Central Sets and a Non-Commutative Roth Theorem

V. Bergelson and R. McCutcheon

Measurable multiple recurrence results for non-nilpotent groups have up to now been limited to an ergodic Roth theorem [BMZ], which states that for any measure preserving actions $\{T_g\}_{g\in G}$ and $\{S_g\}_{g\in G}$ of a countable amenable group G on a probability space $(X,\mathcal{B},\mu)$ that commute in the sense $T_gS_h = S_hT_g$ for all $g,h\in G$, and any $A\in\mathcal{A}$ with $\mu(A)>0$, $\lim_n \frac{1}{|\Phi_n|}\sum_{g\in\phi_n}\mu\big(A\cap T_g^{-1}A\cap(T_gS_g)^{-1}A\big)>0$ for any Følner sequence (Φ_n) for G. This yields, in particular, that $\{g:\mu\big(A\cap T_g^{-1}A\cap(T_gS_g)^{-1}A\big)>0\}$ is syndetic. Here, using novel ultrafilter techniques for doing what might be called "ergodic theory without averaging", we remove the amenability condition in this result while simultaneously strengthening the conclusion.

1. Introduction.

Let T be an invertible measure preserving transformation on a probability space $(X,\mathcal{A},\mu)$ and let $A\in\mathcal{A}$ with $\mu(A)>0$. The classical and widely applicable Poincaré recurrence theorem states that $\mu(A\cap T^{-n}A)>0$ for some $n\in\mathbf{N}$. While nowadays this result is merely an exercise, some of its extensions are highly non-trivial, for instance the following theorem of H. Furstenberg from [F1].

Theorem 1.1. Let $(X,\mathcal{A},\mu)$ be a probability space and suppose $T:X\to X$ is invertible and μ-preserving. For any $k\in\mathbf{N}$ and any $A\in\mathcal{A}$ with $\mu(A)>0$, the set $R=\{n\in\mathbf{Z}: \mu(A\cap T^{-n}A\cap T^{-2n}A\cap\cdots\cap T^{-kn}A)>0\}$ is syndetic[1].

The significance of Theorem 1.1 was twofold. First, it provided an independent proof of Szemerédi's celebrated theorem [Sz] on arithmetic progressions[2]. Perhaps more importantly, its proof proceeded by way of the development of a deep structure theory for measure preserving $\mathbf{Z}$-actions that has proved susceptible to a wide variety of modifications and extensions, leading to many new, far-reaching multiple recurrence results and combinatorial applications. See for example, [FK1-3], [BL], [L], [BM1], [BMZ] and [BLM]. These various results, together with their applications, form the core body of what is today called *Ergodic Ramsey Theory.*

While the accumulated techniques and ideology of Ergodic Ramsey Theory have for some time formed an established, well understood, coherent body of knowledge for abelian or even nilpotent (semi-)groups, multiple recurrence results for more general groups are scarce indeed, being essentially restricted to the following double recurrence theorem for amenable groups from [BMZ].

Theorem 1.2. Let G be a countable amenable group and let $(T_g)_{g\in G}$, $(S_g)_{g\in G}$ be measure preserving G-actions on a probability space $(X,\mathcal{A},\mu)$ satisfying $T_gS_h=S_hT_g$ for all $g,h\in$

[1] If G is a group, a set $S\subset G$ is *left syndetic* (respectively *right syndetic*) if for some finite set $F\subset G$, one has $G=\bigcup_{g\in F}gS$ (respectively $G=\bigcup_{g\in F}Sg$). For commutative groups the notions clearly coincide, hence one says simply *syndetic.*

[2] For every $\delta>0$ and $k\in\mathbf{N}$ there is an $N\in\mathbf{N}$ such that for every set $E\subset\{1,\ldots,N\}$ with $|E|>\delta N$, E contains an arithmetic progression of length k.

G. Then for any $A \in \mathcal{A}$ with $\mu(A) > 0$ there exists $\lambda > 0$ such that

$$\{g \in G : \mu\big(A \cap T_g^{-1}A \cap (T_gS_g)^{-1}A\big) > \lambda\}$$

is both left and right syndetic.

Theorem 1.2 has a variety of applications for density and partition Ramsey theory in amenable groups; we mention here just one, for two others see [BMZ, Corollary 7.2] and [BM2, Theorem 3.4]. (In Section 4 of this paper, we give refinements of all three.)

Theorem 1.3. Suppose that G is a countable amenable group and $E \subset G \times G$ has positive upper density $\overline{d} = \limsup_n \frac{|E \cap \Phi_n|}{|\Phi_n|}$ with respect to a left Følner sequence [3] (Φ_n) for $G \times G$. Then

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

is both left and right syndetic in G.

One feature of Theorem 1.2 that we will be interested in improving here is the "largeness" of the set of double return times $R_A = \big\{g \in G : \mu\big(A \cap T_g^{-1}A \cap (T_gS_g)^{-1}A\big) > 0\big\}$ achieved therein. Although syndeticity is a strong and historically apt property, being the very one Furstenberg obtained for the set of multiple return times $\{n : \mu(A \cap T^{-n}A \cap \cdots \cap T^{-kn}A) > 0\}$ in [F1], there are good reasons to suspect that R_A has even stronger formulable largeness properties. For example, if $R_B = \big\{g \in G : \mu\big(B \cap T_g^{-1}B \cap (T_gS_g)^{-1}B\big) > 0\big\}$ is another set of double return times, then $R_A \cap R_B$ is non-empty (in fact large in the same sense[4]). That is to say, the family of double return times sets R_A has the *filter property*, and it would be nice to know, e.g. for the purposes of stengthening of combinatorial applications, just how exclusive this filter is.

Another sense in which we will be seeking to improve Theorem 1.2 is indicated by the natural question of whether it holds for non-amenable groups. Since the proof of Theorem 1.2 is achieved in [BMZ] by establishing positivity of limits of Cesaro-type averages having the form $\frac{1}{|\Phi_n|}\sum_{g \in \Phi_n} \mu\big(A \cap T_g^{-1}A \cap (T_gS_g)^{-1}A\big)$, where (Φ_n) is a left Følner sequence, it is clear that the question calls for a change in methodology. Aficionados may recognize that one might tackle both this problem and that of the previous paragraph by utilizing the "IP structure theory" developed for abelian groups by Furstenberg and Katznelson in [FK2] and [FK3]. Their strategy was to replace averaging along Følner sets by so-called *IP limits* and to thereby achieve, in the abelian case, the startling result that multiple return times sets are *IP** [5]. The reader will now of course anticipate that the IP* notion

[3] A *left* (respectively *right*) *Følner sequence* for a discrete group H is a sequence of finite sets $\Phi_n \subset G$ having the property that for every $g \in H$, $\frac{|g\Phi_n \cap \Phi_n|}{|\Phi_n|} \to 1$ (respectively $\frac{|\Phi_n g \cap \Phi_n|}{|\Phi_n|} \to 1$) as $n \to \infty$. For groups, admission of a Følner sequence is one characterization of amenability.

[4] This is an easy consequence of the observation that $R_A \cap R_B \subset R_{A\times B}$, where $R_{A\times B} = \big\{g \in G : \mu \times \mu\big((A \times B) \cap (T_g \times T_g)^{-1}(A \times B) \cap (T_gS_g \times T_gS_g)^{-1}(A \times B)\big) > 0\big\}$.

[5] The family of IP* sets is significantly smaller than the set of syndetic sets and does, in fact, possess the filter property.

has meaning for non-abelian groups (*yes*) and that we'll soon prove R_A to be IP* (alas, *no*).

As we shall discuss below, methods based on IP-limits fare even worse than Følner averaging methods at achieving our unique ends. They do translate seamlessly to the non-amenable situation, but this is part of the problem; they can't exploit any of the features that classical ergodic averaging is designed to exploit (in particular, the compactness of orbits in the *Kronecker factor*–see below), and in consequence one is left at the mercy of the non-commutativity of G. Fortunately a composite method, in which one takes limits along certain types of *ultrafilters* (members of the Stone-Čech compactification of G, i.e. βG), carries the day.

We shall give a brief overview of ultrafilters and the algebraic properties of βG below; for the purposes of this introduction, it will be convenient to think of an ultrafilter p on G as a $\{0,1\}$-valued, finitely additive probability measure on the power set of G, that is $\mathcal{P}(G)$. If $A \subset G$ has p-measure 1, we write $A \in p$ and say that A is *p-large*. This notion of largeness gives rise to a natural notion of convergence for G-indexed sequences. Indeed, let $(x_g)_{g\in G}$ be a sequence in a topological space. Given an ultrafilter $p \in \beta G$, one writes $p\text{-lim}_g x_g = y$ if for every neighborhood U of y one has $\{g : x_g \in U\} \in p$. One may easily check that, in a compact Hausdorff space, $p\text{-lim}_g x_g$ always exists and is unique.

An ultrafilter $p \in \beta G$ is *idempotent*, or *almost shift invariant*, if any $A \in p$ has the property that for p-many $g \in G$, the set $\{h \in G : hg \in A\}$ is p-large. An important property of idempotents p that we shall utilize repeatedly is that for a G-indexed sequence (x_g) in a compact Hausdorff space, $p\text{-lim}_g p\text{-lim}_h x_{gh} = p\text{-lim}_g x_g$. For example, if $\mathcal{H}$ is a separable Hilbert space, and (U_g) is a unitary G-action of $\mathcal{H}$, the foregoing fact implies that the weak operator limit $P = p\text{-lim}_g U_g$ is itself idempotent and hence the orthogonal projection onto its range, the space of *p-rigid* functions $\{f \in \mathcal{H} : p\text{-lim}_g U_g f = f\}$.

This construction has much in common so far with the IP methodology of [FK2], with p-limits taking an almost identical role to that which IP-limits had there. Things get really interesting, however, when one requires that the idempotent p have some additional algebraic properties. For example, if p is a *minimal idempotent* (see Section 2 for a definition), then the space of p-rigid functions is equal to the space of functions having pre-compact orbit $\{U_g f : g \in G\}$; i.e. the Kronecker factor. The advantage that working in this more classical object of ergodic theory (normally associated with methods involving averaging along Følner sequences) confers is that this space is easily seen to be invariant under the action (U_g). Consequently, one has $PU_g = U_g P$. Significantly, this equality may fail when working with arbitrary idempotents or with IP-limits. On the other hand, in working with p-limits, one needn't rely on the presence of Følner sequences, so that the method works for general groups. Thus the hybrid nature of our new methodology allows us to navigate freely between the Scylla of non-invariant factors and Charybdis of classical ergodic averaging's sheer meltdown in the face of non-amenability.

We now give a first formulation of our main theorem.

Theorem 1.4. Let G be a countable group, let $(X, \mathcal{A}, \mu)$ be a probability space and suppose that (T_g), (S_g) are μ-preserving actions of G which commute in the sense that $T_g S_h = S_h T_g$ for all $g, h \in G$. Then for any $A \in \mathcal{A}$ with $\mu(A) > 0$ and any minimal idempotent p, $p\text{-lim}_g \mu\big(A \cap T_g^{-1} A \cap (T_g S_g)^{-1} A\big) > 0$.

Since it is not immediately obvious that Theorem 1.4 fully generalizes Theorem 1.2, we will presently offer another, stronger formulation.

Let $\mathcal{S}$ be a family of ultrafilters in βG. We say that a set $E \subset G$ is an $\mathcal{S}^*$*-set* if for any $p \in \mathcal{S}$, $E \in p$. It is easy to see that for any finite collection $\{E_1, \dots, E_k\}$ of $\mathcal{S}^*$-sets and any $p \in \mathcal{S}$, one has $\bigcap_{i=1}^k E_i \in p$ and hence $\bigcap_{i=1}^k E_i$ is also $\mathcal{S}^*$. Hence the family of $\mathcal{S}^*$ sets has the filter property. Let now $\mathcal{M}$ be the family of minimal idempotents in βG. It is a standard fact (see, for example, [B]) that any $\mathcal{M}^*$-set is right syndetic. An essentially equivalent way to get at the $\mathcal{M}^*$ property is as follows. A set $C \subset G$ is said to be a *central set* if for some $p \in \mathcal{M}$, $E \in p$. Now a set $E \subset G$ is said to be *central**, or *C**, if it intersects non-trivially every central set. One easily sees in fact that any C* set intersects every central set in a central set. Of course, the C* sets are just the $\mathcal{M}^*$ sets.

Under the hypotheses of Theorem 1.4, what our methods actually show is that for some $\lambda > 0$, the set $R_{A,\lambda} = \{g \in G : \mu\big(A \cap T_g^{-1}A \cap (T_gS_g)^{-1}A\big) > \lambda\}$ is C* and hence, in particular, right syndetic. Switching the roles of T_g and S_g, one can choose λ with the additional property that $R'_{A,\lambda} = \{g \in G : \mu\big(A \cap S_g^{-1}A \cap (T_gS_g)^{-1}A\big) > \lambda\}$ is also C*. Next observe that $R'_{A,\lambda} = \{g^{-1} : g \in R_{A,\lambda}\} = R_{A,\lambda}^{-1}$. In particular, $R_{A,\lambda}^{-1}$ is C*. We call a set E having the property that E^{-1} is C* an *inverse C* set*. It is an easy exercise that inverse C* sets are left syndetic. Hence the following reformulation of Theorem 1.4 extends Theorem 1.2 as well.

Theorem 1.5. Let G be a countable group, let $(X, \mathcal{A}, \mu)$ be a probability space and suppose that (T_g), (S_g) are μ-preserving actions of G which commute in the sense that $T_gS_h = S_hT_g$ for all $g, h \in G$. Then for any $A \in \mathcal{A}$ with $\mu(A) > 0$ there exists $\lambda > 0$ such that $\{g \in G : \mu\big(A \cap T_g^{-1}A \cap (T_gS_g)^{-1}A\big) > \lambda\}$ is both C* and inverse C*.

As mentioned before, we have three main combinatorial applications, which shall be presented in Section 4. The most basic, a simple consequence of Theorem 1.5 and Furstenberg's correspondence principle (cf. Proposition 4.1 below), refines Theorem 1.3. The following theorem, which implies that for any countable amenable group G, every large enough subset of $G \times G$ contains many "isosceles right triangles" $\{(a, b), (ag, b), (ag, bg)\}$, is restated below as Theorem 4.1.

Theorem 1.6. Let G be a countable amenable group with identity e and suppose $E \subset G \times G$ has positive upper density with respect to some right Følner sequence (Φ_n). Then with respect to (Φ_n), for some $\lambda > 0$ $\{g : \overline{d}\big(E \cap E(g^{-1}, e) \cap E(g^{-1}, g^{-1})\big) > \lambda\}$ is both C* and inverse C*.

We should mention that, although most of the results here are formulated for minimal idempotents, our proof actually establishes that they hold for a wider family of idempotents, which in turn implies that the filter of sets R_A is even smaller and the combinatorial corollaries could be correspondingly enhanced. We don't worry ourselves too much about this issue but do offer some discussion along these lines in Section 4 below.

The structure of the paper is as follows. In Section 2 we collect a few needed lemmas and review basic facts about the algebraic structure of βG. In Section 3 we bring these tools to bear in proving the double recurrence theorem. Section 4 contains the three aforementioned combinatorial applications as well as discussions concerning the strength of our methods. In particular, we discuss here why the proof yields central* results but not IP* results[6]. Finally in a fifth section, we speculate about the wider applicability of the new methodology we are introducing here, offering as a sample of its potential a proof that a kind of k-fold weak mixing occurs for certain collections of measure preserving actions of a general group G.

2. Preliminaries.

An ultrafilter p on a set G is a nonempty family of subsets of G satisfying the following conditions:

(i) $\emptyset \notin p$.

(ii) If $A \in p$ and $A \subset B$ then $B \in p$.

(iii) If $A \in p$ and $B \in p$ then $A \cap B \in p$

(iv) if $r \in \mathbf{N}$ and $G = A_1 \cup A_2 \cup \cdots \cup A_r$ then for some i, $1 \leq i \leq r$, $A_i \in p$.

Assume now that $(G, \cdot)$ is a discrete semigroup. We denote, as is customary, by βG the space of ultrafilters on G. (The reason for this custom is that with the standard topology we are about to define on it, βG is just the classical Stone-Čech compactification of G.) The semigroup operation $\cdot$ on G extends naturally to βG by the rule $A \in p \cdot q \Leftrightarrow \{x : Ax^{-1} \in p\} \in q$, where $Ax^{-1} = \{y \in G : yx \in A\}$. Now, for $A \subset G$, let $\overline{A} = \{p \in \beta G : A \in p\}$. One may check that the family $\mathcal{A} = \{\overline{A} : A \subset G\}$ is a basis for a topology on βG. Under this operation and this topology, βG becomes a compact Hausdorff left topological semigroup. (The last condition means that for any fixed $q \in \beta G$, the map $p \to q \cdot p$ is continuous. For more details, see [B] or [HS].)

By a theorem of Ellis [E], any compact semigroup with a left-continuous operation (in particular, any closed subsemigroup of βG) has an idempotent. If p is idempotent and $A_1 \in p$, then since $A_1 \in p \cdot p$, one may choose straight from the definition of $p \cdot p$ some $g_1 \in A_1$ such that $Ag_1^{-1} \in p$ (so that also $A_2 = A_1 \cap A_1 g_1^{-1} \in p$). One may iterate this process, choosing now $g_2 \in A_2$ with $A_3 = A_2 \cap A_2 g_2^{-1} \in p$. One may now check that $\{g_1, g_2, g_2 g_1\} \subset A_1$. Continuing in this fashion, one may find a sequence (g_i) such that $FP(\langle g_i \rangle_{i=1}^{\infty}) \subset A_1$, where $FP(\langle g_i \rangle_{i=1}^{\infty}) = \{g_{i_1} g_{i_2} \cdots g_{i_k} : k \in \mathbf{N},\ i_1 > i_2 > \cdots > i_k\}$. Since for any finite partition of G, some cell is contained in p, the presence of idempotent ultrafilters gives an immediate proof of Hindman's theorem[7].

A *right ideal* (respectively *left ideal*) of βG is a set $J \subset \beta G$ such that for every $q \in \beta G$ and every $p \in J$, $p \cdot q \in J$ (respectively $q \cdot p \in J$). An *ideal* is a set $I \subset \beta G$ that is both a left and a right ideal. A *minimal right ideal* is a non-empty right ideal J containing no proper, non-empty subset which is itself a right ideal. If J is such and $x \in J$, then xG, being a non-empty right ideal, must equal J and must also, being the continuous image of G, be closed. By a routine application of Zorn's Lemma, any right ideal in βG contains a minimal right ideal.

---

[6] As a matter of fact these difficulties, together with some IP* counterexamples from [BH] for partition Ramsey theory in free groups, make us wonder whether the naturally corresponding IP* assertions are even true.

[7] First proved in [H]. Various formulations exist, one such stating that for any finite partition of a semigroup G, some cell contains $FP(\langle g_i \rangle_{i=1}^{\infty})$ for some sequence (g_i) in G.

Let $\mathcal{K}$ be the union of the minimal right ideals of βG. Then $\mathcal{K}$ is a two-sided ideal, and, in fact, the smallest two-sided ideal. To see this, note first that, being the union of right ideals, $\mathcal{K}$ is trivially a right ideal. To show that $\mathcal{K}$ is a left ideal, let $x \in \mathcal{K}$ and $y \in \beta G$. Let R be a minimal right ideal such that $x \in R$. Now $yx \in yR$. We claim that yR is a minimal right ideal. Indeed, let $\emptyset \neq J \subset yR$ be a right ideal and let $C = \{z \in R: \ yz \in J\}$. Then C is a non-empty right ideal which is contained in R and so $C = R$ and $J = yR$. Finally, one may easily show that any minimal right ideal must be contained in every two sided ideal. Hence, if I is a two-sided ideal then $\mathcal{K} \subset I$.

A *minimal idempotent* is an idempotent ultrafilter p that belongs to the minimal ideal $\mathcal{K}$. A set $A \subset G$ is a *right central set* if there exists a minimal idempotent p such that $A \in p$. Right central sets have various largeness properties that prompt the following digression: A set $A \subset G$ is *right syndetic* if for some finite set $F \subset G$ one has $\bigcup_{t\in F} At^{-1} = G$. A set $T \subset G$ is *left thick* if for every finite set $F \subset G$, there exists some $x \in G$ with $xF \subset T$. Finally, a set $A \subset G$ is *right piecewise syndetic* if for some finite set $F \subset G$, $\bigcup_{t\in F} At^{-1}$ is left thick. It is a standard fact (see e.g. [B]) that right central sets are right piecewise syndetic. In consequence of this, they have the following largeness feature that will be critical for us in the proof of Theorem 2.2 below.

Proposition 2.1 ([B]). If A is right central then $A^{-1}A = \{x \in G : yx \in A \text{ for some } y \in G\}$ is right syndetic.

By a unitary representation of a group G on a separable Hilbert space $\mathcal{H}$ we mean a function $g \to U_g$ taking G to unitary operators on $\mathcal{H}$ in such a way that $U_{gh} = U_gU_h$. In this case a vector $\varphi \in \mathcal{H}$ is said to be *compact* if $\{U_g\varphi : g \in G\}$ is totally bounded in the strong topology of $\mathcal{H}$. The representation $(U_g)_{g\in G}$ is said to be *weakly mixing* if there are no nonzero compact vectors.

Since the unit ball in $\mathcal{H}$ is a compact metrizable space with respect to the weak topology, for any ultrafilter p on G, $p\text{-lim}_g U_g f$ exists weakly for any $f \in \mathcal{H}$. Indeed, it is easy to see that if $p\text{-lim}_g U_g f = f$ weakly then $p\text{-lim}_g ||U_g f - f|| = 0$.

The following theorem is from [B]. We include a proof for completeness.

Theorem 2.2. Let $(U_g)_{g\in G}$ be a unitary representation of a group G on a Hilbert space $\mathcal{H}$. For any $f \in \mathcal{H}$ the following are equivalent:

(i) There exists a minimal idempotent $p \in \beta G$ such that $p\text{-lim}_g U_g f = f$.

(ii) f is a compact vector.

(iii) $p\text{-lim}_g U_g f = f$ for every idempotent $p \in \beta G$.

Proof. (i) $\Rightarrow$ (ii). For any $\epsilon > 0$, $A = \{g \in G : ||U_g f - f|| < \frac{\epsilon}{2}\} \in p$. For any $g_1, g_2 \in A$, $||U_{g_1} f - U_{g_2} f|| < \epsilon$, which implies that for any $g \in A^{-1}A$ one has $||U_g f - f|| < \epsilon$. But by Proposition 2.1, $A^{-1}A$ is syndetic, meaning that finitely many shifts of $A^{-1}A$ cover G. It follows that finitely many balls of radius ϵ cover $\{U_g f\}_{g\in G}$, which means that f is a compact vector.

(ii)$\Rightarrow$ (iii) Let p be an idempotent ultrafilter and set $h = p\text{-lim}_g U_g f$ (strong convergence). For $\epsilon > 0$, $B = \{g \in G : ||U_g f - h|| < \epsilon\} \in p$, so one may choose $x \in B$ such that $Bx^{-1} \in p$. Choose $y \in B \cap Bx^{-1}$. Then $||U_x f - h|| < \epsilon$, $||U_{yx} f - h|| < \epsilon$ and $||U_y f - h|| < \epsilon$.

Now

$$\begin{aligned}||f-h|| =&||U_yf-U_yh|| \le ||U_yf-h||+||h-U_{yx}f||+||U_{yx}f-U_yh||\\ =&||U_yf-h||+||h-U_{yx}f||+||U_xf-h|| \le 3\epsilon.\end{aligned}$$

Since ϵ was arbitrary, one has $f=h$ and p-lim $U_gf=f$.

(iii)$\Rightarrow$ (i) Obvious. □

The following result is motivated by a difference trick of van der Corput and has many cousins, among them [FK2, Lemma 5.3], [F2, Lemma 9.24] and [BM1, Proposition 2.18].

Theorem 2.3. Assume that $(x_g)_{g\in G}$ is a bounded sequence in a Hilbert space $\mathcal{H}$. Let $p\in\beta G$ be an idempotent. If $p\text{-lim}_h p\text{-lim}_g\langle x_{gh},x_g\rangle=0$ then $p\text{-lim}_g x_g=0$ in the weak topology.

Proof. Without loss of generality we will assume that $||x_g||\le 1$, $g\in G$. Suppose to the contrary that $p\text{-lim}_g x_g=\tilde{x}\ne 0$. Let $\delta=\frac{||\tilde{x}||^2}{2}$ and pick $k\in\mathbf{N}$ and $\epsilon>0$ such that $\frac{1}{k}+\epsilon<\delta$. Inductively choose $g_1,\dots,g_k\in G$ such that for all j, $1\le j\le k$, one has

(i) for every $\alpha,\beta\subset\{1,\dots,j\}$ with $\alpha\ne\emptyset$, $\beta\ne\emptyset$ and $\beta<\alpha$, $|\langle x_{g_\alpha g_\beta},x_{g_\alpha}\rangle|\le\epsilon$.

(ii) for every $\alpha,\beta\subset\{1,\dots,j\}$ with $\beta\ne\emptyset$ and $\beta<\alpha$, $p\text{-lim}_g|\langle x_{gg_\alpha g_\beta},x_{gg_\alpha}\rangle|\le\epsilon$.

(iii) for all $r\in FP(g_1,\dots,g_j)$, $\langle x_r,\tilde{x}\rangle>\delta$.

(iv) for all $r\in\{0\}\cup FP(g_1,\dots,g_j)$, $\{g:\langle x_g,\tilde{x}\rangle>\delta\}r^{-1}\in p$.

(v) for all $r\in FP(g_1,\dots,g_j)$, $\{h: p\text{-lim}_g|\langle x_{gh},x_g\rangle|\le\epsilon\}r^{-1}\in p$.

Having done this, we let $y_i=g_kg_{k-1}\cdots g_i$, $1\le i\le k$, and observe that $|\langle y_i,y_j\rangle|\le\epsilon$ and $\langle y_i,\tilde{x}\rangle>\delta$, $1\le i\ne j\le k$. From the former it follows that $\langle\sum_{i=1}^k y_i,\sum_{i=1}^k y_i\rangle<k+k^2\epsilon<k^2\delta$, which implies that $|\langle\sum_{i=1}^k y_i,\tilde{x}\rangle|<k\delta$, contradicting the latter and completing the proof.

Suppose then that $0\le j<k$ and $g_1,\dots,g_j$ have been chosen. By the induction hypothesis,

$$B=\Big(\bigcap_{r\in\{0\}\cup FP(g_1,\dots,g_j)}\{g:\langle x_g,\tilde{x}\rangle>\delta\}r^{-1}\Big)\cap\Big(\bigcap_{\alpha,\beta\subset\{1,\dots,j\},\emptyset\ne\beta<\alpha}\{g:|\langle x_{gg_\alpha g_\beta},x_{gg_\alpha}\rangle|\le\epsilon\}\Big)$$
$$\cap\Big(\bigcap_{r\in\{0\}\cup FP(g_1,\dots,g_j)}\{h: p\text{-lim}_g|\langle x_{gh},x_g\rangle|\le\epsilon\}r^{-1}\Big)$$

is a member of p. Therefore, we may choose $g_{j+1}\in B$ such that $Bg_{j+1}^{-1}\in p$. It is now a routine matter to check that (i)-(v) above hold for j replaced by $j+1$. □

Theorem 2.4. Let $(U_g)_{g\in G}$ be a unitary representation of G on a separable Hilbert space $\mathcal{H}$, suppose $p\in\beta G$ is idempotent and for $f\in\mathcal{H}$, let $Pf=p\text{-lim}_gU_gf$ weakly. Then P is the orthogonal projection onto a closed subspace of $\mathcal{H}$. If p is minimal, then for every $g\in G$ and every $f\in\mathcal{H}$ one has $PU_gf=U_gPf$.

Proof. One may easily check that P is linear and that $||P||\le 1$. These facts imply that if P is idempotent then P is an orthogonal projection. One may also check that P is weakly continuous. Let ρ be a metric for the weak topology on the unit ball of $\mathcal{H}$ and pick some f in the unit ball. We must show that $P^2f=Pf$.

Take $\epsilon > 0$ to be arbitrary and let $B = \{g \in G : \rho(PU_gf, P^2f) < \epsilon$ and $\rho(U_gf, Pf) < \epsilon\}$. Then $B \in p$ so we may choose $x \in B$ such that $Bx^{-1} \in p$. Next choose $y \in B \cap Bx^{-1} \cap \{g \in G : \rho(U_gU_xf, PU_xf) < \epsilon\}$. Then $\rho(P^2f, Pf) \leq \rho(P^2f, PU_xf) + \rho(PU_xf, U_yU_xf) + \rho(U_yU_xf, Pf) < 3\epsilon$. Since ϵ was arbitrary, $P^2f = Pf$ and we are done.

Suppose now that p is minimal. Then by Theorem 2.2, P is the orthogonal projection onto the space of compact functions. Clearly, for any compact function f and any $g \in G$, U_gf is compact. Suppose now that $Ph = 0$, and let f be an arbitrary compact function and $g \in G$. Then $\langle h, U_{g^{-1}}f\rangle = 0$, so that $\langle U_gh, f\rangle = 0$. But since f is an arbitrary compact function, this implies that $PU_gh = 0$. We have shown that $Ph = 0 \Rightarrow PU_gh = 0$. Now let $f \in \mathcal{H}$ be arbitrary. Since Pf is compact, so is U_gPF, hence $PU_gPf = U_gPf$. On the other hand $P(f - Pf) = 0$, so letting $h = f - Pf$ in the above argument yields $PU_gf - PU_gPf = P(U_gf - U_gPf) = 0$. So $U_gPf = PU_gf$. □

As we shall elaborate upon in Section 4, the conclusion $PU_g = U_gP$ may fail if the idempotent p used in the construction of P is not minimal.

The following standard theorem follows [FK3, Lemma 3.1].

Theorem 2.5 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 $\{fg, f + g\} \subset E_0$.

Finally, we shall require basic facts concerning decomposition of measures over sub-σ-algebras. For more details, the reader is referred to, e.g., [F2, Chapter 5], or [F1, Section 4]. If $(X, \mathcal{A}, \mu)$ is a Lebesgue space and $\mathcal{B} \subset \mathcal{A}$ is a sub-σ-algebra, there exists a family of probability measures $\{\mu_x : x \in X\}$ on X such that for any $f \in L^1(X, \mathcal{A}, \mu)$ and any $B \in \mathcal{B}$, $\int_B f\ d\mu = \int_B(\int f\ d\mu_x)\ d\mu(x)$. If we write $E(f|\mathcal{B})(x) = \int f\ d\mu_x$ 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 $\mathcal{B}$ to define a measure $\tilde{\mu}$ on $(X \times X, \mathcal{A} \otimes \mathcal{A})$ as follows: $\int f(x, y)\ d\tilde{\mu}(x, y) = \int(\int f(x, y)\ d\mu_z \times \mu_z(x, y))\ d\mu(z)$. In particular, one has $\int f(x)h(y)\ d\tilde{\mu}(x, y) = \int PfPh\ d\mu$.

3. Main Theorem.

We now apply the ideas of the previous section to the following situation. Let $(X, \mathcal{A}, \mu)$ be a Lebesgue probability space (measurably isomorphic to a compact metric space equipped with a completed regular Borel measure) and suppose that $(T_g)_{g\in G}$ and $(S_g)_{g\in G}$ are commuting measure preserving G-anti-actions on X. That is to say, for each $g, h \in G$, T_g and S_g are invertible μ-preserving point transformations of X, with $T_{gh}x = T_hT_gx$, $S_{gh}x = S_hS_gx$ and $T_gS_hx = S_hT_gx$.

The anti-actions (T_g) and (S_g) naturally give rise to unitary G-actions on $L^2(\mu)$ by the rules $T_gf(x) = f(T_gx)$ and $S_gf(x) = f(S_gx)$, as, for example, $T_{gh}f(x) = f(T_{gh}x) = f(T_hT_gx) = T_hf(T_gx) = T_gT_hf(x)$. In this paper we always take L^2-spaces to consist of real-valued functions only. The assumption that $(X, \mathcal{A}, \mu)$ is a Lebesgue space implies that $L^2(X, \mathcal{A}, \mu)$ is separable and its unit ball is compact and metrizable in the weak topology. Letting p be a minimal idempotent ultrafilter on G, therefore, we can write $p\text{-lim}\ S_gf = Pf$ for all $f \in L^2(\mu)$, where convergence is in the weak topology.

According to Theorem 2.4, P is the orthogonal projection onto a closed subspace $E \subset L^2(X, \mathcal{A}, \mu)$. E contains the constants, and $Pf = f$ if and only if $p\text{-}\lim_g ||S_g f - f|| = 0$. 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 $f(x) \geq l$, $f_l(x) = -l$ if $f(x) \leq -l$ and $f_l(x) = f(x)$ if $-l < f(x) < l$ satisfies $p\text{-}\lim_f ||S_g f_l - f_l|| = 0$, and hence lies in E. Thus E contains a dense subset consisting of bounded functions. Moreover, if f, h are bounded functions in E, we have

$$\begin{aligned} p\text{-}\lim_g \big|\big|S_g f S_g h - fh\big|\big| &\leq p\text{-}\lim_g \big|\big|S_g f S_g h - (S_g f)h\big|\big| + p\text{-}\lim_g \big|\big|(S_g f)h - fh\big|\big| \\ &\leq \big|\big|f\big|\big|_\infty p\text{-}\lim_g \big|\big|S_g h - h\big|\big| + \big|\big|h\big|\big|_\infty p\text{-}\lim_g \big|\big|S_g f - f\big|\big| = 0, \end{aligned}$$

so that $fg \in E$ (and is bounded). Moreover $f + g$ is clearly a bounded member of E, so by Theorem 2.5 $E = L^2(X, \mathcal{B}, \mu)$, where $\mathcal{B} \subset \mathcal{A}$ is a σ-algebra, and consequently $Pf = E(f|\mathcal{B})$. We let $\{\mu_x : x \in X\}$ be the disintegration of μ over $\mathcal{B}$ and define the measure $\tilde{\mu}$ on $X \times X$ as noted in the previous section.

For G, $(X, \mathcal{A}, \mu)$, $(T_g)_{g \in G}$, $(S_g)_{g \in G}$ and p as above, our main theorem may be stated as follows.

Theorem 3.1. For any $A \in \mathcal{A}$ with $\mu(A) > 0$, there exists $\lambda > 0$ (depending on A, (T_g) and (S_g)) such that $\big\{g \in G : \mu\big(A \cap T_g^{-1} A \cap (T_g S_g)^{-1} A\big) > \lambda\big\}$ is C*.

The remainder of this section shall constitute a proof of Theorem 3.1. Fix a minimal idempotent p and a set $A \in \mathcal{A}$ with $\mu(A) > 0$. Let λ be a positive number to be named later and let $L = p\text{-}\lim_g \mu\big(A \cap T_g^{-1} A \cap (T_g S_g)^{-1} A\big)$. We must show that $L > \lambda$.

For $H \in L^2(\tilde{\mu})$, write $Q_1 H = p\text{-}\lim_g (T_g \times T_g) H$ $Q_1 H = p\text{-}\lim_g (T_g S_g \times T_g S_g) H$ (these limits are in the weak topology of $L^2(\tilde{\mu})$). By Theorem 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_g)-*almost periodic over* $\mathcal{B}$ *along* p if for every $\epsilon > 0$ there exists a set $D \in \mathcal{B}$ with $\nu(D) < \epsilon$ and functions $h_1, \cdots, h_N \in L^2(X, \mathcal{A}, \mu)$ having the property that for every $\delta > 0$ there exists a set $C \in p$ such that for every $g \in C$ there is a set $E(g) \in \mathcal{B}$ with $\mu\big(E(g)\big) < \delta$ having the property that for all $x \in X \setminus \big(D \cup E(g)\big)$ there exists $i(x, g)$ with $1 \leq i(x, g) \leq N$ such that $\big|\big|T_g f - h_{i(x,g)}\big|\big|_x < \epsilon$. $(T_g S_g)$-almost periodicity is similarly defined.

An argument similar to the one used above to show that range P has the form $L^2(X, \mathcal{B}, \mu)$ can be used here to show that the closure in $L^2(X, \mathcal{A}, \mu)$ of the space of functions (T_g)-almost periodic over $\mathcal{B}$ has the form $L^2(X, \mathcal{B}_1, \mu)$ for some σ-algebra $\mathcal{B}_1 \subset \mathcal{A}$. (One first shows that the set of functions (T_g)-almost periodic over $\mathcal{B}$ has a dense algebra of bounded functions and then applies Theorem 2.5.) Moreover, one easily sees that $\mathcal{B} \subset \mathcal{B}_1$. Similarly, the closure of the set of $(T_g S_g)$-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 $H \in L^2(\tilde{\mu})$, define $\mathbf{H} * f(x) = \int H(x, t) f(t) \, d\mu_x(t)$. For almost every x, $\mathbf{H}$ is then a compact linear operator on $L^2(\mu_x)$.

Lemma 3.3 Let $H \in L^2(X \times X, A \otimes A, \tilde{\mu})$ and $f \in L^\infty(X, \mathcal{A}, \mu)$.

(a) If $Q_1 H = H$ then $\mathbf{H} * f$ is (T_g)-almost periodic over $\mathcal{B}$.

(b) If $Q_2 H = H$ then $\mathbf{H} * f$ is $(T_g S_g)$-almost periodic over $\mathcal{B}$.

Proof. We will prove (a). Part (b) is similar. Let $\epsilon > 0$ be arbitrary and let $(g_i)_{i=1}^{\infty}$ be an enumberation of G. Since $\mathbf{H}$ is compact on $L^2(\mu_x)$ a.e. there exists a $\mathcal{B}$-measurable function $x \to M(x)$ into the naturals such that $\{\mathbf{H} * (T_{g_i} f) : 1 \le i \le M(x)\}$ is $\frac{\epsilon}{2}$-dense (a.e., for the metric $\rho(g,h) = ||g-h||_x$) in the set $\{\mathbf{H} * (T_g f) : g \in G\}$.

Let N be so large that $D = \{x \in X : M(x) > M\} \in \mathcal{B}$ satisfies $\mu(D) < \epsilon$ and put $h_i = T_{g_i} f$, $1 \le i \le N$. For any $x \in X \setminus D$ and any $g \in G$ there exists $i(x,g)$ with $1 \le i(x,g) \le N$ such that $||\mathbf{H} * (T_g f) - h_{i(x,g)}||_x < \frac{\epsilon}{2}$. Let now $\delta > 0$ be arbitary. One has

$$\begin{aligned}
&p\text{-lim}_g ||T_g(\mathbf{H} * f) - \mathbf{H} * (T_g f)||^2 \\
&= p\text{-lim}_g \int \Big| \int \big(H(T_g x, T_g t) - H(x,t)\big) f(T_g t)\, d\mu_x(t)\Big|^2\, d\mu(x) \\
&\le p\text{-lim}_g \int\int |H(T_g x, T_g t) - H(x,t)|^2 |f(T_g t)|^2\, d\mu_x(t)\, d\mu(x) \\
&\le p\text{-lim} ||(T_g \times T_g) H - H||^2_{L^2(\tilde{\mu})} ||f||^2_\infty = 0.
\end{aligned}$$

Let $C = \{g \in G : ||T_g(\mathbf{H} * f) - \mathbf{H} * (T_g f)||^2 < \delta(\frac{\epsilon}{2})^2$. Then $C \in p$ and for every $g \in C$, $||T_g(\mathbf{H} * f) - \mathbf{H} * (T_g f)||_x^2 < (\frac{\epsilon}{2})^2$ for every x outside of a set $E(g) \in \mathcal{B}$ satisfying $\mu(E(g)) \le \delta$. If now $g \in C$ and $x \in X \setminus \big(D \cup E(g)\big)$ then $||T_g(\mathbf{H} * f) - h_{i(x,g)}||_x < \epsilon$. ☐

Lemma 3.4. Let $f \in L^\infty(X, \mathcal{A}, \mu)$.

(a) If $E(f|\mathcal{B}_1) = 0$, then $p\text{-lim}_g ||P(f T_g f)|| = 0$.

(b) If $E(f|\mathcal{B}_2) = 0$, then $p\text{-lim} ||P(f T_g S_g f)|| = 0$.

Proof. Again (a) and (b) have similar proofs. This time we'll prove (b). Since $E(f|\mathcal{B}_2) = 0$, f is orthogonal to $\mathbf{H} * f$ for every $H \in L^2(X \times X, A \otimes A, \tilde{\mu})$ satisfying $Q_2 H = H$. Therefore, if $Q_2 H = H$ then

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

In consequence, $f \otimes f$ is orthogonal to $Q_2 H$ for all $H \in L^2(\tilde{\mu})$, from which it follows that

$$\begin{aligned}
p\text{-lim}_g ||P(f T_g S_g f)||^2 &= p\text{-lim}_g \int \Big| \int f(t) T_g S_g f(t)\, d\mu_x(t)\Big|^2\, d\mu(x) \\
&= p\text{-lim}_g \int f(x) f(t) T_g S_g f(x) T_g S_g f(t)\, d\tilde{\mu}(x,t) \\
&= \int (f \otimes f) Q_2 (f \otimes f)\, d\tilde{\mu} = 0.
\end{aligned}$$

☐

Lemma 3.5 If $f_1, f_2 \in L^\infty(X, \mathcal{A}, \mu)$ with either $E(f_1|\mathcal{B}_1) = 0$ or $E(f_2|\mathcal{B}_2) = 0$, then $p\text{-lim}_g T_g f_1 T_g S_g f_2 = 0$ in the weak topology.

Proof. We apply Theorem 2.3. For $g \in G$, let $x_g = T_g f_1 T_g S_g f_2$. Then

$$\begin{aligned}
&p\text{-lim}_h p\text{-lim}_g \langle x_{gh}, x_g \rangle \\
&= p\text{-lim}_h p\text{-lim}_g \int T_{gh} f_1 T_{gh} S_{gh} f_2 T_g f_1 T_g S_g f_2 \, d\mu \\
&= p\text{-lim}_h p\text{-lim}_g \int T_g T_h f_1 T_g T_h S_g S_h f_2 T_g f_1 T_g S_g f_2) \, d\mu \\
&= p\text{-lim}_h p\text{-lim}_g \int (f_1 T_h f_1) S_g (f_2 T_h S_h f_2) \, d\mu \\
&= p\text{-lim}_h \int P(f_1 T_h f_1) P(f_2 T_h S_h f_2) \, d\mu = 0, \\
&\leq p\text{-lim}_h ||P(f_1 T_h f_1)|| \cdot ||P(f_2 T_h S_h f_2)|| = 0
\end{aligned}$$

by Lemma 3.4. The desired conclusion follows from Theorem 2.3. □

We're now ready to complete the proof of Theorem 3.1. Let $f = 1_A$, $f_1 = E(f|\mathcal{B}_1)$, and $f_2 = E(f|\mathcal{B}_2)$. Put $h_1 = f - f_1$ and $h_2 = f - f_2$. By Lemma 3.5

$$\begin{aligned}
&p\text{-lim}_g T_g f_1 T_g S_g h_2 \\
=&p\text{-lim}_g T_g h_1 T_g S_g f_2 \\
=&p\text{-lim}_g T_g h_1 T_g S_g h_2 = 0
\end{aligned}$$

in the weak topology, from which it follows that

$$\begin{aligned}
L &= p\text{-lim}_g \mu(A \cap T_g^{-1} A \cap (T_g S_g)^{-1} A) \\
&= p\text{-lim}_g \int f T_g f T_g S_g f \, d\mu \\
&= p\text{-lim}_g \int f T_g (f_1 + h_1) T_g S_g (f_2 + h_2) \, d\mu \\
&= p\text{-lim}_g \int f T_g f_1 T_g S_g f_2 \, d\mu.
\end{aligned}$$

From the decomposition of μ one sees that $f_1(x) f_2(x) > 0$ for a.e. $x \in A$. Choose a number $a > 0$ and a set $A' \subset A$ with $\mu(A') > 0$ such that $f_1(x) f_2(x) > a$ for all $x \in A'$. Next choose $b, \xi > 0$ and $B_1 \in \mathcal{B}$ with $\mu(B_1) = 3\xi > 0$ such that for all $x \in B_1$, $\mu_x(A') > b$. Then $\int f f_1 f_2 \, d\mu_x > ab$ for all $x \in B_1$.

Let $\epsilon = \frac{ab}{18}$. Pick ϕ_1 that is T_g-almost periodic over $\mathcal{B}$ and ϕ_2 that is $T_g S_g$-almost periodic over $\mathcal{B}$ so that for some $B_2 \in \mathcal{B}$ with $B_2 \subset B_1$ and $\mu(B_2) > 2\xi$, $\left|\left|f_1 - \phi_1\right|\right|_x < \epsilon$ and $\left|\left|f_2 - \phi_2\right|\right|_x < \epsilon$ for all $x \in B_2$. Now, by definition there are $M \in \mathbf{N}$, $\{h_1, \ldots, h_M\} \subset L^2(X, \mathcal{A}, \mu)$ and $D \in \mathcal{B}$ with $\mu(D) < \xi$ such that:

(∗) For every $\delta > 0$ there exists a set $C \in p$ such that for every $g \in C$ there is a set $E(g) \in \mathcal{B}$ with $\mu\big(E(g)\big) < \delta$ such that for every $x \in X \setminus \big(D \cup E(g)\big)$ there are $i(x,g)$ and $j(x,g)$ with $1 \leq i(x,g), j(x,g) \leq M$ satisfying $\big|\big|T_g\phi_1 - h_{i(x,g)}\big|\big|_x < \epsilon$ and $\big|\big|T_gS_g\phi_2 - h_{j(x,g)}\big|\big|_x < \epsilon$.

Put $B_3 = B_2 \setminus D$, so that $\mu(B_3) > \xi$, and let $N = M^2 + 1$. We are finally ready to make the value of λ explicit: $\lambda = \frac{ab\xi^{2^N}}{9N^2}$. What we must show, therefore, is that $L = p\text{-}\lim_g \int fT_gf_1T_gS_gf_2 \; d\mu > \lambda$. In order to accomplish this, it suffices to find, for arbitrary $C_1 \in p$, some $g \in C_1$ with $\int fT_gf_1T_gS_gf_2 \; d\mu \geq \frac{ab\xi^{2^N}}{8N^2}$. Accordingly, let $C_1 \in p$. For $\delta = \frac{\epsilon^2\xi^{2^N}}{4N^2}$, let C be as guaranteed by (∗). Replacing C by $C \cap C_1$, we may also assume that $C \subset C_1$.

Let $\eta = \frac{\xi^{2^N}}{6\cdot 4^N}$. We now inductively choose $g_1, \ldots, g_N \in G$ such that for $1 \leq j \leq N$ the following are satisfied:

(i) $A_j = B_3 \cap \bigcap_{g \in FP(g_1,\ldots,g_j)} \big(T_g^{-1}B_3 \cap (T_gS_g)^{-1}B_3\big)$ satisfies $\mu(A_j) > \xi^{2^j}$.

(ii) For all $\alpha, \beta \subset \{1, \ldots, j\}$ with $\alpha \neq \emptyset$, $\beta \neq \emptyset$ and $\alpha < \beta$,

$$\int \Big| \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{g_\beta}x} - \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{g_\beta}S_{g_\beta}x} \Big|^2 d\mu(x) \leq \eta.$$

(iii) For all $\alpha, \beta \subset \{1, \ldots, j\}$ with $\alpha \neq \emptyset$, and $\alpha < \beta$,

$$p\text{-}\lim_g \int \Big| \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{gg_\beta}x} - \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{gg_\beta}S_{gg_\beta}x} \Big|^2 d\mu(x) \leq \eta.$$

(iv) $FP(g_1, \ldots, g_j) \subset C$.

(v) For all $r \in FP(g_1, \ldots, g_j)$, $Cr^{-1} \in p$.

Suppose that $g_1, \ldots, g_j$ have been chosen with (i)-(v) satisfied. Observe that $A_j \in \mathcal{B}$ and choose $\gamma > 0$ so small that $\mu(A_j) > \xi^{2^{j+1}} + \gamma$. Let

$$\begin{aligned} E = &\Big(\bigcap_{r \in FP(g_1,\ldots,g_j)} Cr^{-1}\Big) \cap \{g : \mu(A_j \cap T_g^{-1}A_j) > \xi^{2^{j+1}} + \gamma\} \cap \{g : \mu(A_j \triangle S_g^{-1}A_j) < \gamma\} \\ &\cap C \cap \bigcap_{\alpha,\beta \subset \{1,\ldots,j\}, \emptyset \neq \alpha < \beta} \Big\{g : \int \Big| \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{gg_\beta}x} \\ &\qquad - \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{gg_\beta}S_{gg_\beta}x} \Big|^2 d\mu(x) \leq \eta \Big\}. \end{aligned}$$

Then $E \in p$ so we may choose $g_{j+1} \in G$ such that $Eg_{j+1}^{-1} \in p$. One now routinely checks that (i)-(v) are satisfied for j replaced by $j + 1$.

Supposing that $g_1, \ldots, g_N$ have been chosen, for non-empty $\alpha, \beta \subset \{1, 2, \ldots, N\}$ with $\alpha < \beta$ there is a set $C(\alpha, \beta) \in \mathcal{B}$ with $\mu\big(C(\alpha, \beta)\big) < \frac{\xi^{2^N}}{4N^2}$ such that for all $x \in X \setminus C(\alpha, \beta)$, $\Big| \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{g_\beta}x} - \big|\big|f_2 - T_{g_\alpha}S_{g_\alpha}f_2\big|\big|_{T_{g_\beta}S_{g_\beta}x} \Big| < \epsilon$.

Let $B_4 = B_3 \setminus \left(\bigcup_{g \in FP(g_1,\ldots,g_N)} E(g) \right)$ and put

$$B_5 = B_3 \cap \Big(\bigcap_{g \in FP(g_1,\ldots,g_N)} \big(T_g^{-1} B_3 \cap (T_g S_g)^{-1} B_3 \big) \Big)$$
$$\setminus \Big(\bigcup_{g,h \in FP(g_1,\ldots,g_N)} \big(E(h) \cup T_g^{-1} E(h) \cup (T_g S_g)^{-1} E(h) \big) \Big).$$

Let us now take account of where we stand:

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

(ii) For any $x \in B_4$, $\int f f_1 f_2 \, d\mu_x > ab$, $\big|\big|f_1 - \phi_1\big|\big|_x < \epsilon$ and $\big|\big|f_2 - \phi_2\big|\big|_x < \epsilon$. Moreover, for all $g \in FP(g_1,\ldots,g_N)$, $\big|\big|T_g\phi_1 - h_{i(x,g)}\big|\big|_x < \epsilon$ and $\big|\big|T_g S_g \phi_2 - h_{j(x,g)}\big|\big|_x < \epsilon$.

(iii) For any $x \in B_5$ and any $g \in FP(g_1,\ldots,g_N)$, $\{T_g x, T_g S_g x\} \subset B_4$.

Since $N = M^2 + 1$, for any $x \in B_5$ there exist $l(x)$ and $m(x)$ (which may be chosen $\mathcal{B}$-measurable) with $1 \leq l(x) < m(x) \leq N$ such that $i\big(x, g_N g_{N-1} \cdots g_{l(x)}\big) = i\big(x, g_N g_{N-1} \cdots g_{m(x)}\big)$ and $j\big(x, g_N g_{N-1} \cdots g_{l(x)}\big) = j\big(x, g_N g_{N-1} \cdots g_{m(x)}\big)$. As $\mu(B_5) > \frac{1}{2}\xi^{2^N}$ and there are less than N^2 possibilities for the pair $\big(l(x), m(x)\big)$, we may fix l and m with $1 \leq l < m \leq N$ and a set $B_6 \subset B_5$ with $\mu(B_6) > \frac{\xi^{2^N}}{2N^2}$ such that $l(x) = l$ and $m(x) = m$ for $x \in B_6$. Let $h = g_{m-1} g_{m-2} \cdots g_l$ and $g = g_N g_{N-1} \cdots g_m$.

Let now $x \in B_6$. We have $i(x, gh) = i(x, g)$. Since $x \notin (D \cup E(g) \cup E(gh)$,

$$\begin{aligned}
\big|\big|\phi_1 - T_h\phi_1\big|\big|_{T_g x} &\leq \big|\big|T_g\phi_1 - T_g(T_h\phi_1)\big|\big|_x + \epsilon \\
&= \big|\big|T_g\phi_1 - T_{gh}\phi_1\big|\big|_x \\
&\leq \big|\big|T_g\phi_1 - h_{i(x,g)}\big|\big|_x + \big|\big|T_{gh}\phi_1 - h_{i(x,g)}\big|\big|_x \\
&= \big|\big|T_g\phi_1 - h_{i(x,g)}\big|\big|_x + \big|\big|T_{gh}\phi_1 - h_{i(x,gh)}\big|\big|_x < 2\epsilon.
\end{aligned}$$

Also, since $T_g x \in B_3$, $\big|\big|\phi_1 - f_1\big|\big|_{T_g x} < \epsilon$. On the other hand, since $T_{gh}x \in B_3$ we have $||T_h f_1 - T_h\phi_1||_{T_g x} = ||f_1 - \phi_1||_{T_{gh} x} < \epsilon$. Putting it all together,

$$\big|\big|f_1 - T_h f_1\big|\big|_{T_g x} \leq \big|\big|f_1 - \phi_1\big|\big|_{T_g x} + \big|\big|\phi_1 - T_h\phi_1\big|\big|_{T_g x} + \big|\big|T_h\phi_1 - T_h f_1\big|\big|_{T_g x} < 4\epsilon = \frac{4ab}{18}.$$

Similarly, $\big|\big|f_2 - T_h S_h f_2\big|\big|_{T_g S_g x} < 4\epsilon$.

Let $\alpha = \{l, l+1, \ldots, m-1\}$ and $\beta = \{m, m+1, \ldots, N\}$. Notice that $\alpha, \beta \subset \{1, \ldots, N\}$, both are non-empty and $\alpha < \beta$. Moreover, observe that $h = g_\alpha$ and $g = g_\beta$. We let $B_7 = B_6 \setminus C(\alpha, \beta)$. Then $\mu(B_7) > \frac{\xi^{2^N}}{4N^2}$ and for every $x \in B_7$, $x \notin C(\alpha, \beta)$, which implies that $\Big|\big|\big|f_2 - T_{g_\alpha} S_{g_\alpha} f_2\big|\big|_{T_{g_\beta} x} - \big|\big|f_2 - T_{g_\alpha} S_{g_\alpha} f_2\big|\big|_{T_{g_\beta} S_{g_\beta} x}\Big| < \epsilon$. But $\big|\big|f_2 - T_h S_h f_2\big|\big|_{T_g S_g x} < 4\epsilon$. Therefore, $\big|\big|f_2 - T_h S_h f_2\big|\big|_{T_g x} < 5\epsilon = \frac{5ab}{18}$.

Since $x \in B_5$, $T_g x \in B_4 \subset B_1$. It follows by that $\int f f_1 f_2 \, d\mu_{T_g x} > ab$. Keeping in mind that $0 \leq f, f_1, f_2 \leq 1$, we may conclude $\int f f_1 T_h S_h f_2 \, d\mu_{T_g x} > ab - \frac{5ab}{18} = \frac{13ab}{18}$ and hence

$\int fT_hf_1T_hS_hf_2\,d\mu_{T_gx} > \frac{13ab}{18} - \frac{4ab}{18} = \frac{ab}{2}$. This holds for all $x \in B_7$, so $\int fT_hf_1T_hS_hf_2\,d\mu > \frac{ab\xi^{2^N}}{8N^2}$, as required. □

4. Combinatorial applications.

Although Theorem 3.1 holds for general countable groups, we have density combinatorial applications only for those groups in which Furstenberg's correspondence principle holds, namely *amenable* groups. Recall that a group G is amenable if and only if there exists a *right Følner sequence* for G, that is a sequence (Φ_n) of finite subsets of G such that for every $g \in G$, $\lim_n \frac{|\Phi_n \cap \Phi_n g|}{|\Phi_n|} = 1$. Left and two-sided Følner sequences are defined similarly. (We remark that every amenable group admits a two-sided Følner sequence.) The presence of Følner sequences gives rise to the notion of *density* for subsets of amenable groups. If $E \subset G$, where G is amenable, and (Φ_n) is a (left, right or two-sided) Følner sequence for G, we define the *upper density* of E (with respect to (Φ_n)) by $\overline{d}(E) = \limsup_n \frac{|E\cap\Phi_n|}{|\Phi_n|}$. (If the limit exists, we write $d(E)$.)

One may realize 1_E as an element of the compact metrizable space $\Omega = \{0,1\}^G$ (with product topology). An anti-action $\{T_g\}$ of G may be defined on Ω as follows: for $\xi \in \Omega$, let $(T_g\xi)(h) = \xi(gh)$. Furstenberg's correspondence principle for this context may now be formulated.

Proposition 4.1. Fix some Følner sequence sequence (Φ_n). Let $X = \overline{\{T_h1_E : h \in G\}}$ 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 $\mu(A) = \overline{d}(E)$ and such that for every $g_1, \ldots, g_k \in G$, $\mu(T_{g_1}^{-1}A \cap \cdots \cap T_{g_k}^{-1}A) \le \overline{d}(Eg_1^{-1} \cap \cdots \cap Eg_k^{-1})$.

Proof. (cf. [BM2, Theorem 2.1].) Let $\mathcal{A}$ be the algebra of sets generated by $\{T_g^{-1}A : g \in G\}$. Passing to a subsequence of the original Følner sequence (Φ_n), we may assume that $d(E)$ is as great as possible and that for every $g_1, g_2, \cdots, g_k \in G$, $d(E_1g_1^{-1} \cap \cdots \cap E_kg_k^{-1})$ exists for all choices $E_i \in \{E, E^c\}$, $1 \le i \le k$. Now for $A_1, A_2, \cdots, A_k \in \{A, A^c\}$ and $g_1, \ldots, g_k \in G$, put $\mu(T_{g_1}^{-1}A_1 \cap \cdots \cap T_{g_k}^{-1}A_k) = d(E_1g_1^{-1} \cap \cdots \cap E_kg_k^{-1})$, 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 this measure μ plainly satisfies the desired conclusions. □

If a group G is amenable then clearly $G \times G$ is amenable as well. (One way to see this is by showing that for arbitrary right Følner sequences (Φ_n) and (Ψ_n) on G, $(\Phi_n \times \Psi_n)$ is a right Følner sequence on $G \times G$.)

Theorem 4.2. Let G be a countable amenable group with identity e and suppose $E \subset G \times G$ has positive upper density with respect to some right Følner sequence (Φ_n). Then with respect to (Φ_n), for some $\lambda > 0$ $\{g : \overline{d}\big(E \cap E(g^{-1}, e) \cap E(g^{-1}, g^{-1})\big) > \lambda\}$ is both C* and inverse C*.

Proof. Let $\Omega = \{0,1\}^{G\times G}$, and define an anti-action $(U_{(g,h)})$ on Ω by $U_{(g,h)}\xi(a,b) = \xi(ga, hb)$. Let $\xi = 1_E \in \Omega$ and let $X = \overline{\{U_{(g,h)}\xi : (g,h) \in G \times G\}}$. Put $A = \{\eta \in$

$X : \eta(e,e) = 1\}$. According to Theorem 4.2 there exists a $\{U_{(g,h)}\}$-invariant measure μ on X with $\mu(A) = \overline{d}(E) > 0$ and such that for every $(g_1,h_1),\ldots,(g_k,h_k) \in G \times G$, $\mu(U^{-1}_{(g_1,h_1)}A \cap \cdots \cap U^{-1}_{(g_k,h_k)}A) \le \overline{d}(E(g_1,h_1)^{-1} \cap \cdots \cap E(g_k,h_k)^{-1})$. Let $T_g = U_{(g,e)}$ and $S_g = U_{(e,g)}$, $g \in G$. Then (T_g) and (S_g) are commuting measure preserving anti-actions.

By Theorem 3.1 there exists $\lambda > 0$ such that $\{g \in G : \mu\big(A \cap T_g^{-1} \cap (T_gS_g)^{-1}A\big) > \lambda\}$ is both C* and inverse C*. But $\mu\big(A \cap T_g^{-1} \cap (T_gS_g)^{-1}A\big) = \mu\big(A \cap U^{-1}_{(g,e)} \cap U^{-1}_{(g,g)}A\big) \le \overline{d}\big(E \cap E(g^{-1},e) \cap E(g^{-1},g^{-1})\big)$, so we are done. □

Our next result is a topological multiple recurrence theorem for three commuting anti-actions of a countable amenable group G. We do not know whether the C* conclusion can be upgraded to IP*, nor whether the C* result holds for general countable groups. However, a counterexample in [BH] shows at least that the IP* conclusion is false for general countable groups.

Theorem 4.3. Let G be a countable amenable group. Suppose (X,ρ) is a compact metric space, and let $\{T_g\}$, $\{R_g\}$, and $\{S_g\}$ be commuting anti-actions of G by homeomorphisms of X. Then for any $\epsilon > 0$, the set $\{g \in G :$ there exists $x \in X$ such that $\rho(x, R_gx) < \epsilon, \rho(x, R_gT_gx) < \epsilon$ and $\rho(x, R_gT_gS_gx) < \epsilon\}$ is C*.

Proof. Passing to a closed subset of X if necessary, we may assume that X is minimal with respect to the $G \times G \times G$-anti-action $\{T_gR_jS_h : (g,j,h) \in G \times G \times G\}$.

We claim that for every non-empty open set $U \subset X$ and every minimal idempotent ultrafilter q there exists a set $H \in q$ such that for all $g \in H$ there exists $z \in X$ such that $\{z, T_gz, T_gS_gz\} \subset U$. To prove the claim, let $U \subset X$ be open. Pick $x \in U$ and choose $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. Let $Y \subset X$ be a closed set which is minimal with respect to the $G \times G$-action $\{T_gS_h : (g,h) \in G \times G\}$. One may check that $\overline{\bigcup_{j\in G} R_jY}$ is R_g-, T_g- and S_g-invariant, and is therefore equal to X. It follows that for some $g_0 \in G$, $R^{-1}_{g_0}B_{\frac{\epsilon}{2}}(x) \cap Y \neq \emptyset$. Let $\delta > 0$ be so small that if $y_1, y_2 \in Y$ with $\rho(y_1,y_2) < \delta$ then $\rho(R_{g_0}y_1, R_{g_0}y_2) < \frac{\epsilon}{2}$. Let $U' \subset R^{-1}_{g_0}B_{\frac{\epsilon}{2}}(x) \cap Y$ be a set open in Y and of diameter less that δ. Let $y_0 \in Y$. Since the action $\{T_gS_h : (g,h) \in G \times G\}$ is minimal on Y, the set

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

is right syndetic in $G \times G$, and therefore we have $\overline{d}(E) > 0$ with respect to some (in fact any) right Følner sequence on $G \times G$. It follows from Theorem 4.2 that

$$H = \big\{g : \text{ there exists } (a,b) \in G \times G \text{ such that } \{(a,b),(ag,b),(ag,bg)\} \subset E\big\} \in q.$$

For $g \in H$, set $y = T_aS_by_0 \in U'$, where $\{(a,b),(ag,b),(ag,bg)\} \subset E$. Then $T_gy \in U'$ and $T_gS_gy \in U'$, so that letting $z = R_{g_0}y$, we have $z \in B_{\frac{\epsilon}{2}}(x)$, $\rho(z, T_gz) < \frac{\epsilon}{2}$ and $\rho(z, T_gS_gz) < \frac{\epsilon}{2}$. Therefore $\{z, T_gz, T_gS_gz\} \subset U$, establishing the claim.

We now begin the proof of Theorem 4.3. proper. Let $\epsilon > 0$, choose any minimal idempotent ultrafilter p and let $C \in p$. Choose $x_0 \in X$ arbitrarily and let U_0 be an open set of diameter less that $\frac{\epsilon}{2}$ containing x_0. Let H_0 be the member of p guaranteed by the claim above for the open set U_0. Without loss of generality, we may assume $H_0 \subset C$. Let

$h_0 \in H_0$ such that $H_0h_0^{-1} \in p$ and choose y_0 such that $\{y_0, T_{h_0}y_0, T_{h_0}S_{h_0}y_0\} \subset U_0$. Put $x_1 = R_{h_0}^{-1}y_0$ and let U_1 be an open set of diameter less that $\frac{\epsilon}{2}$ containing x_1 and having the property that for every $x \in U_1$ we have $\{R_{h_0}x, R_{h_0}T_{h_0}x, R_{h_0}T_{h_0}S_{h_0}x\} \subset U_0$.

Suppose now that we have chosen open sets of diameter less that $\frac{\epsilon}{2}$ $U_0, U_1, \cdots, U_t$ containing points $x_0, x_1, \cdots, x_t$, respectively, and $h_0, h_1, \cdots, h_{t-1} \in G$, such that for $0 \le m < n \le t$,

(i) $h_{m,n} = h_{n-1}h_{n-2}\cdots h_m \in C$,

(ii) $Ch_{m,n}^{-1} \in p$, and

(iii) $\{R_{h_{m,n}}x, R_{h_{m,n}}T_{h_{m,n}}x, R_{h_{m,n}}T_{h_{m,n}}S_{h_{m,n}}x\} \subset U_m$ for all $x \in U_n$.

Let H_t be the member of p guaranteed by the claim above for the open set U_t. Without loss of generality, we may assume that H_t is contained in $Ch_{m,n}^{-1}$ for every $0 \le m < n \le t$. Let $h_t \in H_t$ such that $H_th_t^{-1} \in p$. Note now that (i) and (ii) hold for $n = t+1$. Now pick, as we may by the claim, y_t such that $\{y_t, T_{h_t}y_t, T_{h_t}S_{h_t}y_t\} \subset U_t$. Let $x_{t+1} = R_{h_t}^{-1}y_t$ 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 $x \in U_{t+1}$ one has $\{R_{h_t}x, R_{h_t}T_{h_t}x, R_{h_t}T_{h_t}S_{h_t}x\} \subset U_t$. (iii) now follows (recall we deal with anti-actions, not actions) for $n = t+1$; that is to say, for $0 \le m \le t$ and $x \in U_{t+1}$,

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

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

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

Letting $x = x_n$ and $g = h_{m,n}$, one has $g \in C$, $\rho(x, R_gx) < \epsilon$, $\rho(x, R_gT_gx) < \epsilon$ and $\rho(x, R_gT_gS_gx) < \epsilon$. Since p and $C \in p$ were arbitrary, we are done. □

Here now is a multi-dimensional van der Waerden theorem[8] for amenable groups.

Theorem 4.4. *Let G be a countable group and let p be a minimal idempotent ultrafilter on G. For any finite partition $G \times G \times G = \bigcup_{i=1}^r C_i$,*

$$\big\{g:\ \text{there exist } i,\ 1 \le i \le r,\ \text{and } (a,b,c) \in G \times G \times G \text{ such that } \{(a,b,c), (ag,b,c), (ag,bg,c), (ag,bg,cg)\} \subset C_i\big\} \in p.$$

Proof. Let e be the identity of G. Set $\Omega = \{1, 2, \cdots, r\}^{G\times G\times G}$. 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 $\gamma(e,e,e) = \eta(e,e,e)$. We define three commuting anti-actions of G by homeomorphisms of Ω as follows: $R_g\gamma(a,b,c) = \gamma(ga,b,c)$, $T_g\gamma(a,b,c) = \gamma(a,gb,c)$ and $S_g\gamma(a,b,c) = \gamma(a,b,gc)$. Let ξ be the element of Ω defined by $\xi(j,g,h) = i$ when $(j,g,h) \in C_i$. Let $X = \overline{\{R_jT_gS_h\xi : (j,g,h) \in G \times G \times G\}}$. By Theorem 4.3, if one denotes by H the set of $g \in G$ for which there exists some $x \in X$ with $\rho(x, R_gx) < 1$, $\rho(x, R_gT_gx) < \epsilon$ and $\rho(x, R_gT_gS_gx) < \epsilon$, then $H \in p$.

[8] First proved in [vdW]; states that for any $k, r \in \mathbf{N}$ there is some n such that for any r-cell partition of $\{1, 2, \ldots, n\}$, some cell contains a k-term arithmetic progression.

Now, for $g \in H$ and x as guaranteed, one may pick $a, b, c \in G$ such that $y = R_a T_b S_c \xi$ is close enough to x to ensure that $\rho(y, R_g y) < 1$, $\rho(y, R_g T_g y) < 1$ and $\rho(y, R_g T_g S_g y) < 1$ as well. It follows that $\xi(a, b, c) = \xi(ag, b, c) = \xi(ag, bg, c) = \xi(ag, bg, cg)$, which is to say $\{(a, b, c), (ag, b, c), (ag, bg, c), (ag, bg, cg)\} \subset C_i$, where $i = \xi(a, b, c)$. □

At this time we would like to say a word or two about the strength of our results. In Sections 2 and 3, we required our idempotent ultrafilters to be minimal. However, the alert reader tracing the role of minimality to its source in the proof of Lemma 2.2, implication (i)→(ii), may have noticed that any idempotent p having the property that $A^{-1}A$ is right syndetic for all $A \in p$ would have done just as well. For that matter, it would be enough to require $(A^{-1}A)^{-1}(A^{-1}A)$ right syndetic, or even $\big((A^{-1}A)^{-1}(A^{-1}A)\big)^{-1}\big((A^{-1}A)^{-1}(A^{-1}A)\big)$ (and so on). A fully satisfactory account of the strength of the argument given would require a characterization of the class of idempotents having any of these various properties. We are not prepared to do this in general here, however we can easily cite one simple case. If G is amenable and every member A of an idempotent p has positive density with respect a given right Følner sequence, then $A^{-1}A$ will indeed be right syndetic. Below, we show that such ultrafilters are abundant; even minimal ones. More to the point, there exist such ultrafilters p that *aren't* minimal, suggesting that, indeed, the proof of our main theorem actually gives somewhat more than is claimed.

On the other hand, as a way of bounding from above what the argument actually does yield, it is instructive to see why our proof cannot easily be modified to deal with *arbitary* idempotents p. One crucial element of the proof of Theorem 3.1 was the last assertion of Theorem 2.4, namely that for minimal idempotents p, $U_g P = P U_g$, where $Pf = p\text{-}\lim_g U_g f$. (This followed from the fact that p characterized compact functions, per Theorem 2.2.) To see that this condition need not hold for arbitrary idempotents p, consider the following (sketch of an) example.

Let $X = [0,1] \times [0,1]$ with Lebesgue measure and for $n \in \mathbf{N}$ let $T_n : X \to X$ be the linear transformation which rotates each rectangle $[\frac{i}{2^n}, \frac{i+1}{2^n}] \times [0,1]$ by 90 degrees clockwise (with corresponding vertical and horizontal scaling to make it fit), $0 \le i < 2^n$. Let G be the free group on letters y_i, $i \in \mathbf{N}$, and for $g = \prod_{j=1}^{r} y_{i_j}^{e_j}$, where $e_j \in \{-1, 1\}$, put $U_g = \prod_{j=1}^{r} T_{i_j}^{-e_j}$. Then (U_g) is a measure preserving anti-action of G that induces a unitary action on $L^2(X)$. Consider the set $B = [0,1] \times [\frac{1}{2}, 1]$. It is easy to see that for any natural numbers $n < m$, one has $T_m T_n B = T_n B$. It follows that if p is any idempotent ultrafilter supported on the IP set generated by the y_n's, and we write $Pf = p\text{-}\lim_g U_g f$, then $P U_{y_n} 1_B = U_{y_n} 1_B$. On the other hand, it's pretty clear that $P 1_B \neq 1_B$, as every $T_{y_n} B$ is actually independent from B. Hence the identity $P U_g = U_g P$ fails. □

In preparation for our final application, which refines a non-commutative Schur theorem[9] proved in [BM2], and in the spirit of completeness, we will now justify the standard assertion that for any amenable group G and any (right, left or two-sided) Følner sequence

[9] Schur's theorem ([S]) states that for any $r \in \mathbf{N}$ there is some n such that for any r-cell partition of $\{1, 2, \ldots, n\}$, some cell contains a configuration of the form $\{x, y, x + y\}$.

(Φ_n), there exist ultrafilters p on G having the property that for every $A \in p$, $\overline{d}(A) > 0$ with respect to (Φ_n). To this end, fix (Φ_n) and let $\mathcal{H}$ be the family of subsets of G having density 1 with respect to (Φ_n). It is easy to see that $\mathcal{H}$ is closed under intersections and supersets, contains G and does not contain $\emptyset$. In other words, $\mathcal{H}$ is a *filter*. An easy consequence of Zorn's lemma is that any filter is contained in some ultrafilter. Let p be any ultrafilter containing $\mathcal{H}$. Then if $A \in p$, $A^c \notin p$, hence $A^c \notin \mathcal{H}$, which is to say that A^c does not have density 1 with respect to (Φ_n). In other words, $\overline{d}(A) > 0$.

In fact, for any Følner sequence (Φ_n) one may even find minimal idempotents, all of whose members A satisfy $\overline{d}(A) > 0$ with respect to (Φ_n). To see this, let p be as in the previous paragraph. We claim that every member of the right ideal $p(\beta G)$ has the property that all of its members have positive upper density with respect to (Φ_n). Indeed, let $q \in \beta G$ and suppose $A \in p \cdot q$. That is, suppose $\{x : Ax^{-1} \in p\} \in q$. Then, in particular, there is some $x \in G$ such that $Ax^{-1} \in p$, which by assumption implies $\overline{d}(Ax^{-1}) > 0$. But since (Φ_n) is a right Følner sequence, $\overline{d}(A) > 0$. Now simply choose a minimal right ideal in $p(\beta G)$ and an idempotent q in that minimal right ideal.

We shall make use of the following well-known facts as well:

(a) Suppose $E \subset G$, and let (Φ_n) be a (left, right or two-sided) Følner sequence for G. There exists an increasing sequence $(n_k)_{k\in\mathbf{N}} \subset \mathbf{N}$ such that $d(E)$ exists with respect to $\{\Phi_{n_k}\}$.

(b) If G is a countable amenable group then $G\times G$ is as well. Furthermore, if (Φ_n) and $\{\Psi_n\}_{n\in\mathbf{N}}$ are any two (left, right or two-sided) Følner sequences for G then $\{\Phi_n \times \Psi_n\}_{n\in\mathbf{N}}$ is a (left, right or two-sided) Følner sequence for $G \times G$.

(c) If G is a countable amenable group, $A \subset G$ is a subgroup, and (Φ_n) is any (left, right or two-sided) Følner sequence for G then $d(A)$ (with respect to (Φ_n)) exists and is equal to $\frac{1}{[G:A]}$ if $[G : A] < \infty$ and 0 if $[G : A] = \infty$.

(d) If G is a group having two subgroups of finite index A and B then $A \cap B$ is of finite index as well. (If $[G : A] = m$ and $[G : B] = n$, consider the homomorphism $g \to (\phi_1(g), \phi_2(g))$ of G into $S_m \times S_n$ (S_m being the group of permutations of the left cosets of A and S_n being the group of permutations of the left cosets of B), where $\phi_1(g)(hA) = (gh)A$ and $\phi_2(g)(hB) = (gh)B$. The kernel lies in $A \cap B$.)

Theorem 4.5. Suppose that G is a countable amenable group having the property that, letting $A = \{g \in G : [G : C(g)] < \infty\}$, $[G : A] = \infty$. Let (Φ_n) be a two-sided Følner sequence for G and suppose p is an idempotent ultrafilter on G having the property that for every $A \in p$, $\overline{d}(A) > 0$ with respect to (Φ_n). Then for every $C \in p$ there exist $x, y \in G$ with $xy \neq yx$ such that $\{x, y, xy, yx\} \subset C$.

Proof. That A is a subgroup of G is a consequence of (d), since $C(g) \cap C(h)$ is contained in $C(gh)$. Let (Φ_n) be a two-sided Følner sequence for G with regard to which $\overline{d}(C) > 0$. Passing if necessary to a sub-sequence of (Φ_n) we may by (a) assume that $d(C)$ exists. Let $(k_n)_{n\in\mathbf{N}}$ be a sequence in $\mathbf{N}$ having the property that $\frac{|\Phi_{k_n} \cap g^{-1}\Phi_{k_n}|}{|\Phi_{k_n}|} > 1 - \frac{1}{n}$ for all $n \in \mathbf{N}$ and all $g \in \Phi_n$ (this is possible since (Φ_n) is a left Følner sequence). By (b), $\{\Phi_n \times \Phi_{k_n}\}_{n\in\mathbf{N}}$ is a two-sided Følner sequence for $G \times G$.

For any $S \subset G$ let $\tilde{S} = \{(a, b) \in G \times G : a^{-1}b \in S\}$. We claim that if $d(S)$ exists then $d(\tilde{S})$ (measured with respect to $\{\Phi_n \times \Phi_{k_n}\}_{n\in\mathbf{N}}$) exists and equals $d(S)$. To see this, consider that

$$\begin{aligned}\lim_{n\to\infty}\frac{|\tilde{S}\cap(\Phi_n\times\Phi_{k_n})|}{|\Phi_n\times\Phi_{k_n}|} &= \lim_{n\to\infty}\frac{1}{|\Phi_{k_n}||\Phi_n|}\sum_{g\in\Phi_n}|\tilde{S}\cap(\{g\}\times\Phi_{k_n})| \\ &= \frac{1}{|\Phi_{k_n}||\Phi_n|}\sum_{g\in\Phi_n}|\Phi_{k_n}\cap gS| \\ &= \lim_{n\to\infty}\frac{1}{|\Phi_n|}\sum_{g\in\Phi_n}\frac{|g^{-1}\Phi_{k_n}\cap S|}{|\Phi_{k_n}|} \\ &= \lim_{n\to\infty}\frac{1}{|\Phi_n|}\sum_{g\in\Phi_n}\frac{|\Phi_{k_n}\cap S|}{|\Phi_{k_n}|} = d(S).\end{aligned}$$

For any $E\in p$, $\overline{d}(E)>0$ with respect to (Φ_n), a right Følner sequence, which implies that $E^{-1}E$ is right syndetic. Hence (see discussion above) the main theorem and its corollaries (in particular Theorem 4.2) apply to p. So, since $d(\tilde{C})>0$, by Theorem 4.2 there is some $\lambda>0$ such that $B=\{g:\overline{d}\big(\tilde{C}\cap\tilde{C}(g^{-1},e)\cap\tilde{C}(g^{-1},g^{-1})\big)>\lambda\}\in p$. Since every member of p has positive upper density with respect to (Φ_n), $A\notin p$, so $(B\setminus A)\in p$. Choose $g\in(B\setminus A)\cap C$. Since $\overline{d}\big(\tilde{C}\cap\tilde{C}(g^{-1},e)\cap\tilde{C}(g^{-1},g^{-1})\big)>\lambda$, while $\overline{d}\big(\tilde{C(g)}\big)=d\big(C(g)\big)=0$, we may choose $(a,b)\in\big(\tilde{C}\cap\tilde{C}(g^{-1},e)\cap\tilde{C}(g^{-1},g^{-1})\big)\setminus\tilde{C(g)}$. Then $\{(a,b),(ag,b),(ag,bg)\}\subset\tilde{C}$, which implies that $\{g,a^{-1}b,g^{-1}a^{-1}b,g^{-1}a^{-1}bg\}\subset C$, while $a^{-1}b$ does not commute with g. Letting $x=g$ and $y=g^{-1}a^{-1}b$, we are done. □

5. Jointly weak mixing systems of G-anti-actions.

Theorem 3.1 prompts the following natural question. Given k commuting measure preserving G-anti-actions $(T_g^{(i)})$, $1\le i\le k$, on a probability space $(X,\mathcal{A},\mu)$, and given $A\in\mathcal{A}$ with $\mu(A)>0$, does there exist $\lambda>0$ such that

$$\{g\in G:\mu\big(A\cap(T_g^{(1)})^{-1}A\cap(T_g^{(1)}T_g^{(2)})^{-1}A\cap\cdots\cap(T_g^{(1)}T_g^{(2)}\cdots T_g^{(k)})^{-1}A\big)>\lambda\}$$

is C*?

We strongly suspect the answer to this question to be *yes*, and are hopeful that a proof will be not long in coming. Although there are some as-yet unresolved obstacles[10], it is natural to imagine that the proof will incorporate elements now familiar to the ergodic theoretic proofs of various extensions of Szemerédi's theorem on arithmetic progressions, such as multiple weak mixing for "jointly" weak mixing systems of G-anti-actions. [11]

[10] The most serious obstacle is non-invariance of various naturally arising factors under some of the G-actions. For example, if P is the projection onto the space spanned by the $(T_g^{(1)}T_g^{(2)})$-almost periodic functions, we see no reason why $T_g^{(1)}P=PT_g^{(1)}$ should hold, so no successful adaptation of an extant proof is likely to be completely straightforward.

[11] A single measure preserving anti-action (T_g) is weak mixing if the only compact vectors for the associated unitary action on L^2 are the constants. By a jointly weak mixing system we essentially mean one satisfying the hypotheses of Theorem 5.1.

In this short section, we intend to show that this step, at least, works fine. At the same time, we will be giving an affirmative answer to the aforementioned question under the additional joint weak mixing hypothesis.

Theorem 5.1. Let $(X, \mathcal{A}, \mu)$ be a probability space, $k \in \mathbf{N}$ and suppose that $(T_g^{(i)})$, $1 \leq i \leq k$, are anti-actions of G by measure preserving transformations of X. Let $p \in \beta G$ be a minimial idempotent and suppose that for all $1 \leq i \leq j \leq k$, the anti-action $(\prod_{t=i}^{j} T_g^{(t)})$ is weak mixing. Then for $f_i \in L^\infty(X)$, $0 \leq i \leq k$,

$$p\text{-lim}_g \int f_0(T_g^{(1)} f_1)(T_g^{(1)} T_g^{(2)} f_2) \cdots (T_g^{(1)} T_g^{(2)} \cdots T_g^{(k)} f_k)\, d\mu = \prod_{i=0}^{k} \int f_i \, d\mu.$$

Proof. As is standard, we make use under the integral of the identity $\prod_{i=0}^k a_i - \prod_{i=0}^k b_i = (a_0 - b_0) b_1 \cdots b_k + a_0(a_1 - b_1) b_2 \cdots b_k + \cdots + a_0 \cdots a_{k-1}(a_k - b_k)$, with $a_i = \left(\prod_{j=1}^{i} T_g^{(j)}\right) f_i$ and $b_i = \int f_i \, d\mu$. This allows us to assume, without loss of generality, that $\int f_i \, d\mu = 0$ for some i, $0 \leq i \leq k$.

The proof is by induction on k. If $k = 1$, let $P = p\text{-lim}_g T_g^{(1)}$ in the weak operator topology. By Theorem 2.4, P is an orthogonal projection, and by Theorem 2.2 and the fact that $(T_g^{(1)})$ is weak mixing, P is in fact the projection onto the constants. In other words, $p\text{-lim}_g T_g^{(1)} f_1 = \int f_1 \, d\mu$ weakly, from which it follows that $p\text{-lim}_g \int f_0 T_g^{(1)} f_1 \, d\mu = \left(\int f_0 \, d\mu\right)\left(\int f_1 \, d\mu\right)$, as desired.

Suppose now that the result holds for $k-1$. We employ Theorem 2.3. For $g \in G$, let $x_g = (T_g^{(1)} f_1)(T_g^{(1)} T_g^{(2)} f_2) \cdots (T_g^{(1)} T_g^{(2)} \cdots T_g^{(k)} f_k)$. Then

$$\begin{aligned}
&\langle x_{gh}, x_g \rangle \\
&= \int (T_{gh}^{(1)} f_1)(T_{gh}^{(1)} T_{gh}^{(2)} f_2) \cdots (T_{gh}^{(1)} T_{gh}^{(2)} \cdots T_{gh}^{(k)} f_k) \\
&\qquad\qquad (T_g^{(1)} f_1)(T_g^{(1)} T_g^{(2)} f_2) \cdots (T_g^{(1)} T_g^{(2)} \cdots T_g^{(k)} f_k)\, d\mu \\
&= \int (T_h^{(1)} f_1)(T_h^{(1)} T_{gh}^{(2)} f_2) \cdots (T_h^{(1)} T_{gh}^{(2)} \cdots T_{gh}^{(k)} f_k)(f_1)(T_g^{(2)} f_2) \cdots (T_g^{(2)} \cdots T_g^{(k)} f_k)\, d\mu \\
&= \int \left(f_1 T_h^{(1)} f_1\right) T_g^{(2)} \left(f_2 T_h^{(1)} T_h^{(2)} f_2\right) T_g^{(2)} T_g^{(3)} \left(f_3 T_h^{(1)} T_h^{(2)} T_h^{(3)} f_3\right) \cdots \\
&\qquad\qquad T_g^{(2)} T_g^{(3)} \cdots T_g^{(k)} \left(f_k T_h^{(1)} T_h^{(2)} \cdots T_h^{(k)} f_k\right) d\mu.
\end{aligned}$$

Applying now the induction hypothesis in the inner limit and the $k = 1$ case in the outer, using the fact that $\int f_i \, d\mu = 0$ for some i,

$$\begin{aligned}
&p\text{-lim}_h p\text{-lim}_g \langle x_{gh}, x_g \rangle \\
&= p\text{-lim}_h \left(\int f_1 T_h^{(1)} f_1 \, d\mu\right)\left(\int f_2 T_h^{(1)} T_h^{(2)} f_2 \, d\mu\right) \cdots \left(\int f_k T_h^{(1)} T_h^{(2)} \cdots T_h^{(k)} f_k \, d\mu\right) = 0.
\end{aligned}$$

Now the conclusion of Theorem 2.3 gives $p\text{-lim}_g x_g = 0$ weakly, so that in particular $p\text{-lim}_g \int f_0(T_g^{(1)} f_1)(T_g^{(1)} T_g^{(2)} f_2) \cdots (T_g^{(1)} T_g^{(2)} \cdots T_g^{(k)} f_k)\, d\mu = 0$, as required. □

Taking $f_i = 1_A$, $0 \leq i \leq k$, the following is now immediate.

Corollary 5.2. Let $(X, \mathcal{A}, \mu)$ be a probability space, $k \in \mathbf{N}$ and suppose that $(T_g^{(i)})$, $1 \leq i \leq k$, are anti-actions of G by measure preserving transformations of X. Suppose that for all $1 \leq i \leq j \leq k$, the anti-action $(\prod_{t=i}^{j} T_g^{(t)})$ is weak mixing. Then for any $A \in \mathcal{A}$ with $\mu(A) > 0$ and any $\epsilon > 0$, the set

$$\{g \in G : \big|\mu\big(A \cap (T_g^{(1)})^{-1}A \cap (T_g^{(1)}T_g^{(2)})^{-1}A \cap \cdots \cap (T_g^{(1)}T_g^{(2)} \cdots T_g^{(k)})^{-1}A\big) - \mu(A)^{k+1}\big| < \epsilon\}$$

is C*.

REFERENCES

[B] V. Bergelson, Minimal idempotents and ergodic Ramsey theory, *Topics in Dynamics and Ergodic Theory*, 8-39. *London Math. Soc. Lecture Note Ser.*, **310**, Cambridge Univ. Press, Cambridge, 2003.

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

[BM1] V. Bergelson and R. McCutcheon, *An Ergodic IP Polynomial Szemerédi Theorem, Mem. Amer. Math. Soc.* **695** (2000).

[BM2] V. Bergelson and R. McCutcheon, Recurrence for semigroup actions and a non-commutative Schur theorem, *Contemporary Mathematics.* **215** (1998), 205-222.

[BMZ] V. Bergelson, R. McCutcheon and Q. Zhang, A Roth theorem for amenable groups, *Amer. J. Math.* **119** (1997), 1173–1211.

[E] R. Ellis, Distal transformation groups, *Pacif. J. Math.* **8** (1958), 401-405.

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

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

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

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

[FK3] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. d'Analyse Math.* **57** (1991), 64-119.

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

[L] A. Leibman, Multiple recurrence theorem for measure preserving actions of a nilpotent group, *Geom. Funct. Anal.* **8** (1998), 853-931.

[S] Schur, Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$, *Jahresbericht der Deutschen Math.-Ver.* **25** (1916), 114-117.

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

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

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

Randall McCutcheon
Dept. of Math. Sciences
University of Memphis
Memphis, TN 38152 USA
rmcctchn@memphis.edu