MINIMAL IDEMPOTENTS AND ERGODIC RAMSEY THEORY

VITALY BERGELSON

Contents

1. Introduction

What is common between the invertibility of distal maps, partition regularity of diophantine equation $x - y = z^2$, and the notion of mild mixing? The answer is: idempotent ultrafilters, and the goal of this survey is to convince the reader of the unifying role and usefulness of idempotent ultrafilters (and, especially, the minimal ones) in ergodic theory, topological dynamics and Ramsey theory.

We start with reviewing some basic facts about ultrafilters. The reader will find the missing details and more information in the self-contained Section 3 of [B2]. (See also [HiS] for a comprehensive presentation of the material related to topological algebra in the Stone-Čech compactification).

An ultrafilter p on $\mathbb{N} = \{1, 2, ...\}$ is, by definition, a *maximal filter*, namely, a nonempty family of subsets of $\mathbb{N}$ satisfying the following conditions (the first three of which constitute the definition of a *filter*):

(i) $\emptyset \notin p$;

(ii) $A \in p$ and $A \subset B$ imply $B \in p$;

(iii) $A \in p$ and $B \in p$ imply $A \cap B \in p$;

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

The space of ultrafilters, denoted by $\beta\mathbb{N}$, and equipped with appropriately defined topology, is nothing but Stone-Čech compactification of $\mathbb{N}$ and plays

This work was partially supported by NSF under the grants DMS-9706057 and DMS-0070566.

an important role in various areas of mathematics including topology, analysis and ergodic Ramsey theory.

In what follows we will find it useful to view ultrafilters as finitely-additive, {0,1}-valued probability measures on the power set $\mathcal{P}(\mathbb{N})$.

Given an ultrafilter $p \in \beta\mathbb{N}$, define a mapping $\mu_p : \mathcal{P}(\mathbb{N}) \to \{0,1\}$ by $\mu_p(A) = 1 \Leftrightarrow A \in p$. It is easy to see that $\mu_p(\emptyset) = 0$, $\mu_p(\mathbb{N}) = 1$ (follows from (i), (iv) and (ii)), and that for any finite collection of disjoint sets $A_1, A_2, ..., A_r$, one has $\mu_p(\bigcup_{i=1}^r A_i) = \sum_{i=1}^r \mu_p(A_i)$. Indeed, note that if none of A_i belongs to p, then both sides equal zero. Also, it follows from (i) that at most one among the (disjoint!) sets A_i may satisfy $A \in p$, in which case both sides of the above equation equal one.

One of the major advantages of viewing the ultrafilters as measures is that one can naturally define the convolution operation which makes $\beta\mathbb{N}$ a compact semigroup. Given two σ-additive measures μ and ν on a topological group G, the convolution is usually defined as $\mu * \nu(A) = \int_G \mu(Ay^{-1})d\nu(y)$. In particular, $\mu * \nu(A) > 0$ iff for ν-many y one has $\mu(Ay^{-1}) > 0$. Taking into account that a value of ultrafilter measure on a set $A \subseteq \mathbb{N}$ is positive iff it equals one, we make the following definition in which for a reason to be explained in the remark below, we denote the convolution by $+$.

Definition 1.1. *Given $p, q \in \beta\mathbb{N}$, the convolution $p + q$ is defined by*

$$p + q = \{A \subseteq \mathbb{N} : \ \{n : \ (A - n) \in p\} \in q\}.$$

In other words, A is $(p+q)$-large iff the set $A - n = \{n \in \mathbb{N} : \ m + n \in A\}$ is p-large for q-many n.

It is not too hard to check that $p+q$ is an ultrafilter and that the operation defined above is associative (see, for example, [B2], p.27).

Now we shall explain the reason for denoting this operation by $+$. For any $n \in \mathbb{N}$ define an ultrafilter μ_n as a "delta measure" concentrated at point n:

$$\mu_n(A) = \begin{cases} 1, & n \in A \\ 0, & n \notin A. \end{cases}$$

The ultrafilters $\mu_n, n \in \mathbb{N}$, are called *principal* and it is clear that for any $n, k \in \mathbb{N}$ the convolution of μ_n and μ_k equals μ_{n+k}. In other words, the principal ultrafilters μ_n, $n \in \mathbb{N}$, form a semigroup which is isomorphic to $(\mathbb{N}, +)$ and the convolution defined above extends the operation $+$ to the space $\beta\mathbb{N}$, the closure of $\mathbb{N}$. At this point it will be instructive to say a few words about the topology on $\beta\mathbb{N}$. Given $A \subset \mathbb{N}$, let $\overline{A} = \{p \in \beta\mathbb{N} : \ A \in p\}$. It is immediate that for any $A, B \subset \mathbb{N}$ one has $\overline{A} \cap \overline{B} = \overline{A \cap B}$, $\overline{A} \cup \overline{B} = \overline{A \cup B}$. Also, since $\overline{\mathbb{N}} = \beta\mathbb{N}$, one has $\bigcup_{A \in \mathcal{A}} \overline{A} = \beta\mathbb{N}$, where $\mathcal{A} = \{\overline{A} : \ A \subset \mathbb{N}\}$. It follows that the set $\mathcal{A}$ forms the basis for the open sets of $\beta\mathbb{N}$ (and the basis for closed sets too!). One can show that with this topology $\beta\mathbb{N}$ is a compact Hausdorff space and that for any fixed $p \in \beta\mathbb{N}$ the function $\lambda_p(q) = p + q$ is a continuous self map of $\beta\mathbb{N}$ (see for example Theorems 3.1 and 3.2 in [B2]). In view of these facts, $(\beta\mathbb{N},+)$ becomes a compact *left topological* semigroup. We remark in passing that the operation $\rho_p(q)=q+p$ is, unlike $\lambda_p(q)$, continuous only when p is a principal ultrafilter, and that the convolution defined above on $\beta\mathbb{N}$ is the unique extension of the operation $+$ on $\mathbb{N}$ such that $\lambda_p(q)$ and $\rho_p(q)$ have the properties described above.

Before going on to explore additional features of the semigroup $(\beta\mathbb{N},+)$ that are important for us we want to caution the reader that while having various nice and convenient properties, the semigroup $(\beta\mathbb{N},+)$ is in many respects an odd and counterintuitive object. First, the compact Hausdorff space $\beta\mathbb{N}$ is too large to be metrizable: its cardinality is that of $\mathcal{P}(\mathcal{P}(\mathbb{N}))$. Yet, in view of the fact that $\overline{\mathbb{N}}=\beta\mathbb{N}$, it is a closure of a countable set $\mathbb{N}$. Second, the operation $+$ on $\beta\mathbb{N}$ is highly non-commutative: the center of the semigroup $(\beta\mathbb{N},+)$ contains only the principal ultrafilters. (Here the analogy with the convolution of σ-additive measures on locally compact abelian groups fails. The reason: the ultrafilters, being only *finitely* additive measures, do not obey the Fubini theorem which is crucial for the commutativity of the convolution of σ-additive measures).

By a theorem due to Ellis [E1], any compact semigroup with a left-continuous operation has an idempotent. Actually, $(\beta\mathbb{N},+)$ has plenty of them, since any compact subsemigroup in $(\beta\mathbb{N},+)$ should have one and there are 2^{c} disjoint compact subsemigroups in $\beta\mathbb{N}$. As we shall see below, of special importance for combinatorial and ergodic-theoretical applications are *minimal* idempotents, which we will define and apply later in this section. In a way, idempotent ultrafilters in $\beta\mathbb{N}$ are, in a way, just generalized shift-invariant measures. Indeed, if $p+p=p$, it means that any $A\in p=p+p$ has the property that $\{n:\ (A-n)\in p\}\in p$, or, in other words, for p-almost all n, the set $A-n$ is p-large.

It is easy to see that principal ultrafilters are never idempotent and hence, if p is an idempotent ultrafilter, any p-large set A is infinite, as is the p-large set $\{n:\ (A-n)\in p\}$. As we shall presently see, the members of idempotent ultrafilters always contain highly structured subsets which can be viewed as generalized subsemigroups of $\mathbb{N}$.

Let $A\in p$, where $p+p=p$. Since

$$A\cap\{n:\ (A-n)\in p\}\in p,$$

we can choose $n_1\in A$ such that $A_1=A\cap(A-n_1)\in p$. (Note that this is nothing but a version of Poincaré recurrence theorem; the important bonus is that $n_1\in A$. By iterating this procedure one can chose $n_2\in A\cap(A-n_1)$, $n_2>n_1$, such that

$$A_1\cap(A_1-n_2)=A\cap(A-n_1)\cap(A-n_2)\cap(A-n_1-n_2)\in p.$$

Note that $n_1,n_2,n_1+n_2\in A$. Continuing in this fashion, one obtains an increasing sequence $(n_i)_{i=1}^{\infty}$ and inductively defined sets $A=A_0,A_1,A_2,...$, such that $n_1\in A$, $n_{i+2}\in A_{i+1}:=\ A_i\cap(A_i-n_{i+1})$, $i=0,1,2,....$ One readily checks that this construction implies that A contains the set of *finite sums* of $(n_i)_{i=1}^{\infty}$:

$$FS(n_i)_{i=1}^{\infty} = \{n_{i_1} + n_{i_2} + ... + n_{i_k},\ k \in \mathbb{N},\ i_1 < i_2 < ... < i_k\}.$$

Such sets of finite sums are customarily called IP sets (IP stands for IdemPotent) and are featured in the following important theorem due to N.Hindman [Hi1].

Theorem 1.2. (N.Hindman). *For any finite partition $\mathbb{N} = \bigcup_{i=1}^{r} C_i$ one of the cells of partition contains an IP set.*

Proof. Fix any idempotent ultrafilter $p \in \beta\mathbb{N}$ and observe that one (and only one!) of C_i belongs to it. Now use the fact proved above that any member of p contains an IP set. □

Let $\mathcal{F}$ denote the family of non-empty finite subsets of N. Noticing that the mapping $\mathcal{F} \to \mathbb{N}$ defined by $\{i_1, i_2, ..., i_k\} \to 2^{i_1} + 2^{i_2} + ... + 2^{i_k}$ is 1-1 and that elements of IP sets are naturally indexed by elements of $\mathcal{F}$, we have that each of the following two theorems implies Hindman's theorem, each revealing yet another facet of it.

Theorem 1.3. (Finite unions theorem). *For any finite partition $\mathcal{F} = \bigcup_{i=1}^{r} C_i$, one of C_i contains an infinite sequence of non-empty disjoint sets $(U_i)_{i\in\mathbb{N}}$ together with all the unions $U_{i_1} \cup U_{i_2} \cup ... \cup U_{i_k},\ i_1 < i_2 < ... < i_k, k \in \mathbb{N}$. In addition, one can assume without the loss of generality that for all $i \in \mathbb{N}$ one has $\max U_i < \min U_{i+1}$.*

Theorem 1.4. *For any finite partition of an IP set in $\mathbb{N}$ one of the cells of the partition contains an IP set.*

Exercise 1. Prove that Theorems 1.2, 1.3, 1.4 are equivalent.

In the proof of Hindman's theorem above IP sets emerge as subsets of members of idempotent ultrafilters. One may wonder whether given an idempotent p and a set $A \in p$, it is possible to find in A an IP set which is itself p-large. It turns out that this is not always the case. For example, the minimal idempotents which we will define below, can not have this property. The following theorem shows that, nevertheless, any IP set is a support of an idempotent.

Theorem 1.5. *For any sequence $(x_i)_{i\in\mathbb{N}}$ in $\mathbb{N}$ there is an idempotent $p \in \beta\mathbb{N}$ such that $FS((x_i)_{i\in\mathbb{N}}) \in p$.*

Sketch of the proof. Let $\Gamma = \bigcap_{n=1}^{\infty} \overline{FS((x_i)_{i=n}^{\infty})}$. (The closures are taken in the natural topology of $\beta\mathbb{N}$). Clearly, Γ is compact and non-empty. It is not hard to show that Γ is a subsemigroup of $(\beta\mathbb{N}, +)$. Being a compact left-topological semigroup, Γ has an idempotent. If $p \in \Gamma$ is an idempotent, then $\overline{\Gamma} = \Gamma \ni p$

which, in particular, implies $FS((x_i)_{i=1}^{\infty}) \in p$. $\square$

The above definitions and theorems readily extend to general semigroups. Given a semigroup $(G, \cdot)$, one defines βG as the set of ultrafilters on G. The semigroup operation naturally extends to βG by the rule

$$A \in p \cdot q \Leftrightarrow \{x : Ax^{-1} \in p\} \in q$$

(where $Ax^{-1} := \{y \in G : yx \in A\}$).

Exercise 2. Verify that $(\beta G, \cdot)$ is a left-topological compact semigroup.

The IP sets, which, in the case of multiplicative notation, become *finite product sets*, are defined as follows. Given any sequence $(x_n)_{n\in\mathbb{N}} \subset G$ and $F \in \mathcal{F}$ denote by $\prod_{n\in F} x_n$ the product of $x_n, n \in F$ in the *decreasing* order of indices. Then the IP set generated by $(x_n)_{n\in\mathbb{N}}$ is defined as $FP((x_n)_{n=1}^{\infty}) = \{\prod_{n\in F} x_n, F \in \mathcal{F}\}$. As in the case of $(\beta\mathbb{N}, +)$, the IP sets in general semigroups are closely related to idempotent ultrafilters (whose existence follows from Ellis' theorem, alluded to above). In particular, one can check that the proof of Hindman's theorem given above transfers verbatim to give a proof of the following theorem and its corollary. Note that the Corollary 1.7 below can also be obtained from the finite unions theorem (Theorem 1.3 above).

Theorem 1.6. *Let $(G, \cdot)$ be a discrete semigroup and let p be an idempotent in $(\beta G, \cdot)$. Then for any $A \in p$ there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in βS such that $FP((x_n)_{n=1}^{\infty}) \subseteq A$.*

Corollary 1.7. *For any finite partition $G = \bigcup_{i=1}^{r} C_i$ one of the C_i contains an IP set.*

Exercise 3. Give a detailed proof of Theorem 1.6 and Corollary 1.7.

Since $\mathbb{N}$ (and, hence, $\beta\mathbb{N}$) has two natural structures, namely, those of addition and multiplication, it follows that for any finite partition $\mathbb{N} = \bigcup_{i=1}^{r} C_i$ there are $i, j \in \{1, 2, ..., r\}$ such that C_i contains an additive IP set and C_j contains a multiplicative IP set. The following theorem due to Hindman shows that one can always have $i = j$.

Theorem 1.8. ([Hi2]) *For any finite partition $\mathbb{N} = \bigcup_{i=1}^{r} C_i$ there exists $i \in \{1, 2, ..., r\}$ and sequences $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS((x_n)_{n=1}^{\infty}) \cup FP((y_n)_{n=1}^{\infty}) \subseteq C_i$.*

Proof. Let Γ be the closure in $\beta\mathbb{N}$ of the set of additive idempotents. We claim that $p \in \Gamma$ if and only if every p-large set A contains an additive IP set. Indeed, if $A \in p \in \Gamma$, then $\overline{A}$ is a (clopen) neighborhood of p. It follows that there exists $q \in \overline{A}$ with $q + q = q$. Then $A \in q$ and by Theorem 1.2 A contains an IP set. Conversely, if $\overline{A}$ is a basic neighborhood of p and for some $(x_n)_{n=1}^{\infty}$, $FS((x_n)_{n=1}^{\infty}) \subseteq A$, then by Theorem 1.5 above there exists an idempotent q with $FS((x_n)_{n=1}^{\infty}) \in q$, which implies $A \in q$, and hence $p \in \Gamma$.

We will show now that Γ is a right ideal in $(\beta\mathbb{N}, \cdot)$. Let $p \in \Gamma$, $q \in \beta\mathbb{N}$, and let $A \in p \cdot q$. Then $\{x \in \mathbb{N} : \ Ax^{-1} \in p\} \in q$ and, in particular, $\{x \in \mathbb{N} : \ Ax^{-1} \in p\}$ is non-empty. Let x be such that $Ax^{-1} \in p$. Since $p \in \Gamma$, there exists a sequence $(y_n)_{n=1}^{\infty}$ with $FS((y_n)_{n=1}^{\infty}) \subseteq Ax^{-1}$, which implies $FS((xy_n)_{n=1}^{\infty}) \subseteq A$ and so $p \cdot q \in \Gamma$. We see that Γ is a compact subsemigroup in $(\beta\mathbb{N}, \cdot)$ and hence contains a multiplicative idempotent. To finish the proof, let $\cup_{i=1}^{r} C_i = \mathbb{N}$ and let $p \in \Gamma$ satisfy $p \cdot p = p$. Let $i \in \{1, 2, ..., r\}$ be such that $C_i \in p$. Then, since $p \in \Gamma$, C_i contains an additive IP set. Also, since p is a multiplicative idempotent, C_i contains (by Theorem 1.6) a multiplicative IP set. We are done. □

Remarks (i) For an elementary proof of Theorem 1.8 see [BH2].

(ii) Theorem 2.12 below shows that for any finite partition $\bigcup_{i=1}^{r} C_i = \mathbb{N}$ one C_i has interesting additional properties. In particular, one C_i can be shown to contain in addition to an additive and a multiplicative IP sets, also arbitrarily long arithmetic and arbitrarily long geometric progressions.

2. Minimal idempotents, central sets, and combinatorial applications

Let $\sigma : \mathbb{N} \to \mathbb{N}$ denote the shift operation: $\sigma(x) = x + 1$, $x \in \mathbb{N}$. As we saw above, all that it takes to prove Hindman's theorem is to apply a version of Poincaré recurrence theorem to the "measure preserving system" $(\mathbb{N}, p, \sigma)$, where p is an arbitrary idempotent in $(\beta\mathbb{N}, +)$. Indeed, the idempotence of p implies that any p-large set A has the property that for p-many $n \in \mathbb{N}$ the set $A - n = \sigma^{-n}(A)$ is also p-large and hence for some n (which in our situation can be chosen from A) one has $(A \cap \sigma^{-n}(A)) \in p$. The rest of the proof is just a routine iteration.

We shall introduce now an important subclass of idempotents which will allow us to make a connection with another basic dynamical notion, namely, that of a minimal dynamical system.

A topological dynamical system (with "time" $\mathbb{N}$) is a pair (X, T), where X is a compact space and $T : X \to X$ is a continuous map. The system (X, T) is called *minimal*, if for any $x \in X$ one has $\overline{(T^n x)_{n \in \mathbb{N}}} = X$. One can show, by a simple application of Zorn's lemma, that any system (X, T) contains a minimal compact non-empty T-invariant subset Y which, consequently, gives rise to minimal system (Y, T) (by slight abuse of notation we denote the restriction of T to Y by the same symbol). Extending the shift operation σ from $\mathbb{N}$ to $\beta\mathbb{N}$ by the rule $q \to q + 1$ (where 1 denotes the principal ultrafilter of sets containing the integer 1), we obtain a topological dynamical system

$(\beta\mathbb{N}, \sigma)$. The following theorem establishes the connection between minimal subsystems of $(\beta\mathbb{N}, \sigma)$ and minimal right ideals in $(\beta\mathbb{N}, +)$.

Theorem 2.1. *The minimal closed invariant subsets of the dynamical system* $(\beta\mathbb{N}, \sigma)$ *are precisely the minimal right ideals of* $(\beta\mathbb{N}, +)$.

Proof. We first observe that closed σ-invariant sets in $\beta\mathbb{N}$ coincide with right ideals. Indeed if I is a right ideal, i.e. satisfies $I + \beta\mathbb{N} \subseteq I$, then for any $p \in I$ one has $p+1 \in I+\beta\mathbb{N} \subseteq I$, so that I is σ-invariant. On the other hand, if S is a closed σ-invariant set in $\beta\mathbb{N}$ and $p \in S$, then $p+\beta\mathbb{N} = p+\overline{\mathbb{N}} = \overline{p+\mathbb{N}} \subseteq \overline{S} = S$, which implies $S + \beta\mathbb{N} \subseteq S$.

Now the theorem follows from a simple general fact that any minimal right ideal in a compact left-topological semigroup $(G, \cdot)$ is closed. Indeed, if R is a right ideal in $(G, \cdot)$ and $x \in R$, then xG is compact as the continuous image of G and is an ideal. Hence the minimal ideal containing x is compact as well. (The fact that R contains a minimal ideal follows by an application of Zorn's lemma to the non-empty family $\{I : I \text{ is a closed right ideal of } G \text{ and } I \subseteq R\}$). □

Our next step is to observe that any minimal right ideal in $(\beta\mathbb{N}, +)$, being a compact left-topological semigroup, contains, by Ellis' theorem, an idempotent.

Definition 2.2. *An idempotent* p *in* $(\beta\mathbb{N}, +)$ *is called* minimal *if* p *belongs to a minimal ideal.*

Theorem 2.3. *Any minimal subsystem of* $(\beta\mathbb{N}, \sigma)$ *is of the form* $(p + \beta\mathbb{N}, \sigma)$ *where* p *is a minimal idempotent in* $(\beta\mathbb{N}, +)$.

Proof. It is obvious that, for any $p \in (\beta\mathbb{N}, +)$, $p + \beta\mathbb{N}$ is a right ideal. To see that any minimal right ideal is of this form, take any $q \in R$ and observe that $q + \beta\mathbb{N} \subseteq R + \beta\mathbb{N} \subseteq R$. Since R is minimal, we get $q + \beta\mathbb{N} = R$. In particular, one can take q to be an idempotent.

Before moving to some immediate corollaries of Theorem 2.3 we want to remind the reader that a set $A \subseteq \mathbb{N}$ is called *syndetic* if it has bounded gaps, or equivalently, if for some finite $F \subset \mathbb{N}$ one has $\bigcup_{t\in F}(A - t) = \mathbb{N}$. A set $A \subseteq \mathbb{N}$ is called *piecewise syndetic* if it can be represented as an intersection of a syndetic set with an infinite union of intervals $[a_n, b_n]$, where $b_n - a_n \to \infty$.

An equivalent definition of piecewise syndeticity (and the one which makes sense in any semigroup) is given in the following exercise.

Exercise 4. Prove that a set $A \subseteq \mathbb{N}$ is piecewise syndetic if and only if there exists a finite set $F \subset \mathbb{N}$, such that the family $\{\bigcup_{t\in F}(A - t) - n : n \in \mathbb{N}\}$ has the finite intersection property.

Exercise 5. (i) Prove that if (X,T) is a minimal system then every point $x \in X$ has a dense orbit.

(ii) Prove that if (X,T) is a minimal system then for any $x \in X$ and any neighborhood V of x the set $\{n : \ T^n x \in V\}$ is syndetic.

Theorem 2.4. *Let p be a minimal idempotent in $(\beta\mathbb{N},+)$.*
(i) *For any $A \in p$ the set $B = \{n : \ (A-n) \in p\}$ is syndetic.*
(i) *Any $A \in p$ is piecewise syndetic.*
(iii) *For any $A \in p$ the set $A - A = \{n_1 - n_2 : n_1, n_2 \in A\}$ is syndetic.*

Proof. Statement (i) follows immediately from the fact that $(p + \beta\mathbb{N}, \sigma)$ is a minimal system. Indeed, note that the assumption $A \in p$ just means that $p \in \overline{A}$, i.e. $\overline{A}$ is a (clopen) neighborhood of p. Now, by Exercise 5, in a minimal dynamical system every point x is *uniformly recurrent*, i.e. visits any of its neighborhoods V along a syndetic set. This implies that the set $\{n : \ p + n \in \overline{A}\} = \{n : \ A \in p + n\} = \{n : \ A - n \in p\}$ is syndetic.

(ii) Since the set $B = \{n : \ A - n \in p\}$ is syndetic, the union of finitely many shifts of B covers $\mathbb{N}$, i.e. for some finite set $F \subset \mathbb{N}$ one has $\bigcup_{t\in F}(B-t) = \mathbb{N}$. So, for any $n \in \mathbb{N}$ there exists $t \in F$ such that $n \in B - t$, or $n + t \in B$. By the definition of B this implies $(A - (n+t)) \in p$. It follows that for any n the set $\bigcup_{t\in F}(A-t) - n$ belongs to p, and consequently, the family $\{\bigcup_{t\in F}(A-t) - n : \ n \in \mathbb{N}\}$ has the finite intersection property. By Exercise 4 this is equivalent to piecewise syndeticity of A.

(iii) Observe that $n \in A - A$ if and only if $A \cap (A-n) \neq \emptyset$. But then it follows from (i) that the set $\{n : A \cap (A-n) \in p\}$ is syndetic. We are done. □

Exercise 6. Let $(\Gamma, \cdot)$ be a compact left-topological semigroup and let $p \in \Gamma$. Verify that $p \cdot \Gamma$ is a compact right ideal and that any minimal right ideal is representable in this form.

For minimal idempotents in $(\beta G, \cdot)$ (i.e. idempotents belonging to minimal right ideals) one has an analogue of Theorem 2.4. In order to properly formulate it, one has first to define in a general context the notions of syndeticity and piecewise syndeticity since in a non-commutative situation one has left and right versions and needs to make "right" choices (pun not intended!).

Since we work with the left-topological semigroup $(\beta G, \cdot)$ the correct notions, which allow one to painlessly transfer the theorems above to the general set-up, turn out to be the left ones. Recall that given a set A and an element x in a semigroup $(G, \cdot)$ one has, by definition, $Ax^{-1} = \{y \in G : \ yx \in A\}$.

Definition 2.5. *Let G be a semigroup. A set $A \subseteq G$ is called syndetic if for some finite set $F \subset G$ one has $\bigcup_{t\in F} At^{-1} = G$. A set $A \subseteq G$ is piecewise syndetic if for some finite set $F \subset G$ the family $\{(\bigcup_{t\in F} At^{-1})a^{-1} : \ a \in G\}$ has the finite intersection property.*

The following exercise establishes a general form of Theorem 2.4.

Exercise 7. Let G be a discrete semigroup and $p \in (\beta G, \cdot)$ a minimal idempotent. Prove:

(i) For any $A \in p$ the set $B = \{g : Ag^{-1} \in p\}$ is syndetic.

(ii) Any $A \in p$ is piecewise syndetic.

(iii) For any $A \in p$, the set $A^{-1}A = \{x \in G : yx \in A \text{ for some } y \in A\}$ is syndetic. (Note that if G is a group, then $A^{-1}A = \{g_1^{-1}g_2 : g_1, g_2 \in A\}$.

The notion of minimality for idempotents can also be expressed in terms of a natural partial order which we will presently introduce.

Definition 2.6. *Let p, q be idempotents in a semigroup $(G, \cdot)$. We shall say that p is dominated by q and write $p \leq q$ if $pq = qp = p$.*

Exercise 8. Check that $\leq$ is transitive, reflexive and antisymmetric relation on the set of idempotents in G.

Theorem 2.7. *Let $(G, \cdot)$ be a compact left topological semigroup. Then an idempotent p is minimal with respect to the order $\leq$ if and only if it belongs to a minimal right ideal.*

Proof. (i) Assume that $p \leq q$ is minimal with respect to the order $\leq$. By Exercise 6, pG is a right ideal which contains a closed minimal right ideal R. Let $q_0 \in R$ be an idempotent. Since $q_0 \in pG$, $q_0 = pg$ for some $g \in G$. We have $pq_0 = ppg = pg = q_0$. Therefore $q_0pq_0p = q_0q_0p = q_0p$ and q_0p is also an idempotent. It satisfies $pq_0p = q_0pp = q_0p$, and hence $q_0p \leq p$. But p was assumed to be minimal with respect to the order $\leq$, and so $q_0p = p$. This gives $pG = q_0pG \subseteq q_0G$ and shows that pG is itself a minimal right ideal.

(ii) Assume that p is an idempotent in a minimal right ideal R. Note that pG is a right ideal and since $pG \subset R$ we have $pG = R$. Assume now that for some idempotent q one has $q \leq p$ and show that $p \leq q$. Since $q \leq p$ then we have also $qG \subset R$. It follows that for some $g \in G$, $p = qg$. Hence $qp = qqg = qg = p$. But $qp = q$ (because $q \leq p$). So $p \leq q$ and we are done. □

Remark. It follows from the proof of part (i) above that for any idempotent q, there exists a minimal idempotent p with $p \leq q$.

Before moving to some applications of minimal idempotents in Ramsey theory, we want to introduce one more useful algebraic-topological notion.

Definition 2.8. *Let $(G, \cdot)$ be a compact left-topological semigroup. Then $K(G) = \bigcup\{R : R \text{ is a minimal right ideal }\}$.*

Note that in view Exercise 6, $K(G) \neq \emptyset$.

Theorem 2.9. *Let $(G, \cdot)$ be a compact left topological semigroup. Then $K(G)$ is a two-sided ideal, and, in fact, the smallest two-sided ideal.*

Proof. Being the union of right ideals, $K(G)$ is trivially a right ideal. We note also that if I is a two-sided ideal of G then $K(G) \subseteq I$. Indeed, for any minimal right ideal R of G one has $R \cap I \neq \emptyset$ (since for any $x \in R$ and $y \in I$ one has $xy \in R \cap I$) and hence $R \cap I$ is a right ideal which, in view of minimality of R implies $R \cap I = R$. Hence, $K(G) \subseteq I$.

It remains to show that $K(G)$ is a left ideal of G, i.e. $G \cdot K(G) \subseteq K(G)$. Let $x \in K(G)$ and let R be a minimal right ideal of G such that $x \in R$. For arbitrary $y \in G$ one has: $yx \in yR$ and, since yR is a right ideal, it remains to show that it is a minimal right ideal. Indeed, then one has $yx \in yR \subseteq K(G)$ which clearly implies $G \cdot K(G) \subseteq K(G)$.

So let J be a right ideal of G satisfying $J \subseteq yR$ and let $C = \{z \in R : yz \in J\}$. Then C is a right ideal of G which is contained in R and so $C = R$ and $J = yR$. □

We are going to present now a proof, via the minimal idempotents, of the celebrated van der Waerden theorem on arithmetic progressions, which states that for any finite partition $\mathbb{N} = \bigcup_{i=1}^{r} C_i$, one C_i contains arbitrarily long arithmetic progressions. (This proof is a slight modification of the proof from [BFHK], which, in its turn, was a modification of an argument due to Furstenberg and Katznelson that first appeared, in the framework of an Ellis enveloping semigroup, in [FK2]).

Since for any minimal idempotent $p \in (\beta\mathbb{N}, +)$ and any partition $\mathbb{N} = \bigcup_{i=1}^{r} C_i$, one C_i belongs to p, van der Warden's theorem clearly follows from the following result.

Theorem 2.10. *Let $p \in (\beta\mathbb{N}, +)$ be a minimal idempotent and let $A \in p$. Then A contains arbitrarily long arithmetic progressions.*

Proof. Fix $k \in \mathbb{N}$ and let $G = (\beta\mathbb{N})^k$. Clearly, G is a compact left topological semigroup with respect to the product topology and coordinatewise addition. Let

$$E_0 = \{(a, a+d, ..., a+(k-1)d) : \ a \in \mathbb{N}, \ d \in \mathbb{N} \cup \{0\}\},$$

$$I_0 = \{(a, a+d, ..., a+(k-1)d) : \ a, d \in \mathbb{N}\}.$$

Clearly, E_0 is a semigroup in $\mathbb{N}^k$ and I_0 is an ideal of E_0. Let $E = cl_G E_0$ and $I = cl_G I_0$ be, respectively, the closures of E_0 and I_0 in G. It follows by an easy argument, which we leave to the reader, that E is a compact subsemigroup of G and I is a two-sided ideal of E. Let now $p \in K(\beta\mathbb{N}, +)$ be a minimal idempotent and let $\tilde{p} = (p, p, ..., p) \in G$. We claim that $\tilde{p} \in I$ and that this implies that each member of p contains a length k arithmetic progression. Indeed, assume that $\tilde{p} \in I$ and let $A \in p$. Then $\overline{A} \times ... \times \overline{A} = (\overline{A})^k$ is a neighborhood of $\tilde{p}$. Hence $\tilde{p} \in (\overline{A})^k \cap cl_G I_0 = cl_G(A^k \cap I_0)$, which implies

$A^k \cap I_0 \neq \emptyset$. It follows that for some $a, d \in \mathbb{N}$ $(a, a+d, ..., a+(k-1)d) \in A^k$ which finally implies $\{a, a+d, ..., a+(k-1)d\} \subset A$.

So it remains to show that $\tilde{p} \in I$. We check first that $\tilde{p} \in E$. Let $A_1, A_2, ..., A_k \in p$. Then $\overline{A}_1 \times \overline{A}_2 \times ... \times \overline{A}_k \ni \tilde{p}$. If $a \in \bigcap_{i=1}^k A_i$ then $(a, a, ..., a) \in (\overline{A}_1 \times \overline{A}_2 \times ... \times \overline{A}_k) \cap E_0$ which implies $\tilde{p} \in E$.

Now, since $p \in K((\beta\mathbb{N}, +))$, there is a minimal right ideal R of $(\beta\mathbb{N}, +)$ such that $p \in R$. Since $\tilde{p} \in E$, $\tilde{p} + E$ is a right ideal of E and there is a minimal right ideal $\tilde{R}$ of E such that $\tilde{R} \subseteq \tilde{p} + E$. Let $\tilde{q} = (q_1, q_2, ..., q_k)$ be an idempotent in $\tilde{R}$. Then $\tilde{q} \in \tilde{p} + E$ and for some $\tilde{s} = (s_1, s_2, ..., s_k)$ in E we get $\tilde{q} = \tilde{p} + \tilde{s}$. We shall show now that $\tilde{p} = \tilde{q} + \tilde{p}$. Indeed, from $\tilde{q} = \tilde{p} + \tilde{s}$ we get, for each $i = 1, 2, ..., k$, $q_i = p + s_i$. This implies $q_i \in R$ and since R is minimal, $q_i + \beta\mathbb{N} = R$. Hence $p \in q_i + \beta\mathbb{N}$. Let , for each $i = 1, 2, ..., k$, $t_i \in \beta\mathbb{N}$ be such that $p = q_i + t_i$. Then $q_i + p = q_i + q_i + t_i = q_i + t_i = p$ and so we obtained $\tilde{p} = \tilde{q} + \tilde{p}$.

To finish the proof, we observe that $\tilde{p} = \tilde{q} + \tilde{p}$ implies $\tilde{p} \in \tilde{q} + E = \tilde{R}$ which, in its turn, implies $\tilde{p} \in K(E) \subseteq I$ (since $K(E)$ is the smallest ideal in E). We are done. □

Exercise 9. Show that there is an idempotent p in $(\beta\mathbb{N}, +)$ with the property that not every member of p contains a 3-term arithmetic progression.
Hint. Consider $FS(10^n)_{n=1}^{\infty}$ and utilize Theorem 1.5.

Definition 2.11. *Let $(G, \cdot)$ be a discrete semigroup. A set $A \subseteq G$ is called* central *if there exists a minimal idempotent $p \in (\beta G, \cdot)$ such that $A \in p$.*

Exercise 10. Prove that any multiplicatively central set in $\mathbb{N}$ (namely, a member of a minimal idempotent in $(\beta\mathbb{N}, \cdot)$) contains arbitrarily long geometric progressions.

Exercise 11. (i) Let S be a central set in $(\mathbb{N}, +)$ and let $d \in \mathbb{N}$. Let $S/d = \{n : \ nd \in S\}$ and $dS = \{n : \ n/d \in S\}$. Prove that S/d and dS are central in $(\mathbb{N}, +)$.

(ii) Use (i) (for $n = 2$) to show that each of the following sets is additively central:

$$C_1 = \{3^k(3m+1) : \ k, m \in \mathbb{N} \cup \{0\}\},$$

$$C_2 = \{3^k(3m+2) : \ k, m \in \mathbb{N} \cup \{0\}\}.$$

As theorems above indicate, central sets are an ideal object for Ramsey-theoretical applications. For example, central sets in $(\mathbb{N}, +)$ not only are large (i.e. piecewise syndetic) but also are combinatorially rich and, in particular, contain IP sets and arbitrarily long arithmetic progressions. Similarly, the multiplicative central sets in $(\mathbb{N}, \cdot)$ (namely, the members of minimal idempotents in $(\beta\mathbb{N}, \cdot)$) are multiplicatively piecewise syndetic, contain finite products sets (i.e. the multiplicative IP sets), arbitrarily long geometric progressions etc.

The following theorem obtained in collaboration with N.Hindman may be viewed as enhancement of Theorem 1.8 above.

Theorem 2.12. ([BH1], p.312) *For any finite partition* $\mathbb{N} = \bigcup_{i=1}^{r} C_i$ *one of* C_i *is both additively and multiplicatively central.*

Sketch of the proof. Let $M = cl\{p : p$ is a minimal idempotent in $(\beta\mathbb{N}, +)\}$. Then one can show that M is a right ideal in $(\beta\mathbb{N}, \cdot)$ (see [BH1], Theorem 5.4, p.311). Let $R \subseteq M$ be a minimal right ideal and pick an idempotent $q = q \cdot q$ in R. Let $i \in \{1, 2, ..., r\}$ be such that $C_i \in q$. Since q is a minimal idempotent in $(\beta\mathbb{N}, \cdot)$, C_i is central in $(\mathbb{N}, \cdot)$. Since $C_i \in q$ and $q \in M$, there is some minimal idempotent p in $(\beta\mathbb{N}, +)$ with $C_i \in p$. Hence C_i is also central in $(\mathbb{N}, +)$. □

The following theorem supplies a useful family of examples of additively and multiplicatively central sets in $\mathbb{N}$.

Theorem 2.13. (cf. [BH3], Lemma 3.3) *For any sequence* $(a_n)_{n=1}^{\infty}$ *and an increasing sequence* $(b_n)_{n=1}^{\infty}$ *in* $\mathbb{N}$, $\bigcup_{n=1}^{\infty}\{a_n, a_n+1, a_n+2, ..., a_n+b_n\}$ *is additively central and* $\bigcup_{n=1}^{\infty}\{a_n \cdot 1, a_n \cdot 2, ..., a_n \cdot b_n\}$ *is multiplicatively central.*

As we shall see in the subsequent sections, central sets play also an important role in the study of recurrence in topological dynamics and ergodic theory. The original definition of central sets in $(\mathbb{N}, +)$, due to H. Furstenberg, was made in the language of topological dynamics. Before introducing Furstenberg's definition of centrality, we want first to recall some relevant dynamical notions.

Given a compact metric space (X, d), a continuous map $T : X \to X$ and not necessarily distinct points $x_1, x_2 \in X$, one says that x_1, x_2 are *proximal*, if for some sequence $n_k \to \infty$ one has $d(T^{n_k}x_1, T^{n_k}x_2) \to 0$.

A point which is proximal only to itself is called distal. In case all the points of X are distal T is called a distal transformation and (X, T) is called a distal system.

Exercise 12. (i) Show that any isometry is a distal transformation.

(ii) Let α be an irrational number. Show that the *skew product* map defined on 2-torus $\mathbb{T}^2$ by $T(x, y) = (x + \alpha, y + x)$ is distal but cannot be made an isometry in any equivalent metric on $\mathbb{T}^2$.

Remark. The skew product map featured in the Exercise 12 (ii) is an example of an *isometric extension.* A deep structure theorem of distal systems proved by Furstenberg in [F2] states that, for a sufficiently broadly interpreted notion of an isometric extension, any minimal distal system is a (potentially transfinite) tower of successive isometric extensions. We shall see in the next section that the notion of distality is intrinsically linked with the idempotent ultrafilters.

Recall that a point x in a dynamical system (X, T) is called *uniformly recurrent* if for any neighborhood V of x the set $\{n : T^n x \in V\}$ is syndetic.

Exercise 13. Prove that if (X, T) is minimal (i.e. every point has dense orbit) then each point $x \in X$ is uniformly recurrent.

One can show (see for example [A1], [E2], [F4, p.160]) that in a dynamical system on a compact metric space every point is proximal to a uniformly recurrent point. In particular, this implies that any distal point is uniformly recurrent. (See Theorem 3.9 below for a proof of an enhanced version of this fact.)

We are now ready to formulate Furstenberg's original definition of central sets in $(\mathbb{N}, +)$. For the proof of the equivalence of this definition to Definition 2.11 above, see Theorem 3.6

Definition 2.14. (see [F4], p.161) *A subset $S \subseteq \mathbb{N}$ is a central set if there exists a system (X, T), a point $x \in X$, a uniformly recurrent point y proximal to x, and a neighborhood U_y of y such that $S = \{n : T^n x \in U_y\}$.*

We shall conclude this section by introducing and discussing various notions of largeness for subsets of $\mathbb{N}$, which will be utilized in the subsequent sections.

Following the terminology introduced in [F4], given a family $\mathcal{M}$ of non-empty sets in $\mathbb{N}$, let us call a set $E \subseteq \mathbb{N}$ an $\mathcal{M}^*$ set if for any $M \in \mathcal{M}$ one has $E \cap M \neq \emptyset$.

For example, if $\mathcal{M}$ consists of sets containing arbitrarily long blocks of consecutive integers (these sets are called *replete* in [GH] and *thick* in [F4]), the family $\mathcal{M}^*$ consists of syndetic sets. If $\mathcal{M}$ is the collection of all IP sets in $(\mathbb{N}, +)$, then the elements of $\mathcal{M}^*$ will be called IP* sets. Similarly, central* sets, or, simply C* sets are defined as the sets which have nontrivial intersection with any central set.

Theorem 2.15. (i) *A set $E \subseteq \mathbb{N}$ is an IP* set if and only if E is a member of any idempotent $p \in (\beta\mathbb{N}, +)$.*

(ii) *A set E is C* set if and only if it is a member of any minimal idempotent $p \in (\beta\mathbb{N}, +)$.*

Proof. (i) Let p be an idempotent in $(\beta\mathbb{N}, +)$. If $E \notin p$, then $E^c \in p$ and, by Hindman's theorem (Theorem 1.2 above), there exists an IP set in E^c which fails to have a non-trivial intersection with E, which contradicts the assumption that E is an IP* set.

(ii) The proof is similar to that of (i). Assume that E is C* set. If p is minimal idempotent and $E \notin p$, then $E^c \in p$ and hence E^c is a central set which has an empty intersection with E. This contradicts the assumption that E is C*. □

Corollary 2.16. *C* sets and IP* sets have finite intersection property.*

Remark. One can easily check that the definitions of IP* and C* sets, as well as Theorem 2.15 and Corollary 2.16, extend naturally to general semigroups. We record here the following trivial extension of Theorem 2.15 (ii), that will be used in the course of the proof of Theorem 4.4 below:

A set E in a countable semigroup G is C if and only if it is a member of every minimal idempotent $p \in \beta G$.*

It immediately follows from Theorem 2.15 that every IP* set is C*. On the other hand, since, by Theorem 2.10, every central set contains arbitrarily long arithmetic progressions, the complement of $FS(10^n)_{n=1}^{\infty}$ is, in view of Exercise 9, a C* set which is not IP*. We shall indicate now still another possibility of making the distinction between C* and IP* sets.

Definition 2.17. *For a set $E \subseteq \mathbb{N}$ the* upper Banach *density, $d^*(E)$, is defined by the formula*

$$d^*(E) = \limsup_{N-M\to\infty} \frac{|E \cap [M, N]|}{N - M + 1}$$

Proposition 2.18. *If $E \subseteq \mathbb{N}$ is a C* set and $P \subseteq \mathbb{N}$ satisfies $d^*(P) = 0$, then $E \setminus P$ is a C* set.*

Proof. It is obvious that a set having zero upper Banach density cannot be piecewise syndetic. It follows from Theorem 2.4, (ii) that $d^*(P) = 0$ implies that $\mathbb{N} \setminus P$ is a member of any minimal idempotent and hence is a C* set. The claim of the lemma follows now from Corollary 2.16. □

Remark. Proposition 2.18 implies that if E is a C* set and P is an IP set satisfying $d^*(P) = 0$, then $E \setminus P$ is a C* set which is not IP*.

Exercise 14. (i) Show that $2\mathbb{N} + 1$ is a syndetic set which is not C*.
(ii) Show that every C* set is syndetic.

It follows from the above exercise that, unlike the syndetic sets, IP* and C* sets are not stable under a shift: $2\mathbb{N}$ is a C* set but $2\mathbb{N} + 1$ is not. This motivates the following definition.

Definition 2.19. *A set E is called C_+^* if for some $k \in \mathbb{Z}$, $E - k$ is a C* set. Similarly, if for some $k \in \mathbb{Z}$, $E - k$ is IP* set then E is called IP_+^* set.*

It is an immediate observation that a set E is IP_+^* (C_+^*) if and only if for some $k \in \mathbb{Z}$, $E - k$ is a member of any idempotent (any minimal idempotent) in $(\beta\mathbb{N}, +)$. It follows that any IP_+^* set is C_+^* and that any C_+^* set is syndetic. The following result shows that these inclusions are strict.

Theorem 2.20. (i) *Not every syndetic set is* C_+^**;* (ii) *not every* C_+^* *set is* IP_+^*.

Proof. (i) Let us call any set of the form $S = \bigcup_{i=1}^\infty [a_i, b_i]$, where $b_i - a_i \to \infty$, a T-set (T stands for Thick). By Theorem 2.13 and Exercise 11(i), any T-set S is central as well as is the set $2S$. Note now that one can easily construct T-sets S_i, $i \geq 0$, so that the sets $2S_0$, $2S_1 - 1$, $2S_2 + 1$, $2S_3 - 2$, $2S_4 + 2$, ..., $2S_{2n-1} - n$, $2S_{2n} + n$, ... are all disjoint. Let U be the union of these sets. Then $V = \mathbb{N} \setminus U$ is certainly syndetic. At the same time, for any $k \in \mathbb{Z}$, V misses a k-shift of a central set and hence is not C_+^*.

(ii) The proof is similar to that of (i). It is not hard to construct "thin" IP sets $(S_n)_{n=0}^\infty$ such that $S_0, S_1 - 1, S_1 + 1, S_2 - 2, S_2 + 2, ..., S_n - n, S_n + n, ...$ are all disjoint and their union U has zero upper Banach density. Then, by Proposition 2.18, $V = \mathbb{N} \setminus U$ is C* but not IP*, since for every $k \in \mathbb{Z}$, V misses a k-shift of an IP set. □

3. Convergence along ultrafilters, topological dynamics, and some diophantine applications

In this section we shall introduce and exploit the notion of convergence along ultrafilters. As we shall see, this notion, especially the convergence along minimal idempotents, allows one to better understand distality, proximality and reccurence in topological dynamical systems. We want to point out that many proofs to be given below are similar to known proofs which utilize the so called Ellis enveloping semigroup (see [E3], [A2]). This is not too surprising since the Ellis semigroup is a particular type of compactification and is in many respects similar to the universal object, the Stone-Čech compactification. On the other hand, the usage of ultrafilters, especially the idempotent ones, has at least two advantages. First, one can, on many occasions, replace in the proofs the convergence along idempotent ultrafilters by the notion of IP convergence which is based on Hindman's theorem, and hereby eliminate the usage of nonmetrisable objects in metrisable dynamics (there are people who care about such things...). See [B2, p.34] for more discussion. Second, the usage of ultrafilters allows one to much more easily deal with combinatorial applications of topological dynamics. For example, we shall show in this section that the two different notions of central discussed in Section 2 are equivalent.

To keep the presentation more accessible and in order to be able to better emphasize the main ideas we shall be dealing in this section mostly with the topological systems of the form (X, T) where X is a compact metric space and T is a (not necessarily invertible) continuous selfmap of X. All the definitions and results below are more or less routinely transferable to actions of general (countably) infinite semigroups of continuous mappings of compact metric spaces. Also, many results in this section can be extended to the case where

X is not metrisable compact Hausdorff space. We leave all these extensions as an exercise to the reader.

Throughout this and the subsequent sections $\beta\mathbb{N}$ stands for $(\beta\mathbb{N}, +)$.

Given an ultrafilter $p \in \beta\mathbb{N}$ one writes $p\text{-}\lim_{n\in\mathbb{N}} x_n = y$ if for every neighbourhood U of y one has $\{n : x_n \in U\} \in p$.

Exercise 15. (i) Check that $p\text{-}\lim_{n\in\mathbb{N}} x_n$ exists and is unique in any compact Hausdorff space.

(ii) Fix a sequence $(x_n)_{n\in\mathbb{N}}$ in a compact Hausdorff space X. Prove that the map $F : \beta\mathbb{N} \to X$, defined by $F(p) = p\text{-}\lim_{n\in\mathbb{N}} x_n$ is continuous.

(iii) Prove that if x_1, x_2 are proximal points in a topological system (X, T), then there exists $p \in \beta\mathbb{N}$ such that $p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = p\text{-}\lim_{n\in\mathbb{N}} T^n x_2$.

Theorem 3.1. *Let X be a compact Hausdorff space and let $p, q \in \beta\mathbb{N}$. Then for any sequence $(x_n)_{n\in\mathbb{N}}$ in X one has*

$$(q+p)\text{-}\lim_{r\in\mathbb{N}} x_r = p\text{-}\lim_{t\in\mathbb{N}} q\text{-}\lim_{s\in\mathbb{N}} x_{s+t}. \tag{3.1}$$

In particular, if p is an idempotent, and $q = p$, one has

$$p\text{-}\lim_{r\in\mathbb{N}} x_r = p\text{-}\lim_{t\in\mathbb{N}} p\text{-}\lim_{s\in\mathbb{N}} x_{s+t}.$$

Proof. Note that by Exercise 15 both sides of equation (3.1) are well defined. Let $x = (q+p)\text{-}\lim_{r\in\mathbb{N}} x_r$. Given a neighborhood U of x we have $\{r : x_r \in U\} \in q+p$. Recalling that a set $A \subseteq \mathbb{N}$ is a member of ultrafilter $q+p$ if and only if $\{n \in \mathbb{N} : (A-n) \in q\} \in p$, we get

$$\{t : (\{s : x_s \in U\} - t) \in q\} = \{t : \{s : x_{s+t} \in U\} \in q\} \in p$$

This means that, for p-many t, $q\text{-}\lim_{s\in\mathbb{N}} x_{s+t} \in U$ and we are done. $\square$

Exercise 16. Let R be a minimal right ideal in $\beta\mathbb{N}$. Recall that (R, σ), where $\sigma : p \mapsto p+1$, is a minimal (non-metrizable) system (see Theorem 2.3 above). Given a topological system (X, T) and a point $x \in X$ consider the mapping $\varphi : R \to X$, defined by $\varphi(p) = p\text{-}\lim_{n\in\mathbb{N}} T^n x$. Denote by Y the set $\{p\text{-}\lim_{n\in\mathbb{N}} T^n x : p \in R\}$. Prove that (Y, T) is a minimal system by checking that the following diagram is commutative:

$$\begin{array}{ccc} R & \xrightarrow{\sigma} & R \\ \varphi\big\downarrow & & \big\downarrow\varphi \\ Y & \xrightarrow{T} & Y \end{array}$$

Proposition 3.2. *Let (X, T) be a topological system and let $x \in X$ be an arbitrary point. Given an idempotent ultrafilter $p \in \beta\mathbb{N}$, let $p\text{-}\lim_{n\in\mathbb{N}} T^n x = y$. Then $p\text{-}\lim_{n\in\mathbb{N}} T^n y = y$. If x is a distal point (i.e. x is proximal only to itself) then $p\text{-}\lim_{n\in\mathbb{N}} T^n x = x$.*

Proof. Applying Theorem 3.1 (and the fact that $p+p=p$), we have

$$p\text{-}\lim_{n\in\mathbb{N}} T^n y = p\text{-}\lim_{n\in\mathbb{N}} T^n \ p\text{-}\lim_{m\in\mathbb{N}} T^m x$$

$$= p\text{-}\lim_{n\in\mathbb{N}} \ p\text{-}\lim_{m\in\mathbb{N}} T^{m+n} x = p\text{-}\lim_{n\in\mathbb{N}} T^n x = y.$$

If x is a distal point, then the relations $p\text{-}\lim_{n\in\mathbb{N}} T^n x = y = p\text{-}\lim_{n\in\mathbb{N}} T^n y$ clearly imply $x=y$ and we are done. □

Exercise 17. Show that if T is a continuous distal selfmap of a compact metric space then T is invertible and T^{-1} is also distal. *Hint.* It follows from Proposition 3.2 that T is *onto.*

Proposition 3.3. *If (X,T) is a minimal system then for any $x\in X$ and any minimal right ideal R in $\beta\mathbb{N}$ there exists a minimal idempotent $p\in R$ such that $p\text{-}\lim T^n x = x$.*

Proof. By Exercise 16, $X=\{p\text{-}\lim_{n\in\mathbb{N}} T^n x,\ p\in R\}$. It follows that the set $\Gamma=\{p\in R : p\text{-}\lim_{n\in R} T^n x = x\}$ is non-empty and closed. We claim that Γ is a semigroup. Indeed, if $p,q\in\Gamma$, one has :

$$(p+q)\text{-}\lim_{n\in\mathbb{N}} T^n x = q\text{-}\lim_{n\in\mathbb{N}} T^n p\text{-}\lim_{m\in\mathbb{N}} T^m x = x.$$

By Ellis theorem Γ contains an idempotent which has to be minimal since it belongs to R. We are done. □

Exircise 18. Let (X,T) be a topological system, R a minimal right ideal in $\beta\mathbb{N}$, and let $x\in X$ be a point in X. Prove that the following are equivalent:

(i) x is uniformly reccurent;

(ii) there exists a minimal idempotent $p\in R$ such that $p\text{-}\lim_{n\in\mathbb{N}} T^n x = x$.

It follows from Proposition 3.2 that for any topological system (X,T), any $x\in X$, and any idempotent ultrafilter p, the points x and $y=p\text{-}\lim_{n\in\mathbb{N}} T^n x$ are proximal. (If (X,T) is a distal system then $y=x$). The following theorem gives a partial converse of Proposition 3.3.

Theorem 3.4. *If (X,T) is a topological system and x_1,x_2 are proximal, not necessarily distinct points and if x_2 is uniformly reccurent, then there exists a minimal idempotent $p\in\beta\mathbb{N}$ such that $p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = x_2$.*

Proof. Let $I=\{p\in\beta\mathbb{N} : p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = p\text{-}\lim_{n\in\mathbb{N}} T^n x_2\}$. By Exercise 15 (iii), I is a non-empty closed subset of $\beta\mathbb{N}$. One immediately checks that I is a right ideal. Let R be a minimal right ideal in I. Since x_2 is uniformly recurrent, its orbital closure is a minimal system. By Proposition 3.3 there exists a minimal idempotent $p\in R$ such that $p\text{-}\lim T^n x_2 = x_2$. Then $p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = p\text{-}\lim_{n\in\mathbb{N}} T^n x_2 = x_2$ and we are done. □

One can give a similar proof to the following classical result due to J. Auslander [A1] and Ellis [E2].

Theorem 3.5. *Let (X,T) be a topological system. For any $x \in X$ there exists a uniformly recurrent point y in the orbital closure $\overline{\{T^n x\}}_{n\in N}$, such that x is proximal to y. Moreover, for any minimal right ideal $R \subset \beta\mathbb{N}$ there exists a minimal idempotent $p \in R$ such that $p\text{-}\lim_{n\in\mathbb{N}} T^n x = y$.*

Proof. Let R be a minimal ideal in $\beta\mathbb{N}$ and let p be a (minimal) idempotent in R. Let $y = p\text{-}\lim_{n\in\mathbb{N}} T^n x$. Clearly, y belongs to the orbital closure of x. By Proposition 3.2, x and y are proximal. By Exercise 18, y is uniformly recurrent. We are done. □

We are in position now to establish the equivalence of two notions of central that were discussed in Section 2.

Theorem 3.6. *The following properties of a set $A \subseteq \mathbb{N}$ are equivalent:*

(i) (cf. [F4], Definition 8.3) *There exists a topological system (X,T), and a pair of (not necesserily distinct) points $x, y \in X$ where y is uniformly recurrent and proximal to x, such that for some neighborhood U of y one has:*

$$A = \{n \in \mathbb{N} : T^n x \in U\}$$

(ii) ([BH1], Definition 3.1) *There exists a minimal idempotent $p \in \beta\mathbb{N}$ such that $A \in p$.*

Proof. (i) ⇒ (ii) By Theorem 3.5 there exists a minimal idempotent p, such that $p\text{-}\lim_{n\in\mathbb{N}} T^n x = y$. This implies that for any neighborhood U of y the set $\{n \in \mathbb{N} : T^n x \in U\}$ belongs to p.

(ii) ⇒(i) The idea of the following proof is due to B. Weiss. Let A be a member of a minimal idempotent $p \in \beta\mathbb{N}$. Let $X = \{0,1\}$, the space of bilateral 0-1 sequences. Endow X with the standard metric:

$$d(\omega_1, \omega_2) = \min\{\frac{1}{n+1} : \omega_1(i) = \omega_2(i) \text{ for } |i| < n\}$$

It is easy to check that (X,d) is a compact metric space. Let $T : X \to X$ be the shift operator: $T(\omega)(n) = \omega(n+1)$. Then T is a homeomorphism of X and (X,T) is a topological dynamical system. Viewing A as a subset of $\mathbb{Z}$, let $x = 1_A \in X$. Finally, let $y = p\text{-}\lim_{n\in\mathbb{N}} T^n x$. By Proposition 3.2, x and y are proximal. Also, since p is minimal, y is, by Exercise 18, a uniformly recurrent point. We claim that $y(0) = 1$. Indeed, define $U = \{z \in X : z(0) = y(0)\}$, and note that, since $y = p\text{-}\lim_{n\in\mathbb{N}} T^n x$ and $A \in p$, one can find $n \in A$ such that $T^n x \in U$. But since $x = 1_A$, $(T^n x)(0) = 1$. But then, given $n \in \mathbb{Z}$, we have: $T^n x \in U \Leftrightarrow (T^n x)(0) = 1 \Leftrightarrow x(n) = 1 \Leftrightarrow n \in A$. It follows that $A = \{n \in \mathbb{Z} : T^n x \in U\}$ and we are done. □

Remark. One can show that the characterization of central sets given in Theorem 3.6 extends to general semigroups. See [BH1] and [SY].

Let (X, T) be a topological system. In [F4], a point $x \in X$ is called IP* recurrent (C* recurrent) if for any neighborhood U of x, $\{n \in \mathbb{N} : T^n x \in U\}$ is an IP* set (C* set). The following straightforward exercise establishes the connection between these notions of recurrence and convergence along idempotent ultrafilters.

Exercise 19. Let (X, T) be a topological system and let $x \in X$. Prove:

(i) x is IP* recurrent if and only if for any idempotent $p \in \beta\mathbb{N}$ one has $p\text{-}\lim_{n\in\mathbb{N}} T^n x = x$;

(ii) x is C* recurrent if and only if for any minimal idempotent $p \in \beta\mathbb{N}$ one has $p\text{-}\lim_{n\in\mathbb{N}} T^n x = x$.

We saw in Section 2 that the family of C* sets is wider them that of IP* sets (see the discussion after Corollary 2.16 and Remark after Proposition 2.18). It turns out, however, that, somewhat surprisingly, the notions of IP* recurrence and C* recurrence coincide.

Theorem 3.7. (cf. [F4], Proposition 9.17) *Let (X, T) be a dynamical system. A point $x \in X$ is IP* recurrent if and only if it is C* recurrent.*

Proof. We need to show only that C* recurrence implies IP* recurrence. Let $q \in \beta\mathbb{N}$ be an arbitrary idempotent and let us show that $q\text{-}\lim_{n\in\mathbb{N}} T^n x = x$. By the remark after Theorem 2.7, there exists a minimal idempotent p, such that $p \leq q$. Then $p + q = p$ and we have: $x = p\text{-}\lim_{n\in\mathbb{N}} T^n x = (p+q)\text{-}\lim_{n\in\mathbb{N}} T^n x = q\text{-}\lim_{n\in\mathbb{N}} T^n(p\text{-}\lim_{m\in\mathbb{N}} T^m x) = q\text{-}\lim_{n\in\mathbb{N}} T^n x$. We are done. □

We recall that a point is called distal if it is proximal only to itself.

Theorem 3.8. *Let (X, T) be a dynamical system and $x \in X$. The following are equivalent:*

(i) *x is a distal point;*

(ii) *x is IP* recurrent.*

Proof. (i) $\Rightarrow$ (ii). By Proposition 3.2, for any idempotent p, x and $p\text{-}\lim_{n\in\mathbb{N}} T^n x$ are proximal. Since x is distal, this may happen only if $x = p\text{-}\lim T^n x$. By Exercise 19 (i), this means that x is an IP* recurrent point.

(ii) $\Rightarrow$ (i). If x is not distal, then there exists $y \neq x$, such that x and y are proximal. But then, by Theorem 3.4, there exists an idemponent p such that $p\text{-}\lim T^n x = y$. Since $y \neq x$, this contradicts (ii). (We are using again the characterization of IP* recurrence given in Exercise 19). □

Remark. The property of a point x to be IP* recurrent is much stronger than that of uniform recurrence (which, by Exercise 18, is equivalent to the fact that for *some* minimal idempotent $p\text{-}\lim_{n\in\mathbb{N}} T^n x = x$). While, in a minimal system, every point is uniformly recurrent, there are minimal systems having no distal points. In particular, any minimal topologically weakly mixing system is such. (See [F4, Theorem 9.12]).

We shall conclude this section with some diophantine applications which may be viewed as enhancements of classical results due to Hardy–Littlewood and Weyl. But first we need the following variation on the theme of Theorem 3.8.

Theorem 3.9. *Assume that (X,T) is a minimal system. Then it is distal if and only if for any $x \in X$ and any open set $U \subseteq X$ the set $\{n : T^n x \in U\}$ is IP$^*_+$.*

Proof. Assume that (X,T) is distal. By minimality, there exists $n_0 \in \mathbb{N}$ such that $T^{n_0}x \in U$. By Theorem 3.8, the set $\{n : T^n(T^{n_0}x) \in U\}$ is IP* which, of course, implies that the set $\{n : T^n x \in U\}$ is IP$^*_+$.

Assume now that for any x_1, x_2 and a neighborhood U of x_2 the set $\{n : T^n x_1 \in U\}$ is IP$^*_+$. We will find it convenient to call an IP$^*_+$ set $A \subseteq \mathbb{N}$ *proper* if A is not IP* (i.e. A is a nontrivial shift of an IP* set and, moreover, this shifted IP* set is not IP*). If T were not distal, then for some distinct points x_1, x_2 and idempotents p, q one would have: $p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = x_2$, $q\text{-}\lim_{n\in\mathbb{N}} T^n x_2 = x_1$ and also $p\text{-}\lim_{n\in\mathbb{N}} T^n x_2 = x_2$, $q\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = x_1$ (see Theorem 3.4 and Proposition 3.2). Let U be a small enough neighborhood of x_2. Then, since $p\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = x_2$, the set $S = \{n : T^n x_1 \in U\}$ is a member of p, and hence cannot be a proper IP$^*_+$ set. But, since $q\text{-}\lim_{n\in\mathbb{N}} T^n x_1 = x_1$, the set S cannot be an improper IP$^*_+$ set (that is, an IP* set) either: if U is small enough, $S \notin q$. So T has to be distal. We are done. □

We record the following immediate corollary (of the proof) of Theorem 3.9.

Corollary 3.10. *If (X,T) is distal and minimal, and x_1, x_2 are distinct points in X, then if U is a small enough neighbourhood of x_2, the set $\{n : T^n x_1 \in U\}$ is a proper IP$^*_+$ set.*

We move now to diophantine applications. Our starting point is Kronecker's approximation theorem.

Theorem 3.11. *([Kro]) If the numbers $1, \alpha_1, \alpha_2, ..., \alpha_k$ are linearly independent over Q, then for any k subintervals $I_j = (a_j, b_j) \subset [0,1]$ there exists $n \in \mathbb{N}$ such that one has simultaneously $n\alpha_j \mod 1 \in I_j$, $j = 1, 2, ..., k$.*

In 1916 H. Weyl ([Weyl]) revolutionized the field of diophantine approximations by introducing the notion of uniform distribution in terms of exponential sums. One of his celebrated results dealing with the values of polynomials mod 1 will be discussed below. As for the Kronecker's theorem, Weyl's approach gives the fact that the set $\Gamma = \{n \in \mathbb{N} : n\alpha_j \mod 1 \in I_j, j = 1, ..., k\}$

has positive density (which is equal $\prod_{j=1}^{k}(b_j - a_j)$). By considering the uniform Cesàro averages one can actually show that the set Γ is syndetic. Since, as we saw above, the property of a set to be IP_+^* is considerably stronger than syndeticity, the following refinement of Kronecker's theorem is of interest, especially, since it does not seem to follow from considerations involving the Cesàro averages.

Theorem 3.12. *Under the assumptions and notation of Theorem 3.11 the set* $\Gamma = \{n \in \mathbb{N} : n\alpha_j \mod 1 \in I_j,\ j = 1, ..., k\}$ *is* IP_+^*.

Proof. Let $T_j : \mathbb{T} \to \mathbb{T}$ be defined by $x \mapsto x + \alpha_j \mod 1,\ j = 1, ..., k$. Noticing that the product transformation $T_1 \times T_2 \times \cdots \times T_k$, acting on k-dimensional torus $\mathbb{T}^k$, is distal (this is obvious) and minimal (this follows from Kronecker's theorem, but may be proved in a variety of ways), we see that our result immediately follows from Theorem 3.9. □

We are going to discuss now similar refinements of some other classical results. In what follows the crucial role will be played by minimal distal systems. It was H. Furstenberg who introduced and applied the idea of using the unique ergodicity of a class of affine transformations of the torus to obtain a dynamical proof of Weyl's theorem on equidistribution of polynomials (see [F1] and [F4, p.69]). As we shall see, affine transformations of the kind treated by Furstenberg can also be utilized to obtain polynomial results in the spirit of Theorem 3.12.

The following extension of Kronecker's theorem was obtained by Hardy and Littlewood in [HaL].

Theorem 3.13. *If the numbers* $1, \alpha_1, ..., \alpha_k$ *are linearly independent over* $\mathbb{Q}$, *then for any* $d \in \mathbb{N}$ *and any* kd *intervals* $I_{lj} \subset [0, 1],\ l = 1, ..., d; j = 1, ..., k$ *the set*

$$\Gamma_{dk} = \{n \in \mathbb{N} : n^l\alpha_j \mod 1 \in I_{lj},\ l = 1, ..., d;\ j = 1, ..., k\}$$

is infinite.

As with Kronecker's theorem, Weyl was able to show in [Weyl] that the set Γ_{dk} has positive density equal to the product of lengths of I_{lj}. In 1953 P. Szüsz [Sz] proved that the set Γ_{dk} is syndetic. The following theorem shows that Γ_{dk} is actually IP_+^*.

Theorem 3.14. *Under the assumptions and notation of Theorem 3.13, the set* Γ_{dk} *is* IP_+^*.

Proof. To make the formulas more transparent we shall put $d = 3$. It will be clear that the same proof gives the general case.

We start with easily checkable claim that if $T_\alpha : \mathbb{T}^3 \to \mathbb{T}^3$ is defined by $T_\alpha(x, y, z) = (x+\alpha, y+2x+\alpha, z+3x+3y+\alpha)$ then $T_\alpha^n(0, 0, 0) = (n\alpha, n^2\alpha, n^3\alpha)$. This transformation T is distal (easy) and minimal. The last assertion can actually be derived from the case $k = 1$ of Hardy - Littlewood theorem above, but also can be proved directly. (For example, this fact is a special case of Lemma 1.25, p.36 in [F4]). Our next claim is that if the numbers $1, \alpha_1, \alpha_2, ..., \alpha_k$ are linearly independent over Q, then the product map $T = T_{\alpha_1} \times \cdots \times T_{\alpha_k}$ (acting on $\mathbb{T}^{3k}$) is distal and minimal as well. (The distality is obvious, the minimality follows, again, from an appropriately modified Lemma 1.25 in [F4]). By minimality of T, the orbit of zero in $\mathbb{T}^{3k}$ is dense which, together with Theorem 3.9, gives the desired result. □

Remark. In case the intervals I_{lj} contain 0, it can be shown that Γ_{dk} is IP*. This special case is also treated in [FW2].

Exercise 20. (i) Derive from Theorem 3.14 the following fact: for any irrational numbers $\alpha_1, ..., \alpha_k$ and any subintervals $I_j \subset [0,1]$, $j = 1, ..., k$, the set

$$\{n \in \mathbb{N} : n^j \alpha_j \mod 1 \in I_j,\ j = 1, ..., k\}$$

is IP$^*_+$.

(ii) Use (i) to obtain the following fact: for any real polynomial $p(t)$ having at least one coefficient other than the constant term irrational and for any subinterval $I \subset [0,1]$ the set $\{n \in \mathbb{N} : p(n) \bmod 1 \in I\}$ is IP$^*_+$.

Remark. Another possibility of proving the statement (ii) is to apply Theorem 3.9 to the transformation $T : (x_1, x_2, ..., x_k) \mapsto (x_1 + \alpha, x_2 + x_1, ..., x_k + x_{k-1})$ which is used in [F4] to derive Weyl's theorem on uniform distribution.

We conclude this section by formulating a general result which may be proved by refining the techniques used above.

Theorem 3.15. *If real polynomials $p_1(t), p_2(t), ..., p_k(t)$ have the property that for any non-zero vector $(h_1, h_2, ..., h_k) \in \mathbb{Z}^k$ the linear combination $\sum_{i=1}^k h_i p_i(t)$ is a polynomial with at least one irrational coefficient other than the constant term then for any k subintervals $I_j \subset [0,1]$, $j = 1, ..., k$, the set*

$$\{n \in \mathbb{N} : p_j(n) \bmod 1 \in I_j,\ j = 1, ..., k\}$$

is IP$^*_+$.

4. Minimal idempotents and weak mixing

In this chapter we shall connect convergence along minimal idempotents with the theory of weakly mixing unitary actions and weakly mixing measure preserving systems. While the customary definitions of weakly mixing $\mathbb{Z}$-actions can be more or less routinely extended to actions of abelian or even amenable (semi)groups (cf. [D]), the study of the weakly mixing actions of non-amenable groups necessitates introduction of new tools and ideas.

Before moving to a more advanced discussion, we want to illustrate the multifariousness of the notion of weak mixing by listing some equivalent conditions for a system to be weakly mixing. (In most books either (i) or (ii) below is taken as "official" definition of weak mixing).

Theorem 4.1. *Let T be an invertible measure-preserving transformation of a probability measure space $(X, \mathcal{B}, \mu)$. Let U_T denote the operator defined on measurable functions by $(U_T f)(x) = f(Tx)$. The following conditions are equivalent:*

(i) *For any $A, B \in \mathcal{B}$*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu(A \cap T^{-n}B) - \mu(A)\mu(B)| = 0;$$

(ii) *For any $A, B \in \mathcal{B}$ there is a set $P \subset \mathbb{N}$ of density zero such that*

$$\lim_{n\to\infty,\, n\notin P} \mu(A \cap T^{-n}B) = \mu(A)\mu(B);$$

(iii) *$T \times T$ is ergodic on the Cartesian square of $(X, \mathcal{B}, \mu)$;*

(iv) *For any ergodic probability measure preserving system $(Y, \mathcal{D}, \nu, S)$ the transformation $T \times S$ is ergodic on $X \times Y$;*

(v) *If f is a measurable function such that for some $\lambda \in \mathbb{C}$, $U_T f = \lambda f$ a.e., then $f = const$ a.e.;*

(vi) *For $f \in L^2(X, \mathcal{B}, \mu)$ with $\int f = 0$ consider the representation of the positive definite sequence $\langle U_T^n f, f\rangle, n \in \mathbb{Z}$, as a Fourier transform of a measure ν on $\mathbb{T}$:*

$$\langle U_T^n f, f\rangle = \int_{\mathbb{T}} e^{2\pi i n x} d\nu, \quad n \in \mathbb{Z}$$

(this representation is guaranteed by Herglotz theorem, see [He]). Then ν has no atoms.

(vii) *If for some $f \in L^2(X, \mathcal{B}, \mu)$ the set $\{U_T^n f,\ n \in \mathbb{Z}\}$ is totally bounded, then f is a constant a.e.;*

(viii) *Weakly independent sets are dense in $\mathcal{B}$. (A set $A \in \mathcal{B}$ is called* weakly independent *if there exists a sequence $n_1 < n_2 < \cdots$ such that the sets $T^{-n_i}A$, $i = 1, 2, ...,$ are mutually independent);*

(ix) *For any $A \in \mathcal{B}$ and $k \in \mathbb{N}$, $k \geq 2$, one has*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n}A \cap T^{-2n}A \cap \cdots \cap T^{-kn}A) = (\mu(A))^{k+1}$$

(x) *For any $k \in \mathbb{N}$, $k \geq 2$, any $f_1, f_2, ..., f_k \in L^\infty(X, \mathcal{B}, \mu)$ and any non-constant polynomials $p_1(n), p_2(n), ..., p_k(n) \in \mathbb{Z}[n]$ such that for all $i \neq j$, $\deg(p_i - p_j) > 0$, one has*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x) \cdots f_k(T^{p_k(n)}x) = \int f_1 d\mu_1 \int f_2 d\mu_2 \cdots \int f_k d\mu_k$$

in L^2-norm.

Remark. Weakly mixing systems (for measure preserving $\mathbb{R}$-actions) were introduced under the name *dynamical systems of continuous spectra* in [KN]. See also [Hopf1] and [Hopf2]. These papers, as well as Hopf's book [Hopf3], already contain (versions of) the conditions (i) through (vii). Condition (viii) is due to Krengel (see [Kre] for this and other related results). The last two conditions, while being of interest in their own right, are strongly connected with combinatorial and number-theoretical applications. In particular, (ix) plays a crucial role in Furstenberg's ergodic proof of Szemerédi's theorem on arithmetic progressions (see [F3] and [F4]). Similarly, variations on condition (x) (see [B1]) are needed for proofs of polynomial extensions of Szemerédi's theorem (see, for example, [BL], [BM1], [BM2], [L]).

We are moving now to the discussion of weak mixing for general group actions. For sake of simplicity we shall be dealing with countably infinite (but not necessarily amenable) groups. We remark however that the results below can be extended to actions of general locally compact (semi)groups.

One possible approach to weak mixing for general groups is via the theory of invariant means. For example, if the acting group is amenable, one can replace the condition (i) in Theorem 4.1 by the assertion that the averages of the expressions $|\mu(A \cap T_g B) - \mu(A)\mu(B)|$ taken along any Følner sequence converge to zero. If the acting group G is noncommutative, one has, in addition, to replace conditions (v) and (vi) by the assertion that the only finite-dimensional subrepresentation of $(U_g)_{g\in G}$ (where U_g if defined by $(U_g f)(x) = f(T_g x)$) on $L^2(X, \mathcal{B}, \mu)$ is its restriction to the subspace of constant functions.

H. Dye has shown in [D] that under these modifications the conditions (i), (iii) and (v) in Theorem 4.1 are equivalent for measure-preserving actions of general amenable semigroups. By using the Ryll-Nardzewski theorem ([R-N]) which guarantees the existence of unique invariant mean on the space of weakly almost periodic functions on a group, one can extend the notion of weak mixing to actions of general locally compact groups. See [BR] for the details.

We are going now to indicate how to recover and refine some of the main results in [BR] by using minimal idempotents in βG.

We shall deal first with unitary representations of a group G on a Hilbert space $\mathcal{H}$ and will specialize to the case of measure preserving actions later. Since we work with countable groups only, we may and will always assume that the Hilbert spaces we work with are separable. To appreciate the terminology one should think of Hilbert space $\mathcal{H}$ as $L^2_0(X, \mathcal{B}, \mu) = \{f \in L^2(X, \mathcal{B}, \mu) : \int f d\mu = 0\}$.

Definition 4.2. *Let $(U_g)_{g\in G}$ be a unitary representation of a group G on a separable Hilbert space $\mathcal{H}$.*

(i) *A vector* $\varphi \in \mathcal{H}$ *is called* compact *if the set* $\{U_g\varphi : g \in G\}$ *is totally bounded in* $\mathcal{H}$.

(ii) *The representation* $(U_g)_{g\in G}$ *is called* weakly mixing *if there are no nonzero compact vectors.*

In the following theorem we are going to use expressions of the form $p\text{-}\lim_{g\in G} U_g\varphi$. Since the unit ball in $\mathcal{H}$ is a compact metrizable space with respect to the weak topology, these expressions have a well defined meaning. We remark also that if $p\text{-}\lim_{g\in G} U_g\varphi = \varphi$ weakly (which will be the case when φ is a compact vector and p a minimal idempotent), then one actually has $p\text{-}\lim_{g\in G} U_g\varphi = \varphi$ strongly. (The verification of this easy fact is left to the reader).

Theorem 4.3. *Let* $(U_g)_{g\in G}$ *be a unitary representation of a group* G *on a Hilbert space* $\mathcal{H}$. *For any* $\varphi \in \mathcal{H}$ *the following conditions are equivalent:*

(i) *There exists a minimal idempotent* $p \in \beta G$ *such that* $p\text{-}\lim_{g\in G} U_g\varphi = \varphi$;

(ii) φ *is a compact vector;*

(iii) *For any idempotent* $p \in \beta G$ *one has* $p\text{-}\lim_{g\in G} U_g\varphi = \varphi$.

Proof. (i) $\Rightarrow$ (ii). Let $\varepsilon > 0$ and consider the set

$$\Gamma = \{g \in G : \|U_g\varphi - \varphi\| < \varepsilon/2\} \in p.$$

It follows that for any $g_1, g_2 \in \Gamma$ one has

$$\|U_{g_1}\varphi - U_{g_2}\varphi\| < \varepsilon.$$

This implies that for any $g \in \Gamma^{-1}\Gamma$ one has

$$\|U_g\varphi - \varphi\| < \varepsilon.$$

By Exercise 7 (iii), the set $\Gamma^{-1}\Gamma$ is syndetic. This means that a union of finitely many shifts of $\Gamma^{-1}\Gamma$ gives G and, hence, a union of finitely many balls of radius ε cover the set $\{U_g\varphi\}_{g\in G}$, which, in its turn, means that φ is a compact vector.

(ii)$\Rightarrow$ (iii) If φ is a compact vector, the group G acts on compact space $X = \overline{\{U_g\varphi\}}_{g\in G}$ by isometries $U_g, g \in G$, and hence distally. By utilizing the same argument as in the proof of Proposition 3.2 we see that for any idempotent $p \in \beta G$ one has to have $p\text{-}\lim_{g\in G} U_g\varphi = \varphi$. Since the implication (iii) $\Rightarrow$ (i) is trivial, this concludes the proof of Theorem 4.3. □

Theorem 4.4. *Let* $(U_g)_{g\in G}$ *be a unitary representation of a group* G *on a Hilbert space* $\mathcal{H}$. *The following conditions are equivalent:*

(i) $(U_g)_{g\in G}$ *is weakly mixing;*

(ii) *For any minimal idempotent* $p \in \beta G$ *and any* $\varphi \in \mathcal{H}$ *one has* $p\text{-}\lim_{g\in G} U_g\varphi = 0$;

(iii) *For any* $\varepsilon > 0$ *and* $\varphi_1, \varphi_2 \in \mathcal{H}$ *the set* $\{g : |\langle U_g\varphi_1, \varphi_2\rangle| < \varepsilon\}$ *is* C^*;

(iv) *There exists a minimal idempotent* $p \in \beta G$ *such that for any* $\varphi \in \mathcal{H}$ *one has* $p\text{-}\lim_{g\in G} U_g\varphi = 0$.

Proof. (i) $\Rightarrow$ (ii) If for some minimal $p \in \beta G$ and $\varphi \in \mathcal{H}$ one has $p\text{-}\lim_{g \in G} U_g \varphi = \psi \neq 0$, then (by Proposition 3.2) one has $p\text{-}\lim_{g \in G} U_g \psi = \psi$. It follows now from Theorem 4.3 that ψ is a nontrivial compact vector, which contradicts (i).

(ii) $\Rightarrow$ (iii) The equivalence of these two statements follows immediately from the fact that a set $E \subseteq G$ is C* if and only if E is a member of every minimal idempotent $p \in \beta G$ (see Remark after Corollary 2.16).

(iii) $\Rightarrow$ (iv) Trivial.

(iv) $\Rightarrow$ (i) If $(U_g)_{g \in G}$ is not weakly mixing, then there exists a non-zero compact vector $\varphi \in \mathcal{H}$. But then, by Theorem 4.3, one has $p\text{-}\lim_{g \in G} U_g \varphi = \varphi$, which contradicts (iv). We are done. □

Exercise 21. Show that each of the following properties of a unitary representation $(U_g)_{g \in G}$ on a Hilbert space $\mathcal{H}$ is equivalent to weak mixing.

(i) For any $\varepsilon > 0$, $\varphi_1, \varphi_2 \in \mathcal{H}$ the set $E = \{g : |\langle U_g \varphi_1, \varphi_2 \rangle| < \varepsilon\}$ has the following property: for any $n \in \mathbb{N}$ and any $g_1, g_2, ..., g_n \in G$, $\bigcap_{i=1}^n g_i E$ is syndetic;

(ii) For any $\varepsilon > 0, n \in \mathbb{N}$, and $\varphi_1, ..., \varphi_n \in \mathcal{H}$, there exists $g \in G$ such that $|\langle U_g \varphi_i, \varphi_i \rangle| < \varepsilon$ for $i = 1, ..., n$;

(iii) For any unitary representation $(V_g)_{g \in G}$ of G on a Hilbert space $\mathcal{K}$, the tensor product representation $(U_g \otimes V_g)_{g \in G}$ on $\mathcal{H} \otimes \mathcal{K}$ is weakly mixing.

Theorem 4.5. *Given a unitary representation $(U_g)_{g \in G}$ of a group G on a Hilbert space $\mathcal{H}$, let*

$$\mathcal{H}_c = \{f \in \mathcal{H} : f \textit{ is compact with respect to } (U_g)_{g \in G}\}.$$

Then the restriction of $(U_g)_{g \in G}$ to the invariant space $\mathcal{H}_{wm} = \mathcal{H}_c^\perp$ is weakly mixing.

Proof. Let $\varphi \in \mathcal{H}$, $\varphi \perp \mathcal{H}_c$, and let $p \in \beta G$ be a minimal idempotent. Since $\mathcal{H}_c$ is an invariant subspace, the vector $\psi = p\text{-}\lim_{g \in G} U_g \varphi$ is in $\mathcal{H}_{wm}$. If $\psi \neq 0$, then since $\psi = p\text{-}\lim_{g \in G} U_g \psi$, ψ is a non-zero compact vector, which contradicts the fact that $\psi \in \mathcal{H}_c^\perp$. We are done. □

Corollary 4.6. *Let $p \in \beta G$ be a minimal idempotent and for $\varphi \in \mathcal{H}$, let $P\varphi = p\text{-}\lim_{g \in G} U_g \varphi$. Then P is an orthogonal projection onto $\mathcal{H}_c$.*

We shall specialize now to the case of measure-preserving actions. Let us call a measure-preserving action $(T_g)_{g \in G}$ on a probability space $(X, \mathcal{B}, \mu)$ weakly mixing if the unitary action of G defined on $L^2(X, \mathcal{B}, \mu)$ by $(U_g f)(x) = f(T_g x)$ is weakly mixing on the space $L_0^2(X, \mathcal{B}, \mu) = \{f \in L^2(X, \mathcal{B}, \mu) : \int f d\mu = 0\}$.

The following result is an immediate corollary of Theorem 4.4(i) and is of interest even for $\mathbb{Z}$-actions.

Theorem 4.7. *Let $(T_g)_{g\in G}$ be a weakly mixing measure preserving action on a probability space $(X,\mathcal{B},\mu)$. Then for any $A,B\in\mathcal{B}$ and any $\varepsilon>0$, the set*

$$\{g\in G : |\mu(A\cap T_gB)-\mu(A)\mu(B)|<\varepsilon\}$$

is a C^*-set.

The following theorem should be juxtaposed with the polynomial result formulated above as condition (x) in Theorem 4.1.

Theorem 4.8. *Assume that $(X,\mathcal{B},\mu,T)$ is a weakly mixing system. For any $k\in\mathbb{N}$, $f_0,f_1,f_2,...,f_k\in L^\infty(X,\mathcal{B},\mu)$, any non-constant polynomials $p_1(n),p_2(n),...,p_k(n)\in\mathbb{Z}[n]$ such that for all $i\neq j$, $\deg(p_i-p_j)>0$ and for any minimal idempotent $p\in\beta\mathbb{N}$, one has*

$$p\text{-}\lim_{n\in\mathbb{N}}\int f_0(x)f_1(T^{p_1(n)}x)\cdots f_k(T^{p_k(n)}x)d\mu=\int f_0d\mu\int f_1d\mu\cdots\int f_kd\mu.$$

The proof of Theorem 4.8 is too long to give here. However, it should be noted, that the proof has essentially the same structure as the proof of the parallel Cesàro version given in [B1]. The main and, practically, the only distinction with the proof in [B1] is the need to replace the Cesàro van der Corput trick utilized in [B1] by the following useful fact the proof of which is left to the reader.

Proposition 4.9. *Assume that $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in a Hilbert space $\mathcal{H}$. Let $p\in\beta\mathbb{N}$ be an idempotent. If $p\text{-}\lim_{h\in\mathbb{N}}p\text{-}\lim_{n\in\mathbb{N}}\langle x_{n+h},x_n\rangle=0$ then $p\text{-}\lim_{n\in\mathbb{N}}x_n=0$ weakly.*

The following proposition is an immediate corollary of Theorem 4.8.

Theorem 4.10. *If $(X,\mathcal{B},\mu,T)$ is weakly mixing system, then for any $k\in\mathbb{N}$, any sets $A_0,A_1,...,A_k\in\mathcal{B}$, any non-constant polynomials $p_1(n),p_2(n),...,p_k(n)\in\mathbb{Z}[n]$, such that for all $i\neq j$, $\deg(p_i-p_j)>0$, and any $\varepsilon>0$, the set*

$$\{n:\ |\mu(A_0\cap T^{p_1(n)}A_1\cap...\cap T^{p_k(n)}A_k)-\mu(A_0)\mu(A_1)...\mu(A_k)|<\varepsilon\}$$

is a C^* set.

We shall conclude this section by proving a refinement of the so called Khintchine's recurrence theorem under the assumption that our system is ergodic. Khintchine's theorem (proved in [Kh] for measure preserving $\mathbb{R}$-actions) states that for any measure preserving system $(X,\mathcal{B},\mu,T)$, any $A\in\mathcal{B}$, and any $\varepsilon>0$, the set $E=\{n:\ \mu(A\cap T^{-n}A)>(\mu(A))^2-\varepsilon\}$ is syndetic. As a matter of fact, it is quite easy to show that the set E is always an IP* set and, moreover, this result holds for general semigroup actions. (See [B2], section 5 for details). One would like to extend Khintchine's theorem to sets of the form $A\cap T^{-n}B$. It is clear, however, that any such result can hold only under the additional assumption of ergodicity. We have the following theorem which, for simplicity, will be formulated and proved for $\mathbb{Z}$-actions.

Theorem 4.11. *Assume that $(X, \mathcal{B}, \mu, T)$ is an ergodic, invertible, probability measure preserving system. Then for any $\varepsilon > 0$ and any $A, B \in \mathcal{B}$ the set*

$$E = \{n \in \mathbb{Z} : \ \mu(A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\}$$

is a C_+^* set.

Proof. We are going to utilize the splitting $L^2(X, \mathcal{B}, \mu) = \mathcal{H}_c \oplus \mathcal{H}_{wm}$ (cf. Theorem 4.5 above). Let $1_A = f = f_1 + f_2$, $1_B = g = g_1 + g_2$, where f_1, g_1 belong to the space of compact vectors $\mathcal{H}_c$ and $f_2, g_2 \in \mathcal{H}_c^{\perp} = \mathcal{H}_{wm}$. Note that since the constants belong to the space $\mathcal{H}_c$, one has $\int f_1 d\mu = \mu(A)$, $\int g_1 d\mu = \mu(B)$. By ergodicity of T one has

$$\frac{1}{N}\sum_{n=0}^{N-1} \int f_1(T^n x) g_1(x) d\mu \to \int f_1 d\mu \int g_1 d\mu = \mu(A)\mu(B)$$

as $N \to \infty$. It follows that for any $\varepsilon > 0$ one can find n_0 such that

$$\int f_1(T^{n_0} x) g_1(x) d\mu > \mu(A)\mu(B) - \varepsilon.$$

Let $p \in \beta\mathbb{N}$ be a minimal idempotent. Using the fact that $\mathcal{H}_c$ and $\mathcal{H}_{wm}$ are orthogonal spaces, invariant with respect to the unitary operator U defined by $(Uf)(x) = f(Tx)$, we have:

$$p\text{-}\lim_{n\in\mathbb{N}} \mu(T^{n_0} A \cap T^n B) = p\text{-}\lim_{n\in\mathbb{N}} \int f(T^{n_0} x) g(T^n x) d\mu$$

$$= p\text{-}\lim_{n\in\mathbb{N}} \int f_1(T^{n_0} x) g_1(T^n x) d\mu + p\text{-}\lim_{n\in\mathbb{N}} \int f_2(T^{n_0} x) g_2(T^n x) d\mu$$

$$= \int f_1(T^{n_0} x) g_1(x) d\mu > \mu(A)\mu(B) - \varepsilon.$$

(We also used the fact that $p\text{-}\lim_{n\in\mathbb{N}} g_1(T^n x) = g_1(x)$). It follows that

$$E = \{n : \ \mu(T^{n_0} A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\} \in p.$$

Since p was an arbitrary minimal idempotent, E is a C* set and hence the set $\{n : \ \mu(A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\}$ is C_+^*. We are done. □

5. Some concluding remarks

We started this essay with a question: "What is common between the invertibility of distal maps, partition regularity of the diophantine equation $x - y = z^2$, and the notion of mild mixing"? It was claimed in the Introduction that the answer is: idempotent ultrafilters; the purpose of this short concluding section is to convince the reader that this is indeed so.

As for the relevance of idempotents to the invertibility of distal maps, it is apparent from Proposition 3.2 and Exercise 17 above. So it remains to explain the connection of idempotent ultrafilters to the notion of mild mixing and to partition regularity of the equation $x - y = z^2$.

The notion of mild mixing, which lies between weak and strong mixing, was introduced by Walters in 1972 ([Wa1]) and rediscovered by Furstenberg and Weiss in 1978 ([FW1]). Not unlike the weak mixing, mild mixing admits many equivalent, yet diverse, formulations and plays a crucial role in ergodic proofs of some strong combinatorial results. (See, for instance, [FK2] and [BM2].) The following proposition, which we do not prove here, lists some interesting equivalent conditions, each of which may be taken for a definition of mild mixing. For sake of simplicity, we restrict ourself to measure preserving $\mathbb{Z}$-actions, but it should be mentioned that the notion of mild mixing makes perfect sense for unitary actions of locally compact groups and is of importance in the study of non-singular (i.e. not necessarily measure-preserving) actions as well. (See, for example, [ScW], [Sc] and [Wa2] for more information.) The connection of mild mixing to idempotent ultrafilters is perfectly clear from the items (iv) and (v) below.

Theorem 5.1. *Let T be an invertible measure-preserving transformation of a probability measure space $(X, \mathcal{B}, \mu)$. Let U_T denote the operator defined on measurable functions by $(U_T)f(x) = f(Tx)$. The following conditions are equivalent:*

(i) *For every $A \in \mathcal{B}$ with $0 < \mu(A) < 1$, one has*

$$\liminf_{|n|\to\infty} \mu(A \triangle T^n A) > 0.$$

(ii) *For any ergodic measure preserving system $(Y, \mathcal{D}, \nu, S)$, the transformation $T \times S$ is ergodic on $X \times Y$. (Note: it is not assumed that $\nu(Y)$ is finite.)*

(iii) *There are no non-constant rigid functions in $L^2(X, \mathcal{B}, \mu)$. (A function $f \in L^2(X, \mathcal{B}, \mu)$ is called rigid, if for some sequence $(n_k)_{k\in\mathbb{N}}$ one has $U_T^{n_k} f \to f$ in L^2.)*

(iv) *For any idempotent $p \in \beta\mathbb{N}$ and any $f \in L^2(X, \mathcal{B}, \mu)$ one has*

$$p\text{-}\lim_{n\in\mathbb{N}} U_T^n f = \int f\, d\mu \ \textit{(weakly)}.$$

(v) *For any $k \in \mathbb{N}$, any $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ and any non-constant polynomials $p_1(n), p_2(n), \ldots, p_k(n) \in \mathbb{Z}[n]$, such that for all $i \neq j$, $\deg(p_i - p_j) > 0$, and any idempotent $p \in \beta\mathbb{N}$, one has*

$$p\text{-}\lim_{n\in\mathbb{N}} \int f_0(x) f_1(T^{p_1(n)}x) \ldots f_k(T^{p_k(n)}x) d\mu = \int f_0\, d\mu \int f_1\, d\mu \ldots \int f_k\, d\mu.$$

Finally, we shall explain briefly how one proves that for any finite partition of $\mathbb{N}$ there exist x, y, z in the same cell of partition, such that $x - y = z^2$. The reader will find the missing details and more discussion in [B2] and [B3]. The proof hinges in the following fact.

Proposition 5.2. *(Cf. [B2], Propositions 3.11 and 3.12; see also [BFM]) For any probability measure preserving system* $(X, \mathcal{B}, \mu, T)$*, any* $A \in \mathcal{B}$ *and any idempotent* $p \in \beta\mathbb{N}$ *one has* $p\text{-}\lim_{n\in\mathbb{N}} \mu(A \cap T^{n^2}A) \geq \mu^2(A)$*.*

It follows (via Furstenberg's correspondence principle, see, for example, [F4], p.77, or [B3], Theorems 6.4.4 and 6.4.7), that for any set of positive upper density $E \subseteq \mathbb{N}$ and any IP set Γ, there exists $z \in \Gamma$ with $d^*(E \cap (E - z^2)) > 0$. This, in its turn, immediately implies that, for some $x, y \in E$ and $z \in \Gamma$, one has $x - y = z^2$.

Consider now an arbitrary finite partition $\mathbb{N} = \bigcup_{i=1}^r C_i$. Reindexing if necessary, we may assume that those C_i which have positive upper Banach density have indices $1, 2, \ldots, s$, where $s \leq r$. Let $U = \bigcup_{i=1}^s C_i$. It is not hard to see that U contains an IP set (this follows from the almost obvious fact that U has to contain arbitrarily long blocks of consecutive integers). It follows now from Hindman's theorem that there exist i_0, $1 \leq i_0 \leq s$, and an IP set Γ such that $\Gamma \subseteq C_{i_0}$. It follows now from the remarks above that for some $x, y \in C_{i_0}$ and $z \in \Gamma$ one has $x - y = z^2$. We are done.

Acknowledgment. I wish to thank Hillel Furstenberg, Neil Hindman, Sasha Leibman and Christian Schnell for their help in preparing these notes. I also would like to thank Sergey Bezuglyi and Sergiy Kolyada for organizing, under the aegis of INTAS, CRDF and ESF, the Katsiveli-2000 meeting and for their efforts to promote the theory of dynamical systems.

References

[A1] J. Auslander, On the proximal relation in topological dynamics, *Proc. Amer. Math. Soc.*, **11** (1960), 890–895.

[A2] J. Auslander, *Minimal Flows and their Extensions*, North-Holland Mathematics Studies, 153, North-Holland Publishing Co., Amsterdam, 1988. xii+265 pp.

[B1] V. Bergelson, Weakly mixing PET, *Ergodic Theory and Dynamical Systems*, **7** (1987), 337-349.

[B2] V. Bergelson, Ergodic Ramsey theory–an update. *Ergodic theory of* $\mathbb{Z}^d$ *actions* (Warwick, 1993–1994), 1–61, London Math. Soc. Lecture Note Ser., 228, Cambridge Univ. Press, Cambridge, 1996.

[B3] V. Bergelson, Ergodic theory and Diophantine problems. *Topics in symbolic dynamics and applications*, 167–205, London Math. Soc. Lecture Note Ser., 279, Cambridge Univ. Press, Cambridge, 2000.

[BFM] V. Bergelson, H. Furstenberg and R. McCutcheon, IP-sets and polynomial recurrence, *Ergodic Theory Dynam. Systems*, **16** (1996), no. 5, 963–974.

[BFHK] V. Bergelson, H. Furstenberg, N. Hindman, and Y. Katznelson, An algebraic proof of van der Waerden's theorem, *Enseign. Math. (2)*, **35** (1989), no. 3-4, 209–215.

[BH1] V. Bergelson and N. Hindman, Nonmetrizable topological dynamics and Ramsey theory, *Trans. AMS*, **320** (1990), 293-320.

[BH2] V. Bergelson and N. Hindman, Additive and multiplicative Ramsey theorems in $\mathbb{N}$ – some elementary results, *Combin. Probab. Comput.*, **2** (1993), no. 3, 221–241.

[BH3] V. Bergelson and N. Hindman, On IP* sets and central sets, *Combinatorica*, **14** (1994), no. 3, 269–277.

[BL] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *Journal of AMS*, **9** (1996), 725 –753.

[BM1] V. Bergelson and R. McCutcheon, Uniformity in the polynomial Szemerédi theorem. *Ergodic theory of $\mathbb{Z}^d$ actions* (Warwick, 1993–1994), 273–296, London Math. Soc. Lecture Note Ser., 228, Cambridge Univ. Press, Cambridge, 1996.

[BM2] V. Bergelson and R. McCutcheon, An ergodic IP polynomial Szemerédi theorem, *Mem. Amer. Math. Soc.*, **146** (2000), no. 695, viii+106 pp.

[BR] V. Bergelson and J. Rosenblatt, Mixing actions of groups, *Illinois J. Math.*, **32** (1988), no. 1, 65–80.

[D] H. Dye, On the ergodic mixing theorem, *Trans. AMS*, **118** (1965), 123– 30.

[E1] R. Ellis, Distal transformation groups, *Pacific J. Math.*, **8** (1958), 401–405.

[E2] R. Ellis, A semigroup associated with a transformation group, *Trans. AMS*, **94** (1960), 272–281.

[E3] R. Ellis, *Lectures on Topological Dynamics*, Benjamin, New York, 1969.

[F1] H. Furstenberg, Strict ergodicity and transformations of the torus, *Amer. J. Math.*, **83** (1961), 573–601.

[F2] H. Furstenberg, The structure of distal flows, *Amer. J. Math.*, **85** (1963), 477–515.

[F3] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.*, **31** (1977), 204-256.

[F4] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[FK1] H. Furstenberg and Y. Katznelson, An ergodic Szeméredi theorem for IP-systems and combinatorial theory, *J. Analyse Math.* **45** (1985), 117–168.

[FK2] H. Furstenberg and Y. Katznelson, Idempotents in compact semigroups and Ramsey theory, *Israel J. Math.*, **68** (1990), 257-270.

[FW1] H. Furstenberg and B. Weiss, The finite multipliers of infinite ergodic transformations. The structure of attractors in dynamical systems (Proc. Conf., North Dakota State Univ., Fargo, N.D., 1977), 127–132, *Lecture Notes in Math.*, **668**, Springer, Berlin, 1978.

[FW2] H. Furstenberg and B. Weiss, Simultaneous Diophantine approximation and IP-sets. *Acta Arith.*, **49** (1988), 413–426.

[GH] W. Gottschalk and G. Hedlund, *Topological Dynamics*, AMS Coll. Publ., vol. 36, Providence, 1955.

[HaL] G. Hardy and J. Littlewood, Some problems of diophantine approximation, *Acta Mathematica*, **37** (1914), 155–191.

[He] H. Helson, *Harmonic Analysis*, Addison-Wesley, Reading, MA, 1983.

[Hi1] N. Hindman, Finite sums from sequences within cells of a partition of $\mathbb{N}$, *J. Combinatorial Theory* (Series A), **17** (1974), 1–11.

[Hi2] N. Hindman, Partitions and sums and products of integers, *Trans. AMS*, **247** (1979), 227–245.

[HiS] N. Hindman and D. Strauss, *Algebra in the Stone-Čech Compactification. Theory and Applications*, de Gruyter Expositions in Mathematics, 27. Walter de Gruyter & Co., Berlin, 1998. xiv+485 pp.

[Hopf1] E. Hopf, Complete transitivity and the ergodic principle, *Proc. Nat. Acad. Sci. U.S.A.*, **18** (1932), 204–209.

[Hopf2] E. Hopf, Proof of Gibbs' hypothesis on the tendency toward statistical equilibrium, *Proc. Nat. Acad. Sci. U.S.A.*, **18** (1932), 333–340.

[Hopf3] E. Hopf, *Ergodentheorie,* Chelsea, New York, 1948.

[Kh] A.Y. Khintchine, Eine Verschärfung des Poincaréschen Wiederkehrsatzes, *Comp. Math.*, **1** (1934), 177–179.

[KN] B.O. Koopman and J. von Neumann, Dynamical systems of continuous spectra, *Proc. Nat. Acad. Sci. U.S.A.*, **18** (1932), 255–263.

[Kre] U. Krengel, Weakly wandering vectors and weakly independent partitions, *Trans. AMS*, **164** (1972), 199–226.

[Kro] L. Kronecker, Die Periodensysteme von Funktionen Reeller Variablen, *Berliner Sitzungsberichte* (1884), 1071–1080.

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

[R-N] C. Ryll-Nardzewski, On fixed points of semigroups of endomorphisms of linear spaces, *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, 55–61.

[Sc] K. Schmidt, Asymptotic properties of unitary representations and mixing, *Proc. London Math. Soc.* (3) **48** (1984), no. 3, 445–460.

[Sz] P. Szüsz, Verschärfung eines Hardy-Littlewoodschen Satzes, *Acta Math. Acad. Sci. Hungar.*, **4** (1953), 115–118.

[SY] H. Shi and H. Yang, Nonmetrizable topological dynamical characterization of central sets, *Fund. Math.*, **150** (1996), 1–9.

[ScW] K. Schmidt and P. Walters, Mildly mixing actions of locally compact groups, *Proc. London Math. Soc.* (3) **45** (1982), 506–518.

[Wa1] P. Walters, Some invariant σ-algebras for measure-preserving transformations, *Trans. Amer. Math. Soc.* **163** (1972), 357–368.

[Wa2] P. Walters, A mixing property for nonsingular actions, 163–183, Aste'risque, 98-99, Soc. Math. France, Paris, 1982.

[Weyl] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* **77** (1916), 313–352.

Department of Mathematics, The Ohio State University, Columbus, OH 43210

E-mail address: vitaly@math.ohio-state.edu