in \Phi'_{n_t} \cap H_t$. There exists $y_t \in U_t$ with $\{y_t, S_{h_t} y_t, S_{h_t} T_{h_t} y_t\} \subset U_t$. Put $x_{t+1} = R_{h_t}^{-1} y_t$. Let U_{t+1} be a small open set containing x_{t+1} such that for all $x \in U_{t+1}$, we have $h_{i,t+1} = h_i h_{i+1} \cdots h_t \in \Phi_{n_t}$, and

$$\{R_{h_t} x, R_{h_t} S_{h_t} x, R_{h_t} S_{h_t} T_{h_t} x\} \subset U_t.$$

It follows that if $0 \le i < t$ and $x \in U_{t+1}$, then

$$\{R_{h_{i,t+1}}x, R_{h_{i,t+1}}S_{h_{i,t+1}}x, R_{h_{i,t+1}}S_{h_{i,t+1}}T_{h_{i,t+1}}x\} \subset U_i.$$

We have therefore established by induction that this process may be continued indefinitely. By compactness of X, we will eventually have $\rho(x_i, x_j) < \epsilon/2$ for some $i < j$. We then will have $h_{i,j} = h_i h_{i+1} \cdots h_{j-1} \in \Phi_{n_{j-1}}$, and

$$\rho(x_j, R_{h_{i,j}}x_j) < \epsilon, \rho(x_j, R_{h_{i,j}}S_{h_{i,j}}x_j) < \epsilon, \rho(x_j, R_{h_{i,j}}S_{h_{i,j}}T_{h_{i,j}}x_j) < \epsilon,$$

so that $h_{i,j} \in J_\epsilon$. Therefore, as we have shown that the set J_ϵ intersects non-trivially any arbitrarily chosen left Følner sequence $\{\Phi_n\}$, it must be left syndetic. Right syndeticity is proved similarly.

□

Corollary 7.2 *Suppose that G is a countable amenable group, $r \in \mathbf{N}$, and that $G \times G \times G = \bigcup_{i=1}^r C_i$. Then the set*

$$\big\{g \in G : \text{ there exists } i, 1 \le i \le r, \text{ and } (a,b,c) \in G \times G \times G, \text{ such that}$$

$$\{(a,b,c),(ga,b,c),(ga,gb,c),(ga,gb,gc)\} \subset C_i\big\}$$

is both left and right syndetic in G.

Proof. Let $\Omega = \{1, 2, \cdots, r\}^{G^3}$. We may choose a metric ρ on Ω generating the product topology such that for $\gamma, \eta \in \Omega$, $\rho(\gamma, \eta) < 1$ if and only if $\gamma(e,e,e) = \eta(e,e,e)$. Commuting G-actions by homeomorphisms $\{R_g\}, \{S_g\}$ and $\{T_g\}$ can be defined on Ω by $R_g\gamma(g_1, g_2, g_3) = \gamma(g_1 g, g_2, g_3)$, $S_g\gamma(g_1, g_2, g_3) = \gamma(g_1, g_2 g, g_3)$, and $T_g\gamma(g_1, g_2, g_3) = \gamma(g_1, g_2, g_3 g)$. Let ξ be the element of Ω defined by $\xi(g_1, g_2, g_3) = i$ when $(g_1, g_2, g_3) \in C_i$. Let

$$X = \overline{\{R_g S_h T_k \xi : g, h, k \in G\}}.$$

By Theorem 7.1 there is a set H, which is both left and right syndetic in G, such that for all $g \in H$, there exists $x \in X$ with $\rho(x, R_g x) < 1$, $\rho(x, R_g S_g x) < 1$, and $\rho(x, R_g S_g T_g x) < 1$. Suppose that $g \in H$, and let $x \in X$ have the properties just mentioned. There exist $a, b, c \in G$ such that $y = R_a S_b T_c \xi$ is so close to x that $\rho(y, R_g y) < 1$, $\rho(y, R_g S_g y) < 1$, and $\rho(y, R_g S_g T_g y) < 1$. It follows that

$$\xi(a,b,c) = \xi(ga,b,c) = \xi(ga,gb,c) = \xi(ga,gb,gc).$$

In other words,

$$\{(a,b,c),(ga,b,c),(ga,gb,c),(ga,gb,gc)\} \subset C_i,$$

where $i = \xi(a,b,c)$.

□

References

[1] Bergelson, V. and Hindman, N. *Some topological semicommutative van der Waerden type theorems and their combinatorial consequences*, J. London Math. Soc. (2) *45* (1992), 385-403.

[2] Bergelson, V. and Rosenblatt, J. *Joint ergodicity for group actions*, Ergodic Theory & Dynamical Systems, *8* (1988), 351-364.

[3] Brown, T.C. and Buhler, J.P. *A density version of a geometric Ramsey theorem*, Journal of Comb. Theory, Series A *32* (1982), 20-34.

[4] Conze, J. and Lesigne, E. *Théorèmes ergodiques pour des mesures diagonales*, Bull. Soc. Math. Fr., *112* (1984), 143-175.

[5] Conze, J. and Lesigne, E. *Sur un théorème ergodique pour des mesures diagonales*, C.S. Acad. Sci. Paris, *306*, I, (1988), 491-493.

[6] Frankl, P., Graham, R.L. and Rödl, V. *Abelian groups with no 3-term arithmetic progression*, Journal of Comb. Theory, Series A, *45*, (1987), 157-161.

[7] Furstenberg, H. *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math, *31* (1977), 204-256.

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

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

[10] Furstenberg, H. and Katznelson, Y. *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. d'Analyse Math, *45* (1985), 117-168.

[11] Furstenberg, H. and Katznelson, Y. *A density version of the Hales-Jewett theorem*, J. d'Analyse Math, *57* (1991), 64-119.

[12] Furstenberg, H. and Weiss, B. *A mean ergodic theorem for* $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2} x)$, Convergence in Ergodic Theory and Probability (edited by V. Bergelson, P. March and J. Rosenblatt), Ohio State University Mathematical Research Institute Publications **5**, Walter de Gruyter, Berlin, New York, 1996.

[13] Miller, G. A., Blichfeldt, F. and Dickson, L. E. *Theory and Applications of Finite Groups* (reprint). Dover, 1961.

[14] Ramsay, A. *Virtual groups and group actions*, Advances in Math. *6* (1971), 253-322.

[15] Roth, K. *Sur quelques ensembles d'entiers*, C.R. Acad. Sci. Paris *234* (1952), 388-390.

[16] Szemerédi, E. *On sets of integers containing no k elements in arithmetic progression*, Acta. Arith. *27* (1975), 199-245.

[17] Zimmer, R. *Extensions of ergodic group actions*, Illinois J. Math. *20* (1976), 373-409.

[18] Zimmer, R. *Ergodic actions with generalized discrete spectrum* , Illinois J. Math. *20* (1976), 555-588.

V. Bergelson
Department of Mathematics
The Ohio State University
Columbus, OH 43210
vitaly@math.ohio-state.edu

R. McCutcheon
Department of Mathematics
Case Western Reserve University
Cleveland, OH 44106
rxm16@cwru.edu

Q. Zhang
Department of Mathematics
Clark College
Vancouver, WA 98663
qzhang@clark.edu