Acta Math. Univ. Comenianae
Vol. LXIX, 1(2000), pp. 19–39

A HALF–COMMUTATIVE IP ROTH THEOREM

R. McCUTCHEON

1. Introduction

In 1985 H. Furstenberg and Y. Katznelson gave a powerful ergodic-theoretic multiple recurrence theorem for commuting IP-systems of measure preserving transformations. (Let $\mathcal{F} = \{A : A \subset \mathbf{N},\ 0 < |A| < \infty\}$. An IP-system is a sequence $(T_\alpha)_{\alpha\in\mathcal{F}}$ in an abelian semigroup satisfying $T_{\alpha\cup\beta} = T_\alpha T_\beta$ whenever $\alpha \cap \beta = \emptyset$.) Their result follows.

Theorem A ([**FK2**]). *Let $\big((T_\alpha^{(i)})_{\alpha\in\mathcal{F}}\big)_{i=1}^k$ be IP-systems contained in a abelian group of measure preserving transformations of a probability space $(X, \mathcal{A}, \mu)$. Then for all $A \in \mathcal{A}$ with $\mu(A) > 0$ there exists $\alpha \in \mathcal{F}$ such that $\mu\big(A \cap \bigcap_{i=1}^k (T_\alpha^{(i)})^{-1} A\big) > 0$.*

This result strengthens a previous theorem, again due to Furstenberg and Katznelson ([**FK1**]), stating that for commuting measure preserving transformations $T_1, \dots, T_k$ of a probability space $(X, \mathcal{A}, \mu)$ and $A \in \mathcal{A}$ with $\mu(A) > 0$ there exists $n > 0$ such that $\mu\big(A \cap \bigcap_{i=1}^k T_i^{-n} A\big) > 0$. The case of [**FK1**] in which each T_i is a power of the same transformation T was proved by Furstenberg in 1977 ([**F**]), and is equivalent to Szemerédi's theorem ([**Sz**]) on existence of arithmetic progressions in positive density subsets of $\mathbf{N}$. (In particular, Theorem A provides that n may be chosen from any IP-system in $\mathbf{N}$.) For this reason, Furstenberg and Katznelson refer to Theorem A as an "IP Szemerédi theorem". The special case $k = 2$ of Theorem A might therefore be called an "IP Roth theorem", for it may be used to infer the case of Szemerédi's theorem dealing with three-term arithmetic progressions, which is due to K. Roth ([**R**]).

Our goal here is to remove some of the commutativity restrictions on Theorem A in the special case $k = 2$ (hence the title of the paper). There are at least two reasonable choices for the definition of "IP-system" $(T_\alpha)_{\alpha\in\mathcal{F}}$ in a non-abelian semigroup. We may require either $T_{\alpha\cup\beta} = T_\alpha T_\beta$ or $T_{\alpha\cup\beta} = T_\beta T_\alpha$ for $\alpha, \beta \in \mathcal{F}$

Received September 14, 1999; revised January 19, 2000.
1980 *Mathematics Subject Classification* (1991 *Revision*). Primary 28D15; Secondary 54H20, 20M20.
Key words and phrases. Ramsey theory, Szemerédi's theorem, Van der Warden's theorem, IP-system, amenability.

with $\max\alpha < \min\beta$. We adopt the former condition, calling systems satisfying the latter condition **reversed** IP-systems. The typical manner for constructing an IP-system $(T_\alpha)_{\alpha\in\mathcal{F}}$ is to choose "generators" $(S_i)_{i=1}^\infty$ and put $T_\alpha = \prod_{i\in\alpha} S_i$, where the product is taken with increasing indices (eg. $T_{\{1,4,5\}} = S_1S_4S_5$). Taking products with decreasing indices yields a reversed IP-system.

We consider two IP-systems $(T_\alpha)_{\alpha\in\mathcal{F}}$ and $(S_\alpha)_{\alpha\in\mathcal{F}}$ of measure preserving transformations of a probability space $(X,\mathcal{A},\mu)$ which commute with each other in the sense that $T_\alpha S_\beta = S_\beta T_\alpha$ for all $\alpha,\beta\in\mathcal{F}$, and such that $(S_\alpha)_{\alpha\in\mathcal{F}}$ is itself commutative; namely $S_\alpha S_\beta = S_\beta S_\alpha$ for all $\alpha,\beta\in\mathcal{F}$. The conclusion we obtain is that if $\mu(A) > 0$ then there exists $\alpha\in\mathcal{F}$ such that $\mu\big(A\cap T_\alpha^{-1}A\cap(T_\alpha S_\alpha)^{-1}A\big) > 0$. In other words, we have succeeded in removing the commutativity restriction on one of the IP-systems in the $k = 2$ case of Theorem A. The method of proof is a natural outgrowth of the methods of Furstenberg and Katznelson. After a preparatory section (Section 2), we present the result and its proof in Section 3.

Section 4 is devoted to two combinatorial corollaries, one for sets of positive density in amenable semigroups (Theorem 4.2) and one for partition Ramsey theory (Corollary 4.5). Also included in this section is a topological recurrence theorem.

2. Preliminaries

In this section we will give definitions and formulate preliminary results and combinatorial tools needed for the proof of our main theorem, Theorem 3.1. Recall that $\mathcal{F}$ denotes the family of non-empty finite subsets of $\mathbf{N}$.

Definition 2.1. (a) Suppose that $\alpha,\beta\in\mathcal{F}$ have the property that every member of α is less than every member of β, that is, $x\in\alpha$ and $y\in\beta$ implies that $x<y$. In this case we shall say that α **precedes** β and write $\alpha<\beta$.

(b) Given a sequence $(\alpha_i)_{i\in\mathbf{N}}\subset\mathcal{F}$ such that $\alpha_1<\alpha_2<\alpha_3<\dots$, we denote the set of all unions of finitely many elements of the sequence by $FU(\alpha_1,\alpha_2,\dots)$. We also let $FU_\emptyset(\alpha_1,\alpha_2,\dots) = FU(\alpha_1,\alpha_2,\dots)\cup\{\emptyset\}$. Both $FU(\alpha_1,\alpha_2,\dots)$ and $FU_\emptyset(\alpha_1,\alpha_2,\dots)$ are called **IP-rings**, and are usually denoted by symbols such as $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$, etc. If $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ are IP-rings and $\mathcal{F}^{(2)}\subset\mathcal{F}^{(1)}$ then $\mathcal{F}^{(2)}$ is said to be a **subring** of $\mathcal{F}^{(1)}$.

(c) Suppose that $(x_\alpha)_{\alpha\in\mathcal{F}}$ is a sequence indexed by $\mathcal{F}$ (called an **$\mathcal{F}$-sequence**) in a topological space X, $x\in X$ and $\mathcal{F}^{(1)} = FU(\alpha_1,\alpha_2,\dots)$ is an IP-ring. Suppose that for every neighborhood U of x there exists $\beta\in\mathcal{F}$ having the property that for every $\alpha\in\mathcal{F}^{(1)}$ with $\beta<\alpha$ we have $x_\alpha\in U$. Then we shall say that the sequence (x_α) **converges to x along $\mathcal{F}^{(1)}$**, and we shall write

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} x_\alpha = x.$$

(d) An $\mathcal{F}$-sequence $(x_\alpha)_{\alpha\in\mathcal{F}}$ in a semigroup is called an **IP-system** (respectively **reversed IP-system**) if $x_{\alpha\cup\beta} = x_\alpha x_\beta$ (respectively $x_{\alpha\cup\beta} = x_\beta x_\alpha$) for all $\alpha, \beta \in \mathcal{F}$ with $\alpha < \beta$.

A fundamental combinatorial tool useful when dealing with IP-convergence is Hindman's Theorem:

Theorem 2.2 ([**H**]). *Suppose $r \in \mathbf{N}$ and $\bigcup_{i=1}^r C_i = \mathcal{F}^{(1)}$ is a partition of an IP-ring into r cells. Then some cell of the partition contains an IP-ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$.*

Given an IP-ring $\mathcal{F}^{(1)} = FU(\alpha_1, \alpha_2, \dots)$, the correspondence $\beta \leftrightarrow \bigcup_{i\in\beta} \alpha_i$ is a bijection between $\mathcal{F}$ and $\mathcal{F}^{(1)}$ which preserves unions and the partial order relation $<$. Hence IP-rings themselves have the structure of $\mathcal{F}$. It follows that given any $\mathcal{F}$-sequence $(x_\alpha)_{\alpha\in\mathcal{F}}$ and an IP-ring $\mathcal{F}^{(1)}$, the sequence $(x_\alpha)_{\alpha\in\mathcal{F}^{(1)}}$ has the character of an $\mathcal{F}$-sequence. Therefore, we may refer to the restriction of an $\mathcal{F}$-sequence to an IP-ring as an $\mathcal{F}$-sequence as well.

If $\mathcal{F}^{(2)}$ is a subring of $\mathcal{F}^{(1)}$ then the restriction of a given $\mathcal{F}$-sequence to $\mathcal{F}^{(2)}$ is called a **subsequence** of the restriction the $\mathcal{F}$-sequence to $\mathcal{F}^{(1)}$. As is the case with ordinary sequences, any $\mathcal{F}$-sequence in a compact metric space has a convergent subsequence. The following generalization of this fact may be proved using Hindman's Theorem and a standard diagonalization argument.

Corollary 2.3 ([**FK2**, Theorem 1.5]). *Suppose that, for all $n \in \mathbf{N}$, $(x_\alpha^{(n)})_{\alpha\in\mathcal{F}}$ is an $\mathcal{F}$-sequence in a compact metric space X_n. Then for any IP-ring $\mathcal{F}^{(1)}$ there exists a subring $\mathcal{F}^{(2)}$ such that*

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(2)}} x_\alpha^{(n)} = z_n$$

exists for each $n \in \mathbf{N}$.

Any IP-system or reversed IP-system $(U_\alpha)_{\alpha\in\mathcal{F}}$ of isometries on a separable Hilbert space $\mathcal{H}$ is of course also an $\mathcal{F}$-sequence. Since the unit ball of $\mathcal{H}$ is compact and metrizable in the weak topology, one may show as a consequence of Corollary 2.3 that along some subring $\mathcal{F}^{(1)}$,

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} U_\alpha f = Pf \tag{2.1}$$

exists weakly for all $f \in \mathcal{H}$. (One need only show that the limit exists for all f in a countable dense subset of $\mathcal{H}$.) The content of the following proposition is that the operator P of (2.1) is an orthogonal projection.

Proposition 2.4. *Suppose that $(U_\alpha)_{\alpha\in\mathcal{F}}$ is an IP-system or reversed IP-system of isometries on a separable Hilbert space $\mathcal{H}$ and $\mathcal{F}^{(1)}$ is an IP-ring such that*

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} U_\alpha f = Pf$$

exists weakly for all $f \in \mathcal{H}$. Then P is an orthogonal projection onto a closed subspace of $\mathcal{H}$.

Proof. We will prove the assertion in the case that (U_α) is a reversed IP-system. (The IP-system case is actually somewhat easier.) That P is linear is easily checked, and furthermore it is easily established that $\|P\| \leq 1$. These facts imply that if $P^2 = P$ then P must be an orthogonal projection. Therefore we must only show that P is idempotent.

Let ρ be a metric for the weak topology on the closed unit ball of $\mathcal{H}$, let $f \in \mathcal{H}$ with $\|f\| < 1$ and let $\epsilon > 0$ be arbitrary. We begin with the observation that P is weakly continuous. Therefore we may choose $\alpha_0 \in \mathcal{F}$ having the property that for every $\alpha > \alpha_0$ we have

$$\rho(U_\alpha f, Pf) < \frac{\epsilon}{3} \tag{2.2}$$

and $\rho(PU_\alpha f, P^2 f) < \frac{\epsilon}{3}$.

Suppose now that $\alpha, \beta \in \mathcal{F}^{(1)}$ with $\beta > \alpha > \alpha_0$. Then $(\alpha \cup \beta) > \alpha_0$, hence

$$\rho(U_\beta U_\alpha f, Pf) = \rho(U_{\alpha\cup\beta} f, Pf) < \frac{\epsilon}{3}. \tag{2.3}$$

Fixing $\alpha > \alpha_0$, there exists $\beta > \alpha$ far enough out that

$$\rho(U_\beta U_\alpha f, PU_\alpha f) < \frac{\epsilon}{3}. \tag{2.4}$$

It follows from (2.2), (2.3) and (2.4) that

$$\rho(Pf, P^2 f) \leq \rho(Pf, U_\beta U_\alpha f) + \rho(U_\beta U_\alpha f, PU_\alpha f) + \rho(PU_\alpha f, P^2 f) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Since ϵ was arbitrary, $P^2 f = Pf$. Since f was arbitrary, $P^2 = P$. □

Another combinatorial tool we'll use is the following corollary of Hindman's theorem. In order to formulate it, we adopt the following notation: if $n \in \mathbf{N}$ and $\mathcal{F}^{(1)}$ is an IP ring, let

$$\left(\mathcal{F}^{(1)}\right)^n_< = \left\{(\alpha_1, \ldots, \alpha_n) : \alpha_1, \ldots, \alpha_n \in \mathcal{F}^{(1)}, \alpha_1 < \alpha_2 < \cdots < \alpha_n\right\}.$$

Theorem 2.5 ([**M**], [**T**]). *Suppose that $\mathcal{F}^{(1)}$ is an IP-ring, $n, r \in \mathbf{N}$, and that $\left(\mathcal{F}^{(1)}\right)^n_< = \bigcup_{i=1}^r C_i$ is an r-cell partition of $\left(\mathcal{F}^{(1)}\right)^n_<$. Then there exists j, $1 \leq j \leq r$, and an IP-ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that $\left(\mathcal{F}^{(2)}\right)^n_< \subset C_j$.*

The following is a standard trick for showing convergence of $\mathcal{F}$-sequences in a Hilbert space.

Proposition 2.6 ([**FK2**, Lemma 5.3]). *Suppose that $\{x_\alpha\}_{\alpha\in\mathcal{F}}$ is a bounded $\mathcal{F}$-sequence of vectors in a Hilbert space $\mathcal{H}$ and that $\mathcal{F}^{(1)}$ is an IP-ring. If*

$$\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \left\langle x_\alpha,\ x_{\alpha\cup\beta}\right\rangle = 0$$

then for some IP-subring $\mathcal{F}^{(2)}\subset\mathcal{F}^{(1)}$, $\underset{\alpha\in\mathcal{F}^{(2)}}{\text{IP-lim}}\ x_\alpha = 0$ in the weak topology.

The following fact follows from [**FK3**, Lemma 3.1]. We include a proof for completeness.

Theorem 2.7. *Suppose that $(X,\mathcal{A},\mu)$ is a probability space. A closed subspace $E\subset L^2(X,\mathcal{A},\mu)$ has the form $E = L^2(X,\mathcal{B},\mu)$ for a sub-σ-algebra $\mathcal{B}\subset\mathcal{A}$ if and only if E has a dense subset E_0 of bounded functions containing the constants and having the property that if $f,g\in E_0$ then $\{fg, f+g\}\subset E_0$.*

Proof. One direction is obvious. Given $E = L^2(X,\mathcal{B},\mu)$, simply let E_0 consist of those members of E that are bounded. E_0 is clearly dense in E and has the required properties.

For the converse, let E_0 be as stated in the theorem. Let $a,b\in\mathbf{R}$ and let $g\in E_0$. Let $B=\{x: g(x)\in(a,b)\}$. We claim that $1_B\in\overline{E_0}$. Indeed, replacing g by $\frac{g}{N}$, where N is sufficiently large, we may assume without loss of generality that $\sup|g|<1$ and $-1<a<b<1$. Let $\epsilon>0$ be arbitrary and choose $\delta>0$ such that $\mu(B'\triangle B)<\epsilon$, where $B'=\{x: g(x)\in(a+\delta,b-\delta)\}$. By the Weierstrass approximation theorem, choose now a polynomial $p(t)$ with $p([-1,1])\subset[0,1]$, $p(t)>1-\epsilon$ for $t\in(a+\delta,b-\delta)$, and $p(t)<\epsilon$ for $t\in[-1,a]\cup[b,1]$. Then $p\circ g\in E_0$ with $||p\circ g-1_B||<2\epsilon$.

Let $\mathcal{B}_0$ be the algebra of sets generated by sets of the form $\{x: g(x)\in(a,b)\}$, where $g\in E_0$ and $a,b\in\mathbf{R}$. Since E_0 is closed under products (to approximate intersection) and sums (to approximate union), $1_B\in\overline{E_0}$ for all $B\in\mathcal{B}_0$. Since E_0 is closed under products and contains the constants (and is therefore closed under finite linear combinations), $\overline{E_0}$ contains all simple functions $\sum_{i=1}^n 1_{B_i}$, where the B_i's come from $\mathcal{B}_0$.

Let $\mathcal{B}=\{B\in\mathcal{A}: \exists(B_i)_{i=1}^\infty\subset\mathcal{B}_0 \text{ with } \mu(B\triangle B_i)\to 0\}$. One easily checks that $\mathcal{B}$ is a σ-algebra containing $\mathcal{B}_0$. Therefore $\overline{E_0}$ contains all simple functions $\sum_{i=1}^n 1_{B_i}$, where the B_i's come from $\mathcal{B}$. This implies that $L^2(X,\mathcal{B},\mu)\subset\overline{E_0}$. On the other hand all the members of E_0 are clearly $\mathcal{B}$-measurable by construction. It follows that $E_0\subset L^2(X,\mathcal{B},\mu)$ and since $L^2(X,\mathcal{B},\mu)$ is closed, $\overline{E_0}\subset L^2(X,\mathcal{B},\mu)$, so that $E=\overline{E_0}=L^2(X,\mathcal{B},\mu)$. □

The following routine lemma is needed. If $\alpha_1,\dots,\alpha_n$ are sets let $FU\{\alpha_1,\dots,\alpha_n\}=\left\{\bigcup_{i\in\beta}\alpha_i : \emptyset\neq\beta\subset\{1,\dots,n\}\right\}$ and put $FU_\emptyset\{\alpha_1,\dots,\alpha_n\}=FU\{\alpha_1,\dots,\alpha_n\}\cup\{\emptyset\}$.

Lemma 2.8. *Suppose that $\{T_\alpha\}_{\alpha\in\mathcal{F}}$ is an IP-system of measure preserving transformations of a probability space $(X,\mathcal{B},\mu)$. Let $n\in\mathbf{N}$ and $\xi>0$. For any $A\in\mathcal{B}$ with $\mu(A)>\xi$ there exists an IP-ring $\mathcal{G}$ such that for every $(\alpha_1,\dots,\alpha_n)\in(\mathcal{G})^n_<$ we have*

$$\mu\Big(\bigcap_{\alpha\in FU_\emptyset\{\alpha_1,\dots,\alpha_n\}} T_\alpha^{-1}A\Big)>\xi^{2^n}.$$

Proof. Since we are only concerned with the orbit of a single set A under countably many transformations, we may assume without loss of generality that $(X,\mathcal{B},\mu)$ is separable.

Case 1: $n=1$. Letting $U_\alpha f(x)=f(T\alpha x)$ for $f\in L^2(X,\mathcal{B},\mu)$, $\{U_\alpha\}$ is a reversed IP-system of isometries. Let $\mathcal{F}^{(1)}$ be an IP-ring having the property that

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} U_\alpha f=Pf$$

exists for all $f\in L^2(X,\mathcal{B},\mu)$ (see the discussion preceding Proposition 2.4). According to Proposition 2.4, P is an orthogonal projection. Let $\mathcal{F}^{(2)}\subset\mathcal{F}^{(1)}$ be an IP-ring having one of the following two properties:

(a) For every $\alpha\in\mathcal{F}^{(2)}$, $\mu(A\cap T_\alpha^{-1}A)>\xi^2$.

(b) For every $\alpha\in\mathcal{F}^{(2)}$, $\mu(A\cap T_\alpha^{-1}A)\le\xi^2$.

This is possible by Theorem 2.2. Let $f=1_A$. Since P is the orthogonal projection onto a subspace containing the constants we have

$$\begin{aligned}\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(2)}}\ \mu(A\cap T_\alpha^{-1}A)&=\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(2)}}\int fT_\alpha f\,d\mu\\&=\int fPf\,d\mu=\big\|Pf\big\|^2\ge\mu(A)^2>\xi^2.\end{aligned}$$

It follows that (b) is an impossibility, hence (a) holds. Letting $\mathcal{G}=\mathcal{F}^{(2)}$ completes the proof of Case 1.

Case 2: Suppose the result has been established for $n-1$. There exists an IP-ring $\mathcal{F}^{(1)}$ such that for all $(\alpha_1,\dots,\alpha_{n-1})\in(\mathcal{F}^{(1)})^{n-1}_<$ we have

$$\mu\Big(\bigcap_{\alpha\in FU_\emptyset\{\alpha_1,\dots,\alpha_{n-1}\}} T_\alpha^{-1}A\Big)>\xi^{2^{n-1}}.$$

By Theorem 2.5 we may choose an IP-ring $\mathcal{F}^{(2)}\subset\mathcal{F}^{(1)}$ such that one of the following two criteria is met:

(a) For all $(\alpha_1,\dots,\alpha_n)\in(\mathcal{F}^{(2)})^n_<$ we have

$$\mu\Big(\bigcap_{\alpha\in FU_\emptyset\{\alpha_1,\dots,\alpha_n\}} T_\alpha^{-1}A\Big)>\xi^{2^n}.$$

(b) For all $(\alpha_1, \dots, \alpha_n) \in (\mathcal{F}^{(2)})^n_<$ we have

$$\mu\Big(\bigcap_{\alpha \in FU_\emptyset\{\alpha_1,\dots,\alpha_n\}} T_\alpha^{-1} A\Big) \le \xi^{2^n}.$$

Let $(\alpha_1, \dots, \alpha_{n-1}) \in (\mathcal{F}^{(2)})^{n-1}_<$ and let

$$B = \bigcap_{\alpha \in FU_\emptyset\{\alpha_1,\dots,\alpha_{n-1}\}} T_\alpha^{-1} A.$$

Then $\mu(B) > \xi^{2^{n-1}}$. Proceeding as in Case 1 we get

$$\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(2)}} \mu(B \cap T_\alpha^{-1} B) \ge \mu(B)^2 > \xi^{2^n}.$$

In particular for α_n far enough out we have

$$\mu\Big(\bigcap_{\alpha \in FU_\emptyset\{\alpha_1,\dots,\alpha_n\}} T_\alpha^{-1} A\Big) > \xi^{2^n}.$$

Therefore (b) is an impossibility and (a) holds. Let $\mathcal{G} = \mathcal{F}^{(2)}$. □

3. Proof of Main Theorem

Theorem 3.1. *Let $(X, \mathcal{A}, \mu)$ be a probability space, let $\{T_\alpha\}_{\alpha \in \mathcal{F}}$ and $\{S_\alpha\}_{\alpha \in \mathcal{F}}$ be IP-systems of measure preserving transformations of X such that $T_\alpha S_\beta = S_\beta T_\alpha$ and $S_\alpha S_\beta = S_\beta S_\alpha$ for all $\alpha, \beta \in \mathcal{F}$. Then for every $A \in \mathcal{A}$ with $\mu(A) > 0$ there exists an IP-ring $\mathcal{F}^{(1)}$ with*

$$\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \mu(A \cap T_\alpha^{-1} A \cap (T_\alpha S_\alpha)^{-1} A) > 0.$$

Although Theorem 3.1 is stated for an arbitrary probability space, in the proof it will be necessary to place some additional restrictions on $(X, \mathcal{A}, \mu)$. Specifically, we shall require that $(X, \mathcal{A}, \mu)$ be a **regular** space, namely X is compact metric and μ is a Borel measure. To see that this may be done without loss of generality, consider the following standard construction.

Under the conditions of Theorem 3.1, let G be the (countable) semigroup generated by $\{T_\alpha : \alpha \in \mathcal{F}\} \cup \{S_\alpha : \alpha \in \mathcal{F}\}$. If G has an identity, denote it by e. Otherwise, let e be an identity adjoined to G and in either case put $\tilde{X} = \{0,1\}^{G \cup \{e\}}$. Endowed with the product topology, $\tilde{X}$ is compact and metrizable. Let $\tilde{A} = \{\gamma \in \tilde{X} : \gamma(e) = 1\}$. For $g_1, g_2, \dots, g_r \in G \cup \{e\}$, and $h_1, h_2, \dots, h_r \in \{0,1\}$, define

$$\tilde{\mu}\big(\{\gamma \in \tilde{X} : \gamma(g_i) = h_i, 1 \le i \le r\}\big) = \mu\Big(\bigcap_{i=1}^{r} g_i^{-1} A_i\Big),$$

where $A_i = A$ if $h_i = 1$ and $A_i = A^c$ if $h_i = 0$, $1 \le i \le r$. $\tilde{\mu}$ is a premeasure on the algebra of cylinder sets and hence extends uniquely to a measure on the Borel σ-algebra $\tilde{\mathcal{A}}$ of $\tilde{X}$, so that $(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$ is a regular space.

For $\alpha \in \mathcal{F}$ and $\gamma \in X$, put $\tilde{T}_\alpha \gamma(g) = \gamma(gT_\alpha)$ and $\tilde{S}_\alpha \gamma(g) = \gamma(gS_\alpha)$. One may easily check that $\{\tilde{T}_\alpha\}_{\alpha\in\mathcal{F}}$ and $\{\tilde{S}_\alpha\}_{\alpha\in\mathcal{F}}$ are IP-systems of $\tilde{\mu}$-preserving maps with $\tilde{S}_\alpha \tilde{T}_\beta = \tilde{T}_\beta \tilde{S}_\alpha$ and $\tilde{S}_\alpha \tilde{S}_\beta = \tilde{S}_\beta \tilde{S}_\alpha$ for all $\alpha, \beta \in \mathcal{F}$.

Finally, note that $\tilde{\mu}(\tilde{A}) = \mu(A)$ and for all $\alpha \in \mathcal{F}$,

$$\begin{aligned}\tilde{\mu}\big(\tilde{A} \cap \tilde{T}_\alpha^{-1}\tilde{A} \cap (\tilde{T}_\alpha\tilde{S}_\alpha)^{-1}\tilde{A}\big) &= \tilde{\mu}\big(\{\gamma \in \tilde{X} : \gamma(e) = 1,\ \gamma(T_\alpha) = 1, \gamma(T_\alpha S_\alpha) = 1\}\big)\\ &= \mu\big(A \cap T_\alpha^{-1}A \cap (T_\alpha S_\alpha)^{-1}A\big).\end{aligned}$$

This establishes that we may in fact assume without loss of generality that we deal with regular spaces and hence take advantage of the structure they afford. For more details on the following discussion, the reader is referred to [**F2**, Chapter 5], [**F1**, Section 4] or [**FK3**, Section 3]. If $(X, \mathcal{A}, \mu)$ is a regular space and $\mathcal{B} \subset \mathcal{A}$ is a sub-σ-algebra, there exists a space $(Y, \tilde{B}, \nu)$ and a measure preserving transformation $\pi\colon X \to Y$ such that $\mathcal{B} = \pi^{-1}(\tilde{B})$ (modulo null sets). Y is said to be a **factor** of X. Sometimes we abuse notation and say that $\mathcal{B}$ is a factor of X, or use the symbol $\mathcal{B}$ in place of $\tilde{\mathcal{B}}$.

Moreover, there exists a family of probability measures $\{\mu_y : y \in Y\}$ on X such that for any $f \in L^1(X, \mathcal{A}, \mu)$, $\int f\ d\mu = \int(\int f\ d\mu_y)\ d\nu(y)$ and so that $\mu_y\big(\pi^{-1}(y)\big) = 1$ a.e. If we write $E(f|\mathcal{B})(x) = \int f\ d\mu_{\pi(x)}$ then $E(f|\mathcal{B})$ is the **conditional expectation** of f given $\mathcal{B}$, and the map $f \to E(f|\mathcal{B})$ is the orthogonal projection from $L^2(X, \mathcal{A}, \mu)$ to $L^2(X, \mathcal{B}, \mu)$.

We can use the disintegration of μ over Y to define a measure $\mu \times_{\mathcal{B}} \mu$ on $(X \times X, \mathcal{A} \otimes \mathcal{A})$ as follows: $\int f(x_1, x_2) d\mu \times_{\mathcal{B}} \mu(x_1, x_2) = \int(\int f(x_1, x_2)\ d\mu_y \times \mu_y(x_1, x_2))\ d\nu(y)$. One may check that $\mu \times_{\mathcal{B}} \mu$ is supported on the set $X \times_Y X = \{(x_1, x_2) : \pi(x_1) = \pi(x_2)\}$, and we sometimes denote this **conditional product space** as $(X \times_Y X, \mathcal{A} \otimes_{\mathcal{B}} \mathcal{A}, \mu \times_{\mathcal{B}} \mu)$, or by $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu})$ for convenience. One may easily check that if $T : X \to X$ is μ-invariant, then $\tilde{T} = T \times T$ defined by $\tilde{T}(x_1, x_2) = (Tx_1, Tx_2)$ is $\tilde{\mu}$-invariant.

For the remainder of this section, $(X, \mathcal{A}, \mu)$, $\{T_\alpha\}_{\alpha\in\mathcal{F}}$, $\{S_\alpha\}_{\alpha\in\mathcal{F}}$, and $A \in \mathcal{A}$ with $\mu(A) > 0$ will be fixed, with $(X, \mathcal{A}, \mu)$ regular. We take $L^2(X, \mathcal{A}, \mu)$ to consist of real-valued functions. In particular, since $(X, \mathcal{A}, \mu)$ is regular, $L^2(X, \mathcal{A}, \mu)$ is separable and its unit ball is compact and metrizable in the weak topology.

Let $\mathcal{F}^{(1)}$ be an IP-ring with the property that

$$\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \mu(A \cap T_\alpha^{-1}A \cap (T_\alpha S_\alpha)^{-1}A) = L \tag{3.1}$$

exists, with the additional requirement that for all $f \in L^2(X, \mathcal{A}, \mu)$,

$$\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ S_\alpha f = Pf$$

exists weakly (possible by Corollary 2.3, as one need only assure convergence for a countable dense set of f's in the unit ball). We must show that $L > 0$.

By Proposition 2.4, P is an orthogonal projection onto a closed subspace E of $L^2(X, \mathcal{A}, \mu)$. E contains the constants, and since P is idempotent we have $Pf = f$ if and only if $f \in E$. In other words, $S_\alpha f \to f$ weakly along $\mathcal{F}^{(1)}$. But $\|S_\alpha f\| = \|f\|$, hence $S_\alpha f \to f$ strongly as well along $\mathcal{F}^{(1)}$. It follows that for any $f \in E$ and any $l \in \mathbf{R}$ the function f_l defined by $f_l(x) = l$ if $f(x) > l$, $f_l(x) = -l$ if $f(x) < -l$ and $f_l(x) = f(x)$ if $-l < f(x) < l$ satisfies $S_\alpha f_l \to f_l$ along $\mathcal{F}^{(1)}$ and hence lies in E. It follows that E contains a dense subset consisting of bounded functions. Moreover, if f, g are bounded functions in E, we have

$$\begin{aligned}\|S_\alpha f S_\alpha g - fg\| &\leq \|S_\alpha f S_\alpha g - (S_\alpha f)g\| + \|(S_\alpha f)g - fg\| \\ &\leq \|f\|_\infty \|S_\alpha g - g\| + \|g\|_\infty \|S_\alpha f - f\| \to 0\end{aligned}$$

along $\mathcal{F}^{(1)}$, so that $fg \in E$ (and bounded). Moreover $f + g$ is bounded and in E (this is obvious), so by Theorem 2.7 $E = L^2(X, \mathcal{B}, \mu)$, where $\mathcal{B} \subset \mathcal{A}$ is a σ-algebra, and thus $Pf = E(f|\mathcal{B})$, that is, P is the orthogonal projection onto $L^2(X, \mathcal{B}, \mu)$. One may further check that $\mathcal{B}$ must be T_α- and S_α-invariant for all $\alpha \in \mathcal{F}$. (It is here, and only here, that we use the fact that $S_\alpha S_\beta = S_\beta S_\alpha$ for all $\alpha, \beta \in \mathcal{F}$.)

The factor determined by $\mathcal{B}$ will be denoted $(Y, \mathcal{B}, \nu)$. Note that $\{T_\alpha\}_{\alpha \in \mathcal{F}}$ and $\{S_\alpha\}_{\alpha \in \mathcal{F}}$ project to measure preserving IP-systems on $(Y, \mathcal{B}, \nu)$. $(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$ will denote the conditional product probability space $(X \times_Y X, \mathcal{A} \otimes_{\mathcal{B}} \mathcal{A}, \mu \times_{\mathcal{B}} \mu)$. $\{\tilde{T}_\alpha\}_{\alpha \in \mathcal{F}}$ and $\{\tilde{S}_\alpha\}_{\alpha \in \mathcal{F}}$ will denote the $\tilde{\mu}$-measure preserving IP systems on $\tilde{X}$ defined by $\tilde{T}_\alpha(x_1, x_2) = (Tx_1, Tx_2)$ and $\tilde{S}(x_1, x_2) = (Sx_1, Sx_2)$, respectively.

Remark. Throughout the course of this section we will periodically obtain IP-subrings of $\mathcal{F}^{(1)}$ (using Theorem 2.5, for example) having some desirable properties. One may check that previous properties observed for $\mathcal{F}^{(1)}$ are possessed by the subring. Therefore, we will simply replace $\mathcal{F}^{(1)}$ by its subring and continue using the name $\mathcal{F}^{(1)}$ (for the subring). We may say simply "without loss of generality, $\mathcal{F}^{(1)}$ possesses such-and-such property."

By Corollary 2.3, without loss of generality for all $H \in L^2(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$,

$$\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \tilde{T}_\alpha H = Q_1 H$$

and

$$\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \tilde{T}_\alpha \tilde{S}_\alpha H = Q_2 H$$

exist in the weak topology. By Proposition 2.4, Q_1 and Q_2 are orthogonal projections.

Definition 3.2. A function $f \in L^\infty(X, \mathcal{A}, \mu)$ is T_α-*almost periodic over* $\mathcal{B}$ **along** $\mathcal{F}^{(1)}$ if for every $\epsilon > 0$ there exists a set $D \in \mathcal{B}$ with $\nu(D) < \epsilon$ and functions

$g_1, \ldots, g_N \in L^2(X, \mathcal{A}, \mu)$ having the property that for every $\delta > 0$ there exists $\alpha_0 \in \mathcal{F}^{(1)}$ such that for every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ there is a set $E(\alpha) \in \mathcal{B}$ with $\nu\big(E(\alpha)\big) < \delta$ having the property that for all $y \in Y \setminus \big(D \cup E(\alpha)\big)$ there exists $i(y, \alpha)$ with $1 \leq i(y, \alpha) \leq N$ such that $\big\|T_\alpha f - g_{i(y,\alpha)}\big\|_y < \epsilon$. $T_\alpha S_\alpha$-almost periodicity is similarly defined.

One may easily check that linear combinations of T_α-almost periodic functions are T_α-almost periodic, as are products. Furthermore the constants are T_α-almost periodic. By Theorem 2.7, the closure of the set of T_α-almost periodic functions has the form $L^2(X, \mathcal{B}_1, \mu)$ for some σ-algebra $\mathcal{B}_1 \subset \mathcal{A}$.

One easily sees that $\mathcal{B} \subset \mathcal{B}_1$. Similarly, the closure of the set of $T_\alpha S_\alpha$-almost periodic functions has the form $L^2(X, \mathcal{B}_2, \mu)$ for some σ-algebra $\mathcal{B}_2$ with $\mathcal{B} \subset \mathcal{B}_2 \subset \mathcal{A}$.

For $H \in L^2(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$, we define an operator $\mathbf{H} : \phi \rightarrow \mathbf{H} * \phi$ by

$$\mathbf{H} * \phi(x) = \int H(x, x')\phi(x')\, d\mu_{\pi(x)}(x'). \tag{3.2}$$

For a.e. $y \in Y$, this equation defines a compact (in fact Hilbert-Schmidt) operator $\mathbf{H}$ on $L^2(X, \mathcal{A}, \mu_y)$. Alternatively, of course, $\mathbf{H}$ is an operator on $L^2(X, \mathcal{A}, \mu)$.

Lemma 3.3. *(a) If $H \in L^2(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$ satisfies $Q_1 H = H$ and $\phi \in L^\infty(X, \mathcal{A}, \mu)$ then $\mathbf{H} * \phi$ is T_α-almost periodic over $\mathcal{B}$ along $\mathcal{F}^{(1)}$.*

(b) If $H \in L^2(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$ satisfies $Q_2 H = H$ and $\phi \in L^\infty(X, \mathcal{A}, \mu)$ then $\mathbf{H} * \phi$ is $T_\alpha S_\alpha$-almost periodic over $\mathcal{B}$ along $\mathcal{F}^{(1)}$.

Proof. We will prove (a). Part (b) is similar. Let $\alpha_1, \alpha_2, \ldots$ be an enumeration of the atoms which comprise $\mathcal{F}^{(1)}$ (that is, $\mathcal{F}^{(1)} = FU\{\alpha_1, \alpha_2, \ldots\}$). Suppose that $H \in L^2(\tilde{X}, \tilde{\mathcal{A}}, \tilde{\mu})$ satisfies $Q_1 H = H$. Let $\epsilon > 0$ be arbitrary. For a.e. $y \in Y$, the operator $\mathbf{H}$ defined by (3.2) is a compact operator on $L^2(X, \mathcal{A}, \mu_y)$, hence there exists a number $M(y) \in \mathbf{N}$ such that

$$\big\{\mathbf{H} * (T_\alpha f) : \alpha \in FU_\emptyset\{\alpha_1, \ldots, \alpha_{M(y)}\}\big\}$$

is $\frac{\epsilon}{2}$-dense (under the metric $\rho(g, h) = \big\|g - h\big\|_y$) in the set $\{\mathbf{H} * (T_\alpha f) : \alpha \in \mathcal{F}_\emptyset\}$. Let M be so large that there exists a set $D \in \mathcal{B}$ with $\nu(D) < \epsilon$ such that $M > M(y)$ for all $y \in Y \setminus D$. Now let $\{g_1, \ldots, g_N\}$ be an enumeration of $\big\{\mathbf{H} * (T_\alpha f) : \alpha \in FU_\emptyset\{\alpha_1, \ldots, \alpha_M\}\big\}$.

For any $y \in Y \setminus D$ and any $\alpha \in \mathcal{F}$ there exists $i(y,\alpha) \in \mathbf{N}$ with $1 \le i(y,\alpha) \le M$ such that $\big\|\mathbf{H} * (T_\alpha \phi) - g_{i(y,\alpha)}\big\|_y < \frac{\epsilon}{2}$. Letting $\delta > 0$ be arbitrary,

$$\begin{aligned}
\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ & \big\|T_\alpha(\mathbf{H} * \phi) - \mathbf{H} * (T_\alpha \phi)\big\|^2 \\
&= \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \Big| \int \big(H(T_\alpha x, T_\alpha x') - H(x,x')\big)\phi(T_\alpha x')\, d\mu_{\pi(x)}(x') \Big|^2\, d\mu(x) \\
&\le \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \int \big|H(T_\alpha x, T_\alpha x') - H(x,x')\big|^2 \big|\phi(T_\alpha x')\big|^2\, d\mu_{\pi(x)}(x')\, d\mu(x) \\
&\le \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|\tilde{T}_\alpha H - H\big\|^2_{L^2(\tilde{X},\tilde{\mathcal{A}},\tilde{\mu})} \big\|\phi\big\|^2_\infty = 0.
\end{aligned}$$

Therefore, there exists $\alpha_0 \in \mathcal{F}^{(1)}$ having the property that for every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$, $\big\|T_\alpha(\mathbf{H} * \phi) - \mathbf{H} * (T_\alpha \phi)\big\|$ is so small that there exists a set $E(\alpha) \in \mathcal{B}$ with $\nu\big(E(\alpha)\big) < \delta$ such that $\big\|T_\alpha(\mathbf{H} * \phi) - \mathbf{H} * (T_\alpha \phi)\big\|_y < \frac{\epsilon}{2}$ for all $y \in Y \setminus E(\alpha)$. If now $y \in Y \setminus \big(D \cup E(\alpha)\big)$, then $\big\|T_\alpha(\mathbf{H} * \phi) - g_{i(y,n_\alpha)}\big\|_y < \epsilon$. □

Lemma 3.4. *(a) If $f \in L^\infty(X,\mathcal{A},\mu)$ satisfies $E(f|\mathcal{B}_1) = 0$, then*

$$\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|P(fT_\alpha f)\big\| = 0.$$

(b) If $f \in L^\infty(X,\mathcal{A},\mu)$ satisfies $E(f|\mathcal{B}_2) = 0$, then

$$\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|P(fT_\alpha S_\alpha f)\big\| = 0.$$

Proof. Again we prove only (a), as (b) is similar. By Lemma 3.3 and the fact that $E(f|\mathcal{B}_1) = 0$, f is orthogonal to $\mathbf{H} * f$ for every $H \in L^2(\tilde{X},\tilde{\mathcal{A}},\tilde{\mu})$ satisfying $Q_1 H = H$. It follows that the function $f \otimes f(x,x') = f(x)f(x') \in L^2(\tilde{X},\tilde{\mathcal{A}},\tilde{\mu})$ is orthogonal to all $H \in L^2(\tilde{X},\tilde{\mathcal{A}},\tilde{\mu})$ which satisfy $Q_1 H = H$. To see this, note that

$$\begin{aligned}
\int & f(x)f(x')H(x,x')\, d\tilde{\mu}(x,x') \\
&= \int f(x) \int H(x,x')f(x')\, d\mu_{\pi(x)}(x')\, d\mu(x) \\
&= \int f(x)\big(\mathbf{H} * f(x)\big)\, d\mu(x) = \langle \mathbf{H} * f, f\rangle = 0.
\end{aligned}$$

Since $Q_1(Q_1 H) = Q_1^2 H = Q_1 H$, it follows that $f \otimes f$ is orthogonal to $Q_1 H$ for all $H \in L^2(\tilde{X},\tilde{\mathcal{A}},\tilde{\mu})$. Hence,

$$\begin{aligned}
\underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|P(fT_\alpha f)\big\|^2 &= \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \Big| \int f(x) T_\alpha f(x)\, d\mu_y(x)\Big|^2\, d\nu(y) \\
&= \underset{\alpha\in\mathcal{F}^{(2)}}{\text{IP-lim}} \int f(x)f(x')T_\alpha f(x) T_\alpha f(x')\, d\tilde{\mu}(x,x') \\
&= \int (f \otimes f) Q_1 (f \otimes f)\, d\tilde{\mu} = 0.
\end{aligned}$$ □

Lemma 3.5. *If $f, g \in L^\infty(X, \mathcal{A}, \mu)$ with either $E(f|\mathcal{B}_1) = 0$ or $E(g|\mathcal{B}_2) = 0$, then there exists an IP-ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that*

$$\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(2)}} \ T_\alpha f T_\alpha S_\alpha g = 0$$

in the weak topology.

Proof. Let $x_\alpha = T_\alpha f T_\alpha S_\alpha g$. Recall that for $\beta > \alpha$, $T_{\alpha\cup\beta} f = T_\beta(T_\alpha f)$. Therefore

$$\begin{aligned}
&\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ \operatorname*{IP-lim}_{\beta \in \mathcal{F}^{(1)}} \ \langle x_\beta, x_{\alpha\cup\beta}\rangle \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ \operatorname*{IP-lim}_{\beta \in \mathcal{F}^{(1)}} \int T_\beta f T_\beta S_\beta g T_{\alpha\cup\beta} f T_{\alpha\cup\beta} S_{\alpha\cup\beta} g \, d\mu \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ \operatorname*{IP-lim}_{\beta \in \mathcal{F}^{(1)}} \int T_\beta f T_\beta S_\beta g T_\beta(T_\alpha f) T_\beta S_\beta(T_\alpha S_\alpha g) \, d\mu \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ \operatorname*{IP-lim}_{\beta \in \mathcal{F}^{(1)}} \int (f T_\alpha f) S_\beta (g T_\alpha S_\alpha g) \, d\mu \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \int P(f T_\alpha f) P(g T_\alpha S_\alpha g) \, d\mu = 0.
\end{aligned}$$

The desired conclusion is now a consequence of Proposition 2.6. □

We now proceed to show that $L > 0$ (see (3.1)). Let $f = 1_A$, $f_1 = E(f|\mathcal{B}_1)$, and $f_2 = E(f|\mathcal{B}_2)$. Also let $h_1 = f - f_1$ and $h_2 = f - f_2$. By Lemma 3.5 we may by passing to a subring assume that

$$\begin{aligned}
&\operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ T_\alpha f_1 T_\alpha S_\alpha h_2 \\
&\qquad = \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ T_\alpha h_1 T_\alpha S_\alpha f_2 \\
&\qquad = \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ T_\alpha h_1 T_\alpha S_\alpha h_2 = 0
\end{aligned}$$

in the weak topology. Then

$$\begin{aligned}
L &= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \ \mu(A \cap T_\alpha^{-1} A \cap (T_\alpha S_\alpha)^{-1} A) \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \int f T_\alpha f T_\alpha S_\alpha f \, d\mu \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \int f T_\alpha (f_1 + h_1) T_\alpha S_\alpha (f_2 + h_2) \, d\mu \\
&= \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu.
\end{aligned}$$

It is therefore sufficient for the proof of Theorem 3.1 to show that

$$L = \operatorname*{IP-lim}_{\alpha \in \mathcal{F}^{(1)}} \int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu > 0. \tag{3.3}$$

It is clear from the decomposition of measures that $f_1(x)f_2(x) > 0$ for a.e. $x \in A$. Therefore, there exists some $a > 0$ and a set $A' \subset A$ with $\mu(A') > 0$ such that $f_1(x)f_2(x) > a$ for all $x \in A'$. Furthermore, there exist numbers $b, \xi > 0$ and a set $B_1 \in \mathcal{B}$ with $\nu(B_1) = 3\xi > 0$ such that for all $y \in B_1$, $\mu_y(A') > b$. It follows that

$$\int f f_1 f_2 \, d\mu_y > ab \tag{3.4}$$

for all $y \in B_1$.

Let $\epsilon = \frac{ab}{18}$. We may approximate f_1 arbitrarily closely by a function ϕ_1 which is T_α-almost periodic over $\mathcal{B}$ along $\mathcal{F}^{(1)}$. Likewise, we may approximate f_2 arbitrarily closely by a function ϕ_2 which is $T_\alpha S_\alpha$-almost periodic over $\mathcal{B}$ along $\mathcal{F}^{(1)}$. Since $\nu(B_1) > 3\xi$ we may therefore fix such ϕ_1 and ϕ_2 so that there exists a set $B_2 \subset B_1$ with $\nu(B_2) > 2\xi$ having the property that for all $y \in B_2$ we have

$$\left\| f_1 - \phi_1 \right\|_y < \epsilon \tag{3.5}$$

and

$$\left\| f_2 - \phi_2 \right\|_y < \epsilon. \tag{3.6}$$

By the definition of almost periodicity, there exists $\{g_1, \ldots, g_M\} \subset L^2(X, \mathcal{A}, \mu)$ and $D \in \mathcal{B}$ with $\nu(D) < \xi$ such that:

$(*)$ For every $\delta > 0$ there exists $\alpha_0 \in \mathcal{F}^{(1)}$ having the property that for every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ there exists a set $E(\alpha) \in \mathcal{B}$ with $\nu(E(\alpha)) < \delta$ such that for every $y \in Y \backslash \big(D \cup E(\alpha)\big)$ there exist $i(y, \alpha)$ and $j(y, \alpha)$ with $1 \le i(y, \alpha), j(y, \alpha) \le M$ such that

$$\left\| T_\alpha \phi_1 - g_{i(y,\alpha)} \right\|_y < \epsilon$$

and

$$\left\| T_\alpha S_\alpha \phi_2 - g_{j(y,\alpha)} \right\|_y < \epsilon.$$

We claim that

$$L = \underset{\alpha \in \mathcal{F}^{(1)}}{\text{IP-lim}} \int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu \ge \frac{ab\xi^{2^N}}{8(M^2+1)^2}.$$

If this were not the case, we could, by passing to a subring, assume that for all $\alpha \in \mathcal{F}^{(1)}$,

$$\int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu < \frac{ab\xi^{2^N}}{8(M^2+1)^2}.$$

We will show that this is impossible by producing an $\alpha \in \mathcal{F}^{(1)}$ for which

$$\int fT_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu > \frac{ab\xi^{2^N}}{8(M^2+1)^2}\,, \tag{3.7}$$

whereupon the proof of Theorem 3.1 will have been completed.

Let $N = M^2+1$. By setting $\delta = \frac{1}{6}\xi^{2^N}2^{-2N}$ in $(*)$ we may, by deleting finitely many atoms of $\mathcal{F}^{(1)}$, assume that for every $\alpha \in \mathcal{F}^{(1)}$ there exists a set $E(\alpha) \in \mathcal{B}$ with $\nu\big(E(\alpha)\big) < \frac{1}{6}\xi^{2^N}2^{-N}\xi$ having the property that for every $y \in Y \setminus \big(D \cup E(\alpha)\big)$ there exist $i(y,\alpha)$ and $j(y,\alpha)$ with $1 \le i(y,\alpha), j(y,\alpha) \le M$ such that

$$\big\|T_\alpha\phi_1 - g_{i(y,\alpha)}\big\|_y < \epsilon \tag{3.8}$$

and

$$\big\|T_\alpha S_\alpha \phi_2 - g_{j(y,\alpha)}\big\|_y < \epsilon. \tag{3.9}$$

Let $B_3 = (B_2 \setminus D)$. Since $\nu(B_2) > 2\xi$ and $\nu(D) < \xi$, we have $\nu(B_3) > \xi$. According to Theorem 2.8 we may by passing to a subring assume that for every $(\alpha_1,\dots,\alpha_N) \in (\mathcal{F}^{(1)})^N_<$ we have

$$\nu\Big(\bigcap_{\alpha \in FU_\emptyset\{\alpha_1,\dots,\alpha_N\}} T_\alpha^{-1}B_3\Big) > \xi^{2^N}.$$

Furthermore, since $B_3 \in \mathcal{B}$ we have

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} \nu(T_\alpha B_3 \triangle T_\alpha S_\alpha B_3) = \operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(1)}} \nu(B_3 \triangle S_\alpha B_3) = 0.$$

Therefore we may in fact assume that for all $(\alpha_1,\dots,\alpha_N) \in (\mathcal{F}^{(1)})^N_<$ we have

$$\nu\Big(\bigcap_{\alpha \in FU_\emptyset\{\alpha_1,\dots,\alpha_N\}} \big(T_\alpha^{-1}B_3 \cap (T_\alpha S_\alpha)^{-1}B_3\big)\Big) > \xi^{2^N}. \tag{3.10}$$

Let $\delta > 0$ be small (how small we shall say shortly). We claim that by passing to a subring we may assume that for all $(\alpha,\beta) \in (\mathcal{F}^{(1)})^2_<$ we have

$$\int \Big| \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta y} - \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta S_\beta y}\Big|^2 d\nu(y) < \delta. \tag{3.11}$$

If this were not the case, then by the Theorem 2.6 we could by passing to a subring assume that

$$\int \Big| \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta y} - \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta S_\beta y}\Big|^2 d\nu(y) \ge \delta \tag{3.12}$$

for all $(\alpha,\beta)\in(\mathcal{F}^{(1)})^2_<$. However, recall that

$$\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|T_\beta h - T_\beta S_\beta h\big\| = \underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \big\|h - S_\beta h\big\| = 0$$

for all $h\in L^2(X,\mathcal{B},\mu)$. It follows that for any $\alpha\in\mathcal{F}^{(1)}$ we have

$$\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \int\Big|\big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta y} - \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta S_\beta y}\Big|^2\,d\nu(y) = 0,$$

a contradiction to (3.12). This establishes that we may assume that (3.11) holds for all $(\alpha,\beta)\in(\mathcal{F}^{(1)})^2_<$.

Recall that δ was chosen small. It is chosen small enough that (3.11) implies that for all $(\alpha,\beta)\in(\mathcal{F}^{(1)})^2_<$ there exists a set $C(\alpha,\beta)\in\mathcal{B}$ with $\nu\big(C(\alpha,\beta)\big)<\frac{\xi^{2^N}}{4N^2}$ such that for all $y\in Y\setminus C(\alpha,\beta)$ we have

$$\Big|\big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta y} - \big\|f_2 - T_\alpha S_\alpha f_2\big\|_{T_\beta S_\beta y}\Big| < \epsilon. \tag{3.13}$$

We now fix some N-tuple $(\alpha_1,\ldots,\alpha_N)\in\big(\mathcal{F}^{(1)}\big)^N_<$. Let

$$B_4 = B_3\setminus\Big(\bigcup_{\alpha\in FU_\emptyset\{\alpha_1,\ldots,\alpha_N\}} E(\alpha)\Big).$$

We now put

$$\begin{aligned} B_5 = \bigg(\Big(&\bigcap_{\beta\in FU_\emptyset\{\alpha_1,\ldots,\alpha_N\}}\big(T_\beta^{-1}B_3\cap(T_\beta S_\beta)^{-1}B_3\big)\Big)\bigg)\\ &\setminus\Big(\bigcup_{\alpha,\beta\in FU_\emptyset\{\alpha_1,\ldots,\alpha_N\}}\big(E(\alpha)\cup T_\beta^{-1}E(\alpha)\cup(T_\beta S_\beta)^{-1}E(\alpha)\big)\Big).\end{aligned}$$

Then

(i) $\nu(B_5)>\frac{1}{2}\xi^{2^N}$.

(ii) For any $y\in B_4$, (3.4), (3.5), (3.6), (3.8) and (3.9) hold.

(iii) For any $y\in B_5$ and any $\beta\in FU_\emptyset\{\alpha_1,\ldots,\alpha_N\}$ we have $T_\beta y\in B_4$ and $T_\beta S_\beta y\in B_4$.

Since $N=M^2+1$, for any $y\in B_5$ there exist $l(y)$ and $m(y)$ with $1\le l(y)<m(y)\le N$ such that

$$i\big(y,(\alpha_{l(y)}\cup\cdots\cup\alpha_N)\big) = i\big(y,(\alpha_{m(y)}\cup\cdots\cup\alpha_N)\big) \tag{3.14}$$

and

$$j\big(y,(\alpha_{l(y)}\cup\cdots\cup\alpha_N)\big) = j\big(y,(\alpha_{m(y)}\cup\cdots\cup\alpha_N)\big). \tag{3.15}$$

There are less than N^2 possibilities for the pair $\big(l(y), m(y)\big)$. Therefore, since $\nu(B_5) > \frac{1}{2}\xi^{2^N}$, we may choose l and m with $1 \le l < m \le N$ and a set $B_6 \subset B_5$ with $\nu(B_6) > \frac{\xi^{2^N}}{2N^2}$ on which $l(y) = l$ and $m(y) = m$ are constant. Let $\alpha = (\alpha_l \cup \cdots \cup \alpha_{m-1})$ and let $\beta = (\alpha_m \cup \cdots \cup \alpha_N)$. Then $\alpha < \beta$, so $(\alpha, \beta) \in \big(\mathcal{F}^{(1)}\big)^2_<$.

Let $y \in B_6$. (3.14) says that $i(y, \alpha \cup \beta) = i(y, \beta)$. Since $B_6 \subset Y \setminus \big(D \cup E(\beta) \cup E(\alpha \cup \beta)\big)$ we have by (3.8) that

$$
\begin{aligned}
\|\phi_1 - T_\alpha \phi_1\|_{T_\beta y} &= \|T_\beta \phi_1 - T_\beta (T_\alpha \phi_1)\|_y \\
&= \|T_\beta \phi_1 - T_{\alpha\cup\beta} \phi_1\|_y \\
&\le \|T_\beta \phi_1 - g_{i(y,\beta)}\|_y + \|T_{\alpha\cup\beta}\phi_1 - g_{i(y,\beta)}\|_y \\
&= \|T_\beta \phi_1 - g_{i(y,\beta)}\|_y + \|T_{\alpha\cup\beta}\phi_1 - g_{i(y,\alpha\cup\beta)}\|_y \\
&< 2\epsilon.
\end{aligned} \tag{3.16}
$$

Also, since $T_\beta y \in B_3$, we have

$$\|\phi_1 - f_1\|_{T_\beta y} < \epsilon.$$

On the other hand, since $T_{\alpha\cup\beta} y \in B_3$ we have

$$\|T_\alpha f_1 - T_\alpha \phi_1\|_{T_\beta y} = \|f_1 - \phi_1\|_{T_{\alpha\cup\beta} y} < \epsilon. \tag{3.17}$$

(3.16), (3.17) and (3.18) give

$$\|f_1 - T_\alpha f_1\|_{T_\beta y} \le \|f_1 - \phi_1\|_{T_\beta y} + \|\phi_1 - T_\alpha \phi_1\|_{T_\beta y} + \|T_\alpha \phi_1 - T_\alpha f_1\|_{T_\beta y} < 4\epsilon. \tag{3.18}$$

In a completely analogous fashion we may show as well that

$$\|f_2 - T_\alpha S_\alpha f_2\|_{T_\beta S_\beta y} < 4\epsilon. \tag{3.19}$$

Recall that since $(\alpha, \beta) \in \big(\mathcal{F}^{(1)}\big)^2_<$ there exists a set $C(\alpha, \beta) \in \mathcal{B}$ with $\nu\big(C(\alpha, \beta)\big) < \frac{\xi^{2^N}}{4N^2}$ such that for all $z \in Y \setminus C(\alpha, \beta)$ we have

$$\Big| \|f_2 - T_\alpha S_\alpha f_2\|_{T_\beta z} - \|f_2 - T_\alpha S_\alpha f_2\|_{T_\beta S_\beta z} \Big| < \epsilon. \tag{3.20}$$

Let $B_7 = B_6 \setminus C(\alpha, \beta)$. Then $\nu(B_7) > \frac{\xi^{2^N}}{4N^2}$. Suppose now that $y \in B_7$, (3.20) and (3.21) combine to give

$$\|f_2 - T_\alpha S_\alpha f_2\|_{T_\beta y} < 5\epsilon. \tag{3.21}$$

Since $y \in B_5$, $T_\beta y \in B_4 \subset B_1$. It follows by (3.4) that

$$\int f f_1 f_2 \, d\mu_{T_\beta y} > ab.$$

Now from (3.19) and (3.22), together with the fact that the functions f, f_1 and f_2 have ranges in $[0,1]$, we get that

$$\int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu_{T_\beta y} > \frac{ab}{2}\,.$$

This holds for all $y \in B_7$. It follows that

$$\int f T_\alpha f_1 T_\alpha S_\alpha f_2 \, d\mu > \frac{ab\xi^{2^N}}{8(M^2+1)^2}\,.$$

This is (3.7). In light of previous comments, this completes the proof of Theorem 3.1. □

4. Combinatorial Consequences

This section contains a few combinatorial consequences of Theorem 3.1. They are new, and indeed unusual enough not to have any direct predecessors in the literature. The reader may wish to compare them to the results of [**BH**] and [**BMZ**].

Recall that a semigroup S is left amenable if there exists a left invariant mean m on $l^\infty(S)$. Namely, $m \in l^\infty(S)^*$ with $m(\mathbf{1}) = 1$ and $m(f) \geq 0$ if $f(s) \geq 0$ for all $s \in S$, and with $m({}_sf) = m(f)$ for all $s \in S$ and $f \in l^\infty(S)$, where ${}_sf(t) = f(st)$. If S is a semigroup, a subset $A \subset S$ is said to be **left syndetic** if there exists a finite set $H \subset S$ such that $\bigcup_{h\in H} h^{-1}A = S$, where $h^{-1}A = \{s \in S : hs \in A\}$.

Suppose that S is a countable, left amenable semigroup with identity and let $\Omega = \{0,1\}^S$. Ω is a compact metrizable space under the product topology. An S-action $\{T_g\}$ may be defined on Ω as follows: for $\xi \in \Omega$, let $(T_g\xi)(h) = \xi(hg)$. The following proposition follows Furstenberg.

Proposition 4.1. *Let S be a countable left amenable semigroup with identity e and let m be a left invariant mean. Let $X = \overline{\{T_h 1_E : h \in S\}}$ and put $A = \{\eta \in X : \eta(e) = 1\}$. For any $E \subset G$ there exists a $\{T_h\}$-invariant probability measure μ on X such that $\mu(A) = m(E)$.*

Proof. Let $\mathcal{A}$ be the algebra of sets generated by $\{T_g^{-1}A : g \in S\}$. If $g_1, g_2, \ldots, g_k \in S$ and $A_1, A_2, \ldots, A_k \in \{A, A^c\}$, put $\mu(T_{g_1}^{-1}A_1 \cap \cdots \cap T_{g_k}^{-1}A_k) = m(g_1^{-1}E_1 \cap \cdots \cap g_k^{-1}E_k)$, where $E_i = E$ if $A_i = A$ and $E_i = E^c$ if $A_i = A^c$. One easily checks that μ extends to an additive, $\{T_g\}$-invariant set-function on $\mathcal{A}$ which, by compactness of X and the fact that members of $\mathcal{A}$ are open, is a pre-measure. Hence μ extends to a measure on the Borel σ-algebra, and plainly $\mu(A) = m(E)$. □

Let G and H be left amenable semigroups with identities e and e', respectively. $G \times H$ is left amenable as well, and any $G \times H$-action $\{U_{(g,h)} : (g,h) \in G \times H\}$ has the form $U_{(g,h)} = T_g S_h$, where $T_g = U_{(g,e')}$ and $S_g = U_{(e,g)}$. If H is abelian, we can apply Theorem 3.1 to obtain the following density combinatorial result, a "half commutative IP Roth theorem for amenable semigroups".

Theorem 4.2. *Suppose that G is a countable, left amenable semigroup, H is a countable abelian semigroup, and $E \subset G \times H$ has positive upper density. Let $(g_\alpha)_{\alpha \in \mathcal{F}} \subset G$ be an IP-system and let $(h_\alpha)_{\alpha \in \mathcal{F}} \subset H$ be an IP-system. There exists $a \in G$, $b \in H$, and $\alpha \in \mathcal{F}$ such that $\{(a,b), (g_\alpha a, b), (g_\alpha a, h_\alpha b)\} \subset E$.*

Proof. Let $G' = G \cup \{e\}$ and let $H' = H \cup \{e'\}$, where e and e' are identities. Let $\Omega = \{0,1\}^{G' \times H'}$, and define a $G' \times H'$-action $\{U_{(g,h)}\}$ on Ω by $\big(U_{(g,h)}\xi\big)(a,b) = \xi(ag, bh)$. Let $\xi = 1_E \in \Omega$ and let $X = \overline{\{U_{(g,h)}\xi :\ g,h \in G\}}$. Put $A = \{\eta \in X : \eta(e,e') = 1\}$. According to Theorem 4.1 there exists a $\{U_{(g,h)}\}$-invariant measure μ on X with $\mu(A) > 0$. Let $T_\alpha = U_{(g_\alpha, e')}$ and $S_\alpha = U_{(e,h_\alpha)}$, $\alpha \in \mathcal{F}$. Then $\{T_\alpha\}$ is an IP-system and $\{S_\alpha\}$ is a commutative IP-system of measure preserving transformations on X. By Theorem 3.1 there exists $\alpha \in \mathcal{F}$ such that $\mu(A \cap T_\alpha^{-1}A \cap (T_\alpha S_\alpha)^{-1}A) > 0$. Choose $\eta \in A \cap T_\alpha^{-1}A \cap (T_\alpha S_\alpha)^{-1}A$. Since A is open and $\eta \in \overline{\{T_g S_h \xi :\ (g,h) \in G \times H\}}$, there exists $(a,b) \in G \times H$ such that $T_a S_\beta \xi \in \big(A \cap T_\alpha^{-1}A \cap (T_\alpha S_\alpha)^{-1}A\big)$. Therefore

$$\xi(a,b) = \xi(g_\alpha a, b) = \xi(g_\alpha a, h_\alpha b) = 1.$$

In other words, $\{(a,b), (g_\alpha a, b), (g_\alpha a, h_\alpha b)\} \subset E$. $\square$

Remark. In the above theorem, it is not absolutely necessary that H be abelian. It is enough to have H left amenable, and that the IP-system (h_α) be contained in an abelian subgroup H_0 of H. The conclusion in this case is mildly stronger because a positive density subset $E \subset G \times H$ need not intersect $G \times H_0$. For example, if H is the (amenable) group of finite permutations of $\mathbf{N}$, then a subgroup of H is abelian if and only if it is generated by disjoint cycles, and all such subgroups have zero upper density. Similar observations apply to the results to follow, but we shall content ourselves with assuming H to be abelian as the statements are then more natural and the difference is minor.

In the event that G is a group, we can modify Theorem 4.2 such as to give the result a somewhat more natural look.

Corollary 4.3. *Suppose that G is a countable amenable group, H is a countable abelian semigroup, and $E \subset G \times H$ has positive upper density. Let $(g_\alpha)_{\alpha \in \mathcal{F}} \subset G$ be a reversed IP-system and let $(h_\alpha)_{\alpha \in \mathcal{F}} \subset H$ be an IP-system. There exists $a \in G$, $b \in H$, and $\alpha \in \mathcal{F}$ such that $\{(a,b), (g_\alpha a, b), (a, h_\alpha b)\} \subset E$.*

Proof. Note that $(g_\alpha^{-1})_{\alpha \in \mathcal{F}}$ is an IP-system. Hence by Theorem 4.2 there exists $(c,b) \in G \times H$ and $\alpha \in \mathcal{F}$ such that $(c,b), (g_\alpha^{-1}c, b), (g_\alpha^{-1}c, h_\alpha b)\} \subset E$. Let $a = g_\alpha^{-1}c$. $\square$

Our next result is a "one-third commutative" topological multiple recurrence theorem.

Theorem 4.4. *Let J be an arbitrary countable group, let G be a countable left amenable semigroup, let H be a countable abelian semigroup, let (X,ρ) be a compact metric space, and let $\{T_g\}$, $\{R_g\}$, and $\{S_g\}$ be actions of G, J, and H respectively by homeomorphisms of X, such that $T_gR_j = R_jT_g$, $R_jS_h = S_hR_j$, and $T_gS_h = S_hT_g$ for all $g \in G$, $j \in J$, and $h \in H$. Let $(g_\alpha)_{\alpha\in\mathcal{F}} \subset G$ and $(j_\alpha)_{\alpha\in\mathcal{F}} \subset J$ be IP-systems, and let $(h_\alpha)_{\alpha\in\mathcal{F}} \subset H$ be an IP-system. Then for every $\epsilon > 0$ there exists $\alpha \in \mathcal{F}$ and $x \in X$ such that $\rho(x, R_{j_\alpha}x) < \epsilon$, $\rho(x, R_{j_\alpha}T_{g_\alpha}x) < \epsilon$, and $\rho(x, R_{j_\alpha}T_{g_\alpha}S_{h_\alpha}x) < \epsilon$.*

Proof. Passing to a closed subset of X if necessary, we will assume that X is minimal with respect to the $G\times J\times H$-action $\{T_gR_jS_h : (g,j,h) \in G\times J\times H\}$. We claim that for every non-empty open set $U \subset X$ and every $\alpha_0 \in \mathcal{F}$ there exists $\alpha > \alpha_0$ and $z \in X$ such that $\{z, T_{g_\alpha}x, T_{g_\alpha}S_{h_\alpha}x\} \subset U$.

To prove the claim, let $U \subset X$ be open. Pick $x \in U$ 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 $G\times H$-action $\{T_gS_h : (g,h) \in G\times H\}$. One may check that $\overline{\bigcup_{j\in J} R_jY}$ is $\{R_j\}$-, $\{T_g\}$-, and $\{S_h\}$-invariant, and is therefore equal to X. It follows that for some $j_0 \in J$, $R_{j_0}^{-1}B_{\epsilon/2}(x)\cap Y \neq \emptyset$. Let $\delta > 0$ be so small that if $y, y' \in Y$ with $\rho(y,y') < \delta$ then $\rho(R_{j_0}y, R_{j_0}y') < \frac{\epsilon}{2}$. Let $U' \subset R_{j_0}^{-1}B_{\frac{\epsilon}{2}}(x)\cap Y$ be an open set (open in Y) of diameter less that δ. Let $y_0 \in Y$. Since the action $\{T_gS_h : (g,h) \in G\times H\}$ is minimal on Y, the set

$$E = \{(g,h) : T_gS_hy_0 \in U'\}$$

is left syndetic in $G\times H$, and therefore we have $m(E) > 0$ for every left-invariant mean m on $G\times H$. It follows from Theorem 4.2 that for some $\alpha > \alpha_0$ and $(a,b) \in G\times H$, $\{(a,b), (g_\alpha a, b), (g_\alpha a, h_\alpha b)\} \subset E$. Set $y = T_aS_by_0 \in U'$. Then $T_{g_\alpha}y \in U'$ and $T_{g_\alpha}S_{h_\alpha}y \in U'$, so that, letting $z = R_{j_0}y$, we have $z \in B_{\epsilon/2}(x)$, $\rho(z, T_{g_\alpha}z) < \epsilon/2$, and $\rho(z, T_{g_\alpha}S_{h_\alpha}z) < \epsilon/2$. Therefore $\{z, T_{g_\alpha}z, T_{g_\alpha}S_{h_\alpha}z\} \subset U$.

Let $\epsilon > 0$ and choose $x_0 \in X$ arbitrarily. Let U_0 be an open set of diameter less that $\frac{\epsilon}{2}$ containing x_0. According to the claim, there exists $\alpha_0 \in \mathcal{F}$ and $y_0 \in X$ such that $\{y_0, T_{g_{\alpha_0}}y_0, T_{g_{\alpha_0}}S_{h_{\alpha_0}}y_0\} \subset U_0$. Put $x_1 = R_{j_{\alpha_0}}^{-1}y_0$ and let U_1 be an open set of diameter less that $\frac{\epsilon}{2}$ containing x_1 having the property that for every $x \in U_1$ we have $\{R_{j_{\alpha_0}}x, T_{j_{\alpha_0}}T_{g_{\alpha_0}}x, R_{j_{\alpha_0}}T_{g_{\alpha_0}}S_{h_{\alpha_0}}x\} \subset U_0$.

Suppose now that we have chosen open sets of diameter less that $\frac{\epsilon}{2}$ $U_0, U_1, \ldots, U_t$ containing points $x_0, x_1, \ldots, x_t$, and $\alpha_0, \alpha_1, \ldots, \alpha_{t-1} \in \mathcal{F}$, with $\alpha_0 < \alpha_1 < \cdots < \alpha_{t-1}$, such that whenever $0 \leq m < n \leq t$ we have $\{R_{j_{m,n}}x, R_{j_{m,n}}T_{g_{m,n}}x, R_{j_{m,n}}T_{g_{m,n}}S_{h_{m,n}}x\} \subset U_i$ for all $x \in U_j$, where here $j_{m,n} = j_{\alpha_m\cup\alpha_{m+1}\cup\cdots\cup\alpha_{n-1}}$ (and similarly for $g_{m,n}$ and $h_{m,n}$). By the previous clain there exists $\alpha_t > \alpha_{t-1}$ and $y_t \in U_t$ with $\{y_t, T_{g_{\alpha_t}}y_t, T_{g_{\alpha_t}}S_{h_{\alpha_t}}y_t\} \subset U_t$. Let $x_{t+1} = R_{j_{\alpha_t}}^{-1}y_t$ and let U_{t+1} be an open set of diameter less that $\frac{\epsilon}{2}$ containing x_{t+1} and having the property that for all $x \in U_{t+1}$ we have $\{R_{j_{\alpha_t}}x, R_{j_{\alpha_t}}T_{g_{\alpha_t}}x, R_{j_{\alpha_t}}T_{g_{\alpha_t}}S_{h_{\alpha_t}}x\} \subset U_t$. It follows that for $0 \le i < t$ and $x \in U_{t+1}$,

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

Continue until for some $m < n$, $\rho(x_m, x_n) < \frac{\epsilon}{2}$. Then

$$\rho(x_n, R_{j_{m,n}}x_n) < \epsilon,\ \rho(x_n, R_{j_{m,n}}T_{g_{m,n}}x_n) < \epsilon,\ \text{ and } \rho(x_n, R_{j_{i,j}}T_{g_{i,j}}S_{h_{i,j}}x_n) < \epsilon.$$

Letting $x = x_n$ and $\alpha = \alpha_m \cup \alpha_{m+1} \cup \cdots \cup \alpha_{n-1}$ finishes the proof. □

Van der Waerden's theorem ([**vdW**]; see also [**F2**]) states that for any finite coloring of $\mathbf{N}$, at least one cell contains arbitrarily long arithmetic progressions. The following theorem can be viewed as an exotic version of this result.

Corollary 4.5. *Let J be an arbitrary countable group, let G be a countable left amenable semigroup, and let H be a countable abelian semigroup. Let $(g_\alpha)_{\alpha\in\mathcal{F}} \subset G$ and $(j_\alpha)_{\alpha\in\mathcal{F}} \subset J$ be IP-systems, and let $(h_\alpha)_{\alpha\in\mathcal{F}} \subset H$ be an IP-system. Then for any finite partition $J \times G \times H = \bigcup_{i=1}^r C_i$, there exists t, with $1 \le t \le r$, $\alpha \in \mathcal{F}$, and $(a, b, c) \in J\times G\times H$ such that $\{(a, b, c), (j_\alpha a, b, c), (j_\alpha a, g_\alpha b, c), (j_\alpha a, g_\alpha b, h_\alpha c)\} \subset C_t$.*

Proof. Let e be an identity for J, let e' be an identity for G, and let e'' be an identity for H (e' and e'' are supplied, if necessary). Put $G' = G \cup \{e'\}$ and $H' = H \cup \{e''\}$. Let $\Omega = \{1, 2, \ldots, r\}^{J\times G'\times H'}$. 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 of homeomorphisms $\{R_g\}, \{S_g\}$ and $\{T_g\}$ can be defined on Ω by $R_g\gamma(a, b, c) = \gamma(ag, b, c)$, $S_g\gamma(a, b, c) = \gamma(a, bg, c)$, and $T_g\gamma(a, b, c) = \gamma(a, b, cg)$. Let ξ be the element of Ω defined by $\xi(j, g, h) = i$ when $(j, g, h) \in C_i$. Let $X = \overline{\{R_jT_g\S_hxi : (j, g, h) \in J \times G \times H\}}$. By Theorem 4.4 there exists $x \in X$ and and $\alpha \in \mathcal{F}$ such that $\rho(x, R_{j_\alpha}x) < \epsilon$, $\rho(x, R_{j_\alpha}T_{g_\alpha}x) < \epsilon$, and $\rho(x, R_{j_\alpha}T_{g_\alpha}S_{h_\alpha}x) < \epsilon$.

There exist $a, b, c \in G$ such that $y = R_aT_bS_c\xi$ is close enough to x that $\rho(y, R_{j_\alpha}y) < 1$, $\rho(y, R_{j_\alpha}T_{g_\alpha}y) < 1$, and $\rho(y, R_{j_\alpha}T_{g_\alpha}S_{h_\alpha}y) < 1$. It follows that $\xi(a, b, c) = \xi(j_\alpha a, b, c) = \xi(j_\alpha a, g_\alpha b, c) = \xi(j_\alpha a, g_\alpha b, h_\alpha c)$. In other words,

$$\{(a, b, c), (j_\alpha a, b, c), (j_\alpha a, g_\alpha b, c), (j_\alpha a, g_\alpha b, h_\alpha c)\} \subset C_i,$$

where $i = \xi(a, b, c)$. □

Other versions of Corollary 4.5 are possible. Taking $(j_\alpha)_{\alpha\in\mathcal{F}}$ to be a reversed IP-system, for example, yields a monochromatic configuration of the form

$$\{(a, b, c), (j_\alpha a, b, c), (a, g_\alpha b, c), (a, g_\alpha b, h_\alpha c)\}.$$

Taking G to be an amenable group and assuming that both $(j_\alpha)_{\alpha\in\mathcal{F}}$ and $(g_\alpha)_{\alpha\in\mathcal{F}}$ are reversed IP-systems yields a configuration of the form

$$\{(a, b, c), (j_\alpha a, g_\alpha b, c), (a, g_\alpha b, c), (a, b, h_\alpha c)\}.$$

References

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

[**BMZ**] Bergelson V., McCutcheon R. and Zhang Q., *A Roth theorem for amenable groups*, American Journal of Mathematics **119** (1997), 1173–1211.

[**F1**] Furstenberg H., *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math. **31** (1977), 204–256.

[**F2**] ———, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[**FK1**] Furstenberg H. and Katznelson Y., *An ergodic Szemerédi theorem for commuting transformations*, J. d'Analyse Math. **34** (1978), 275–291.

[**FK2**] ———, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. d'Analyse Math. **45** (1985), 117–168.

[**FK3**] ———, *A density version of the Hales-Jewett theorem*, J. d'Analyse Math. **57** (1991), 64–119.

[**H**] Hindman N., *Finite sums from sequences within cells of a partition of* **N**, J. Combinatorial Theory (Series A) **17** (1974), 1–11.

[**M**] Milliken K., *Ramsey's Theorem with sums or unions*, J. Combinatorial Theory (Series A) **18** (1975), 276–290.

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

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

[**T**] Taylor A., *A canonical partition relation for finite subsets of ω*, J. Combinatorial Theory (Series A) **21** (1976), 137–146.

[**vdW**] van der Waerden B. L., *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.

R. McCutcheon, Department of Mathematics, University of Maryland, College Park MD 20742, USA; *e-mail*: randall@math.umd.edu