A proof of a sumset conjecture of Erdős

Joel Moreira* Florian K. Richter† Donald Robertson‡

February 9, 2022

Abstract

In this paper we show that every set $A \subset \mathbb{N}$ with positive density contains $B + C$ for some pair B, C of infinite subsets of $\mathbb{N}$, settling a conjecture of Erdős. The proof features two different decompositions of an arbitrary bounded sequence into a structured component and a pseudo-random component. Our methods are quite general, allowing us to prove a version of this conjecture for countable amenable groups.

Contents

*Northwestern University – joel.moreira@northwestern.edu

†Northwestern University – fkr@northwestern.edu

‡University of Utah – robertso@math.utah.edu

1. Introduction

History and previous results. Sumsets $B+C \coloneqq \{b+c : b \in B, c \in C\}$ for $B, C \subset \mathbb{N}$ are a central object of study in additive combinatorics. In particular, it is natural to ask which sets $A \subset \mathbb{N}$ contain a sumset $B+C$ with B and C infinite. It follows from Hindman's theorem [Hin79a] that, whenever $\mathbb{N}$ is finitely partitioned, one of the cells contains $B+C$ for $B, C \subset \mathbb{N}$ infinite. The following conjectured density analogue, attributed to Erdős in [Nat80], is called an "old problem" in [EG80, p. 85].

Conjecture 1.1 (Erdős sumset conjecture)**.** *If $A \subset \mathbb{N}$ satisfies*

$$\limsup_{N\to\infty} \frac{|A \cap \{1,\ldots,N\}|}{N} > 0$$

then A contains $B+C \coloneqq \{b+c : b \in B, c \in C\}$, where B and C are infinite subsets of $\mathbb{N}$.

Nathanson [Nat80] showed that a set A with positive upper density contains a sum $B+C$ for a set B of positive density and a set C of any finite cardinality. More recently, a major breakthrough was made by Di Nasso, Goldbring, Jin, Leth, Lupini and Mahlburg [DN+15] who employed nonstandard analysis and ideas from ergodic theory to show that a set $A \subset \mathbb{N}$ with upper density greater than $1/2$ contains a sum $B+C$ where B and C are infinite sets. As a corollary, derived using Ramsey's theorem and a result of Hindman, it follows that if A has positive upper density, then for some $t \in \mathbb{N}$ the union $A \cup (A-t)$ contains a sum $B+C$ where B and C are infinite sets. Some further progress on a variant of Conjecture 1.1 was also made in [ACG17].

Main results. The goal of this paper is to verify Conjecture 1.1. In fact we prove a stronger result. Recall that a **Følner sequence** in $\mathbb{N}$ is any sequence $\Phi\colon N \mapsto \Phi_N$ of finite, non-empty subsets of $\mathbb{N}$ satisfying

$$\lim_{N\to\infty} \frac{\big|(\Phi_N + m) \triangle \Phi_N\big|}{|\Phi_N|} = 0$$

for all $m \in \mathbb{N}$. For example, any sequence $N \mapsto \{a_N+1, a_N+2, \ldots, b_N\}$ of intervals in $\mathbb{N}$ with length $b_N - a_N$ tending to infinity is a Følner sequence. Given a Følner sequence Φ and a set $A \subset \mathbb{N}$ the quantity

$$\overline{\mathsf{d}}_\Phi(A) \coloneqq \limsup_{N\to\infty} \frac{|A \cap \Phi_N|}{|\Phi_N|}$$

is the **upper density** of A with respect to Φ. If

$$\lim_{N\to\infty}\frac{|A\cap\Phi_N|}{|\Phi_N|}$$

exists we denote it by $\mathsf{d}_\Phi(A)$ and call it the **density** of A with respect to Φ. The following is our main result, which verifies a generalization of Conjecture 1.1 to Følner sequences.

Theorem 1.2. *For every $A\subset\mathbb{N}$ that satisfies $\overline{\mathsf{d}}_\Phi(A)>0$ for some Følner sequence Φ one can find infinite sets $B,C\subset\mathbb{N}$ with $B+C\subset A$.*

In fact, our methods are flexible enough to prove a version of Theorem 1.2 in countable amenable groups. A **two-sided Følner sequence** on a discrete countable group G is any sequence $\Phi\colon N\mapsto\Phi_N$ of finite, non-empty subsets of G satisfying

$$\lim_{N\to\infty}\frac{|(\Phi_N g)\bigtriangleup\Phi_N|}{|\Phi_N|}=0=\lim_{N\to\infty}\frac{|\Phi_N\bigtriangleup(g\Phi_N)|}{|\Phi_N|} \tag{1}$$

for all $g\in G$. A countable group G is called **amenable** if and only if it admits a two-sided Følner sequence (cf. [Gre69; TW16]). Given a two-sided Følner sequence Φ on G and a set $A\subset G$, the quantity

$$\overline{\mathsf{d}}_\Phi(A):=\limsup_{N\to\infty}\frac{|A\cap\Phi_N|}{|\Phi_N|} \tag{2}$$

is the **upper density** of A with respect to Φ. If

$$\lim_{N\to\infty}\frac{|A\cap\Phi_N|}{|\Phi_N|}$$

exists then we denote it by $\mathsf{d}_\Phi(A)$ and call it the **density** of A with respect to Φ.

Theorem 1.3. *Let G be a countable group, let Φ be a two-sided Følner sequence on G and let $A\subset G$ be such that $\overline{\mathsf{d}}_\Phi(A)>0$. Then there are infinite sets $B,C\subset G$ with $BC=\{bc:b\in B,c\in C\}\subset A$.*

Strategy of the proof. We outline here quite broadly the main ideas in the proof of Theorem 1.2. We freely make use of terminology which is only defined later in the paper. In particular, the relevant background on ultrafilters is given at the beginning of Section 2.

To begin with, we borrow ideas from [DN+15] to show that whenever one has

$$\lim_{m\to\mathsf{p}} \mathsf{d}_\Psi((A-m)\cap(A-\mathsf{p}))>0 \tag{3}$$

for some Følner sequence Φ and some non-principal ultrafilter p, necessarily A contains a sum $B+C$ with $B, C\subset\mathbb{N}$ infinite. Here we write $A-\mathsf{p}$ for the set $\{n\in\mathbb{N} : A-n\in\mathsf{p}\}$. Thus the main part of our proof of Theorem 1.2 consists of finding, for every Følner sequence Φ and every $A\subset\mathbb{N}$ with $\overline{\mathsf{d}}_\Phi(A)>0$, a non-principal ultrafilter p and a Følner subsequence Ψ of Φ such that (3) is satisfied.

Given $f\colon\mathbb{N}\to\mathbb{C}$ and $m\in\mathbb{N}$, write $\mathsf{R}^m f$ for the function $n\mapsto f(m+n)$. If in addition p is an ultrafilter on $\mathbb{N}$ we write $\mathsf{R}^{\mathsf{p}} f$ for the function

$$n\mapsto \lim_{m\to\mathsf{p}} f(n+m)$$

for all $n\in\mathbb{N}$. In doing so one can rewrite 1_{A-m} as $\mathsf{R}^m 1_A$ and $1_{A-\mathsf{p}}$ as $\mathsf{R}^{\mathsf{p}} 1_A$. We can therefore rewrite (3) in the form

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m 1_A,\ \mathsf{R}^{\mathsf{p}} 1_A\rangle_\Psi>0 \tag{4}$$

where, for two bounded functions $f, g\colon\mathbb{N}\to\mathbb{C}$, the inner product $\langle\cdot,\ \cdot\rangle_\Psi$ is defined as

$$\langle f,\ g\rangle_\Psi := \lim_{N\to\infty}\frac{1}{|\Psi_N|}\sum_{n\in\Psi_N} f(n)\,\overline{g(n)}.$$

The utility of ultrafilters in our proof is two-fold. On the one hand, the language of ultrafilters leads us to (3) and (4), which are similar to expressions encountered in other problems of additive combinatorics. In fact, having reduced the proof of Theorem 1.2 to a statement involving the bilinear functional $(f,g)\mapsto \lim_{m\to\mathsf{p}}\langle \mathsf{R}^m f,\ \mathsf{R}^{\mathsf{p}} g\rangle_\Psi$ is particularly useful, since it opens the door for using tools and ideas from functional analysis and ergodic Ramsey theory. On the other hand, shifts by ultrafilters are more versatile than shifts by natural numbers, which we exploit at numerous different places in the proof of Theorem 1.2.

In [DN+15, Theorem 5.5] the language of non-standard analysis was used to verify (4) when A is "pseudo-random". Roughly speaking, the set A is pseudo-random if it is close to independent from most of its shifts. It is natural to ask [DN+15, Questions 5.6, 5.7] what happens when A is not pseudo-random. In this case, it is beneficial to employ a decomposition of 1_A into structured and pseudo-random components. Inspired by the Jacobs–de Leeuw–Glicksberg splitting on Hilbert spaces [Jac56; LG61], we prove that

1_A can always be decomposed as a sum $f_{\mathsf{wm}} + f_{\mathsf{c}}$ of a weak mixing function f_{wm} and a compact function f_{c}. We think of f_{wm} as being the "pseudo-random" component of 1_A and of f_{c} as the "structured" component of 1_A.

The decomposition $1_A = f_{\mathsf{wm}} + f_{\mathsf{c}}$ is stable under shifts by $m \in \mathbb{N}$ in the sense that $\mathsf{R}^m f_{\mathsf{wm}} + \mathsf{R}^m f_{\mathsf{c}}$ is the decomposition of $\mathsf{R}^m 1_A = 1_{A-m}$ into weak mixing and compact functions. In light of this fact, we can consider the left hand side of (4) as a sum of two terms, one with $\mathsf{R}^m 1_A$ replaced by the weak mixing function $\mathsf{R}^m f_{\mathsf{wm}}$, the other with $\mathsf{R}^m 1_A$ replaced by the compact function $\mathsf{R}^m f_{\mathsf{c}}$:

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m 1_A,\ \mathsf{R}^{\mathsf{p}} 1_A \rangle_{\Psi,\mathsf{p}} = \lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} 1_A \rangle_{\Psi} + \lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} 1_A \rangle_{\Psi}. \quad (5)$$

Unfortunately, the decomposition into compact and weak mixing components is not stable under shifts by ultrafilters, so we are unable to use it to understand $\mathsf{R}^{\mathsf{p}} 1_A$. For this reason we devise a second splitting whose interaction with ultrafilters we are able to control. This second splitting asserts that $1_A = f_{\mathsf{anti}} + f_{\mathsf{Bes}}$, where the "structured" component f_{Bes} is a Besicovitch almost periodic function, which is a stronger property then being a compact function, and the complement f_{anti} is characterized by being orthogonal to $e^{2\pi i n\theta}$ for all $\theta \in [0, 1)$, which is a weaker form of "pseudo-randomness" than weak mixing. It is the specialized nature of f_{Bes} that reacts well with ultrafilters.

Applying our second splitting to $\mathsf{R}^{\mathsf{p}} 1_A$ in the last term of (5) leaves us with a sum of the following three terms.

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}} \rangle_{\Psi} \quad (6)$$

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}} \rangle_{\Psi} \quad (7)$$

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} 1_A \rangle_{\Psi} \quad (8)$$

We show that (8) is zero using the pseudo-randomness of weak mixing. Positivity of the term (6) follows from the close relationship between f_{Bes} and its shifts by ultrafilters. The remaining term, (7), which involves f_{c} and f_{anti}, is the most delicate. To show it is non-negative we adapt an argument of Beiglböck [Bei11]. All together, this proves that the sum of the three terms in (6), (7), and (8) is positive, which implies (4).

It is reasonable to ask why we do not apply the splitting $f_{\mathsf{Bes}} + f_{\mathsf{anti}}$ to both occurrences of 1_A in (4). The reason lies in the strength of the pseudo-randomness that weak mixing provides. We would not be able to handle the hypothetical term

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f_{\mathsf{anti}},\ \mathsf{R}^{\mathsf{p}} 1_A \rangle_{\Psi}$$

pairing f_{anti} with 1_A, whereas we are able to handle (8).

Structure of the paper. The purpose of Section 2 is to review the relevant material on ultrafilters and then to prove that (3) implies Theorem 1.2. In Section 3 we prove our two splitting results. The proof of Theorem 1.2 is concluded in Section 4. In Section 5 we explain the few steps where the proof of Theorem 1.3 differs from that of Theorem 1.2. Finally, in Section 6 we discuss some relevant open questions.

Acknowledgements. The first author is grateful for the support of the NSF via grant DMS-1700147. The third author is grateful for the support of the NSF via grant DMS-1703597. We thank John H. Johnson for providing useful references and Vitaly Bergelson for reading and commenting on an early draft of this paper. We would also like to thank the anonymous referees for their careful reading of the manuscript and their detailed suggestions. Their efforts have improved the readability of the paper.

The first author became interested in the Erdős sumset conjecture while visiting Martino Lupini at CalTech in June 2017 and is thankful for the hospitality provided. We are also thankful to Michael Björklund, Alexander Fish and Anush Tserunyan for interesting discussions and to AIM for hosting the workshop "Nonstandard methods in combinatorial number theory" in August 2017 at which these discussions took place.

2. Ultrafilter reformulation

For the proofs of Theorem 1.2 and Theorem 1.3 we found it crucial to rely on the theory of ultrafilters, which has proven to be very effective in solving problems in Ramsey theory in the past. In this section we recall briefly some of the basic definitions and facts that we will utilize in this paper and then reduce Theorem 1.2 to a statement of the form (3). Readers in want of a friendly introduction to ultrafilters may well enjoy [Ber96, Section 3]; for a comprehensive treatment see [HS12].

An **ultrafilter** on $\mathbb{N}$ is any non-empty collection p of subsets of $\mathbb{N}$ that is closed under finite intersections and supersets and satisfies

$$A \in \mathsf{p} \iff \mathbb{N}\backslash A \notin \mathsf{p}$$

for every $A \subset \mathbb{N}$. Given $n \in \mathbb{N}$, the collection $\mathsf{p}_n \coloneqq \{A \subset \mathbb{N} : n \in A\}$ is an ultrafilter; ultrafilters of this kind are called **principal**. We embed $\mathbb{N}$ in $\beta\mathbb{N}$

using the map $n \mapsto \mathsf{p}_n$. For the existence of non-principal ultrafilters, which follows from the axiom of choice, see [HS12, Theorem 3.8].

The set of all ultrafilters on $\mathbb{N}$ is denoted by $\beta\mathbb{N}$. Given $A \subset \mathbb{N}$ and using the above embedding of $\mathbb{N}$ in $\beta\mathbb{N}$, write $\mathsf{cl}(A) \coloneqq \{\mathsf{p} \in \beta\mathbb{N} : A \in \mathsf{p}\}$ for the **closure** of A in $\beta\mathbb{N}$. The family $\{\mathsf{cl}(A) : A \subset \mathbb{N}\}$ forms a base for a topology on $\beta\mathbb{N}$ with respect to which $\beta\mathbb{N}$ is a compact Hausdorff space. We note that $\mathsf{cl}(A) \cap \mathsf{cl}(B) = \mathsf{cl}(A \cap B)$ for all $A, B \subset \mathbb{N}$. The map $n \mapsto \mathsf{p}_n$ embeds $\mathbb{N}$ densely in $\beta\mathbb{N}$. Endowed with this topology, $\beta\mathbb{N}$ can be identified with the Stone–Čech compactification of $\mathbb{N}$, which means that it has the following universal property: for any function $f\colon \mathbb{N} \to K$ into a compact Hausdorff space K there is a unique continuous function $\beta f\colon \beta\mathbb{N} \to K$ such that $(\beta f)(\mathsf{p}_n) = f(n)$ for all $n \in \mathbb{N}$. When no confusion may arise we denote p_n simply by n.

Given a function $f\colon \mathbb{N} \to K$ with K a compact Hausdorff space and given an ultrafilter $\mathsf{p} \in \beta\mathbb{N}$, one can characterize $(\beta f)(\mathsf{p})$ as the unique point x in K such that, for any neighborhood U of x, the set $\{n \in \mathbb{N} : f(n) \in U\}$ belongs to p. For this reason we use the notation

$$\lim_{n \to \mathsf{p}} f(n) \coloneqq (\beta f)(\mathsf{p}).$$

Given a set $A \subset \mathbb{N}$ we define

$$A - \mathsf{p} \coloneqq \{n \in \mathbb{N} : A - n \in \mathsf{p}\}$$

for all ultrafilters p on $\mathbb{N}$. Addition on $\mathbb{N}$ can be extended to a binary operation $+$ on $\beta\mathbb{N}$ by

$$\mathsf{p} + \mathsf{q} = \{A \subset \mathbb{N} : A - \mathsf{q} \in \mathsf{p}\} = \lim_{n \to \mathsf{p}} \lim_{m \to \mathsf{q}} n + m$$

for all p, q in $\beta\mathbb{N}$. We remark that despite being represented with the symbol $+$, this operation is *not* commutative. We mention this operation only to present the following lemma giving a criterion for a set of natural numbers to contain $B + C$; it will not be used throughout in the proof of Theorem 1.2. This lemma was independently discovered by Di Nasso and a proof was presented in [ACG17, Proposition 3.1].

Lemma 2.1 (cf. Lemma 5.1)**.** *Fix $A \subset \mathbb{N}$. There are non-principal ultrafilters p and q with the property that $A \in \mathsf{p} + \mathsf{q}$ and $A \in \mathsf{q} + \mathsf{p}$ if and only if there are infinite sets $B, C \subset \mathbb{N}$ with $B + C \subset A$.*

Here is the main theorem of this section, which is inspired by the proof of [DN+15, Theorem 3.2].

Theorem 2.2. *Let $A \subset \mathbb{N}$. If there exist a Følner sequence Φ in $\mathbb{N}$ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\mathsf{d}_\Phi\big((A-n) \cap (A-\mathsf{p})\big)$ exists for all $n \in \mathbb{N}$ and*

$$\lim_{n\to\mathsf{p}} \mathsf{d}_\Phi\big((A-n) \cap (A-\mathsf{p})\big) > 0 \tag{9}$$

then there exist infinite sets $B, C \subset \mathbb{N}$ such that $A \supset B + C$.

The following result of Bergelson [Ber85] will be crucial for the proof of Theorem 2.2. We present a short proof of it for completeness.

Lemma 2.3 (cf. [Ber85, Theorem 1.1])**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $n \mapsto B_n$ be a sequence in $\mathcal{B}$. Assume that there exists $\epsilon > 0$ such that $\mu(B_n) \geq \epsilon$ for all $n \in \mathbb{N}$. Then there exists an injective map $\sigma \colon \mathbb{N} \to \mathbb{N}$ such that*

$$\mu\left(B_{\sigma(1)} \cap \cdots \cap B_{\sigma(n)}\right) > 0 \tag{10}$$

for every $n \in \mathbb{N}$.

Proof. The collection $\mathcal{F}$ of all finite sets $F \subset \mathbb{N}$ with the property that $\mu(\bigcap_{n\in F} B_n) = 0$ is countable, and therefore the union $X_0 = \bigcup_{F\in\mathcal{F}}(\bigcap_{n\in F} B_n)$ has $\mu(X_0) = 0$.

For each $N \in \mathbb{N}$ let $f_N := \frac{1}{N}\sum_{n=1}^{N} 1_{B_n}$. It is clear that $\int_X f_N \, \mathsf{d}\mu \geq \epsilon$ for every $N \in \mathbb{N}$. By Fatou's lemma, the function $f := \limsup_{N\to\infty} f_N$ also satisfies $\int_X f \, \mathsf{d}\mu \geq \epsilon$. Therefore there exists a point $x \in X \backslash X_0$ with $f(x) > 0$, and in particular the set $\{n \in \mathbb{N} : x \in B_n\}$ is infinite. Let $\sigma(n)$ be an enumeration of that set.

To show that (10) holds notice that, for every $n \in \mathbb{N}$, the set $\{\sigma(1), \ldots, \sigma(n)\}$ can not be in $\mathcal{F}$ because $x \in B_{\sigma(1)} \cap \ldots \cap B_{\sigma(n)}$ but $x \notin X_0$. □

Given a Følner sequence Φ on $\mathbb{N}$ write $\mathcal{M}(\Phi)$ for the set of Radon probability measures on $\beta\mathbb{N}$ that are weak* accumulation points of the set $\{\mu_N : N \in \mathbb{N}\}$, where

$$\mu_N := \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} \delta_n \tag{11}$$

and δ_n is the unit mass at the principal ultrafilter p_n.

Corollary 2.4. *Let Φ be a Følner sequence on $\mathbb{N}$ and, for each $n \in \mathbb{N}$, let $A_n \subset \mathbb{N}$. Assume $\mathsf{d}_\Phi(A_n)$ exists for all $n \in \mathbb{N}$ and that there exists $\epsilon > 0$ such that $\mathsf{d}_\Phi(A_n) \geq \epsilon$ for all $n \in \mathbb{N}$. Then there exists an injective sequence $\sigma \colon \mathbb{N} \to \mathbb{N}$ such that*

$$\overline{\mathsf{d}}_\Phi\left(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)}\right) > 0$$

for every $n \in \mathbb{N}$.

Proof. Let $\mu \in \mathcal{M}(\Phi)$ and let $B_n = \mathsf{cl}(A_n)$. The set B_n is clopen and the density of A_n along Φ exists so $\mu(B_n) = \mathsf{d}_\Phi(A_n)$ for all $n \in \mathbb{N}$. Apply Lemma 2.3 to the probability space $(\beta\mathbb{N}, \mathcal{B}, \mu)$, where $\mathcal{B}$ is the Borel σ-algebra on $\beta\mathbb{N}$, to find an injective map $\sigma\colon \mathbb{N} \to \mathbb{N}$ such that (10) holds for every $n \in \mathbb{N}$. Since $B_{\sigma(1)} \cap \cdots \cap B_{\sigma(n)} = \mathsf{cl}(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)})$, this implies that $\overline{\mathsf{d}}_\Phi\left(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)}\right) \geq \mu(B_{\sigma(1)} \cap \cdots \cap B_{\sigma(n)}) > 0$ as desired. □

The next proposition, whose statement (and proof) is heavily influenced by the paper [DN+15], can be seen as an ultrafilter-free version of Theorem 2.2.

Proposition 2.5. *Let $A \subset \mathbb{N}$. If there exist a Følner sequence Φ in $\mathbb{N}$, a set $L \subset \mathbb{N}$ and $\epsilon > 0$ such that $\mathsf{d}_\Phi\big((A-m) \cap L\big)$ exists for every $m \in \mathbb{N}$, and for every finite subset $F \subset L$*

$$\bigcap_{\ell \in F} (A - \ell) \ \cap \ \Big\{m \in \mathbb{N} : \mathsf{d}_\Phi\big((A-m) \cap L\big) > \epsilon\Big\} \text{ is infinite} \qquad (12)$$

then there exist infinite sets B, C such that $A \supset B + C$.

Proof. Let $F_1 \subset F_2 \subset \cdots$ be an increasing exhaustion of L by finite subsets. Construct a sequence $n \mapsto e_n$ in $\mathbb{N}$ of distinct elements such that

$$e_n \in \bigcap_{\ell \in F_n} (A - \ell) \cap \Big\{m \in \mathbb{N} : \mathsf{d}_\Phi\big((A-m) \cap L\big) > \epsilon\Big\}$$

for each $n \in \mathbb{N}$. This can be done because each of the sets above is infinite by hypothesis.

In particular $\mathsf{d}_\Phi\big((A - e_n) \cap L\big) > \epsilon$ for all $n \in \mathbb{N}$. The Bergelson intersectivity lemma (Corollary 2.4) implies that, for some subsequence $n \mapsto e_{\sigma(n)}$ of e the intersection

$$\Big((A - e_{\sigma(1)}) \cap L\Big) \cap \cdots \cap \Big((A - e_{\sigma(n)}) \cap L\Big)$$

is infinite for all $n \in \mathbb{N}$.

Choose $b_1 \in F_{\sigma(1)}$ and put $j_1 = 1$. Choose $c_1 = e_{\sigma(1)}$. Thus $c_1 \in A - b_1$. Next choose $b_2 \in (A - c_1) \cap L$ outside $F_{\sigma(1)}$ and let j_2 be minimal with $b_2 \in F_{\sigma(j_2)}$. (In particular b_2 is not equal to b_1.) Then choose $c_2 = e_{\sigma(j_2)} \in (A - b_1) \cap (A - b_2)$. Continue this process inductively, choosing

$$b_{n+1} \in (A - c_1) \cap \cdots \cap (A - c_n) \cap L = (A - e_{\sigma(j_1)}) \cap \cdots \cap (A - e_{\sigma(j_n)}) \cap L$$

outside $F_{\sigma(j_n)}$ and choosing j_{n+1} minimal with $b_{n+1} \in F_{\sigma(j_{n+1})}$ and then choosing

$$c_{n+1} = e_{\sigma(j_{n+1})} \in (A - b_1) \cap \cdots \cap (A - b_{n+1})$$

which is distinct from $c_1, \ldots, c_n$ because e is injective. Take $B = \{b_n : n \in \mathbb{N}\}$ and $C = \{c_n : n \in \mathbb{N}\}$ to conclude the proof. $\square$

The proof of Theorem 2.2 is now quite straightforward.

Proof of Theorem 2.2. Let $L = A - \mathsf{p} = \{\ell \in \mathbb{N} : A - \ell \in \mathsf{p}\}$ and let

$$\epsilon = \lim_{n \to \mathsf{p}} \mathsf{d}\big((A - n) \cap (A - \mathsf{p})\big)/2.$$

Then the set $\{n \in \mathbb{N} : \mathsf{d}\big((A - n) \cap L\big) > \epsilon\}$ is in p and hence, for any finite set $F \subset L$, also the intersection

$$\bigcap_{\ell \in F} (A - \ell) \ \cap \ \Big\{m \in \mathbb{N} : \mathsf{d}_\Phi\big((A - m) \cap L\big) > \epsilon\Big\}$$

is in p. Since p is non-principal, this intersection can not be finite. The desired conclusion now follows from Proposition 2.5. $\square$

In view of Theorem 2.2, the proof of Theorem 1.2 now follows from the following theorem.

Theorem 2.6. *Let $A \subset \mathbb{N}$ and let Φ be a Følner sequence on $\mathbb{N}$ with $\mathsf{d}_\Phi(A)$ existing. For every $\epsilon > 0$ there exists a Følner subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\mathsf{d}_\Psi((A - m) \cap (A - \mathsf{p}))$ exists for all $m \in \mathbb{N}$ and*

$$\lim_{m \to \mathsf{p}} \mathsf{d}_\Psi\big((A - m) \cap (A - \mathsf{p})\big) \geq \mathsf{d}_\Psi(A)^2 - \epsilon \tag{13}$$

holds.

Proof of Theorem 1.2 assuming Theorem 2.6. Fix $A \subset \mathbb{N}$ with $\overline{\mathsf{d}}_\Phi(A) > 0$ for some Følner sequence Φ. By passing to a subsequence of Φ we may assume that $\mathsf{d}_\Phi(A)$ is defined and positive. Apply Theorem 2.6 with $\epsilon = \mathsf{d}_\Phi(A)^2/2$. Since $\mathsf{d}_\Phi(A) = \mathsf{d}_\Psi(A)$ for every further subsequence Ψ of Φ the inequality (13) implies the hypothesis (9) of Theorem 2.2, so A indeed contains $B + C$ for infinite sets $B, C \subset \mathbb{N}$. $\square$

We conclude this section by reformulating Theorem 2.6 in a functional analytic language as in (4). Given a bounded function $f\colon \mathbb{N}\to\mathbb{C}$ define, for all $m\in\mathbb{N}$, the shift $\mathsf{R}^m f\colon \mathbb{N}\to\mathbb{C}$ by

$$(\mathsf{R}^m f)(n) := f(n+m)$$

for all $n\in\mathbb{N}$. We extend this to all $\mathsf{p}\in\beta\mathbb{N}$ by defining the function $\mathsf{R}^{\mathsf{p}} f\colon \mathbb{N}\to\mathbb{C}$ by

$$(\mathsf{R}^{\mathsf{p}} f)(n) := \lim_{m\to\mathsf{p}} f(n+m)$$

for all $n\in\mathbb{N}$. Observe that $\mathsf{R}^{\mathsf{p}_m} f = \mathsf{R}^m f$ for all principal ultrafilters p_m. Also, the indicator function of the set $A-\mathsf{p}$ is the function $\mathsf{R}^{\mathsf{p}}1_A$, where 1_A is the indicator function of A.

Given a Følner sequence Φ in $\mathbb{N}$ and functions $f,h\colon \mathbb{N}\to\mathbb{C}$, define the **Besicovitch seminorm** of f along Φ to be

$$\|f\|_\Phi = \left(\limsup_{N\to\infty} \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}|f(n)|^2\right)^{1/2} \tag{14}$$

and the inner product

$$\langle f,\ h\rangle_\Phi = \lim_{N\to\infty} \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} f(n)\overline{h(n)}$$

whenever the limit exists. Minkowski's inequality

$$\left(\sum_{n\in\Phi_N}|f(n)+h(n)|^2\right)^{\frac12} \le \left(\sum_{n\in\Phi_N}|f(n)|^2\right)^{\frac12} + \left(\sum_{n\in\Phi_N}|h(n)|^2\right)^{\frac12} \tag{15}$$

implies that $\|f+h\|_\Phi \le \|f\|_\Phi + \|h\|_\Phi$, and hence $\|\cdot\|_\Phi$ is indeed a seminorm on the set of functions $f:\mathbb{N}\to\mathbb{C}$ for which $\|f\|$ is finite. The following facts will be used throughout the paper.

1. If Ψ eventually agrees with a subsequence of Φ then $\|f\|_\Psi \le \|f\|_\Phi$ for all $f:\mathbb{N}\to\mathbb{C}$;

2. (Cauchy-Schwarz) $|\langle f,\ h\rangle_\Phi| \le \|f\|_\Phi\|h\|_\Phi$ whenever $\langle f,\ h\rangle_\Phi$ exists and both $\|f\|_\Phi$, $\|h\|_\Phi$ are finite.

3. If $\|f\|_\Phi$ is finite then there is a subsequence Ψ of Φ such that $\|f\|_\Xi = \|f\|_\Phi$ for every subsequence Ξ of Ψ.

4. If $\|f\|_\Phi$ and $\|h\|_\Phi$ are both finite then there is a subsequence Ψ of Φ such that $\langle f,\ h\rangle_\Psi$ exists.

The following result, whose proof is given in Section 4 using the material of Section 3, implies Theorem 2.6 by choosing $f = 1_A$.

Theorem 2.7. *Let f be a non-negative bounded function on $\mathbb{N}$ and let Φ be a Følner sequence on $\mathbb{N}$ such that $\langle 1,\ f\rangle_\Phi$ exists. For every $\epsilon > 0$ there exists a subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\langle \mathsf{R}^m f,\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi$ exists for all $m \in \mathbb{N}$ and*

$$\lim_{m\to\mathsf{p}} \langle \mathsf{R}^m f,\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi \geq \langle 1,\ f\rangle_\Psi^2 - \epsilon \tag{16}$$

holds.

3. Two decompositions for functions in $\mathsf{L}^2(\mathbb{N},\Phi)$

In this section we establish several structural results about the space

$$\mathsf{L}^2(\mathbb{N},\Phi) := \{f\colon \mathbb{N}\to\mathbb{C} : \|f\|_\Phi < \infty\}$$

where $\|\cdot\|_\Phi$ is the seminorm defined in (14). In particular, we prove two ways in which elements of $\mathsf{L}^2(\mathbb{N},\Phi)$ can be decomposed into pseudo-random and structured components. These decomposition theorems will play crucial roles in the proof of Theorem 2.7.

Related decompositions of functions on $\mathbb{N}$ into orthogonal components have been studied in [HK09] and [Fra15]. However, those decompositions required some additional regularity on the function being decomposed and do not apply to all bounded functions on $\mathbb{N}$. Also, similar but more quantitative decompositions are known for complex-valued functions over finite intervals $\{1,\ldots,N\}$ (cf. [GT10]), but they don't possess qualitative (i.e. infinitary) analogues for functions over $\mathbb{N}$.

In Section 3.1 we prove a completeness result for the space $\mathsf{L}^2(\mathbb{N},\Phi)$. Then in Section 3.2 we introduce the space $\mathsf{Bes}(\mathbb{N},\Phi)$ of Besicovitch almost periodic functions along a Følner sequence Φ. Members of $\mathsf{Bes}(\mathbb{N},\Phi)$ play the role of the structured part in our first decomposition result, Theorem 3.6.

Our second splitting, of functions from $\mathsf{L}^2(\mathbb{N},\Phi)$ into compact and weak mixing functions, is based on the Jacobs–de Leeuw–Glicksberg splitting and is the topic of Section 3.3.

3.1. A completeness lemma for $\mathsf{L}^2(\mathbb{N},\Phi)$

Minkowski's inequality (15) implies that the space $\mathsf{L}^2(\mathbb{N},\Phi)$ is a vector space over $\mathbb{C}$. However $\mathsf{L}^2(\mathbb{N},\Phi)$ is not a Hilbert space. Indeed, $\|\cdot\|_\Phi$ is not a norm: the limit defining the inner product $\langle f,\ h\rangle_\Phi$ need not exist for all $f,h\in\mathsf{L}^2(\mathbb{N},\Phi)$, and the space $\mathsf{L}^2(\mathbb{N},\Phi)$ need not be complete with respect to $\|\cdot\|_\Phi$. To address the latter issue, we make use of the following proposition. We say that a sequence $j\mapsto f_j:\mathbb{N}\to\mathbb{C}$ of functions is **Cauchy** with respect to $\|\cdot\|_\Phi$ if, for all $\epsilon>0$, there exists $N\in\mathbb{N}$ such that for all $j,k\geq N$ one has $\|f_k-f_j\|_\Phi\leq\epsilon$.

Proposition 3.1. *Let $j\mapsto f_j$ be a sequence in $\mathsf{L}^2(\mathbb{N},\Phi)$ that is Cauchy with respect to $\|\cdot\|_\Phi$. Then there exists a subsequence Ψ of Φ and $f\in\mathsf{L}^2(\mathbb{N},\Psi)$ such that $\|f-f_j\|_\Psi\to 0$ as $j\to\infty$. Moreover, if all the f_j take values in an interval $[a,b]$, then so does f.*

Remark 3.2. If the Følner sequence Φ satisfies $\Phi_N\subset\Phi_{N+1}$ for all $N\in\mathbb{N}$, then one can adapt the proof of [BF45, II §2] to show that $\mathsf{L}^2(\mathbb{N},\Phi)$ is complete with respect to $\|\cdot\|_\Phi$, meaning that any sequence of functions in $\mathsf{L}^2(\mathbb{N},\Phi)$ that is Cauchy with respect to $\|\cdot\|_\Phi$ has a limit in $\mathsf{L}^2(\mathbb{N},\Phi)$. In particular, in this case it is not necessary to pass to a subsequence of Φ. We do not pursue this here for two reasons: on the one hand, the proof of Proposition 3.1 is much shorter. On the other hand, we find it necessary to pass to subsequences of Følner sequences frequently for many reasons, so we see no reason not to do so here as well.

Proof of Proposition 3.1. Since $j\mapsto f_j$ is Cauchy and all Besicovitch seminorms (14) satisfy the triangle inequality, it suffices to find a subsequence Ψ of Φ and a subsequence $j\mapsto f_{\sigma(j)}$ such that $\|f-f_{\sigma(j)}\|_\Psi\to 0$ as $j\to\infty$. To this end we assume, by passing to a subsequence if necessary, that for all $j\in\mathbb{N}$ and all $k\geq j$ we have $\|f_k-f_j\|_\Phi^2\leq\frac{1}{j}$. In particular, with $C:=(\|f_1\|_\Phi+1)^2$, the estimate $\|f_k\|_\Phi^2\leq C$ is valid for all $k\in\mathbb{N}$. Now, for every $k\in\mathbb{N}$, pick $N(k)\in\mathbb{N}$ such that $N(k+1)>N(k)$ for all $k\in\mathbb{N}$ and that, for all $N\geq N(k)$ and all $j\in\{1,\dots,k\}$, one has

$$\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}|f_j(n)-f_k(n)|^2\leq\frac{2}{j}\qquad\text{and}\qquad\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N}|f_j(n)|^2\leq 2C.$$

Also, by further refining the subsequence $k\mapsto N(k)$ if necessary, we can assume that

$$|\Phi_{N(k)}| > k^2 \max\left\{ \sum_{n\in\Phi_{N(i)}} \left|f_k(n) - f_i(n)\right|^2 : 1 \le i < k \right\}$$

for all $k > 1$. Define the Følner sequence Ψ by $\Psi_k \coloneqq \Phi_{N(k)}$ for all $k \in \mathbb{N}$.

Let $\Xi_M \coloneqq \Psi_M \setminus \left(\bigcup_{k=1}^{M-1} \Psi_k\right)$ and set $\zeta_M \coloneqq \Psi_M \setminus \Xi_M$, the latter being a subset of $\bigcup_{i=1}^{M-1} \Psi_i$. Define $f\colon \mathbb{N} \to \mathbb{C}$ by

$$f(n) \coloneqq \sum_{M=1}^{\infty} 1_{\Xi_M}(n) f_M(n) = \begin{cases} 0, & \text{if } n \notin \bigcup\limits_{K=1}^{\infty} \Psi_K \\ f_M(n), & \text{if } M = \min\{K \in \mathbb{N} : n \in \Psi_K\} \end{cases}$$

for all $n \in \mathbb{N}$. By construction, f takes values in an interval $[a, b]$ if all the functions f_M do. Using $|x+y|^2/2 \le |x|^2 + |y|^2$, for each $j \le M \in \mathbb{N}$ we have the estimate

$$\begin{aligned} \frac{1}{2} \sum_{n\in\Psi_M} |f_j(n) - f(n)|^2 &\le \sum_{n\in\Psi_M} |f_j(n) - f_M(n)|^2 + \sum_{n\in\zeta_M} |f_M(n) - f(n)|^2 \\ &\le \frac{2|\Psi_M|}{j} + \sum_{i=1}^{M-1} \sum_{n\in\Xi_i} |f_M(n) - f_i(n)|^2 \\ &\le \frac{2|\Psi_M|}{j} + \frac{|\Psi_M|}{M} \end{aligned}$$

which proves that $\|f - f_j\|_\Psi \le 4/j$, which tends to 0 as $j \to \infty$. □

We will also make use of the following version of Bessel's inequality.

Lemma 3.3 (Bessel's inequality)**.** *Let $u_1, u_2, \ldots$ be a sequence in $\mathsf{L}^2(\mathbb{N}, \Phi)$ such that $\|u_j\|_\Phi = 1$ for all $j \in \mathbb{N}$ and $\langle u_j,\ u_k\rangle_\Phi$ exists and is 0 for all $j \ne k$. If $u \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is such that $\langle u,\ u_j\rangle_\Phi$ exists for all $j \in \mathbb{N}$, then*

$$\sum_{j=1}^{\infty} \left|\langle u,\ u_j\rangle_\Phi\right|^2 \le \|u\|_\Phi^2$$

holds.

Proof. It suffices to show that

$$\sum_{j=1}^{J} \left|\langle u,\ u_j\rangle_\Phi\right|^2 \le \|u\|_\Phi^2 \tag{17}$$

for every $J \in \mathbb{N}$. Fix $N \in \mathbb{N}$ and write

$$[f,\ h]_N = \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} f(n)\, \overline{h(n)}$$

for all $f, h \colon \mathbb{N} \to \mathbb{C}$. Since $[f,\ f]_N \geq 0$ for all $f \colon \mathbb{N} \to \mathbb{C}$ we have

$$\begin{aligned} 0 &\leq \left[u - \sum_{j=1}^{J} u_j\, [u,\ u_j]_N,\ u - \sum_{k=1}^{J} u_k\, [u,\ u_k]_N \right]_N \\ &= [u,\ u]_N - 2 \sum_{j=1}^{J} \big|\, [u,\ u_j]_N \,\big|^2 + \sum_{j,k=1}^{J} [u,\ u_j]_N\, \overline{[u,\ u_k]_N}\, [u_j,\ u_k]_N\,. \end{aligned}$$

Whence

$$2 \sum_{j=1}^{J} \big|\, [u,\ u_j]_N \,\big|^2 \leq [u,\ u]_N + \sum_{j,k=1}^{J} [u,\ u_j]_N\, \overline{[u,\ u_k]_N}\, [u_j,\ u_k]_N \tag{18}$$

holds. Since the u_j are pairwise orthogonal,

$$\lim_{N \to \infty} \sum_{j,k=1}^{J} [u,\ u_j]_N\, \overline{[u,\ u_k]_N}\, [u_j,\ u_k]_N - \sum_{j=1}^{J} \big|\, [u,\ u_j]_N \,\big|^2 = 0.$$

Taking the limit $N \to \infty$ in (18) gives (17) as desired. □

3.2. A general splitting technique for $\mathsf{L}^2(\mathbb{N}, \Phi)$

Our first decomposition result involves a notion of almost periodicity introduced over $\mathbb{R}$ by Besicovitch in [Bes26]. We refer the reader to [Bes55; BL85] and the references therein for more on what have become known as Besicovitch almost periodic functions. Over $\mathbb{N}$ they are defined as follows.

Definition 3.4. By a **trigonometric polynomial** we mean any function $a \colon \mathbb{N} \to \mathbb{C}$ of the form

$$a(n) = \sum_{j=1}^{J} c_j e^{2\pi i \theta_j n} \tag{19}$$

for some $c_1, \dots, c_J \in \mathbb{C}$ and some **frequencies** $0 \leq \theta_1, \dots, \theta_J < 1$. A function $f \colon \mathbb{N} \to \mathbb{C}$ is **Besicovitch almost periodic** along Φ if, for every $\epsilon > 0$, one can find a trigonometric polynomial a with $\| f - a \|_\Phi < \epsilon$.

Write $\mathsf{Bes}(\mathbb{N}, \Phi)$ for the set of all Besicovitch almost periodic functions along Φ and notice that $\mathsf{Bes}(\mathbb{N}, \Phi) \subset \mathsf{L}^2(\mathbb{N}, \Phi)$. The notion of pseudo-randomness complementary to Besicovitch almost periodicity is defined next.

Definition 3.5. The set $\mathsf{Bes}(\mathbb{N},\Phi)^\perp$ is defined to consist of those functions $f \in \mathsf{L}^2(\mathbb{N},\Phi)$ such that

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} f(n)e^{2\pi i n\theta} = 0$$

for all frequencies $\theta \in [0,1)$.

One can show directly from the definitions that $\langle f,\ h\rangle_\Phi = 0$ whenever $f \in \mathsf{Bes}(\mathbb{N},\Phi)$ and $h \in \mathsf{Bes}(\mathbb{N},\Phi)^\perp$. Our main focus is the following splitting result. Throughout this paper we will use f_{anti} to denote elements in $\mathsf{Bes}(\mathbb{N},\Phi)^\perp$.

Theorem 3.6. *For every Følner sequence Φ on $\mathbb{N}$ and any $f \in \mathsf{L}^2(\mathbb{N},\Phi)$ there is a subsequence Ψ of Φ and functions $f_{\mathsf{Bes}} \in \mathsf{Bes}(\mathbb{N},\Psi)$ and $f_{\mathsf{anti}} \in \mathsf{Bes}(\mathbb{N},\Psi)^\perp$ such that $f = f_{\mathsf{Bes}} + f_{\mathsf{anti}}$. Moreover, f_{Bes} minimizes the distance between f and $\mathsf{Bes}(\mathbb{N},\Psi)$ in the sense that $\|f - f_{\mathsf{Bes}}\|_\Psi = \inf\{\|f-g\|_\Psi : g \in \mathsf{Bes}(\mathbb{N},\Psi)\}$, and if f takes values in an interval $[a,b]$, then so does f_{Bes}.*

Proof. Combine Theorem 3.8 and Theorem 3.9 below. □

Instead of directly proving Theorem 3.6, we establish a general framework for decomposition results in $\mathsf{L}^2(\mathbb{N},\Phi)$ that will in particular imply Theorem 3.6. In fact, Theorem 3.6 follows immediately from combining Theorem 3.8 and Theorem 3.9 below.

Suppose that for every Følner sequence Φ we are given a $U(\Phi)$ of $\mathsf{L}^2(\mathbb{N},\Phi)$ satisfying the following properties:

- $U(\Phi)$ is a vector subspace of $\mathsf{L}^2(\mathbb{N},\Phi)$;
- $U(\Phi)$ contains the constant functions and is closed under pointwise complex conjugation;
- for all $u, v \in U(\Phi)$ the inner product $\langle u,\ v\rangle_\Phi$ exists;
- If $u, v \in U(\Phi)$ are real valued, then the function $n \mapsto \max\{u(n), v(n)\}$ is in $U(\Phi)$;
- $U(\Phi)$ is closed with respect to the topology on $\mathsf{L}^2(\mathbb{N},\Phi)$ induced by the semi-norm $\|\cdot\|_\Phi$;
- if Ψ eventually agrees with a subsequence of Φ then $U(\Psi) \supset U(\Phi)$.

Call any such assignment U of subspaces to Følner sequences a **projection family**. Given a projection family one can consider, for each Følner sequence Φ, the subspace

$$U(\Phi)^\perp := \{v \in \mathsf{L}^2(\mathbb{N}, \Phi) : \langle u,\ v\rangle_\Phi \text{ exists and equals 0 for all } u \in U(\Phi)\}$$

of $\mathsf{L}^2(\mathbb{N}, \Phi)$. With a view towards proving Theorem 3.6 we first verify that $\Phi \mapsto \mathsf{Bes}(\mathbb{N}, \Phi)$ is a projection family. The following fact can be viewed as the one-dimensional case of the von Neumann ergodic theorem; we provide a short proof for the sake of completeness.

Lemma 3.7. *Let $\theta \in (0, 1)$ and let Φ be a Følner sequence. Then*

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} e^{2\pi i n\theta} = 0. \tag{20}$$

In particular $\langle a,\ b\rangle_\Phi$ exists for all trigonometric polynomials a and b.

Proof. Let $N \in \mathbb{N}$ be large and let $\epsilon_N = \big|(\Phi_N + 1)\triangle\Phi_N\big|/|\Phi_N|$, $A_N := \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} e^{2\pi i n\theta}$ and $B_N := \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N+1} e^{2\pi i n\theta}$. On the one hand $|A_N - B_N| \le \epsilon_N$ but on the other hand $B_N = e^{2\pi i\theta}A_N$, which implies that $|A_N| < \epsilon_N/|1 - e^{2\pi i\theta}|$. Since $\epsilon_N \to 0$ we conclude that $A_N \to 0$ as desired.

Now, if a and b are trigonometric polynomials then so is $n \mapsto a(n)\overline{b(n)}$ and the limit $\langle a,\ b\rangle_\Phi$ exists as it is a linear combination of constants and of limits of the form (20). □

Theorem 3.8. *The assignment $\Phi \mapsto \mathsf{Bes}(\mathbb{N}, \Phi)$ is a projection family.*

Proof. It follows from the triangle inequality that $\mathsf{Bes}(\mathbb{N}, \Phi)$ is a subspace of $\mathsf{L}^2(\mathbb{N}, \Phi)$. Since constant functions are trigonometric polynomials, and since the complex conjugation of a trigonometric polynomial remains such, it is immediate that $\mathsf{Bes}(\mathbb{N}, \Phi)$ contains the constant functions and is closed under pointwise complex conjugation.

The fact that the space $\mathsf{Bes}(\mathbb{N}, \Phi)$ is closed with respect to $\|\cdot\|_\Phi$ is an immediate consequence of the definition of $\mathsf{Bes}(\mathbb{N}, \Phi)$ as the closure in $\mathsf{L}^2(\mathbb{N}, \Phi)$ of the space of trigonometric polynomials with respect to $\|\cdot\|_\Phi$.

Fix now u, v in $\mathsf{Bes}(\mathbb{N}, \Phi)$ both real-valued. From the relation

$$\max\{u, v\} = \frac{1}{2}(u + v + |u - v|)$$

and linearity, the fact that $\max\{u, v\}$ belongs to $\mathsf{Bes}(\mathbb{N}, \Phi)$ would follow from the knowledge that $|w|$ belongs to $\mathsf{Bes}(\mathbb{N}, \Phi)$ whenever w does. That knowledge is the content of [Bes55, Lemma 5° in Chapter II, §5]; see also [Boh25a; Boh25b]. We give here a proof for completeness. Fix $w \in \mathsf{Bes}(\Phi, \mathbb{N})$ and $\epsilon > 0$. Let a be a trigonometric polynomial with $\|u - a\|_\Phi < \epsilon/2$. The reverse triangle inequality gives $\||u| - |a|\|_\Phi < \epsilon/2$. Apply the Stone-Weierstrass theorem to find a polynomial $b \in \mathbb{C}[z]$ with $|b(z) - |z|| < \epsilon/2$ for all $z \leq \sup\{|a(n)| : n \in \mathbb{N}\}$. (This is possible because trigonometric polynomials have bounded range.) The trigonometric polynomial $n \mapsto b(a(n))$ is then within ϵ of $|u|$ with respect to the $\| \cdot \|_\Phi$ semi-norm.

Next, we prove that $\langle u,\ v\rangle_\Phi$ exists for any $u, v \in \mathsf{Bes}(\mathbb{N}, \Phi)$. For this we use Lemma 3.7 and the inequality

$$\limsup_{N\to\infty} \left| \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} \Big(u(n) - w(n)\Big)\overline{v(n)} \right| \\ \leq \|u - w\|_\Phi \sup\Big\{ \Big(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |v(n)|^2 \Big)^{1/2} : N \in \mathbb{N} \Big\}$$

which is true for all $u, v, w \in \mathsf{L}^2(\mathbb{N}, \Phi)$ and implies continuity of $\langle \cdot,\ \cdot\rangle_\Phi$ in the first variable. Fix $u \in \mathsf{Bes}(\Phi, \mathbb{N})$ and a trigonometric polynomial a. Fix a sequence $n \mapsto b_n$ of trigonometric polynomials converging to u with respect to $\| \cdot \|_\Phi$. The sequence $n \mapsto b_n$ is Cauchy for $\| \cdot \|_\Phi$ so $n \mapsto \langle b_n,\ a\rangle_\Phi$ is Cauchy by the Cauchy-Schwarz inequality. Denote by α its limit. The above inequality implies that $\langle u,\ a\rangle_\Phi = \alpha$.

A similar inequality gives continuity of the form $\langle \cdot,\ \cdot\rangle_\Phi$ in the second variable, and the above argument can be repeated to prove that if $u, v \in \mathsf{Bes}(\mathbb{N}, \Phi)$ and c_n are trigonometric polynomials converging to v with respect to $\| \cdot \|_\Phi$ then $\langle u,\ v\rangle_\Phi$ is the limit of the Cauchy sequence $n \mapsto \langle u,\ c_n\rangle_\Phi$.

Lastly, since $\|f\|_\Psi \leq \|f\|_\Phi$ for all $f : \mathbb{N} \to \mathbb{C}$ whenever Ψ eventually agrees with a subsequence of Φ, it is immediate that $\mathsf{Bes}(\Psi, \mathbb{N}) \supset \mathsf{Bes}(\Phi, \mathbb{N})$ whenever Ψ eventually agrees with a subsequence of Φ. ☐

In view of Theorem 3.8, the following general decomposition result extends Theorem 3.6.

Theorem 3.9. *Let U be a projection family and let Φ be a Følner sequence. For every $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ there exists a subsequence Ψ of Φ and there is $f_U \in U(\Psi)$ such that:*

1. $f - f_U \in U(\Psi)^\perp$,
2. *f_U minimizes the distance between f and $U(\Psi)$ in the sense that* $\|f - f_U\|_\Psi = \inf\{\|f - g\|_\Psi : g \in U(\Psi)\}$,

3. if f takes values in an interval $[a,b]$ then f_U takes values in $[a,b]$.

Theorem 3.9 would be immediate if $\mathsf{L}^2(\mathbb{N},\Phi)$ were a Hilbert space and $U(\Phi)$ were a closed subspace, because then one could simply define f_U as the orthogonal projection of f onto $U(\Phi)$. However, $\mathsf{L}^2(\mathbb{N},\Phi)$ is not a Hilbert space, which requires us to overcome some difficulties. In particular, it is problematic that $\langle f,\ u\rangle_\Phi$ may not exist for all $u \in U(\Phi)$. The following technical lemma offers a way around this issue.

Lemma 3.10. *Let U be a projection family and let Φ be a Følner sequence. For every $f \in \mathsf{L}^2(\mathbb{N},\Phi)$ there exists a subsequence Ψ of Φ such that the inner product $\langle f,\ u\rangle_\Psi$ exists whenever $u \in U(\Psi)$.*

Proof. Fix $f \in \mathsf{L}^2(\mathbb{N},\Phi)$. We start with an inductive construction. Put $u_0 \coloneqq 0$ and $\Phi^{(0)} \coloneqq \Phi$. Certainly $u_0 \in U(\Phi^{(0)})$ and $\langle f,\ u_0\rangle_{\Phi^{(0)}}$ exists. Suppose for some $k \in \mathbb{N}$ that we have defined functions $u_0,\dots,u_{k-1} \in U(\Phi^{(k-1)})$ and a Følner sequences $\Phi^{(0)},\dots,\Phi^{(k-1)}$, each a subsequence of the previous one, such that $\langle f,\ u_i\rangle_{\Phi^{(k-1)}}$ exists for all $0 \le i \le k-1$. For each Følner subsequence Φ' of $\Phi^{(k-1)}$, let

$$O_{k-1}(\Phi') \coloneqq \big\{u \in U(\Phi') : \langle u,\ u_i\rangle_{\Phi'} = 0,\ \forall i \in \{0,\dots,k-1\}\big\}$$

which is a linear subspace of $U(\Phi')$ that contains the constant functions.

We now distinguish two cases depending on whether or not there are a subsequence Φ' of $\Phi^{(k-1)}$ and a member u of $O_{k-1}(\Phi')$ with $\|u\|_{\Phi'} \neq 0$.

In the first case we assume, for every subsequence Φ' of $\Phi^{(k-1)}$, that every $u \in U(\Phi')$ satisfying $\langle u,\ u_i\rangle_{\Phi'}$ for all $0 \le i \le k-1$ has the property $\|u\|_{\Phi'} = 0$. If this happens we terminate our inductive construction, the result being a Følner sequence $\Phi^{(k-1)}$ and a collection $u_0,\dots,u_{k-1}$ of members of $U(\Phi^{(k-1)})$ such that $\langle f,\ u_i\rangle_{\Phi^{(k-1)}}$ exists for all $0 \le i \le k-1$.

We claim in this first case that the conclusion of the lemma is true with $\Psi = \Phi^{(k-1)}$. Fix $u \in U(\Psi)$. The function

$$v = u - \sum_{i=0}^{k-1} u_i \langle u,\ u_i\rangle_\Psi$$

belongs to $O_{k-1}(\Psi)$ and therefore has a $\|\cdot\|_\Psi$ norm of zero. It follows that

$$\begin{aligned}&\frac{1}{|\Psi_N|}\sum_{n\in\Psi_N} f(n)\overline{u(n)}\\ &= \frac{1}{|\Psi_N|}\sum_{n\in\Psi_N} f(n)\overline{v(n)} + \sum_{i=0}^{k-1}\frac{1}{|\Psi_N|}\sum_{n\in\Psi_N} f(n)\overline{u_i(n)}\langle u,\ u_i\rangle_\Psi\end{aligned}$$

converges as $N \to \infty$ as desired.

In the second case we assume there is a subsequence Φ' of $\Phi^{(k-1)}$ and a member u of $O_{k-1}(\Phi')$ with $\|u\|_{\Phi'} \neq 0$. If this happens then the set

$$Q_k \coloneqq \left\{ |\langle f,\ u\rangle_{\Phi'}| : \begin{array}{l} \Phi' \text{ is a Følner subsequence of } \Phi^{(k-1)} \\ u \in O_{k-1}(\Phi') \text{ with } \|u\|_{\Phi'} = 1 \\ \langle f,\ u\rangle_{\Phi'} \text{ exists} \end{array} \right\}$$

is non-empty. Indeed if, for some subsequence Φ' of $\Phi^{(k-1)}$, one can find some member u of $O_{k-1}(\Phi')$ with $\|u\|_{\Phi'} \neq 0$, note that $u/\|u\|$ belongs to $O_{k-1}(\Xi)$ for every subsequence Ξ of Φ' and that $\langle f,\ u\rangle_\Xi$ will exist for a suitable choice of Ξ.

Write δ_k for the supremum of Q_k, which will be at most $\|f\|_\Phi$ by Cauchy-Schwarz. Choose a Følner subsequence $\Phi^{(k)}$ of $\Phi^{(k-1)}$ and $u_k \in O_{k-1}(\Phi^{(k)})$ with $\|u_k\|_{\Phi^{(k)}} = 1$ such that $\langle f,\ u_k\rangle_{\Phi^{(k)}}$ exists and $|\langle f,\ u_k\rangle_{\Phi^{(k)}}| > \delta_k - \frac{1}{k}$. Then $\langle f,\ u_i\rangle_{\Phi^{(k)}}$ exists for all $0 \le i \le k$.

This concludes the consideration of the second case, and the inductive construction. If, at any stage, we find ourselves in the first case discussed above then the proof is complete. We therefore find ourselves with a sequence $u_0, u_1, \dots$ of functions, a sequence $\Phi^{(0)}, \Phi^{(1)}, \dots$ of Følner sequences, and a sequence $\delta_1, \delta_2, \dots$ of suprema, as described in the second case.

Define $\Psi_N \coloneqq \Phi^{(N)}_N$. The sequence Ψ is a subsequence of $\Phi^{(1)}$ and is therefore itself a Følner sequence. We claim that for every $u \in U(\Psi)$ the inner product $\langle f,\ u\rangle_\Psi$ exists. More precisely, we claim that

$$\langle f,\ u\rangle_\Psi = \sum_{i=1}^{\infty} \langle f,\ u_i\rangle_\Psi \overline{\langle u,\ u_i\rangle_\Psi}.$$

Note that the terms in the above series are well defined, since $\langle u,\ u_i\rangle_\Psi$ exists because $u, u_i \in U(\Psi)$ and $\langle f,\ u_i\rangle_\Psi$ exists by construction of Ψ. Moreover, this series is absolutely convergent, because Lemma 3.3 implies that the sequences $i \mapsto \langle f,\ u_i\rangle_\Psi$ and $i \mapsto \langle u,\ u_i\rangle_\Psi$ are in $\ell^2(\mathbb{N})$.

For each $k \in \mathbb{N}$, define

$$v_k \coloneqq u - \sum_{i=1}^{k-1} u_i \langle u,\ u_i\rangle_\Psi$$

and observe that $v_k \in O_{k-1}(\Psi)$ and that $\|v_k\|_\Psi \le \|u\|_\Psi$. Therefore

$$\limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} f(n)\overline{u(n)} - \sum_{i=1}^{\infty} \langle f,\ u_i\rangle_\Psi \overline{\langle u,\ u_i\rangle_\Psi} \right|$$

$$
\begin{aligned}
&\le \limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} f(n)\overline{v_k(n)} \right| + \left| \sum_{i=k}^{\infty} \langle f,\ u_i\rangle_\Psi \overline{\langle u,\ u_i\rangle_\Psi} \right| \\
&\le \delta_k \|v_k\|_\Psi + \left| \sum_{i=k}^{\infty} \langle f,\ u_i\rangle_\Psi \overline{\langle u,\ u_i\rangle_\Psi} \right|.
\end{aligned}
$$

It thus suffices to show that $\delta_k \to 0$ as $k \to \infty$. But by Lemma 3.3, we get

$$
\|f\|_\Psi^2 \ge \sum_{k=1}^{\infty} |\langle f,\ u_k\rangle_\Psi|^2 \ge \sum_{k=1}^{\infty} \left(\delta_k - \tfrac{1}{k}\right)^2
$$

and since $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$, the series converges, which implies that indeed $\delta_k \to 0$ as $k \to \infty$. □

Proof of Theorem 3.9. As guaranteed by Lemma 3.10, let Ψ be a Følner subsequence of Φ such that for every $u \in U(\Psi)$ the limit $\langle f,\ u\rangle_\Psi$ exists. Define

$$
\delta \coloneqq \inf\left\{ \|f - u\|_\Psi^2 : u \in U(\Psi) \right\}.
$$

For each $k \in \mathbb{N}$ choose $u_k \in U(\Psi)$ with $\|f - u_k\|_\Psi^2 < \delta + \frac{1}{k}$.

If f takes values in $[a, b]$, then we can replace u_k with the function

$$
v_k : n \mapsto \begin{cases} a, & \text{if } \Re u_k(n) < a, \\ \Re u_k(n), & \text{if } a \le \Re u_k(n) \le b, \\ b, & \text{if } \Re u_k(n) > b, \end{cases}
$$

where $\Re z$ denotes the real part of a complex number z. Indeed, it is clear that $\|f - v_k\|_\Psi^2 \le \|f - u_k\|_\Psi^2 < \delta + \frac{1}{k}$. On the other hand, since $U(\Psi)$ contains constants, is closed under pointwise complex conjugation, and under taking the pointwise maximum, and therefore also under taking the pointwise minimum, the function v_k still belongs to $U(\Psi)$. Therefore we can assume without loss of generality that when f takes values in $[a, b]$, then so do the functions u_k.

Next, an application of the parallelogram law to the vectors $f - u_j$ and $f - u_k$ shows that $\|u_j - u_k\|_\Psi^2 \le \frac{2}{j} + \frac{2}{k}$, which implies that $(u_k)_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_\Psi$. Using Proposition 3.1 and by refining Ψ if necessary, we can find $f_U \in \mathsf{L}^2(\mathbb{N}, \Psi)$ such that $\lim_{k\to\infty} \|f_U - u_k\|_\Psi = 0$. If f takes values in $[a, b]$ (and hence so do all the u_k), then f_U also takes values in $[a, b]$. Since $U(\Psi)$ is closed, it follows that f_U belongs to $U(\Psi)$. Minkowski's inequality implies that $\|f - f_U\|_\Psi^2 = \delta$. In particular, f_U minimizes the distance between f and $U(\Psi)$.

Write $h \coloneqq f - f_U$. We claim that h belongs to $U(\Psi)^\perp$. First note that $\langle h,\ u\rangle_\Psi$ exists for all $u \in U(\Psi)$ because both $\langle f,\ u\rangle_\Psi$ and $\langle f_U,\ u\rangle_\Psi$ exist. Next, fix $u \in U(\Psi)$ with $\|u\|_\Psi \leq 1$ and define $I \coloneqq \langle h,\ u\rangle_\Psi$. We have

$$\begin{aligned}
&\|h - Iu\|_\Psi^2 \\
=&\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} |h(n)|^2 - h(n)\overline{Iu(n)} - \overline{h(n)}Iu(n) + |I|^2|u(n)|^2 \\
\leq&\|h\|_\Psi^2 - |I|^2(2 - \|u\|_\Psi^2).
\end{aligned}$$

Since $\|u\|_\Psi^2 \leq 1$ and $\|h\|_\Psi^2 = \delta$, we conclude that $\|h\|_\Psi^2 - |I|^2(2 - \|u\|_\Psi^2) \leq \delta - |I|^2$. Therefore

$$\|h - Iu\|_\Psi^2 \leq \delta - |I|^2. \tag{21}$$

On the other hand, $h - Iu = f - (f_U + Iu)$ and $f_U + Iu \in U(\Psi)$. So

$$\|h - Iu\|_\Psi^2 \geq \delta. \tag{22}$$

Combining (21) and (22) proves that $I = 0$. □

Remark 3.11. Under the assumptions of Theorem 3.9, the function $f_U \in U(\Psi)$ is unique in the following two senses:

(a) If $f'_U \in U(\Psi)$ is such that $f - f'_U \in U(\Psi)^\perp$ then $\|f_U - f'_U\|_\Psi = 0$.

(b) If $f'_U \in U(\Psi)$ also minimizes the distance between f and $U(\Psi)$ (i.e. $\|f - f'_U\|_\Psi = \inf\{\|f - g\|_\Psi : g \in U(\Psi)\}$), then $\|f_U - f'_U\|_\Psi = 0$.

In the second half of the proof of Theorem 3.9 we show that a function $f'_U \in U(\Psi)$ that minimizes the distance between f and $U(\Psi)$ must satisfy $f - f'_U \in U(\Psi)^\perp$; therefore part (b) follows from part (a).

To verify part (a), note that $f - f_U, f - f'_U \in U(\Psi)^\perp$ implies that $f_U - f'_U \in U(\Psi)^\perp$, while $f_U, f'_U \in U(\Psi)$ implies that $f_U - f'_U$, and therefore $\|f_U - f'_U\|^2 = \langle f_U - f'_U,\ f_U - f'_U\rangle = 0$.

We conclude this subsection with a small detour on the further applicability of Theorem 3.9; this remarks are unrelated to the proof of Theorem 1.2.

By a **nilsystem** we mean a pair $(G/\Gamma, g)$ where G is a nilpotent Lie group, Γ is a discrete, co-compact subgroup of G, and $g \in G$ acts on G/Γ by left multiplication. A function $\alpha\colon \mathbb{N} \to \mathbb{C}$ is a **basic nilsequence** if there exists a nilsystem $(G/\Gamma, g)$ and a continuous function $F\colon G/\Gamma \to \mathbb{C}$ such that $\alpha(n) = F(g^n\Gamma)$. Call a function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ a **Besicovitch nilsequence** along Φ if for every $\epsilon > 0$ there exists a basic nilsequence $\alpha\colon \mathbb{N} \to \mathbb{C}$ such that $\|f - \alpha\|_\Phi < \epsilon$.

Denote by $U(\Phi)$ the family of all Besicovitch nilsequences with respect to Φ. Since the Cesàro average of a basic nilsequence along any Følner sequence exists (cf. [Lei05]) one can easily adapt the proof of Theorem 3.8 to show that the assignment $\Phi \mapsto U(\Phi)$ is a projection family.

A function $f\colon \mathbb{N} \to \mathbb{C}$ is a **good weight** for the polynomial multiple ergodic theorem if, for every probability space $(X, \mathcal{B}, \mu)$ and any commuting, measure-preserving transformations $T_1, \ldots, T_k\colon X \to X$ the quantity

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} f_{\mathsf{w}}(n)\, T_1^{p_1(n)} h_1 \cdots T_k^{p_k(n)} h_k \,\mathsf{d}\mu$$

exists and equals

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} T_1^{p_1(n)} h_1 \cdots T_k^{p_k(n)} h_k \,\mathsf{d}\mu$$

for any polynomials $p_1, \ldots, p_k \in \mathbb{Z}[x]$ and any $h_1, \ldots, h_k \in \mathsf{L}^\infty(X, \mathcal{B}, \mu)$.

Combining the fact that $U(\Phi)$ is a projection family with Theorem 3.9 and [Fra15, Theorem 1.2] we deduce the following result.

Theorem 3.12. *Let Φ be a Følner sequence on $\mathbb{N}$ and let $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$. Then there exists a subsequence Ψ of Φ and a decomposition $f = f_{\mathsf{nil}} + f_{\mathsf{w}}$ such that f_{nil} is a Besicovitch nilsequence with respect to Ψ and f_{w} is a good weight for the polynomial multiple ergodic theorem.*

3.3. A version of the Jacobs–de Leeuw–Glicksberg splitting for $\mathsf{L}^2(\mathbb{N}, \Phi)$

The second decomposition theorem that we use in the proof of Theorem 2.7, which represents 1_A as a sum of a weak mixing function and a compact function, can be viewed as a discrete version of the Jacobs–de Leeuw–Glicksberg splitting on Hilbert spaces. After recalling this splitting and introducing versions of weak mixing and compactness for functions in $\mathsf{L}^2(\mathbb{N}, \Phi)$ we prove the main result of this section, Theorem 3.22.

Fix an isometry U on a Hilbert space $(\mathscr{H}, \|\cdot\|_{\mathscr{H}})$.

Definition 3.13. An element $x \in \mathscr{H}$ is **compact** if $\{\mathsf{U}^n x : n \in \mathbb{N}\}$ is a pre-compact subset of $(\mathscr{H}, \|\cdot\|_{\mathscr{H}})$. Equivalently, x is compact if for all $\epsilon > 0$ there exists $K \in \mathbb{N}$ such that

$$\min\{\|\mathsf{U}^m x - \mathsf{U}^k x\|_{\mathscr{H}} : 1 \le k \le K\} \le \epsilon$$

for all $m \in \mathbb{N}$.

Definition 3.14. An element $x \in \mathscr{H}$ is called **weak mixing** if for all $\epsilon > 0$ and all $y \in \mathscr{H}$ the set $\{n \in \mathbb{N} : |\langle \mathsf{U}^n x,\ y\rangle| \geq \epsilon\}$ has zero density with respect to every Følner sequence on $\mathbb{N}$.

The set of all compact elements in $\mathscr{H}$, denoted $\mathscr{H}_{\mathsf{c}}$, is a closed and U invariant subspace of $\mathscr{H}$, as is the set $\mathscr{H}_{\mathsf{wm}}$ of weak mixing elements. The principle that $\mathscr{H}$ splits into the direct sum of $\mathscr{H}_{\mathsf{c}}$ and $\mathscr{H}_{\mathsf{wm}}$ traces back as far as the works of Koopman and von Neumann [KN32] (see also [Ber96, Theorem 2.3]) and was later pushed to greater generality by work of Jacobs [Jac56] and de Leeuw, Glicksberg [LG61] (see also [Kre85, Chapter 2.4] and [Eis+15, Example 16.25]).

Theorem 3.15 (The Jacobs-de Leeuw-Glicksberg splitting)**.** *Let U be an isometry on a Hilbert space $\mathscr{H}$. Then $\mathscr{H}_{\mathsf{c}}$ and $\mathscr{H}_{\mathsf{wm}}$ are orthogonal spaces and $\mathscr{H} = \mathscr{H}_{\mathsf{c}} \oplus \mathscr{H}_{\mathsf{wm}}$. In particular, for any $x \in \mathscr{H}$ there exist $x_{\mathsf{c}} \in \mathscr{H}_{\mathsf{c}}$ and $x_{\mathsf{wm}} \in \mathscr{H}_{\mathsf{wm}}$ such that $x = x_{\mathsf{c}} + x_{\mathsf{wm}}$.*

Let us introduce now the analogous notions of compact and weak mixing for elements in $\mathsf{L}^2(\mathbb{N}, \Phi)$. Recall that, given $f \colon \mathbb{N} \to \mathbb{C}$, we write $\mathsf{R}^m f$ for the function $n \mapsto f(m+n)$. One should think of R^1 acting on $\mathsf{L}^2(\mathbb{N}, \Phi)$ as playing the role of the isometry U on $\mathscr{H}$ in Theorem 3.15.

Definition 3.16. A function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is **compact** along Φ if, for every $\epsilon > 0$, one can find $K \in \mathbb{N}$ such that

$$\min\{\|\mathsf{R}^m f - \mathsf{R}^k f\|_{\Phi} : 1 \leq k \leq K\} < \epsilon$$

for all $m \in \mathbb{N}$.

Observe that any trigonometric polynomial is compact along any Φ. Since compact functions form a closed subset of $\mathsf{L}^2(\mathbb{N}, \Phi)$, every $f \in \mathsf{Bes}(\mathbb{N}, \Phi)$ is compact along Φ. We remark that one can show the set of functions compact along Φ is in fact a subspace of $\mathsf{L}^2(\mathbb{N}, \Phi)$.

Definition 3.17. A function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is **weak mixing** along Φ if, for every bounded function $h \colon \mathbb{N} \to \mathbb{C}$ and every subsequence Ψ of Φ such that $\langle \mathsf{R}^n f,\ h\rangle_{\Psi}$ exists for all $n \in \mathbb{N}$, one has

$$\overline{\mathsf{d}}_{\Psi}\Big(\Big\{n \in \mathbb{N} : |\langle \mathsf{R}^n f,\ h\rangle_{\Psi}| > \epsilon\Big\}\Big) = 0$$

for all $\epsilon > 0$.

Lemma 3.18. *If $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is weak mixing along Φ then*

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} |\langle \mathsf{R}^n f,\ h\rangle_{\Psi}| = 0$$

for all subsequences Ψ of Φ and all $h \in \mathsf{L}^2(\mathbb{N}, \Psi)$ such that $\langle \mathsf{R}^n f,\ h \rangle_\Psi$ exists for all $n \in \mathbb{N}$.

Proof. Fix $f \in \mathsf{L}^2(\mathbb{N}, \mathbb{C})$ that is weak mixing along Φ. Fix also a subsequence Ψ of Φ and $h \in \mathsf{L}^2(\mathbb{N}, \Psi)$ such that $\langle \mathsf{R}^n f,\ h \rangle_\Psi$ exists for all $n \in \mathbb{N}$. The sequence $a(n) = \langle \mathsf{R}^n f,\ h \rangle_\Psi$ is bounded. The implication $(ii) \Rightarrow (i)$ of [Wal82, Theorem 1.20] and its proof are valid for averages along any Følner sequence. But (ii) therein follows from our hypothesis on f. □

Lemma 3.19. *Let Φ be a Følner sequence and let $f, h \in \mathsf{L}^2(\mathbb{N}, \Phi)$ be compact and weak mixing along Φ, respectively. Then $\langle f,\ h \rangle_\Phi = 0$.*

Proof. If $\|f\|_\Phi = 0$ or $\|h\|_\Phi = 0$ then the result follows from Cauchy-Schwarz. Otherwise, choose a subsequence Ψ of Φ such that $\langle f,\ h \rangle_\Psi$ exists. Passing to a further subsequence if needed, we will also assume that all the inner products $\langle \mathsf{R}^n f,\ \mathsf{R}^m h \rangle_\Psi$ exist. After scaling if needed, we will further assume that $\|f\|_\Psi = \|h\|_\Psi = 1$.

Fix $\epsilon > 0$ and choose K so that for every $m \in \mathbb{N}$, there is some $1 \le k \le K$ with $\|\mathsf{R}^m f - \mathsf{R}^k f\|_\Phi < \epsilon$. Therefore

$$|\langle f,\ h \rangle_\Psi| = |\langle \mathsf{R}^m f,\ \mathsf{R}^m h \rangle_\Psi| \le \epsilon + |\langle \mathsf{R}^k f,\ \mathsf{R}^m h \rangle_\Psi| \le \epsilon + \sum_{k=1}^{K} |\langle \mathsf{R}^k f,\ \mathsf{R}^m h \rangle_\Psi|$$

holds. Since h is weak mixing, we conclude that

$$|\langle f,\ h \rangle_\Psi| \le \epsilon + \sum_{k=1}^{K} \limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{m \in \Psi_N} |\langle \mathsf{R}^k f,\ \mathsf{R}^m h \rangle_\Psi| = \epsilon$$

via Lemma 3.18. Since ϵ was arbitrary, we obtain $\langle f,\ h \rangle_\Psi = 0$. Since we chose Ψ as an arbitrary subsequence of Φ for which all $\langle \mathsf{R}^n f,\ \mathsf{R}^m h \rangle_\Psi$ exist, it follows that $\langle f,\ h \rangle_\Phi = 0$. □

Any Besicovitch almost periodic function is compact and therefore, if h is weak mixing along Φ, then $\langle h,\ f \rangle_\Phi = 0$ for all $f \in \mathsf{Bes}(\mathbb{N}, \Phi)$ and hence $h \in \mathsf{Bes}(\mathbb{N}, \Phi)^\perp$.

Remark 3.20. The condition of a function f being weak mixing is very similar to the condition that the Host–Kra local seminorm $\|f\|_{\Phi,2}$ of f equals 0 in the sense of [HK09, Definition 2.3]. We stress that this is weaker than the uniformity seminorm $\|f\|_{U(2)}$ of f equaling 0 in the sense of [HK09, Definition 2.6]. In fact, [HK09, Corollary 2.18] implies that $\|f\|_{U(2)} = 0$ is equivalent to $f \in \mathsf{Bes}(\mathbb{N}, \Phi)^\perp$ for every Følner sequence Φ.

As the following example shows (see also the example in [HK09, Section 2.4.3]) there are functions in $\mathsf{Bes}(\mathbb{N},\Phi)^\perp$ which are compact.

Example 3.21. We will now construct a bounded function $f\colon \mathbb{N}\to\mathbb{C}$ and a Følner sequence Φ such that f is simultaneously compact along Φ and a member of $\mathsf{Bes}(\mathbb{N},\Phi)^\perp$. Let $k\mapsto N_k$ be an increasing sequence of natural numbers with $N_{k-1}/N_k\to 0$ as $k\to\infty$. Assume f has already been defined on the interval $[1,N_k)$. Then we define f on the interval $[N_k,N_{k+1})$ by

$$f(n) \coloneqq \begin{cases} (-1)^n, & \text{if } n\in \left[N_k, \lfloor \frac{N_{k+1}}{2}\rfloor\right) \\ \\ -(-1)^n, & \text{if } n\in \left[\lfloor \frac{N_{k+1}}{2}\rfloor, N_{k+1}\right) \end{cases}$$

for all $N_k\le n<N_{k+1}$. Also, let Φ denote the Følner sequence given by $\Phi_k\coloneqq[1,N_k]$ for all $k\in\mathbb{N}$. It is then easy to verify that $\|T^2f-f\|_\Phi=0$ and hence f is compact with respect to Φ. However, using Lemma 3.7 when $\theta\neq\frac12$ and direct calculation when $\theta=\frac12$, one can show that $\langle f,\ e_\theta\rangle_\Phi=0$ for all $\theta\in\mathbb{T}$, where $e_\theta(n)\coloneqq e^{2\pi i n\theta}$, which implies that $f\in\mathsf{Bes}(\mathbb{N},\Phi)^\perp$.

Our second splitting theorem is as follows.

Theorem 3.22. *For every $f\in\mathsf{L}^2(\mathbb{N},\Phi)$ there is a subsequence Ψ of Φ and functions $f_{\mathsf{c}},f_{\mathsf{wm}}\in\mathsf{L}^2(\mathbb{N},\Psi)$ with f_{c} compact along Ψ, f_{wm} weak mixing along Ψ, and $f=f_{\mathsf{c}}+f_{\mathsf{wm}}$. Moreover, if f is real-valued with $a\le f\le b$ for some $a\le b$ then f_{c} is real-valued and satisfies $a\le f_{\mathsf{c}}\le b$.*

Remark 3.23. The conclusion of Theorem 3.22 is similar to that of Theorem 3.9. We remark that, in fact, f_{c} minimizes the distance between f and the closed subspace of compact functions in $\mathsf{L}^2(\mathbb{N},\Phi)$ but will not make use of this. It is also true that f_{c} can be shown to be unique in the sense of parts (a) and (b) of Remark 3.11.

The proof of Theorem 3.22 requires some lemmas, the first of which is essentially [Fur81, Lemma 4.23]. Recall that a triple (X,μ,T) is a **measure preserving system** if X is a compact space equipped with a Borel probability measure μ and $T:X\to X$ is a measurable map that preserves μ. Given a measure preserving system (X,μ,T) one can consider the Hilbert space $\mathsf{L}^2(X,\mu)$ whose norm is denoted $\|\cdot\|_\mu$. The map T induces an isometry U on $\mathsf{L}^2(X,\mu)$ defined by $Uf=f\circ T$ for all $f\in\mathsf{L}^2(X,\mu)$.

Lemma 3.24. *Let (X,μ,T) be a measure preserving system. For the isometry $\mathsf{U}f=f\circ T$ of the Hilbert space $\mathsf{L}^2(X,\mu)$ the constant functions are* *compact, $|\phi|$ is compact whenever ϕ is, and both $\min\{\phi,\psi\}$ and $\max\{\phi,\psi\}$ are compact whenever ϕ,ψ are compact and real-valued.*

Proof. Since the constant functions are fixed points of U they certainly satisfy Definition 3.13. The reverse triangle inequality gives

$$\begin{aligned}\|\mathsf{U}^m(|\phi|)-\mathsf{U}^k(|\phi|)\|_\mu^2 &= \int_X \Big||\phi(\mathsf{T}^m x)|-|\phi(\mathsf{T}^k x)|\Big|^2 \,\mathsf{d}\mu(x)\\ &\le \int_X \Big|\phi(\mathsf{T}^m x)-\phi(\mathsf{T}^k x)\Big|^2 \,\mathsf{d}\mu(x) = \|\mathsf{U}^m(\phi)-\mathsf{U}^k(\phi)\|_\mu^2\end{aligned}$$

so compactness of ϕ implies compactness of $|\phi|$. For the last claim write

$$\min\{\phi,\psi\} = \frac{\phi+\psi-|\phi-\psi|}{2} \quad \text{and} \quad \max\{\phi,\psi\} = \frac{\phi+\psi+|\phi-\psi|}{2}$$

pointwise. □

Corollary 3.25. *Under the hypothesis of Lemma 3.24 if $a \le \phi \le b$ for some $a \le b$ then $a \le \phi_{\mathsf{c}} \le b$.*

Proof. Since ϕ_{c} is the orthogonal projection of ϕ on $\mathscr{H}_{\mathsf{c}}$ it is characterized as the unique element of $\mathscr{H}_{\mathsf{c}}$ closest to ϕ. Since the real part of ϕ_{c} is compact and at least as close to ϕ as ϕ_{c} is, it must be the case that ϕ_{c} is real-valued. Since $\min\{\phi_{\mathsf{c}}, b\}$ is compact and at least as close to ϕ as ϕ_{c} is, we must have $\phi_{\mathsf{c}} \le b$. A similar argument proves that $a \le \phi_{\mathsf{c}}$. □

The next lemma, which realizes an arbitrary bounded sequence as a continuous function evaluated along the orbit of a point in a transitive topological dynamical system, can be seen as a version of the Furstenberg correspondence principle [Fur81, Lemma 3.17]. In fact, it allows one to realize a countable collection of bounded sequences with the help of the same transitive topological dynamical system; in this strengthened form it will contribute to the proof of Theorem 4.15 below.

Lemma 3.26. *Let J be a finite or countably infinite set and let $\{a_i : i \in J\}$ be a collection of bounded functions from $\mathbb{N}$ to $\mathbb{C}$. Then there exists a compact metric space X, a continuous map $\mathsf{S}\colon X \to X$, functions $F_i \in \mathsf{C}(X)$ for each $i \in J$, and a point $x \in X$ with a dense orbit under S such that*

$$a_i(n) = F_i(\mathsf{S}^n x) \qquad \forall n \in \mathbb{N},\ \forall i \in J. \tag{23}$$

Proof. Let $D_i \subset \mathbb{C}$ be a compact set containing the image of a_i. The space

$$Y \coloneqq \prod_{i \in J} D_i^{\mathbb{N} \cup \{0\}}$$

is a countable product of compact metric spaces and therefore a compact metric space itself. We can identify Y with the collection of all sequences $y \colon J \times (\mathbb{N} \cup \{0\}) \to \mathbb{C}$ that satisfy $y(i,n) \in D_i$ for all $n \in \mathbb{N} \cup \{0\}$ and $i \in J$.

Given a point $y \in Y$ we define $\mathsf{S}(y)$ as

$$(\mathsf{S}y)(i,n) = y(i,n+1)$$

which gives a continuous map $\mathsf{S} : Y \to Y$. Let x be the point $x(i,n) \coloneqq a_i(n)$ and let X be the orbit closure of x under the action of S. Then X is a compact metric space. Moreover, if we define $F_i(y) \coloneqq y(i,0)$ then (23) is satisfied. □

We are finally ready to prove Theorem 3.22.

Proof of Theorem 3.22. We will first deal with the case where $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is bounded and then derive from it the general case.

Using Lemma 3.26 we can find a compact metric space X, a continuous map $\mathsf{S} \colon X \to X$, a function $F \in \mathsf{C}(X)$ and a point $x \in X$ with a dense orbit under S such that $F(\mathsf{S}^n(x)) = f(n)$ for all $n \in \mathbb{N}$. Since X is a compact metric space, we can find (using eg. [Gla03, Theorem A.4]) a subsequence Ψ of Φ such that the measures

$$\mu_N \coloneqq \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} \delta_{\mathsf{S}^n x}$$

weak* converge to an S invariant Borel probability measure μ on X. We therefore have a measure preserving system (X, μ, S). The transformation S induces an isometry U on the Hilbert space $\mathsf{L}^2(X,\mu)$ via $\mathsf{U}(H) = H \circ \mathsf{S}$ for all $H \in \mathsf{L}^2(X,\mu)$. Let $F = F_{\mathsf{c}} + F_{\mathsf{wm}}$ be the Jacobs–de Leeuw–Glicksberg decomposition of F given by Theorem 3.15.

Next for each $j \in \mathbb{N}$, let $H_j \in \mathsf{C}(X)$ be such that $\|F_{\mathsf{c}} - H_j\|_\mu < 1/j$. Let $h_j(n) = H_j(\mathsf{S}^n x)$ for all $n \in \mathbb{N}$ and observe that

$$\begin{aligned}\|h_j - h_\ell\|_\Psi^2 &= \limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} \left|H_j(\mathsf{S}^n x) - H_\ell(\mathsf{S}^n x)\right|^2 \\ &= \int_X |H_j - H_\ell|^2 \,\mathsf{d}\mu = \|H_j - H_\ell\|_\mu^2,\end{aligned}$$

which implies, in particular, that $j \mapsto h_j$ is a Cauchy sequence in $\mathsf{L}^2(\mathbb{N},\Psi)$. Using Proposition 3.1, after refining Ψ if necessary, we can find a function $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N},\Psi)$ such that $\|h_j - f_c\|_\Psi \to 0$ as $j \to \infty$. We also define f_{wm} to be $f - f_{\mathsf{c}}$.

To show that f_{c} is compact along Ψ, fix $\epsilon > 0$ and let $K \in \mathbb{N}$ be such that

$$\min\left\{\|\mathsf{S}^m F_{\mathsf{c}} - \mathsf{S}^k F_{\mathsf{c}}\|_\mu : 1 \leq k \leq K\right\} < \epsilon$$

for every $m \in \mathbb{N}$. Then, taking $j > 1/\epsilon$ large enough so that $\|h_j - f_{\mathsf{c}}\|_\Psi < \epsilon$, we have

$$\begin{aligned}\|\mathsf{R}^m f_{\mathsf{c}} - \mathsf{R}^k f_{\mathsf{c}}\|_\Psi &\leq \|\mathsf{R}^m h_j - \mathsf{R}^k h_j\|_\Psi + 2\epsilon \\ &= \|\mathsf{S}^m H_j - \mathsf{S}^k H_j\|_\mu + 2\epsilon \\ &\leq \|\mathsf{S}^m F_{\mathsf{c}} - \mathsf{S}^k F_{\mathsf{c}}\|_\mu + 4\epsilon,\end{aligned}$$

and hence $\min\left\{\|\mathsf{R}^m f_{\mathsf{c}} - \mathsf{R}^k f_{\mathsf{c}}\|_\Psi : 1 \leq k \leq K\right\} < 5\epsilon$. If f takes values in $[a,b]$ then so does F. By Corollary 3.25 it follows that F_{c} also takes values in $[a,b]$. In this case, we can choose H_j to take values in $[a,b]$ and hence h_j takes values in $[a,b]$ for every $j \in \mathbb{N}$. Finally, since f_{c} is the limit of h_j as $j \to \infty$, we have from Proposition 3.1 that it takes values in $[a,b]$ too.

To prove that f_{wm} is weak mixing along Ψ, let $h\colon \mathbb{N} \to \mathbb{C}$ be bounded and let Ψ' be a Følner subsequence of Ψ such that the correlations $\langle \mathsf{R}^n f,\ h\rangle_{\Psi'}$ exist for every $n \in \mathbb{N}$. Using Lemma 3.26 again, we can find another compact metric space $\tilde{X}$, a continuous map $\tilde{\mathsf{S}}\colon \tilde{X} \to \tilde{X}$, a function $\tilde{F} \in \mathsf{C}(\tilde{X})$ and a point $\tilde{x} \in \tilde{X}$ with a dense orbit under S such that $\tilde{F}(\tilde{\mathsf{S}}^n(\tilde{x})) = h(n)$ for all $n \in \mathbb{N}$.

Let $Z \subset X \times \tilde{X}$ be the orbit closure of $(x,\tilde{x})$ under $\mathsf{S} \times \tilde{\mathsf{S}}$. Since Z is a compact metric space, we can find a subsequence Ψ'' of Ψ' such that the measures

$$\nu_N \coloneqq \frac{1}{|\Psi''_N|} \sum_{n \in \Psi''_N} \delta_{(\mathsf{S}\times\tilde{\mathsf{S}})^n(x,\tilde{x})}$$

converge in the weak* topology to an invariant probability measure ν on Z. For all $\epsilon > 0$, if j is sufficiently large, then

$$\begin{aligned}\left|\langle \mathsf{R}^m f_{\mathsf{wm}}, h\rangle_{\Psi'}\right| &\leq |\langle \mathsf{R}^m(f - h_j), h\rangle_{\Psi''}| + \epsilon \\ &= \left|\lim_{N\to\infty} \frac{1}{|\Psi''_N|} \sum_{n \in \Psi''_N} (f - h_j)(n+m)\overline{h(n)}\right| + \epsilon \\ &= \left|\lim_{N\to\infty} \frac{1}{|\Psi''_N|} \sum_{n \in \Psi''_N} (F - H_j)(\mathsf{S}^{n+m}x)\overline{\tilde{F}(\tilde{\mathsf{S}}^n\tilde{x})}\right| + \epsilon\end{aligned}$$

$$= \left| \int_Z (\mathsf{S} \times \tilde{\mathsf{S}})^m ((F - H_j) \otimes 1) \overline{(1 \otimes \tilde{F})} \, \mathsf{d}\nu \right| + \epsilon$$

$$\leq \left| \int_Z (\mathsf{S} \times \tilde{\mathsf{S}})^m (F_{\mathsf{wm}} \otimes 1)(1 \otimes \tilde{F}) \, \mathsf{d}\nu \right| + 2\epsilon.$$

For every $\phi \in \mathsf{C}(X)$ and every $\psi \in \mathsf{C}(\tilde{X})$ we have

$$|\langle F_{\mathsf{wm}} \otimes 1,\ \phi \otimes \psi \rangle_\nu| \leq |\langle F_{\mathsf{wm}},\ \phi \rangle_\mu| \sup_{z \in \tilde{X}} |\psi(z)|$$

which implies $F_{\mathsf{wm}} \otimes 1$ in $\mathsf{L}^2(Z, \nu)$ is a weak mixing function. This implies that the set

$$\left\{ n \in \mathbb{N} : \left| \int_Z (\mathsf{S} \times \tilde{\mathsf{S}})^m (F_{\mathsf{wm}} \otimes 1)(1 \otimes \tilde{F}) \, \mathsf{d}\nu \right| > \epsilon \right\}$$

has zero density with respect to every Følner sequence. Hence the set

$$\left\{ n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}},\ h \rangle_\Psi| > 3\epsilon \right\}$$

has zero density with respect to every Følner sequence, finishing the proof in the case f is bounded.

Finally, we deal with the case where f is not necessarily bounded. Suppose $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is arbitrary and let $j \mapsto f_j$ be a sequence of bounded functions such that $\|f - f_j\|_\Phi \to 0$ as $j \to \infty$. Define $\Psi^{(0)} \coloneqq \Phi$. For every $j \in \mathbb{N}$, apply the decomposition to f_j to obtain a Følner sequence $\Psi^{(j)}$, which is a subsequence of $\Psi^{(j-1)}$, and a decomposition $f_j = f_{j,\mathsf{c}} + f_{j,\mathsf{wm}}$, where $f_{j,\mathsf{c}}$ is compact along $\Psi^{(j)}$ and $f_{j,\mathsf{wm}}$ is weak mixing along $\Psi^{(j)}$.

Define Ψ as $\Psi_N \coloneqq \Psi_N^{(N)}$ for all $N \in \mathbb{N}$. Then, for every $j \in \mathbb{N}$, since Ψ is eventually a Følner subsequence of $\Psi^{(j)}$, the function $f_{j,\mathsf{c}}$ is compact along Ψ and the function $f_{j,\mathsf{wm}}$ is weak mixing along Ψ. In particular $\langle f_{j,\mathsf{c}},\ f_{\ell,\mathsf{wm}} \rangle_\Psi = 0$ for every j, ℓ and hence $\|f_j - f_\ell\|_\Psi^2 = \|f_{j,\mathsf{c}} - f_{\ell,\mathsf{c}}\|_\Psi^2 + \|f_{j,\mathsf{wm}} - f_{\ell,\mathsf{wm}}\|_\Psi^2$. Since $j \mapsto f_j$ is a Cauchy sequence with respect to Φ (and hence with respect to Ψ), it follows that $j \mapsto f_{j,\mathsf{c}}$ is also a Cauchy sequence with respect to Ψ. Using Proposition 3.1, and after refining Ψ if needed, we can find a function f_c in $\mathsf{L}^2(\mathbb{N}, \Psi)$ such that $\|f_{j,\mathsf{c}} - f_\mathsf{c}\|_\Psi \to 0$ as $j \to \infty$. It follows that f_c is compact with respect to Ψ. Then let $f_{\mathsf{wm}} = f - f_\mathsf{c}$ and observe that $\|f_{\mathsf{wm}} - f_{j,\mathsf{wm}}\|_\Psi \to 0$ as $j \to \infty$, which implies that f_{wm} is weak mixing. □

4. Proof of Theorem 2.7

In Section 2 we reduced the proof of Theorem 1.2 to Theorem 2.7. In this section we use the splittings coming from Theorems 3.6 and 3.22 of Section 3 to finish the proof of Theorem 2.7.

The main result of this section is the following theorem, which gives us an ultrafilter satisfying several convenient properties.

Theorem 4.1. *Fix $\epsilon > 0$ and a Følner sequence Φ on $\mathbb{N}$. Given $f_{\mathsf{Bes}} \in \mathsf{Bes}(\mathbb{N}, \Phi)$ bounded and non-negative, $f_{\mathsf{anti}} \in \mathsf{Bes}(\mathbb{N}, \Phi)^{\perp}$ bounded and real-valued, and $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N}, \Phi)$ bounded, non-negative and compact along Φ, one can find a subsequence Ψ of Φ and an ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that:*

U1. $\overline{\mathsf{d}}_{\Psi}(E) > 0$ *for all* $E \in \mathsf{p}$*;*

U2. $\{n \in \mathbb{N} : \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_{\Psi} < \frac{\epsilon}{3}\} \in \mathsf{p}$*;*

U3. $\|\mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}} - f_{\mathsf{Bes}}\|_{\Psi} < \frac{\epsilon}{3}$*;*

U4. $\langle f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_{\Psi}$ *is non-negative.*

The proof of Theorem 4.1 is given in Section 4.1. For now we show how, together with the decompositions provided by Theorems 3.6 and 3.22, it implies Theorem 2.7.

Proof of Theorem 2.7 assuming Theorem 4.1. Fix a bounded, non-negative function $f\colon \mathbb{N} \to \mathbb{R}$ and a Følner sequence Φ on $\mathbb{N}$ with $\langle 1,\ f\rangle_{\Phi}$ existing. The statement is trivial if $\|f\|_{\Phi} = 0$, so let us assume that $\|f\|_{\Phi} > 0$. Fix also $\epsilon > 0$. Our goal is to find a subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that

$$\lim_{n \to \mathsf{p}} \langle \mathsf{R}^n f,\ \mathsf{R}^{\mathsf{p}} f\rangle_{\Psi} \geq \langle 1,\ f\rangle_{\Psi}^2 - \epsilon$$

holds.

Apply Theorem 3.6 and Theorem 3.22 to obtain, after passing to a subsequence Ψ of Φ, decompositions $f = f_{\mathsf{Bes}} + f_{\mathsf{anti}}$ and $f = f_{\mathsf{c}} + f_{\mathsf{wm}}$. Since f is bounded and non-negative, according to the second part of Theorem 3.22, the function f_{c} is also bounded and non-negative. Similarly, f_{Bes} is bounded and real-valued as well. Since $f_{\mathsf{anti}} = f - f_{\mathsf{c}}$, it also follows that f_{anti} is bounded and real-valued, which is another fact that we will use later in the proof. In fact, after passing to a subsequence of Ψ if necessary, all of $\|f_{\mathsf{c}}\|_{\Psi}$, $\|f_{\mathsf{Bes}}\|_{\Psi}$, $\|f_{\mathsf{wm}}\|_{\Psi}$ and $\|f_{\mathsf{anti}}\|_{\Psi}$ are at most $\|f\|_{\Psi}$ by orthogonality and the Pythagoras theorem.

Next we can apply Theorem 4.1 with $\epsilon/\|f\|_{\Phi}$ in place of ϵ to get a finer subsequence Ψ and an ultrafilter p satisfying **U1** through **U4** with $\epsilon/\|f\|_{\Phi}$ in place of ϵ. Finally, pass once more to a subsequence of Ψ such that the inner products $\langle f_{\mathsf{c}},\ f_{\mathsf{Bes}}\rangle_{\Psi}$, $\langle \mathsf{R}^n f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} f\rangle_{\Psi}$, $\langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_{\Psi}$ and $\langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_{\Psi}$

exist for all $n \in \mathbb{N} \cup \{0\}$. Note that $\mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}$ and $\mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}$ are well defined since f_{Bes} and f_{anti} are bounded.

We then have

$$\langle \mathsf{R}^n f,\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi = \langle \mathsf{R}^n f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi + \langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi + \langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi$$

for all $n \in \mathbb{N}$. We claim that

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi = 0 \tag{24}$$

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq -\tfrac{\epsilon}{3} \tag{25}$$

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi \geq \langle 1,\ f\rangle_\Psi^2 - \tfrac{2\epsilon}{3} \tag{26}$$

are all true for our choice of p. Once (24), (25) and (26) have been established, (16) follows immediately and the proof is complete.

Let us first show (24). Since f_{wm} is weak mixing along Ψ, we have, for every $\delta > 0$, that the set $\{n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi| \geq \delta\}$ has zero density with respect to Ψ. It therefore does not belong to p by **U1**. It follows that $\{n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}},\ \mathsf{R}^{\mathsf{p}} f\rangle_\Psi| < \delta\}$ belongs to p for all $\delta > 0$ giving (24).

For the proof of (25) note that

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq \langle f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi - \tfrac{\epsilon}{3} \tag{27}$$

in light of **U2** because of

$$|\langle \mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi| \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f_{\mathsf{anti}}\|_\Psi \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f\|_\Phi$$

by Cauchy-Schwarz. Thus (25) follows from (27) and **U4**.

Utilizing **U2** once more this time combined with

$$|\langle \mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi| \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f_{\mathsf{Bes}}\|_\Psi \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f\|_\Phi$$

via a similar application of Cauchy-Schwarz, we see that

$$\langle f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi \geq \langle 1,\ f\rangle_\Psi^2 - \tfrac{\epsilon}{3} \tag{28}$$

implies (26). To prove (28) use **U3** and Cauchy-Schwarz once more to establish

$$\langle f_{\mathsf{c}},\ \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi \geq \langle f_{\mathsf{c}},\ f_{\mathsf{Bes}}\rangle_\Psi - \tfrac{\epsilon}{3}$$

and then we observe that $\langle f_{\mathsf{c}},\ f_{\mathsf{Bes}}\rangle_\Psi = \|f_{\mathsf{Bes}}\|_\Psi^2 + \langle f_{\mathsf{c}} - f_{\mathsf{Bes}},\ f_{\mathsf{Bes}}\rangle_\Psi$. Since $f_{\mathsf{c}} - f_{\mathsf{Bes}} = f_{\mathsf{anti}} - f_{\mathsf{wm}}$ and every weak mixing function belongs to $\mathsf{Bes}(\mathbb{N}, \Psi)^\perp$, it follows that $\langle f_{\mathsf{c}} - f_{\mathsf{Bes}},\ f_{\mathsf{Bes}}\rangle_\Psi = \langle f_{\mathsf{anti}} - f_{\mathsf{wm}},\ f_{\mathsf{Bes}}\rangle_\Psi = 0$ and hence $\langle f_{\mathsf{c}},\ f_{\mathsf{Bes}}\rangle_\Psi = \|f_{\mathsf{Bes}}\|_\Psi^2$. Finally, we apply the Cauchy-Schwarz inequality to deduce that $\|f_{\mathsf{Bes}}\|_\Psi^2 \geq \langle 1,\ f_{\mathsf{Bes}}\rangle_\Psi^2$ and, using $\langle 1,\ f_{\mathsf{anti}}\rangle_\Psi = 0$, we get $\langle 1,\ f_{\mathsf{Bes}}\rangle_\Psi^2 = \langle 1,\ f\rangle_\Psi^2$. This implies (28) and finishes the proof. □

4.1. Proof of Theorem 4.1

We begin with some preparatory definitions.

Definition 4.2. Given a Følner sequence Φ on $\mathbb{N}$ we say an ultrafilter p is Φ **essential** if $\overline{\mathsf{d}}_\Phi(E) > 0$ for every $E \in \mathsf{p}$. Write $\mathsf{Ess}(\Phi)$ for the set of Φ essential ultrafilters on $\mathbb{N}$.

Observe that property **U1** in Theorem 4.1 means exactly that p is a Ψ essential ultrafilter.

Recall from Section 2 the definition of $\mathcal{M}(\Phi)$.

Definition 4.3. A Borel measurable property of ultrafilters is said to hold Φ **almost everywhere** if the set of ultrafilters p with the property has full measure with respect to every $\mu \in \mathcal{M}(\Phi)$.

Lemma 4.4. *Let Φ be a Følner sequence on $\mathbb{N}$. Then Φ almost every p belongs to $\mathsf{Ess}(\Phi)$.*

Proof. First, observe that

$$\mathsf{Ess}(\Phi) = \bigcap_{E\subset\mathbb{N}:\overline{\mathsf{d}}_\Phi(E)=0} \mathsf{cl}(\mathbb{N}\backslash E) = \beta\mathbb{N}\backslash \bigcup_{E\subset\mathbb{N}:\overline{\mathsf{d}}_\Phi(E)=0} \mathsf{cl}(E)$$

so that it is a closed set (and hence Borel). Fix $\mu \in \mathcal{M}(\Phi)$. We claim that the support of μ is contained in $\mathsf{Ess}(\Phi)$. Since μ is Radon this implies $\mu(\mathsf{Ess}(\Phi)) = 1$ as desired.

To prove the claim, fix $\mathsf{p} \in \beta\mathbb{N}\backslash\mathsf{Ess}(\Phi)$. We need to show that there exists an open set $U \subset \beta\mathbb{N}$ containing p such that $\mu(U) = 0$. But since $\mathsf{p} \in \beta\mathbb{N}\backslash\mathsf{Ess}(\Phi)$, there exists $E \subset \mathbb{N}$ with $\overline{\mathsf{d}}_\Phi(E) = 0$ and $\mathsf{p} \in \mathsf{cl}(E)$. The set $\mathsf{cl}(E)$ is then an open subset of $\beta\mathbb{N}$ containing p and with $\mu(\mathsf{cl}(E)) \leq \overline{\mathsf{d}}_\Phi(E) = 0$. □

Definition 4.5. A **Bohr set** on $\mathbb{N}$ is any set of the form $a^{-1}(U)$ where a is a homomorphism from $\mathbb{N}$ into a compact metrizable abelian group K and U is a non-empty open subset of K whose topological boundary ∂U has zero Haar measure. A Bohr set is a **Bohr$_0$ set** if U contains the identity element of K.

There are various minor variations on the definition of Bohr sets appearing in the literature. For example, sometimes authors restrict attention to the case where K is a product of finitely many copies of the circle group and U is a product of arcs. Alternatively, one could define Bohr sets and Bohr$_0$ sets with the help of the Bohr topology on the integers, which is the topology induced by the embedding of $\mathbb{Z}$ into its Bohr compactification (cf. [Ruz82], [BFW06, Section 1] and [HK11]). Definition 4.5 is the most convenient for our needs because with it the following lemmas are straightforward to prove.

Lemma 4.6. *If A and B are Bohr sets then so is $A \cap B$.*

Proof. Write $A = a^{-1}(U)$ and $B = b^{-1}(V)$ where $a : \mathbb{N} \to K$ and $b : \mathbb{N} \to L$ are homomorphisms to compact metrizable topological groups K and L respectively. Then $A \cap B = c^{-1}(U \times V)$ where $c : \mathbb{N} \to K \times L$ is the homomorphism $c(n) = (a(n), b(n))$. □

The following lemma is folklore; we reproduce a short proof from [GKR18, Lemma 2.7].

Lemma 4.7. *Let $a : \mathbb{N} \to G$ be a homomorphism from $\mathbb{N}$ to a compact abelian topological group G. Then the closure of the image of a is a subgroup of G.*

Proof. Define $S \coloneqq \overline{\{a(n) : n \in \mathbb{N}\}}$ and

$$H \coloneqq \overline{\{a(n) : n \in \mathbb{N}\}} \cup \{0\} \cup \overline{\{-a(n) : n \in \mathbb{N}\}}.$$

We have to show that $S = H$. Define $A \coloneqq \bigcap_{N \in \mathbb{N}} \overline{\{a(n) : n \geq N\}}$. Since A is the intersection of a nested family of non-empty compact sets, it is non-empty. Pick any $x \in A$. Since A is H-invariant, we have $H + x \subset A$ and hence $A = H$. But $A \subset S$, which now implies $H \subset S$. □

Lemma 4.8. *If $B \subset \mathbb{N}$ is a Bohr set then for every Følner sequence Φ its indicator function 1_B is in $\mathsf{Bes}(\mathbb{N}, \Phi)$ and $\mathsf{d}_\Phi(B) > 0$. In fact, if $B = a^{-1}(U)$ then $\mathsf{d}_\Phi(B) = \mathsf{m}(U)$ where m is Haar measure on the implicit compact metrizable group K.*

Proof. Let K be a compact abelian group, let $a \colon \mathbb{N} \to K$ be a homomorphism and let $U \subset K$ be an open set with zero measure boundary and such that $B = a^{-1}(U)$. Replacing K with the closure $\overline{a(\mathbb{N})}$ we can assume that a has a dense image.

For each $N \in \mathbb{N}$, let μ_N be the probability measure on K obtained as the average of the Dirac point masses at the points $\{a(n) : n \in \Phi_N\}$. Since Φ is a Følner sequence, any weak* limit point μ of $(\mu_N)_{N \in \mathbb{N}}$ is invariant under

$a(\mathbb{N})$. By Lemma 4.7 it follows that $\mu = \mathsf{m}$ is the Haar measure on K. Since U is open we have

$$0 < \mathsf{m}(U) = \lim_{N\to\infty} \mu_N(U) = \lim_{N\to\infty} \frac{|B \cap \Phi_N|}{|\Phi_N|} = \mathsf{d}_\Phi(B)$$

in view of [Gla03, Theorem A.5]. Finally, since finite linear combinations of characters (i.e., continuous homomorphisms from K to the circle group S^1) are dense in $\mathsf{L}^2(K, \mathsf{m})$, we can find for every $\epsilon > 0$ a linear combination f of characters such that $\|f - 1_U\|_{\mathsf{m}} < \epsilon$. Since $\mu = \mathsf{m}$ and $f - 1_U$ is m-almost everywhere continuous, it follows that $\|f \circ a - 1_B\|_\Phi = \|f - 1_U\|_{\mathsf{m}} < \epsilon$. Since $f \circ a$ is a trigonometric polynomial and ϵ was arbitrary, we conclude that $1_B \in \mathsf{Bes}(\mathbb{N}, \Phi)$. □

Lemma 4.9. *For every function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ which is compact along Φ and every $\epsilon > 0$, the set $\{n \in \mathbb{N} : \|\mathsf{R}^n f - f\|_\Phi < \epsilon\}$ contains a Bohr$_0$ set B_{c}.*

Proof. Let $g(n) = \|\mathsf{R}^{|n|} f - f\|_\Phi$ for every $n \in \mathbb{Z}$. Since f is compact along Φ it follows that the closure Ω of the set $\{\mathsf{R}^k g : k \in \mathbb{Z}\}$ has a finite ϵ-dense subset with respect to the uniform metric for every $\epsilon > 0$. It therefore has compact closure.

We can make Ω into a compact topological group by defining

$$(\mathsf{R}^n g) \star (\mathsf{R}^k g) = \mathsf{R}^{k+n} g$$

for all $n, k \in \mathbb{Z}$ and extending $\star$ to a binary operation on all of Ω by continuity. Define $U_\eta \coloneqq \{\phi : \mathbb{Z} \to [0, \infty) : \phi(0) < \eta\}$. Using the homomorphism $a(n) = \mathsf{R}^n g$ from $\mathbb{N}$ to our topological group $(\Omega, \star)$ we see that $\{n \in \mathbb{N} : \|\mathsf{R}^n f - f\|_\Phi < \epsilon\} = \{n \in \mathbb{N} : a(n) \in U_\epsilon\}$. Moreover, $\{n \in \mathbb{N} : a(n) \in U_\eta\} \subset \{n \in \mathbb{N} : \|\mathsf{R}^n f - f\|_\Phi < \epsilon\}$ for every $\eta < \epsilon$. Since Haar measure on Ω is finite and the boundaries of the sets U_η are pairwise disjoint, for all but countably many $\eta > 0$ the boundary of the set U_η has zero Haar measure.

Pick any $\eta < \epsilon$ for which ∂U_η has measure 0 and define $B_{\mathsf{c}} \coloneqq \{n \in \mathbb{N} : a(n) \in U_\eta\}$.

□

The following two theorems, proved in subsequent subsections, will be used in the proof of Theorem 4.1. The first, which will be used to guarantee **U3**, relies on the pointwise ergodic theorem. Its proof can be found in Section 4.2.

Theorem 4.10. *Let Φ be a Følner sequence on $\mathbb{N}$ and let $f \in \mathsf{Bes}(\mathbb{N},\Phi)$. For every $\epsilon > 0$ there exists a Bohr$_0$ set B and a subsequence Ψ of Φ such that for Ψ almost every ultrafilter $\mathsf{p} \in \mathsf{cl}(B)$ we have $\|\mathsf{R}^{\mathsf{p}} f - f\|_{\Psi} < \epsilon$.*

The second is a modification of an argument due to Beiglböck [Bei11, Lemma 2] and will be used to guarantee **U4**. Its proof is given in Section 4.3.

Theorem 4.11. *Suppose f is a real-valued bounded function that belongs to $\mathsf{Bes}(\mathbb{N},\Psi)^{\perp}$. Then for every non-empty Bohr set $B \subset \mathbb{N}$ and every bounded function $h\colon \mathbb{N} \to \mathbb{R}$ the set*

$$\left\{ \mathsf{p} \in \mathsf{cl}(B) : \limsup_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} h(m)\,(\mathsf{R}^{\mathsf{p}} f)(m) \geq 0 \right\} \tag{29}$$

is Borel measurable and has positive measure with respect to every $\mu \in \mathcal{M}(\Psi)$.

With these theorems we can give the proof of Theorem 4.1.

Proof of Theorem 4.1. Fix $\epsilon > 0$ and a Følner sequence Φ on $\mathbb{N}$ along which $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N},\Phi)$ is compact, $f_{\mathsf{Bes}} \in \mathsf{Bes}(\mathbb{N},\Phi)$, and $f_{\mathsf{anti}} \in \mathsf{Bes}(\mathbb{N},\Phi)^{\perp}$. We need to find a subsequence Ψ of Φ and an ultrafilter p such that **U1** through **U4** are satisfied.

Lemma 4.9 gives that there exists a Bohr$_0$ set $B_{\mathsf{c}} \subset \{n \in \mathbb{N} : \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_{\Phi} < \frac{\epsilon}{3}\}$. Theorem 4.10 implies that, passing to a subsequence Ψ of Φ, there exists a Bohr$_0$ set B_{Bes} such that for Ψ almost every $\mathsf{p} \in \mathsf{cl}(B_{\mathsf{Bes}})$ we have $\|\mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}} - f_{\mathsf{Bes}}\|_{\Psi} < \epsilon/3$. The set $B \coloneqq B_{\mathsf{c}} \cap B_{\mathsf{Bes}}$ is a Bohr set by Lemma 4.6. Note that B is a Bohr$_0$ set and Ψ almost any $\mathsf{p} \in \mathsf{cl}(B)$ satisfies **U2** and **U3**. Applying Theorem 4.11 with $f = f_{\mathsf{anti}}$ and $h = f_{\mathsf{c}}$ we deduce that the set

$$\left\{ \mathsf{p} \in \mathsf{cl}(B) : \limsup_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} f_{\mathsf{c}}(m)\,(\mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}(m)) \geq 0 \right\} \tag{30}$$

has positive measure for any $\mu \in \mathcal{M}(\Psi)$. Notice that any p in the set (30) satisfies **U4**. Since any such p belongs to $\mathsf{cl}(B)$ it follows that Ψ almost every p in the set (30) satisfies **U2**, **U3** and **U4**.

Finally, in view of Lemma 4.4, Ψ almost every $\mathsf{p} \in \beta\mathbb{N}$ satisfies **U1**. This means that Ψ almost every p in the set (30) satisfies **U1**, **U2**, **U3**, and **U4**. $\square$

4.2. Proof of Theorem 4.10

In this section we present a proof of Theorem 4.10. We start with the following lemma.

Lemma 4.12. *Let Φ be a Følner sequence on $\mathbb{N}$. If a is a trigonometric polynomial and $\mathsf{p} \in \beta\mathbb{N}$ then $\mathsf{R}^{\mathsf{p}}a$ is a trigonometric polynomial and $\|\mathsf{R}^{\mathsf{p}}a\|_{\Phi} = \|a\|_{\Phi}$.*

Proof. Choose $c_1, \ldots, c_J \in \mathbb{C}$ and $\theta_1, \ldots, \theta_J \in \mathbb{R}$ such that a has the form (19). Define $d_j \coloneqq \lim_{m \to \mathsf{p}} c_j e^{2\pi i \theta_j m}$. Notice that

$$(\mathsf{R}^{\mathsf{p}}a)(n) = \sum_{j=1}^{J} d_j e^{2\pi i \theta_j n}$$

and, since $|c_j| = |d_j|$, it follows from Lemma 3.7 that $\|\mathsf{R}^{\mathsf{p}}a\|_{\Phi} = \|a\|_{\Phi}$. □

We will also need a version of the pointwise ergodic theorem. There are Følner sequences for which the pointwise ergodic theorem does not hold [AJ75]. However, every Følner sequence has a subsequence along which the pointwise ergodic theorem holds.

Definition 4.13. A Følner sequence Φ is called **tempered** if there exists $C > 0$ such that

$$\left| \bigcup_{k=1}^{N} \Phi_{N+1} - \Phi_k \right| \leq C|\Phi_{N+1}|$$

for every $N \in \mathbb{N}$, where $\Phi_{N+1} - \Phi_k$ is the set of differences.

According to [Lin01, Proposition 1.4], every Følner sequence has a tempered subsequence. Here is the pointwise ergodic theorem for tempered Følner sequences.

Theorem 4.14 (see [Lin01, Theorem 1.2])**.** *Let (X, ν, T) be a measure preserving system and let Φ be a tempered Følner sequence. Then for every $f \in \mathsf{L}^1(X, \nu)$ the limit*

$$\lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} f(T^n x)$$

exists for ν almost every $x \in X$ and defines a T invariant function in $\mathsf{L}^1(X, \mu)$.

Theorem 4.15. *Let Φ be a Følner sequence on $\mathbb{N}$ and let $h \in \mathsf{Bes}(\mathbb{N}, \Phi)$ be bounded. Then there is a subsequence Ψ of Φ with $\|\mathsf{R}^{\mathsf{p}}h\|_{\Psi} = \|h\|_{\Psi}$ for Ψ almost every p.*

Proof. First we pass to a tempered subsequence Ψ of Φ. Let $j \mapsto a_j$ be a sequence of trigonometric polynomials such that $\|h - a_j\|_\Psi \to 0$ as $j \to \infty$. Apply Lemma 3.26 to the collection $\{h, a_1, a_2, \dots\}$ to find a compact metric space X, a continuous map $\mathsf{S}\colon X \to X$, a point $x \in X$ with a dense orbit under S and functions $H, F_1, F_2, \dots$ in $\mathsf{C}(X)$ such that $a_j(n) = F_j(\mathsf{S}^n x)$ and $h(n) = H(\mathsf{S}^n x)$ for all $j, n \in \mathbb{N}$.

For each $\mathsf{p} \in \beta\mathbb{N}$ define the map $\mathsf{S}^\mathsf{p}\colon X \to X$ by

$$\mathsf{S}^\mathsf{p} x = \lim_{n \to \mathsf{p}} \mathsf{S}^n x$$

and notice that

$$(\mathsf{R}^\mathsf{p} a_j)(n) = \lim_{m \to \mathsf{p}} a_j(n+m) = \lim_{m \to \mathsf{p}} F_j\left(\mathsf{S}^n \mathsf{S}^m x\right) = F_j(\mathsf{S}^n \mathsf{S}^\mathsf{p} x) \tag{31}$$

for every $j, n \in \mathbb{N}$ and every $\mathsf{p} \in \beta\mathbb{N}$. We similarly have

$$(\mathsf{R}^\mathsf{p} h)(n) = H(\mathsf{S}^n \mathsf{S}^\mathsf{p} x) \tag{32}$$

for all $n \in \mathbb{N}$ and every $\mathsf{p} \in \beta\mathbb{N}$.

The map $\pi\colon \beta\mathbb{N} \to X$ defined by $\mathsf{p} \mapsto \mathsf{S}^\mathsf{p} x$ is continuous and surjective by the universal property of $\beta\mathbb{N}$ and the fact that $\{\mathsf{S}^n x : n \in \mathbb{N}\}$ is dense in X respectively.

We next wish to prove that

$$\lim_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} |F_j(\mathsf{S}^n y)|^2 = \|a_j\|_\Psi^2 \tag{33}$$

for all $y \in X$ and all $j \in \mathbb{N}$. Fix $y \in X$ and $j \in \mathbb{N}$. Since π is surjective there is $\mathsf{p} \in \beta\mathbb{N}$ with $\mathsf{S}^\mathsf{p} x = y$. We then have

$$\frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} |F_j(\mathsf{S}^n y)|^2 = \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} |(\mathsf{R}^\mathsf{p} a_j)(n)|^2$$

from (31). By Lemma 4.12 the function $\mathsf{R}^\mathsf{p} a_j$ is also a trigonometric polynomial so

$$\lim_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} |(\mathsf{R}^\mathsf{p} a_j)(n)|^2 = \|\mathsf{R}^\mathsf{p} a_j\|_\Psi^2$$

holds by Lemma 3.7. Lemma 4.12 also gives $\|\mathsf{R}^\mathsf{p} a_j\|_\Psi = \|a_j\|_\Psi$ establishing (33).

Write U for the isometry of $\mathsf{L}^2(X,\nu)$ defined by $\mathsf{U}(f) = f \circ \mathsf{S}$ for all $f \in \mathsf{L}^2(X,\mu)$. By a version of the mean ergodic theorem of von Neumann (cf. [Gla03, Theorem 3.33]) the limit

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} \mathsf{U}^n \left(|F_j|^2\right)$$

exists in $\mathsf{L}^2(X,\nu)$ for all $j \in \mathbb{N}$ and is equal to the orthogonal projection in $\mathsf{L}^2(X,\nu)$ of $|F_j|^2$ onto the closed subspace of U invariant functions. Since constant functions are U invariant, the above combined with (33) implies for all $j \in \mathbb{N}$ that

$$\int |F_j|^2 \,\mathsf{d}\nu = \|a_j\|_\Psi^2$$

is the orthogonal projection in $\mathsf{L}^2(X,\nu)$ of $|F_j|^2$ onto the closed subspace of U invariant functions.

We are now ready to prove that $\|\mathsf{R}^{\mathsf{p}} h\|_\Psi = \|h\|_\Psi$ for Ψ almost every p. To this end fix $\mu \in \mathcal{M}(\Psi)$ and let $\nu = \pi\mu$ for the push-forward of μ under the map π. Since μ is by definition a weak* limit point of the set $\{\mu_N : N \in \mathbb{N}\}$, where μ_N is as in (11), it follows that ν is a weak* accumulation point of the set $\{\pi\mu_N : N \in \mathbb{N}\}$. Since X is a compact metric space, the space of probability measures on X is metrizable, and hence there exists a subsequence Ξ of Ψ such that

$$\nu = \lim_{N\to\infty} \frac{1}{|\Xi_N|} \sum_{n\in\Xi_N} \delta_{\mathsf{S}^n x} \tag{34}$$

in the weak* topology in X, where $\delta_{\mathsf{S}^n x}$ is the point mass on X at the point $\mathsf{S}^n x$. We remark that while every measure $\mu \in \mathcal{M}(\Psi)$ is the limit of a sub-net of $(\mu_N)_{N\in\mathbb{N}}$, there is in general no subsequence of $(\mu_N)_{N\in\mathbb{N}}$ which converges to μ because $\beta\mathbb{N}$ is not metrizable.

Since the functions H_j and F are continuous on X we may calculate from (34) that

$$\begin{aligned} \|F_j - H\|_\nu^2 &= \lim_{N\to\infty} \frac{1}{|\Xi_N|} \sum_{n\in\Xi_N} |F_j(\mathsf{S}^n x) - H(\mathsf{S}^n x)|^2 \\ &= \lim_{N\to\infty} \frac{1}{|\Xi_N|} \sum_{n\in\Xi_N} |a_j(n) - h(n)|^2 = \|a_j - h\|_\Psi^2 \end{aligned}$$

for all $j \in \mathbb{N}$, with the last equality holding because h and all a_j belong to $\mathsf{Bes}(\mathbb{N},\Psi)$. The hypothesis that $\|a_j - h\|_\Psi \to 0$ as $j \to \infty$ therefore implies

$\|F_j - H\|_\nu \to 0$ as $j \to \infty$. Since orthogonal projections on Hilbert spaces are continuous we conclude that

$$\int |H|^2 \, \mathsf{d}\nu = \lim_{j\to\infty} \|a_j\|_\Psi^2 = \|h\|_\Psi^2 \tag{35}$$

is the orthogonal projection of $|H|^2$ to the closed subspace of U invariant functions.

Next, we apply Theorem 4.14 to deduce that the limit

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} |H(\mathsf{S}^n y)|^2$$

exists for ν almost every $y \in X$ and defines a U invariant function in $\mathsf{L}^2(X,\nu)$. Since H is bounded, this limit is also bounded. This limit must therefore be the projection (35) of $|H|^2$ to the closed subspace of U invariant functions. In other words

$$\lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} |H(\mathsf{S}^n y)|^2 = \|h\|_\Psi^2$$

for ν almost every y. Finally, since ν is the push-forward of μ under π, it follows from (32) that $\|\mathsf{R}^\mathsf{p} h\|_\Phi = \|h\|_\Phi$ for μ almost every $\mathsf{p} \in \beta\mathbb{N}$. Since $\mu \in \mathcal{M}(\Psi)$ was arbitrary we are done.

□

We are now ready to finish the proof of Theorem 4.10

Proof of Theorem 4.10. Let Φ be a Følner sequence on $\mathbb{N}$, let $f \in \mathsf{Bes}(\mathbb{N},\Phi)$ and let $\epsilon > 0$. Let a be a trigonometric polynomial such that $\|f-a\|_\Phi < \epsilon/3$. Notice that $f - a \in \mathsf{Bes}(\mathbb{N},\Phi)$ and hence, using Theorem 4.15, we can find a subsequence Ψ of Φ such that for Ψ almost every $\mathsf{p} \in \beta\mathbb{N}$

$$\|\mathsf{R}^\mathsf{p} f - f\|_\Psi \le \|\mathsf{R}^\mathsf{p}(f-a)\|_\Psi + \|\mathsf{R}^\mathsf{p} a - a\|_\Psi + \|a - f\|_\Psi \le \|\mathsf{R}^\mathsf{p} a - a\|_\Psi + \frac{2\epsilon}{3}.$$

It now suffices to find a Bohr$_0$ set B such that for every $\mathsf{p} \in \mathsf{cl}(B)$ we have $\|\mathsf{R}^\mathsf{p} a - a\|_\Psi \le \epsilon/3$.

Write $a(n) = \sum_{j=1}^J c_j e^{2\pi i n\theta_j}$ for some $c_1,\ldots,c_J \in \mathbb{C}$ and $0 \le \theta_1,\ldots,\theta_J < 1$. Let $M = \max_j |c_j|$ and let $\alpha\colon \mathbb{N} \to \mathbb{T}^J$ be the homomorphism $\alpha(n) = (n\theta_1,\ldots,n\theta_J)$ (where $\mathbb{T}^J$ is the torus $\mathbb{R}^J/\mathbb{Z}^J$ as usual). Consider the open set $U = \left(-\frac{\epsilon}{3MJ}, \frac{\epsilon}{3MJ}\right)^J \subset \mathbb{T}^J$ and let $B = \alpha^{-1}(U)$. Certainly the boundary of U has zero Haar measure in $\mathbb{T}^J$ so B is a Bohr_0 set. Notice that for every $m \in B$ and every $n \in \mathbb{N}$,

$$\left|(\mathsf{R}^m a)(n) - a(n)\right| = \left|\sum_{j=1}^{J} c_j e^{2\pi i n\theta_j} \left(e^{2\pi i m\theta_j} - 1\right)\right| < \frac{\epsilon}{3} \tag{36}$$

holds. Finally, let $\mathsf{p} \in \mathsf{cl}(B)$. In view of (36), $|(\mathsf{R}^{\mathsf{p}} a)(n) - a(n)| < \epsilon/3$ for every $n \in \mathbb{N}$, and therefore also $\left\|\mathsf{R}^{\mathsf{p}} a - a\right\|_{\Psi} \leq \epsilon/3$. □

4.3. Proof of Theorem 4.11

This subsection is devoted to the proof of Theorem 4.11. The ideas used in this proof were motivated by the proof of [Bei11, Lemma 2].

Proof of Theorem 4.11. Let $\mu \in \mathcal{M}(\Psi)$. Since B is a Bohr set, we have by Lemma 4.8 that $\mathsf{d}_\Psi(B)$ exists and is positive. It follows that $\mu(\mathsf{cl}(B)) = \mathsf{d}_\Psi(B) > 0$. Define a new probability measure μ_B on $\beta\mathbb{N}$ by

$$\mu_B(\Omega) := \frac{\mu(\Omega \cap \mathsf{cl}(B))}{\mu(\mathsf{cl}(B))}$$

for all Borel sets $\Omega \subset \beta\mathbb{N}$.

For each $n \in \mathbb{N}$ the map $\mathsf{p} \mapsto (\mathsf{R}^{\mathsf{p}} f)(n) = \lim_{m \to \mathsf{p}} f(n+m)$ from $\beta\mathbb{N} \to \mathbb{R}$ is continuous, and hence measurable. Therefore, so is the map

$$\mathsf{p} \mapsto \limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} h(n)\, (\mathsf{R}^{\mathsf{p}} f)(n),$$

which shows that the set defined in (29) is also measurable. In order to show that the set in (29) has positive measure, it suffices to establish the inequality

$$\int_{\beta\mathbb{N}} \limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} h(n)\, (\mathsf{R}^{\mathsf{p}} f)(n) \, \mathsf{d}\mu_B(\mathsf{p}) \geq 0.$$

Using Fatou's lemma it thus suffices to prove that

$$\limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} h(n) \int_{\beta\mathbb{N}} (\mathsf{R}^{\mathsf{p}} f)(n) \, \mathsf{d}\mu_B(\mathsf{p}) \geq 0. \tag{37}$$

Notice that

$$\left|\int_{\beta\mathbb{N}} (\mathsf{R}^{\mathsf{p}} f)(n) \, \mathsf{d}\mu_B(\mathsf{p})\right| = \frac{1}{\mu(\mathsf{cl}(B))} \left|\int_{\beta\mathbb{N}} 1_{\mathsf{cl}(B)}(\mathsf{p}) (\mathsf{R}^{\mathsf{p}} f)(n) \, \mathsf{d}\mu(\mathsf{p})\right|$$

$$\le \limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} 1_B(m) f(n+m) \right|$$
$$= \limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} 1_{B+n}(m) f(m) \right|.$$

Since $f \in \mathsf{Bes}(\mathbb{N},\Psi)^\perp$ and $m \mapsto 1_{B+n}(m)$ is Besicovitch almost periodic along Ψ by Lemma 4.8, we conclude that

$$\limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} 1_{B+n}(m) f(m) \right| = 0$$

and therefore

$$\left| \int_{\beta\mathbb{N}} (\mathsf{R}^{\mathsf{p}} f)(n) \, \mathsf{d}\mu_B(\mathsf{p}) \right| = 0$$

for every $n \in \mathbb{N}$. This implies (37) and finishes the proof. $\square$

5. The proof over countable amenable groups

The proof of Theorem 1.3 is in broad strokes the same as that for $\mathbb{N}$ given in the previous sections. In this section we discuss the salient differences.

We begin with a discussion of ultrafilters on countable groups. Just as over $\mathbb{N}$, or any other set, an **ultrafilter** on a countable group G is any non-empty family p of non-empty subsets of G that is closed under intersections and supersets, and contains either A or $G\backslash A$ for every $A \subset G$. For each $g \in G$ the collection $\mathsf{p}_g \coloneqq \{A \subset G : g \in A\}$ is an ultrafilter, called the **principal** ultrafilter at g.

Denote by βG the set of all ultrafilters on G. The sets $\mathsf{cl}(A) = \{\mathsf{p} \in \beta G : A \in \mathsf{p}\}$ form a base for a topology on βG that is compact and Hausdorff. Moreover, with this topology βG becomes universal for maps f from G to compact, Hausdorff spaces K in the sense that any such map extends to a continuous map $\beta f \colon \beta G \to K$ with $(\beta f)(\mathsf{p}_g) = f(g)$ for all $g \in G$. We usually write

$$\lim_{g\to\mathsf{p}} f(g) \coloneqq (\beta f)(\mathsf{p})$$

for convenience.

Write $Ag^{-1} = \{h \in G : hg \in A\}$ and $g^{-1}A = \{h \in G : gh \in A\}$ whenever $g \in G$ and $A \subset G$. Write also $A\mathsf{p}^{-1} = \{g \in G : g^{-1}A \in \mathsf{p}\}$ for all $A \subset G$

and all $\mathsf{p} \in \beta G$. With these definitions we have $Ag^{-1} = A\mathsf{p}_g^{-1}$ for all $g \in G$. Multiplication on G extends to βG in two ways. For all p, q in βG both of

$$\mathsf{p} \ltimes \mathsf{q} = \{A \subset G : \{g \in G : g^{-1}A \in \mathsf{q}\} \in \mathsf{p}\}$$
$$\mathsf{p} \rtimes \mathsf{q} = \{A \subset G : \{g \in G : Ag^{-1} \in \mathsf{p}\} \in \mathsf{q}\}$$

define associative binary operations on βG. Using both allows us to generalize Lemma 2.1 to countable groups.

Lemma 5.1. *Fix $A \subset G$. There are non-principal ultrafilters p and q with the property that $A \in \mathsf{p} \ltimes \mathsf{q}$ and $A \in \mathsf{p} \rtimes \mathsf{q}$ if and only if there are infinite sets $B, C \subset G$ with $BC \subset A$.*

Proof. First suppose that $BC \subset A$ for infinite sets $B, C \subset G$. Let p and q be non-principal ultrafilters containing B and C respectively. For all $c \in C$ we have $B \subset Ac^{-1}$ so A belongs to $\mathsf{p} \rtimes \mathsf{q}$. For all $b \in B$ we have $C \subset b^{-1}A$ so A also belongs to $\mathsf{p} \ltimes \mathsf{q}$.

Conversely, suppose that we can find non-principal ultrafilters p and q with A belonging to both $\mathsf{p} \ltimes \mathsf{q}$ and $\mathsf{q} \rtimes \mathsf{q}$. Thus $\{g \in G : g^{-1}A \in \mathsf{q}\} \in \mathsf{p}$ and $\{g \in G : Ag^{-1} \in \mathsf{p}\} \in \mathsf{q}$. We construct injective sequences $n \mapsto b_n$ and $n \mapsto c_n$ in G such that $b_i c_j \in A$ for all $i, j \in \mathbb{N}$. First choose $b_1 \in G$ with $b_1^{-1}A \in \mathsf{q}$. Next, choose $c_1 \in G$ from

$$b_1^{-1}A \cap \{y \in G : Ay^{-1} \in \mathsf{p}\}$$

which is possible since both sets above belong to q. Next, choose $b_2 \in G$ from

$$Ac_1^{-1} \cap \{g \in G : b_2^{-1}A \in \mathsf{q}\}$$

and not equal to b_1, choose $c_2 \in G$ from

$$b_1^{-1}A \cap b_2^{-1}A \cap \{g \in G : Ag^{-1} \in \mathsf{p}\}$$

not equal to c_1 and so on. We can choose at each step a never before chosen element of G because all intersections belong to non-principal ultrafilters and are therefore infinite. □

The first step in the proof of Theorem 1.3 is the following reformulation, which involves multiplication by elements of G from both the left and the right. Because of this we need to work with two-sided Følner sequences. We would like to know whether Theorem 1.3 also holds for one-sided Følner sequences.

Theorem 5.2. *Let G be a countable, amenable group and fix $A \subset G$. If there exist a two-sided Følner sequence Φ on G and a non-principal ultrafilter $\mathsf{p} \in \beta G$ such that $\mathsf{d}_\Phi(Ag^{-1} \cap A\mathsf{p}^{-1})$ exists for all $g \in G$ and*

$$\lim_{g \to \mathsf{p}} \mathsf{d}_\Phi(Ag^{-1} \cap A\mathsf{p}^{-1}) > 0 \tag{38}$$

then there exist infinite sets B, C such that $A \supset BC$.

Proof. Suppose that Φ and p are as in the hypothesis with (38) true. Take $L = A\mathsf{p}^{-1}$. Then $g^{-1}A \in \mathsf{p}$ for every $g \in L$. We can find $\epsilon > 0$ such that

$$\{g \in G : \mathsf{d}_\Phi(Ag^{-1} \cap L) > \epsilon\}$$

belongs to p and is therefore infinite. It follows that

$$\{g \in G : \mathsf{d}_\Phi(Ag^{-1} \cap L) > \epsilon\} \cap \bigcap_{h \in F} h^{-1}A$$

is infinite for any finite set $F \subset L$.

Let $F_1 \subset F_2 \subset \cdots$ be an increasing exhaustion of L by finite subsets. Construct a sequence $n \mapsto e_n$ in G of distinct elements such that

$$e_n \in \{g \in G : \mathsf{d}_\Phi(Ag^{-1} \cap L) > \epsilon\} \cap \bigcap_{h \in F_n} h^{-1}A$$

for each $n \in \mathbb{N}$. This can be done because each of the sets above is infinite by hypothesis.

In particular $\mathsf{d}_\Phi(Ae_n^{-1} \cap L) > \epsilon$ for all $n \in \mathbb{N}$. The Bergelson intersectivity lemma (Corollary 2.4) implies that, for some subsequence $n \mapsto e_{\sigma(n)}$ of e the intersection

$$\left(Ae_{\sigma(1)}^{-1} \cap L\right) \cap \cdots \cap \left(Ae_{\sigma(n)}^{-1} \cap L\right)$$

is infinite for all $n \in \mathbb{N}$.

Choose $b_1 \in F_{\sigma(1)}$ and put $j_1 = 1$. Choose $c_1 = e_{\sigma(1)}$. Thus $c_1 \in b_1^{-1}A$. Next choose $b_2 \in Ac_1^{-1} \cap L$ outside $F_{\sigma(1)}$ and let j_2 be minimal with $b_2 \in F_{\sigma(j_2)}$. (In particular b_2 is not equal to b_1.) Then choose $c_2 = e_{\sigma(j_2)} \in b_1^{-1}A \cap b_2^{-1}A$. Continue this process inductively, choosing

$$b_{n+1} \in Ac_1^{-1} \cap \cdots \cap Ac_n^{-1} \cap L = Ae_{\sigma(j_1)}^{-1} \cap \cdots \cap Ac_{\sigma(j_n)}^{-1} \cap L$$

outside $F_{\sigma(j_n)}$ and choosing j_{n+1} minimal with $b_{n+1} \in F_{\sigma(j_{n+1})}$ and then choosing

$$c_{n+1} = e_{\sigma(j_{n+1})} \in b_1^{-1}A \cap \cdots \cap b_{n+1}^{-1}A$$

which is distinct from $c_1, \ldots, c_n$ because e is injective. Take $B = \{b_n : n \in \mathbb{N}\}$ and $C = \{c_n : n \in \mathbb{N}\}$ to conclude the proof. $\square$

Our goal, given $A \subset G$ with positive upper density, is to find an ultrafilter p and a two-sided Følner sequence Φ satisfying (38). To do this we work in the space

$$\mathsf{L}^2(G,\Phi) = \{f\colon G \to \mathbb{C} : \|f\|_\Phi < \infty\}$$

where $\|f\|_\Phi$ is the **Besicovitch seminorm** of f along a two-sided Følner sequence Φ on G defined as

$$\|f\|_\Phi = \left(\limsup_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} |f(g)|^2 \right)^{1/2}$$

for all $f\colon G \to \mathbb{C}$. Given $f, h \in \mathsf{L}^2(G,\Phi)$ write also

$$\langle f,\ h\rangle_\Phi = \lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} f(g)\overline{h(g)}$$

whenever the limit exists. Given a bounded function $f\colon G \to \mathbb{C}$ define, for all $g \in G$, the shift $\mathsf{R}^g f\colon G \to \mathbb{C}$ by $(\mathsf{R}^g f)(h) \coloneqq f(hg)$ for all $g \in G$ and, for all $\mathsf{p} \in \beta G$, the function $\mathsf{R}^{\mathsf{p}} f\colon G \to \mathbb{C}$ by $(\mathsf{R}^{\mathsf{p}} f)(h) \coloneqq \lim_{g\to\mathsf{p}} f(hg)$ for all $h \in G$. One can check that the function $\mathsf{R}^{\mathsf{p}} 1_A$ is the indicator function of $A\mathsf{p}^{-1}$. Our ultimate goal is now reformulated in terms of $\mathsf{L}^2(G,\Phi)$ and R in the following theorem, which is analogous to Theorem 2.7.

Theorem 5.3. *Let G be a countable amenable group and fix $A \subset G$. Let Φ be a two-sided Følner sequence on G such that $\mathsf{d}_\Phi(A)$ exists. For every $\epsilon > 0$ there exists a subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta G$ such that $\langle \mathsf{R}^g 1_A,\ \mathsf{R}^{\mathsf{p}} 1_A\rangle_\Psi$ exists for all $g \in G$ and*

$$\lim_{g\to\mathsf{p}} \langle \mathsf{R}^g 1_A,\ \mathsf{R}^{\mathsf{p}} 1_A\rangle_\Psi \geq \langle 1,\ 1_A\rangle_\Psi^2 - \epsilon \tag{39}$$

holds.

As over $\mathbb{N}$ we will need to split 1_A into structured and pseudo-random components in two ways. For the first we use finite dimensional representations to define an analogue of trigonometric polynomials.

Definition 5.4. By a **matrix coefficient** of a countable group G we mean any map $a\colon G \to \mathbb{C}$ of the form $a(g) = \langle v,\ M(g)w\rangle$ for some homomorphism M from G to the unitary group $U(n)$ over $\mathbb{C}^n$ and some vectors $v, w \in \mathbb{C}^n$ for some $n \in \mathbb{N}$. A function $f\colon G \to \mathbb{C}$ is **Besicovitch almost periodic** along a two-sided Følner sequence Φ on G if, for every $\epsilon > 0$, one can find a matrix coefficient a with $\|f - a\|_\Phi < \epsilon$.

Definition 5.5. The set $\mathsf{Bes}(G,\Phi)^\perp$ is defined to consist of those functions $f \in \mathsf{L}^2(G,\Phi)$ such that

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} f(g)a(g) = 0$$

for all matrix coefficients a.

Write $\mathsf{Bes}(G,\Phi)$ for the set of functions f in $\mathsf{L}^2(G,\Phi)$ that are Besicovitch almost periodic along Φ. We have the following splitting result.

Theorem 5.6. *For every two-sided Følner sequence Φ on G and any $f \in \mathsf{L}^2(G,\Phi)$ there is a subsequence Ψ of Φ and a function f_{Bes} in $\mathsf{L}^2(G,\Psi)$ which is Besicovitch almost periodic along Ψ, and such that $f - f_{\mathsf{Bes}} \in \mathsf{Bes}(G,\Psi)^\perp$. Moreover, if f takes values in an interval $[a,b] \subset \mathbb{R}$ then so does f_{Bes}.*

Proof. The definition of a projection family makes sense, and the proof of Theorem 3.9 goes through, without complication with $\mathbb{N}$ replaced by G. It therefore suffices, in order to prove the result in question, to show that $\Phi \mapsto \mathsf{Bes}(G,\Phi)$ is a projection family.

The only property that is not immediate is that the inner product $\langle a,b\rangle_\Phi$ exists whenever a,b are matrix coefficients. This follows from an application of the mean ergodic theorem; alternatively we provide the following short self contained proof. Write $a(g) = \langle v,\ M(g)w\rangle$ and $b(g) = \langle r,\ \tilde{M}(g)s\rangle$ for homomorphisms $M\colon G \to U(n)$ and $\tilde{M}\colon G \to U(m)$ and appropriate vectors r,s,u,v. Then $a(g)b(g)$ is a matrix coefficient for the tensor product representation $M \otimes \tilde{M}$ on $\mathbb{C}^{nm}$.

Now, if $a(g) = \langle v,\ M(g)w\rangle$ is any matrix coefficient the average

$$\frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} a(g) = \left\langle v, \frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} M(g)w \right\rangle$$

converges because, for all two-sided Følner sequences Φ the sequence

$$N \mapsto \frac{1}{|\Phi_N|} \sum_{g\in\Phi_N} \delta_{M(g)}$$

of probability measures on $U(n)$ converges in the weak topology to Haar measure on the closure of the image of M. $\square$

The second splitting theorem is proved exactly as in Section 3.3. We formulate here the appropriate generalizations of compact and weak mixing function.

Definition 5.7. A function $f \in \mathsf{L}^2(G,\Phi)$ is **compact** along Φ if, for every $\epsilon > 0$, one can find $F \subset G$ finite with $\min\{\|\mathsf{R}^g f - \mathsf{R}^h f\|_\Phi : h \in F\} < \epsilon$ for all $g \in G$.

Definition 5.8. A function $f \in \mathsf{L}^2(G,\Phi)$ is **weak mixing** along Φ if, for every bounded function $h\colon G \to \mathbb{C}$ and every subsequence Ψ of Φ such that $\langle \mathsf{R}^g f,\ h\rangle_\Psi$ exists for all $g \in G$, the set $\{g \in G : |\langle \mathsf{R}^g f,\ h\rangle_\Psi| > \epsilon\}$ has zero density with respect to every two-sided Følner sequence on G.

The proof of the following theorem is exactly as in Section 3.3. See [Eis+15, Chapter 16] for an appropriate version of the Jacobs–de Leeuw–Glicksberg splitting for unitary representations of groups.

Theorem 5.9. *For every two-sided Følner sequence Φ on G and any $f \in \mathsf{L}^2(G,\Phi)$ there is a subsequence Ψ of Φ and functions $f_{\mathsf{c}}, f_{\mathsf{wm}} \in \mathsf{L}^2(G,\Psi)$ with f_{c} compact along Ψ, f_{wm} weak mixing along Ψ, and $f = f_{\mathsf{c}} + f_{\mathsf{wm}}$. Moreover, if f is real-valued and $a \le f \le b$ for some $a \le b$ then f_{c} is also real valued and satisfies $a \le f_{\mathsf{c}} \le b$.*

The next ingredient in the proof of Theorem 5.3 is an analogue of Theorem 4.1. Its statement over G and how it, together with Theorem 5.9 and Theorem 5.6, imply Theorem 5.3, is exactly the same as the proof of Theorem 2.7 at the end of Section 2. Its proof, also, is just as in Section 4.1 but using the following ingredients.

Definitions 4.2 and 4.3 as well as Lemmas 4.4 and 4.8 make sense in arbitrary countable groups. The next three results – versions of Lemma 4.9, Theorem 4.15 and Theorem 4.11 for countable, amenable groups – fill the remaining gaps in the proof of Theorem 5.3. First we recast Definition 4.5 for countable groups.

Definition 5.10. A **Bohr set** in a group G is any set of the form $a^{-1}(U)$ where a is a homomorphism from G into a compact group K and $U \subset K$ is a non-empty open set whose boundary has Haar measure 0. A Bohr set is a **Bohr$_0$ set** if U contains the identity of K.

For more details on Bohr sets in amenable groups see [BBF10, Subsection 1.3].

Lemma 5.11. *For every $f \in \mathsf{L}^2(G,\Phi)$ that is compact along Φ and every $\epsilon > 0$ the set $\{g \in G : \|\mathsf{R}^g f - f\|_\Phi < \epsilon\}$ contains a Bohr$_0$ set.*

Proof. Since f is compact along Φ the function $\phi\colon g \mapsto \|\mathsf{R}^g f - f\|_\Phi$ has the property that the set $\{\mathsf{R}^h \phi : h \in G\}$ has compact closure with respect to the uniform norm on bounded functions $G \to \mathbb{C}$. By [BJM78, Remark 9.8] there is a compact topological group K and a continuous homomorphism $\xi\colon G \to K$ and a continuous function $\psi\colon K \to \mathbb{C}$ such that $\phi(g) = \psi(\xi(g))$. Therefore the set $\{g \in G : \|\mathsf{R}^g f - f\|_\Phi < \epsilon\}$ contains a Bohr$_0$ set. □

Theorem 5.12. *If $h\colon G \to \mathbb{C}$ is bounded and Besicovitch along Φ then there is a subsequence Ψ of Φ such that $\|\mathsf{R}^{\mathsf{p}} h\|_\Psi = \|h\|_\Psi$ for Ψ almost all p.*

Proof. The proof is unchanged from the $\mathbb{N}$ case, except that we need to verify $\|\mathsf{R}^{\mathsf{p}} a\|_\Psi = \|a\|_\Psi$ for all ultrafilters p, all two-sided Følner sequences Ψ and all matrix coefficients $a\colon G \to \mathbb{C}$. Fix $A\colon G \to \mathsf{U}(n)$ and $v, w \in \mathbb{C}^n$ with $a(g) = \langle v,\ A(g)w\rangle$ for all $g \in G$. Let K be the closure of the image of A in $\mathsf{U}(n)$ and let m be its normalized Haar measure. Writing $\psi(k) = \langle v,\ kw\rangle$ for all $k \in K$ we have, as in the proof of Theorem 5.6, that

$$\|a\|_\Psi^2 = \int |\psi|^2 \,\mathsf{dm}$$

for all two-sided Følner sequences Ψ. Since

$$(\mathsf{R}^{\mathsf{p}} a)(h) = \lim_{g \to \mathsf{p}} \langle v,\ A(h)A(g)w\rangle = \langle v,\ A(h)\ell w\rangle$$

for some $\ell \in K$ we have

$$\|\mathsf{R}^{\mathsf{p}} a\|_\Psi^2 = \int |\psi(k\ell)|^2 \,\mathsf{dm}(k) = \int |\psi(k)|^2 \,\mathsf{dm}(k) = \|a\|_\Psi^2$$

by invariance of Haar measure as desired. □

The last theorem – a version of Theorem 4.11 for countable, amenable groups – is proved exactly as in Section 4.3.

Theorem 5.13. *Suppose $f\colon G \to \mathbb{R}$ is a bounded function that is orthogonal to $\mathsf{Bes}(G, \Psi)$. Then for every Bohr set $B \subset G$ and every bounded function $h\colon G \to \mathbb{R}$ the set*

$$\left\{ \mathsf{p} \in \mathsf{Ess}(\Phi) : B \in \mathsf{p} \text{ and } \limsup_{N \to \infty} \frac{1}{|\Psi_N|} \sum_{g \in \Psi_N} h(g)\,(\mathsf{R}^{\mathsf{p}} f)(g) \ge 0 \right\}$$

has positive measure with respect to every $\mu \in \mathcal{M}(\Psi)$.

6. Open questions

Two natural questions, which arise from questions asked by Erdős in [Erd77, Section 6] and [Erd80, p. 105], are as follows.

Question 6.1. Does every set $A \subset \mathbb{N}$ satisfying

$$\limsup_{N\to\infty} \frac{|A \cap \{1,\dots,N\}|}{N} > 0$$

contain a set of the form $t + B + B$ where $t \in \mathbb{N}$ and $B \subset \mathbb{N}$ is infinite?

Question 6.2. Does every set $A \subset \mathbb{N}$ satisfying

$$\limsup_{N\to\infty} \frac{|A \cap \{1,\dots,N\}|}{N} > 0$$

contain a set of the form $t + (B \oplus B)$ where $t \in \mathbb{N}$, $B \subset \mathbb{N}$ is infinite, and $B \oplus B \coloneqq \{b_1 + b_2 : b_1, b_2 \in B,\ b_1 \neq b_2\}$?

It was pointed out to us by Steven Leth that there exists a set of positive upper density which does not contain any set of the form $B+B+t$ for $t \in \mathbb{N}$ and infinite $B \subset \mathbb{N}$. In particular, the answer to question Question 6.1 is negative. An example of such a set is $A = \bigcup_{n=1}^{\infty} \left[4^n, \frac{3}{2}4^n\right]$.

We do not know the answer to Question 6.2. An ultrafilter reformulation of this question was obtained by Hindman in [Hin79b, Section 11]. We also refer the reader to another paper of Hindman [Hin82] which treats this question. Note that an affirmative answer to Question 6.2 implies Conjecture 1.1.

Question 6.3. Suppose $A \subset \mathbb{N}$ has positive upper density. Do there exist infinite sets $B, C, D \subset \mathbb{N}$ such that the sum $B + C + D$ is contained in A? Is it true that for every $k \in \mathbb{N}$ there exist infinite sets $B_1, \dots, B_k \subset \mathbb{N}$ such that $B_1 + \cdots + B_k \subset A$?

The Green–Tao theorem on arithmetic progressions [GT08] gives a version of Szemerédi's theorem in the primes. It is natural to ask (cf. [Gra90]) whether a version of the Erdős sumset conjecture holds for the primes.

Question 6.4. Let $\mathbb{P}$ denote the set of prime numbers. Are there infinite sets $B, C \subset \mathbb{N}$ such that $B + C \subset \mathbb{P}$?

A positive answer to Question 6.4, conditional on the Hardy-Littlewood prime tuples conjecture, was obtained by Granville [Gra90]. (The authors thank Karl Mahlburg for this reference.)

Lastly we pose a more open-ended question which was asked by Jon Chaika.

Question 6.5. Is there a version of Theorem 1.2 over $\mathbb{R}$ or more general locally compact topological groups?

Bibliography

[ACG17] U. Andrews, G. Conant, and I. Goldbring. *Definable sets containing product sets in expansions of groups*. 2017. arXiv: 1701.07791 [math.LO].

[AJ75] M. A. Akcoglu and A. del Junco. "Convergence of averages of point transformations". In: *Proc. Amer. Math. Soc.* 49 (1975), pp. 265–266. ISSN: 0002-9939.

[BBF10] Mathias Beiglböck, Vitaly Bergelson, and Alexander Fish. "Sumset phenomenon in countable amenable groups". In: *Adv. Math.* 223.2 (2010), pp. 416–432. ISSN: 0001-8708.

[Bei11] M. Beiglböck. "An ultrafilter approach to Jin's theorem". In: *Israel J. Math.* 185 (2011), pp. 369–374. ISSN: 0021-2172.

[Ber85] V. Bergelson. "Sets of recurrence of $\mathbf{Z}^m$-actions and properties of sets of differences in $\mathbf{Z}^m$". In: *J. London Math. Soc. (2)* 31.2 (1985), pp. 295–304. ISSN: 0024-6107.

[Ber96] V. Bergelson. "Ergodic Ramsey theory–an update". In: *Ergodic theory of $\mathbb{Z}^d$ actions*. Vol. 228. London Math. Soc. Lecture Note Ser. Cambridge: Cambridge Univ. Press, 1996, pp. 1–61.

[Bes26] A. S. Besicovitch. "On Generalized Almost Periodic Functions". In: *Proc. London Math. Soc. (2)* 25 (1926), pp. 495–512. ISSN: 0024-6115.

[Bes55] A. S. Besicovitch. *Almost periodic functions*. Dover Publications, Inc., New York, 1955, pp. xiii+180.

[BF45] H. Bohr and E. Følner. "On some types of functional spaces. A contribution to the theory of almost periodic functions". In: *Acta Math.* 76 (1945), pp. 31–155. ISSN: 0001-5962.

[BFW06] V. Bergelson, H. Furstenberg, and B. Weiss. "Piecewise-Bohr sets of integers and combinatorial number theory". In: *Topics in discrete mathematics*. Vol. 26. Algorithms Combin. Springer, Berlin, 2006, pp. 13–37.

[BJM78] J. F. Berglund, H. D. Junghenn, and P. Milnes. *Compact right topological semigroups and generalizations of almost periodicity*. Vol. 663. Lecture Notes in Mathematics. Springer, Berlin, 1978, pp. x+243. ISBN: 3-540-08919-5.

[BL85] A. Bellow and V. Losert. "The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences". In: *Trans. Amer. Math. Soc.* 288.1 (1985), pp. 307–345. ISSN: 0002-9947.

[Boh25a] Harald Bohr. "Zur Theorie der Fastperiodischen Funktionen". In: *Acta. Math.* 45.1 (1925). I. Einer Verallgemeinerung der Theorie der Fourierreihen, pp. 29–127.

[Boh25b] Harald Bohr. “Zur Theorie der Fastperiodischen Funktionen”. In: *Acta. Math.* 46.1-2 (1925). II. Zusammenhang der fastperiodischen Funktionen mit Funktionen von unendlich vielen Variabeln; gleichmässige Approximation durch trigonometrische Summen, pp. 101–214.

[DN+15] M. Di Nasso, I. Goldbring, R. Jin, S. Leth, M. Lupini, and K. Mahlburg. “On a sumset conjecture of Erdős”. In: *Canad. J. Math.* 67.4 (2015), pp. 795–809. ISSN: 0008-414X.

[EG80] P. Erdős and R. L. Graham. *Old and new problems and results in combinatorial number theory.* Vol. 28. Monographies de L’Enseignement Mathématique [Monographs of L’Enseignement Mathématique]. Université de Genève, L’Enseignement Mathématique, Geneva, 1980, p. 128.

[Eis+15] T. Eisner, B. Farkas, M. Haase, and R. Nagel. *Operator theoretic aspects of ergodic theory.* Vol. 272. Graduate Texts in Mathematics. Springer, Cham, 2015, pp. xviii+628. ISBN: 978-3-319-16898-2.

[Erd77] P. Erdős. “Problems and results on combinatorial number theory. III”. In: (1977), 43–72. Lecture Notes in Math., Vol. 626.

[Erd80] P. Erdős. “A survey of problems in combinatorial number theory”. In: *Ann. Discrete Math.* 6 (1980). Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978), pp. 89–115.

[Fra15] N. Frantzikinakis. “Multiple correlation sequences and nilsequences”. In: *Invent. Math.* 202.2 (2015), pp. 875–892. ISSN: 0020-9910.

[Fur81] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* M. B. Porter Lectures. Princeton University Press, Princeton, N.J., 1981, pp. xi+203. ISBN: 0-691-08269-3.

[GKR18] Daniel Glasscock, Andreas Koutsogiannis, and Florian K. Richter. *Multiplicative combinatorial properties of return time sets in minimal dynamical systems.* 2018. arXiv: 1809.08702 [math.DS].

[Gla03] Eli Glasner. *Ergodic theory via joinings.* Vol. 101. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2003, pp. xii+384. ISBN: 0-8218-3372-3.

[Gra90] Andrew Granville. “A note on sums of primes”. In: *Canad. Math. Bull.* 33.4 (1990), pp. 452–454. ISSN: 0008-4395.

[Gre69] Frederick P. Greenleaf. *Invariant means on topological groups and their applications.* Van Nostrand Mathematical Studies, No. 16. Van Nostrand Reinhold Co., New York-Toronto, Ont.-London, 1969, pp. ix+113.

[GT08] Ben Green and Terence Tao. “The primes contain arbitrarily long arithmetic progressions”. In: *Ann. of Math. (2)* 167.2 (2008), pp. 481–547. ISSN: 0003-486X.

[GT10] B. Green and T. Tao. “An arithmetic regularity lemma, an associated counting lemma, and applications”. In: *An irregular mind.* Vol. 21. Bolyai Soc. Math. Stud. János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.

[Hin79a] N. Hindman. “Ultrafilters and combinatorial number theory”. In: *Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*. Vol. 751. Lecture Notes in Math. Springer, Berlin, 1979, pp. 119–184.

[Hin79b] Neil Hindman. “Ultrafilters and combinatorial number theory”. In: *Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*. Vol. 751. Lecture Notes in Math. Springer, Berlin, 1979, pp. 119–184.

[Hin82] Neil Hindman. “On density, translates, and pairwise sums of integers”. In: *J. Combin. Theory Ser. A* 33.2 (1982), pp. 147–157. ISSN: 0097-3165.

[HK09] B. Host and B. Kra. “Uniformity seminorms on ℓ^∞ and applications”. In: *J. Anal. Math.* 108 (2009), pp. 219–276. ISSN: 0021-7670.

[HK11] B. Host and B. Kra. “Nil-Bohr sets of integers”. In: *Ergodic Theory Dynam. Systems* 31.1 (2011), pp. 113–142. ISSN: 0143-3857.

[HS12] N. Hindman and D. Strauss. *Algebra in the Stone-Čech compactification*. De Gruyter Textbook. Theory and applications, Second revised and extended edition [of MR1642231]. Walter de Gruyter & Co., Berlin, 2012, pp. xviii+591. ISBN: 978-3-11-025623-9.

[Jac56] K. Jacobs. “Ergodentheorie und fastperiodische Funktionen auf Halbgruppen”. In: *Math. Z.* 64 (1956), pp. 298–338. ISSN: 0025-5874.

[KN32] B. O. Koopman and J. von Neumann. “Dynamical Systems of Continuous Spectra”. In: *Proc. Natl. Acad. Sci.* 18.3 (1932), pp. 255–263.

[Kre85] U. Krengel. *Ergodic theorems*. Vol. 6. De Gruyter Studies in Mathematics. With a supplement by Antoine Brunel. Walter de Gruyter & Co., Berlin, 1985, pp. viii+357. ISBN: 3-11-008478-3.

[Lei05] A. Leibman. “Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold”. In: *Ergodic Theory Dynam. Systems* 25.1 (2005), pp. 201–213. ISSN: 0143-3857.

[LG61] K. de Leeuw and I. Glicksberg. “Applications of almost periodic compactifications”. In: *Acta Math.* 105 (1961), pp. 63–97. ISSN: 0001-5962.

[Lin01] E. Lindenstrauss. “Pointwise theorems for amenable groups”. In: *Invent. Math.* 146.2 (2001), pp. 259–295. ISSN: 0020-9910.

[Nat80] M. B. Nathanson. “Sumsets contained in infinite sets of integers”. In: *J. Combin. Theory Ser. A* 28.2 (1980), pp. 150–155. ISSN: 0097-3165.

[Ruz82] Imre Z. Ruzsa. “Uniform distribution, positive trigonometric polynomials and difference sets”. In: *Seminar on Number Theory, 1981/1982*. Univ. Bordeaux I, Talence, 1982, Exp. No. 18, 18.

[TW16] Grzegorz Tomkowicz and Stan Wagon. *The Banach-Tarski paradox*. Second. Vol. 163. Encyclopedia of Mathematics and its Applications. With a foreword by Jan Mycielski. Cambridge University Press, New York, 2016, pp. xviii+348. ISBN: 978-1-107-04259-9.

[Wal82] Peter Walters. *An introduction to ergodic theory*. Vol. 79. Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1982, pp. ix+250. ISBN: 0-387-90599-5.