Applications of Ultrafilters to Ramsey Theory

by

Cody Johnson

Submitted to the Department of Mathematical Sciences
on May 2, 2019, in partial fulfillment of the
requirements for the degree of
Master of Science in Mathematical Sciences

Abstract

Erdos conjectured in 1980 that any subset $A \subseteq \mathbb{N}$ of positive upper density, i.e.,

$$\limsup_{N\to\infty} \frac{|A \cap \{1, 2, \ldots, N\}|}{N} > 0$$

must contain $B + C$ for somet infinite $B, C \subseteq$, and Joel Moreira, Florian Richter, and Donald Robertson presented an affirmative proof of this fact in 2019. Their proof to this Ramsey-type problem relied on a topological dynamical and functional analytical argument concerning ultrafilters and Følner sequences. One potential extension of their result is whether all sets of positive upper density must contain $B + C + D$ for infinite $B, C, D \subseteq \mathbb{N}$. In this paper, we synthesize Moreira's proof of the Erdos sumset conjecture and then propose some techniques for extending his techniques to higher order versions of the theorem.

Thesis Supervisor: Clinton Conley
Title: Associate Professor, Department of Mathematical Sciences

Acknowledgements

I would first and foremost like to give my warmest thanks to Professor Clinton Conley, who made this work possible, for his passion and skill as my algebra professor and for his inspiration and guidance as my research advisor. Without him, my reseach and enthusiasm for algebra would have never been possible. I also wish to express my gratitude to Professor James Cummings and Professor Florian Frick for serving on my thesis committee. I also thank Alex Zheng for presenting me this interesting problem and Professor John Mackey and Professor Boris Bukh for initial guidance on the problem. Finally, I would like to thank my mother for supporting all of my pursuits and always bringing out the best in me.

Cody Johnson
May 2, 2019

Contents

Chapter 1

Introduction

1.1 Motivations

The Erdos sumset conjecture asks whether every set $A \subseteq \mathbb{N}$ with positive upper density contains a $B+C$ for some infinite $B, C \subseteq \mathbb{N}$. In [MRR19], the author proves the affirmative by using techniques in topological dynamics and ultrafilters, which is interesting due to its seemingly distant relationship with additive combinatorics and Ramsey theory. The application of topological dynamics to Ramsey theory initially became of interest due to another paper of Moreira, [Mor17], concerning whether every coloring of $\mathbb{N}$ contains a monochromatic $\{x, y, x+y, xy\}$. In this paper, it was surprising to see his partial solution to this problem involving the correspondence of so-called *piecewise syndetic* sets, a combinatorial object, with minimal subsystems of $(\beta\mathbb{N}, T)$ (under the action $T : \beta\mathbb{N} \to \beta\mathbb{N},\ \mathsf{p} \mapsto \mathsf{p}+1$), a topological dynamical system. The application of topological dynamics to Ramsey theory was intriguing, especially since he even directly converted his topological dynamical result into a combinatorial construction in the paper. This motivates seeking other applications of topological dynamics to Ramsey theory, which leads to this problem. The Erdos sumset paper was particularly interesting because it immediately reduced the combinatorial Ramsey theory problem into a measure theoretic bounding problem in the ultrafilter reformulation, which is more suitable for an analytic background.

1.2 Ultrafilter Prerequisites

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

$$A \in \mathsf{p} \iff \mathbb{N} \setminus A \in \mathsf{p}$$

for every $A \subseteq \mathbb{N}$. Ultrafilters appear in various different contexts:

- *Algebraic*: A *filter* on a set X is a collection $\mathcal{F}$ of subsets of X such that:
    1. $X \in \mathcal{F}$.
    2. $\varnothing \notin \mathcal{F}$.
    3. If $A \in \mathcal{F}$ and $A \subseteq B$, then $B \in \mathcal{F}$.
    4. If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

Then an ultrafilter p is a maximal filter on X, that is, a filter not contained in any other filter on $\mathcal{F}$. By Zorn's lemma, every filter is contained in an ultrafilter.

- *Probability theoretic*: An ultrafilter p on a set X is a finitely additive probability measure in which every event has probability either 0 or 1. Intuitively, it is a consistent choice of which subsets of X are "large."

- *Category theoretic*: A *Boolean algebra* is a poset B equipped with distinguished elements 0 and 1, binary operations $\vee$ ("join") and $\wedge$ ("meet"), and a unary operation $\neg$ ("complementation") satisfying six axioms to mimic the familiar $\cup$, $\cap$, and $X\setminus$ operations on the poset $(\mathcal{P}(X), \subseteq)$ for some set X. In the category **Bool** of Boolean algebras, the maps $b : B \to \mathbf{2}$ into the initial Boolean algebra $\mathbf{2}$ correspond to the ultrafilters on B, with $\mathsf{p} = b^{-1}(1)$ being an ultrafilter given $b : B \to \mathbf{2}$ and $A \mapsto 1_{A\in\mathsf{p}} : B \to \mathbf{2}$ being a map into the initial Boolean algebra given p.

- *Topological*: The *Stone-Čech compactification* of a topological space X is a compact Hausdorff space βX together with a continuous map with the following universal property: for any compact Hausdorff space K and any continuous map $f : X \to K$, f extends uniquely to a continuous map $\beta f : \beta X \to K$. Then the set of ultrafilters $\beta\mathbb{N}$ on $\mathbb{N}$ corresponds to the Stone-Čech compactification of $\mathbb{N}$ under the discrete topology.

For the purposes of this expository, we only consider ultrafilters on $\mathbb{N}$. For each $n \in \mathbb{N}$, the collection $\mathsf{p}_n := \{A \subseteq \mathbb{N} : n \in A\}$ is an ultrafilter; ultrafilters of this form are called *principal* and are usually denoted by n when it is not ambiguous. Non-principal ultrafilters are of particular interest because they are precisely the ultrafilters that contain no finite sets. To develop the topology on $\beta\mathbb{N}$, let $\mathsf{cl}(A) = \{\mathsf{p} \in \beta\mathbb{N} : A \in \mathsf{p}\}$ be the *closure* of A in $\beta\mathbb{N}$. Then the set $\{\mathsf{cl}(A) : A \subseteq \mathbb{N}\}$ forms a base for a topology on $\beta\mathbb{N}$ since cl preserves $\cup$ and $\cap$, and $\beta\mathbb{N} = \mathsf{cl}(\mathbb{N}) \subseteq \bigcup_{A\subseteq\mathbb{N}} \mathsf{cl}(A)$. As suggested above, we have the following crucial fact about $\beta\mathbb{N}$:

Theorem 1.2.1. *$\beta\mathbb{N}$ forms a compact Hausdorff space.*

Proof. First, we show that $\beta\mathbb{N}$ is compact. Let $\mathcal{K} = \{U_\alpha\}_{\alpha\in I}$ be an open covering of $\beta\mathbb{N}$, and assume for the sake of contradiction that $\mathcal{K}$ has no finite subcover. Assume without loss of generality that $U_\alpha = \mathsf{cl}(A_\alpha)$ for some $A_\alpha \subseteq \mathbb{N}$ for all $\alpha \in I$. Let $\mathcal{F} = \{\mathbb{N} \setminus A_\alpha : \alpha \in I\}$, and case on whether or not $\mathcal{F}$ has the finite intersection property.

Suppose that $\mathcal{F}$ has the finite intersection property. We claim that there exists an ultrafilter p such that $\mathbb{N} \setminus A_\alpha \in \mathsf{p}$ for all $\alpha \in I$. Consider the poset $\tilde{\mathcal{F}}$ of all families $\mathcal{F}' \subseteq \mathcal{P}(\mathbb{N})$ that have the finite intersection property and that contain $\mathcal{F}$. Then $\tilde{\mathcal{F}}$ is nonempty and closed under unions of chains, so by Zorn's lemma there exists a maximal member p of $\tilde{\mathcal{F}}$. Since p is an ultrafilter containing $\mathcal{F}$, we know $A_\alpha \notin \mathsf{p}$ for all $\alpha \in I$. This contradicts that $\mathsf{p} \in \mathsf{cl}(A_\alpha)$ for some $\alpha \in I$.

On the other hand, suppose that $\mathcal{F}$ does not have the finite intersection property. Then let $\varnothing \neq J \subseteq I$ be finite such that $\varnothing = \bigcap_{\alpha\in J}(\mathbb{N} \setminus A_\alpha) = \mathbb{N} \setminus \bigcup_{\alpha\in J} A_\alpha$. Then $\bigcup_{\alpha\in J} A_\alpha = \mathbb{N}$, so

$$\beta\mathbb{N} = \mathsf{cl}(\mathbb{N}) = \mathsf{cl}\left(\bigcup_{\alpha\in J} A_\alpha\right) = \bigcup_{\alpha\in J} \mathsf{cl}(A_\alpha),$$

which contradicts that $\mathcal{K}$ has no finite subcover.

Now, we show that $\mathbb{N}$ is Hausdorff. Let $\mathsf{p}, \mathsf{q} \in \beta\mathbb{N}$ be distinct ultrafilters. Then there exists some $A \in \mathsf{p} \setminus \mathsf{q}$, so $\mathbb{N} \setminus A \in \mathsf{q} \setminus \mathsf{p}$. Then $\mathsf{cl}(A)$ and $\mathsf{cl}(\mathbb{N} \setminus A)$ are disjoint neighborhoods of p and q, respectively. ☐

Corollary 1. *There exist non-principal ultrafilters.*

Proof. $\mathbb{N}$ under the discrete topology embeds into $\beta\mathbb{N}$ via the map $n \mapsto \mathsf{p}_n$, and $\mathbb{N}$ is not compact. ☐

The study of ultrafilters naturally leads to the study of Ramsey theory given the following property.

Lemma 1.2.1. *Given an ultrafilter $\mathsf{p} \in \beta\mathbb{N}$, if $r \in \mathbb{N}$ and $X = A_1 \cup \dots \cup A_r$ for some $X \in \mathsf{p}$ and $A_1, \dots, A_r \subseteq X$, then $A_i \in \mathsf{p}$ for some $1 \le i \le r$.*

Proof. Suppose $A_i \notin \mathsf{p}$ for all $1 \le i \le r$, so $X \setminus A_i \in \mathsf{p}$ for all $1 \le i \le r$. Hence, $\mathsf{p} \ni \bigcap_{i=1}^n (X \setminus A_i) = X \setminus \bigcup_{i=1}^n A_i = \varnothing$, contradiction. ☐

As an easy example, consider the following two proofs of the infinite Ramsey theorem.

Theorem 1.2.2 (Infinite Ramsey)**.** *Let $r \in \mathbb{N}$ and $\chi : E(K_{\mathbb{N}}) \to \{1, \dots, r\}$ be a coloring of the edges of $K_{\mathbb{N}}$. Then there exists an infinite $X \subseteq \mathbb{N}$ and a color $1 \le c \le r$ such that $\chi(e) = c$ for all $e \in E(K_X)$.*

Proof 1 (standard). We will construct three sequences inductively: $\{C_n\}_{n\ge0} \subseteq \mathcal{P}(\mathbb{N})$, $\{x_n\}_{n\ge1} \subseteq \mathbb{N}$, and $\{c_n\}_{n\ge1} \subseteq \{1, \dots, r\}$. Let $C_0 := \mathbb{N}$. Given $n \ge 1$, choose $x_n \in C_{n-1}$ arbitrarily. By the infinite pigeonhole principle, there exists some color $1 \le c_n \le r$ such that $\{y \in C_{n-1} : y > x_n \wedge \chi(\{x_n, y\}) = c_n\} =: C_n$ is infinite. Note that by construction, $C_0 \supseteq C_1 \supseteq \dots$ and $x_1 < x_2 < \dots$.

By the infinite pigeonhole principle again, there exists some color $1 \le c \le r$ that appears infinitely many times in $\{c_n\}_{n\ge1}$. Let $\{n_i\}_{i\ge1}$ be the indices for which $c_{n_i} = c$, and let $X = \{x_{n_i} : i \ge 1\}$. Given $\{x, y\} \in E(K_X)$, we have $x = x_{n_i}$ and $y = x_{n_j}$ for $i < j$, without loss of generality. Then $n_i < n_j$, so $x_{n_j} \in C_{n_j} \subseteq C_{n_i}$. Therefore, $\chi(\{x_{n_i}, x_{n_j}\}) = c$. Therefore, X verifies the theorem statement. ☐

Proof 2 (using ultrafilters). Let p be a non-principal ultrafilter on $\mathbb{N}$, which we have already shown exists. For each $x \in \mathbb{N}$, we construct a color $1 \le c_x \le r$ and a set $A_x \in \mathsf{p}$ in the following way: for each $1 \le i \le r$, define $A_i = \{y \in \mathbb{N} \setminus \{x\} : \chi(\{x, y\}) = i\}$. Then $\mathbb{N} = \{x\} \cup (\mathbb{N} \setminus \{x\})$, so either $\{x\} \in \mathsf{p}$ or $\mathbb{N} \setminus \{x\} \in \mathsf{p}$. Since p is principal, the latter must be true. Then, note that $\mathbb{N} \setminus \{x\} = A_1 \cup \dots \cup A_r$, so $A_{c_x} \in \mathsf{p}$ for some $1 \le c_x \le r$.

Now for each $1 \le i \le r$, let $B_i = \{x \in \mathbb{N} : c_x = i\}$. Note that $\mathbb{N} = B_1 \cup \dots \cup B_r$, so $B_c \in \mathsf{p}$ for some $1 \le i \le r$. From here, we can construct $\{C_n\}_{n\ge1} \subseteq \mathsf{p}$ and $\{x_n\}_{n\ge1} \subseteq \mathbb{N}$, inductively using the recurrence $C_1 = B_c$, $x_n = \min C_n$, and $C_{n+1} = (C_n \setminus \{x_n\}) \cap A_{x_n} \in \mathsf{p}$. By construction, $\chi(\{x_i, x_j\}) = c$ for all $1 \le i < j$, so $X = \{x_n\}_{n\ge1}$ verifies the theorem statement. ☐

In light of the universal property given by Stone-Čech compactification, given a function $f : \mathbb{N} \to K$ into a compact metric space K, there exists a unique continuous $\beta f : \beta\mathbb{N} \to K$ such that $\beta f(\mathsf{p}_n) = f(n)$ for all $n \in \mathbb{N}$. We can characterize $(\beta f)(\mathsf{p})$ as the unique point $x \in K$ such that $f^{-1}(U) \in \mathsf{p}$ for all neighborhoods U of p. Therefore, we will denote

$$\lim_{n \to \mathsf{p}} f(n) := (\beta f)(\mathsf{p})$$

throughout this expository. Further, define

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

This allows us to extend $+$ to a binary operation on $\beta\mathbb{N}$ given by

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

It's easy to verify that this operation corresponds to regular addition for principal ultrafilters, namely that $\mathsf{p}_n + \mathsf{p}_m = \mathsf{p}_{n+m}$ for all $n, m \in \mathbb{N}$.

Chapter 2

Proof of the Erdos sumset conjecture

The following section is a synthesis of the work of Joel Moreira in [MRR19].

2.1 Setup and ultrafilter reformulation

Moreira's paper set out to prove Erdos' sumset conjecture:

Conjecture 1. If $A \subseteq \mathbb{N}$ has positive upper density, i.e.,

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

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

In order to do this, Moreira showed a generalization of this conjecture with respect to Følner sequences. A *Følner sequence* is a sequence $\Phi : N \mapsto \Phi_N$ of finite, non-empty subsets of $\mathbb{N}$ satisfying

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

For instance, any sequence $N \mapsto \{a_N + 1, a_N + 2, \dots, b_N\}$ of intervals in $\mathbb{N}$ with $\lim_{N\to\infty}(b_N - A_N) \to \infty$ is a Følner sequence since

$$\begin{aligned} b_N - a_N > m \implies &|\{a_N + m, \dots, b_N + m\}\triangle\{a_N, \dots, b_N\}| \\ =&|\{a_N, \dots, a_N + m - 1\} \cup \{b_N + 1, \dots, b_N + m\}| = 2m = o(b_N - a_N) \end{aligned}$$

Then the *upper density* of a set $A \subseteq \mathbb{N}$ with respect to a Følner sequence Φ is defined to be the quantity

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

and denoted $\mathsf{d}_\Phi(A)$ when the limit

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

exists. The main result of Moreira's paper is the following theorem, which directly implies the Erdos sumset conjecture applied to the Følner sequence $N \mapsto \{1, 2, \dots, N\}$.

Theorem 2.1.1. *For every $A \subseteq \mathbb{N}$ that satisfies $\overline{\mathsf{d}}_\Phi(A) > 0$ for some Følner sequence Φ, one can find infinite sets $B, C \subseteq \mathbb{N}$ with $B + C \subseteq A$.*

The general approach taken to solving this problem is first to reformulate this in terms of finding an ultrafilter satisfying a specific property, which will then be reduced to lower-bounding a certain inner product, which can then be tackled using tools from functional analysis, specifically analyzing the pseudo-random and structured components of functions in $\mathsf{L}^2(\mathbb{N}, \Phi)$. To begin with, we will prove the following ultrafilter reformulation of the problem:

Theorem 2.1.2. *Let $A \subseteq \mathbb{N}$. If there exists a Følner sequence Φ in $\mathbb{N}$ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\mathsf{d}_\Phi((A-n) \cap (A-\mathsf{p}))$ exists for all $n \in \mathbb{N}$ and*

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

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

With this ultrafilter reformulation, given $A \subseteq \mathbb{N}$ with $\overline{d}_\Phi(A) > 0$, we have reduced Theorem 2.1.1 to finding a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ and a Følner subsequence Ψ of Φ verifying the statement of Theorem 2.1.2. We will then use the following theorem to find such a Ψ, p:

Theorem 2.1.3. *Let $A \subseteq \mathbb{N}$ and let Φ be a Følner sequence on $\mathbb{N}$ with $\mathsf{d}_\Phi(A)$ existing. Then for every $\varepsilon > 0$, 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*

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

And finally, from this, we will reduce the proof of Theorem 2.1.3 to a functional analysis result. First, given a bounded function $f : \mathbb{N} \to \mathbb{C}$, define the shift operator

$$\mathsf{R}^m f : \mathbb{N} \to \mathbb{C}, \quad (\mathsf{R}^m f)(n) := f(n+m)$$

which is readily extended to all ultrafilters

$$\mathsf{R}^\mathsf{p} f : \mathbb{N} \to \mathbb{C}, \quad (\mathsf{R}^\mathsf{p} f)(n) := \lim_{m \to \mathsf{p}} f(n+m)$$

where, not surprisingly, $\mathsf{R}^{\mathsf{p}_m} f = \mathsf{R}^m f$ for all principal ultrafilters p_m. Then, given a Følner sequence Φ in $\mathbb{N}$, define the *Besicovitch seminorm* of $f : \mathbb{N} \to \mathbb{C}$ along Φ via

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

and the inner product of $f, h : \mathbb{N} \to \mathbb{C}$ along Φ via

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

whenever the limit exists. Then we reduce Theorem 2.1.3 to proving the following theorem.

Theorem 2.1.4. *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. Then for every $\varepsilon > 0$, there exists a Følner 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*

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

2.1.1 Proof of Theorem 2.1.2

The goal of Theorem 2.1.2 is to inductively construct a sequence $b_1, c_1, b_2, c_2, \ldots$ such that

$$b_{n+1} \in (A - c_1) \cap \cdots \cap (A - c_n) \quad \text{and} \quad c_{n+1} \in (A - b_1) \cap \cdots \cap (A - b_n)$$

for all $n \in \mathbb{N}$. We will find a set $L \subseteq \mathbb{N}$ of target values for $\{b_n\}_{n\in\mathbb{N}}$ and a sequence $\{e_n\}_{n\in\mathbb{N}} \subseteq \mathbb{N}$ from which to choose $\{c_n\}_{n\in\mathbb{N}}$ as a subsequence where $c_n = e_{j_n}$. The idea behind the construction is to make $L \subseteq \mathbb{N}$ and $\{e_n\}_{n\in\mathbb{N}}$ satisfy:

- $(A - e_1) \cap \cdots \cap (A - e_n) \cap L$ is infinite so that it's easy to choose

$$b_{n+1} \in (A - e_1) \cap \cdots \cap (A - e_{j_n}) \cap L \supseteq (A - c_1) \cap \cdots \cap (A - c_n) \cap L$$

- Given an exhaustion of L by finite sets $F_1 \subset F_2 \subset \ldots$, $e_n \in \bigcap_{\ell \in F_n} (A - \ell)$ so that it's easy to choose $c_n = e_{j_n}$ with $j_n \in \mathbb{N}$ minimal such that $b_n \in F_{j_n}$.

The only thing we have to be careful is about choosing L and choosing e_n so that $(A - e_1) \cap \cdots \cap (A - e_n) \cap L$ is infinite. For the former, the construction will use $L = A - \mathsf{p}$; for the latter, we will devise a way to choose e_n so that $(A - e_1) \cap \cdots \cap (A - e_n) \cap L$ has positive upper density which implies it's infinite. We will do this in two parts.

Proposition 2.1.1. *Suppose $A_n \in \mathbb{N}$ such that $\mathsf{d}_\Phi(A_n)$ exists for all $n \in \mathbb{N}$, and $\mathsf{d}_\Phi(A_n) \geq \varepsilon$. Then there exists an injective $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

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

Proposition 2.1.2. *Suppose $A \subseteq \mathbb{N}$ and there exists $L \in \mathbb{N}$ such that $\mathsf{d}_\Phi((A - m) \cap L)$ exists for all $m \in \mathbb{N}$, and*

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

is infinite for all finite $F \subseteq L$. Then there exist infinite sets $B, C \subseteq \mathbb{N}$ such that $B + C \subseteq A$.

Proposition 2.1.1 acts to guarantee that the interesection in Proposition 2.1.2 is infinite, and Proposition 2.1.2 acts to prove Theorem 2.1.2 by taking $L = A - \mathsf{p}$.

Proof of Proposition 2.1.1. We borrow techniques from measure theory. Let $B_n = \mathsf{cl}(A_n)$, and we will define a probability measure μ on $\beta\mathbb{N}$ such $\mathsf{d}_\Phi(A_n) = \mu(B_n)$ for all $n \in \mathbb{N}$. This can be achieved by taking μ to be a Radon probability measure on $\beta\mathbb{N}$ that is a weak limit of some subsequence of the measures $\frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \delta_n$ where δ_n is the unit mass on p_n. We will show that for some injective $\sigma : \mathbb{N} \to \mathbb{N}$, we have $\mu(B_{\sigma(1)} \cap B_{\sigma(n)}) > 0$, which proves the proposition since

$$\overline{\mathsf{d}}(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)}) \geq \mu(\mathsf{cl}(A_{\sigma(1)} \cap \cdots \cap A_{\sigma(n)})) = \mu(B_{\sigma(1)} \cap \cdots \cap B_{\sigma(n)}) > 0$$

Consider the probability space $(\beta\mathbb{N}, \mathcal{B}, \mu)$, where $\mathcal{B}$ is the Borel σ-algebra on $\beta\mathbb{N}$. Let $\mathcal{F}$ be the collection of all $F \subseteq \mathbb{N}$ such that $|F|$ is finite and $\mu(\bigcap_{n \in F} B_n) = 0$. Then there are countably many elements of $\mathcal{F}$, and hence the set $\bigcup_{F \in \mathcal{F}} \bigcap_{n \in F} B_n$ has measure 0. Therefore, it suffices to show that the set

$$\{x \in \beta\mathbb{N} : x \in B_n \text{ for infinitely many } n \in \mathbb{N}\}$$

has positive measure. By Fatou's lemma, we have

$$\int_{\beta\mathbb{N}} \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_{B_n}\ d\mu \geq \limsup_{N\to\infty} \int_{\beta\mathbb{N}} \frac{1}{N}\sum_{n=1}^{N} 1_{B_n}\ d\mu = \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \int_{\beta\mathbb{N}} 1_{B_n}\ d\mu \geq \varepsilon > 0$$

and hence the set

$$\left\{x \in \beta\mathbb{N} : \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_{B_n} > 0\right\} \subseteq \{x \in \beta\mathbb{N} : x \in B_n \text{ for infinitely many } n \in \mathbb{N}\}$$

has positive measure. This completes the proof. □

Proof of Proposition 2.1.2. Let $F_1 \subset F_2 \subset \cdots \subseteq L$ be an increasing exhaustion of L by finite subsets. Construct $\{e_n\}_{n\in\mathbb{N}}$ inductively by choosing

$$e_n \in \bigcap_{\ell\in F_n} (A-\ell) \cap \{m \in \mathbb{N} : \mathsf{d}_\Phi((A-m)\cap L) > \varepsilon\} \setminus \{e_1, \ldots, e_{n-1}\}$$

which exists because the right-hand side is infinite. Since $\mathsf{d}_\Phi((A-e_n)\cap L) > \varepsilon$ for all $n \in \mathbb{N}$, Proposition 2.1.1 allows us to find an injective $\sigma : \mathbb{N} \to \mathbb{N}$ such that

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

has positive upper density with respect to Φ, and hence it is infinite. Now consider the following construction:

1. Choose $b_1 \in F_{\sigma(1)}$ arbitrarily, $j_1 = 1$, and $c_1 = e_{\sigma(j_1)}$.
2. Choose $b_{n+1} \in (A-c_1)\cap\cdots\cap(A-c_n)\cap L \setminus F_{\sigma(j_n)}$ arbitrarily, which exists since
$$(A-c_1)\cap\cdots\cap(A-c_n)\cap L \supseteq (A-e_{\sigma(1)})\cap\cdots\cap(A-e_{\sigma(j_n)})\cap L$$
is infinite and $F_{\sigma(j_n)}$ is finite.
3. Choose $j_{n+1} > j_n$ minimal such that $b_{n+1} \in F_{\sigma(j_{n+1})}$.
4. Choose $c_{n+1} = e_{\sigma(j_{n+1})} \in \bigcap_{\ell\in F_{\sigma(j_{n+1})}} (A-\ell) \subseteq (A-b_1)\cap\cdots\cap(A-b_{n+1})$.

By this construction, we guarantee that $B + C \subseteq A$. □

Proof of Theorem 2.1.2. Let $L = A - \mathsf{p} \subseteq \mathbb{N}$ and

$$\varepsilon = \frac{1}{2}\lim_{n\to\mathsf{p}} \mathsf{d}_\Phi((A-n)\cap(A-\mathsf{p}))$$

Then $\mathsf{d}_\Phi((A-n)\cap L)$ exists for all $n \in \mathbb{N}$, so we need to show that for all finite $F \subseteq A - \mathsf{p}$ that

$$\bigcap_{\ell\in F}(A-\ell) \cap \{n \in \mathbb{N} : \mathsf{d}_\Phi((A-n)\cap L) > \varepsilon\}$$

is infinite. If $\ell \in F \subseteq A - \mathsf{p}$, then $A - \ell \in \mathsf{p}$ by definition, and

$$\{n \in \mathbb{N} : \mathsf{d}_\Phi((A-n)\cap L) > \varepsilon\} \in \mathsf{p} = \{n \in \mathbb{N} : \mathsf{d}_\Phi((A-n)\cap(A-\mathsf{p})) > 0\} \in \mathsf{p}$$

so

$$\bigcap_{\ell\in F}(A-\ell) \cap \{n \in \mathbb{N} : \mathsf{d}_\Phi((A-n)\cap L) > \varepsilon\} \in \mathsf{p}$$

Since p is non-principal, we conclude that this set is an infinite set. □

2.1.2 Theorem 2.1.2 + Theorem 2.1.3 $\implies$ Theorem 2.1.1

Proof. Given is a subset $A \subseteq \mathbb{N}$ and a Følner sequence Φ such that $\overline{\mathsf{d}}_\Phi(A) > 0$, and we want to find $B, C \subseteq \mathbb{N}$ such that $B + C \subseteq A$. In order to apply Theorem 2.1.2, we will need to find a Følner subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\mathsf{d}_\Psi((A-n)\cap(A-\mathsf{p}))$ exists for all $n \in \mathbb{N}$ and

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

which is suggested to exist by Theorem 2.1.3. First, we will need to assume that $\mathsf{d}_\Phi(A)$ itself exists and is positive to apply Theorem 2.1.3, which is a valid assumption by simply passing to a Følner subsequence Φ' of Φ such that $\mathsf{d}_{\Phi'}(A)$ exists and is positive. In particular, $\mathsf{d}_\Psi(A) = \mathsf{d}_{\Phi'}(A)$ for all Følner subsequences Ψ of Φ'. Therefore, by taking $\varepsilon = \frac{1}{2}\mathsf{d}_{\Phi'}(A)^2$ and applying Theorem 2.1.3, we recover a Følner subsequence Ψ of Φ' (of Φ) and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that $\mathsf{d}_\Psi((A-n)\cap(A-\mathsf{p}))$ exists and

$$\lim_{n\to\mathsf{p}} \mathsf{d}_\Psi((A-n)\cap(A-\mathsf{p})) \geq d_\Psi(A)^2 - \frac{1}{2}d_{\Phi'}(A)^2 = \frac{1}{2}d_{\Phi'}(A)^2 > 0$$

from which we conclude Theorem 2.1.1 by Theorem 2.1.2. $\square$

2.1.3 Theorem 2.1.4 $\implies$ Theorem 2.1.3 and Besicovitch seminorm algebra

Proof of Theorem 2.1.4 $\implies$ Theorem 2.1.3. To see that Theorem 2.1.4 $\implies$ Theorem 2.1.3, let $A \subseteq \mathbb{N}$ and Φ be a Følner sequence on $\mathbb{N}$ with $\mathsf{d}_\Phi(A)$ existing. Then let $f = 1_A$, which is non-negative and bounded. Furthermore,

$$\langle 1, f\rangle_\Phi = \lim_{N\to\infty} \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} 1_A(n) = \lim_{N\to\infty}\frac{|A\cap\Phi_N|}{|\Phi_N|} = \mathsf{d}_\Phi(A)$$

exists. Then for every $\varepsilon > 0$, there exists a subsequence Ψ of Φ and a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ such that

$$\begin{aligned}\langle \mathsf{R}^m 1_A, \mathsf{R}^\mathsf{p} 1_A\rangle_\Psi &= \lim_{N\to\infty}\frac{1}{|\Psi_N|}\sum_{n\in\Phi_N}(\mathsf{R}^m 1_A)(n)(\mathsf{R}^\mathsf{p} 1_A)(n)\\ &= \lim_{N\to\infty}\frac{|(A-m)\cap(A-\mathsf{p})\cap\Psi_N|}{|\Psi_N|} = \mathsf{d}_\Psi((A-m)\cap(A-\mathsf{p}))\end{aligned}$$

exists and

$$\lim_{m\to\mathsf{p}} \mathsf{d}_\Psi((A-m)\cap(A-\mathsf{p})) = \lim_{m\to\mathsf{p}}\langle \mathsf{R}^m 1_A, \mathsf{R}^\mathsf{p} 1_A\rangle_\Psi \geq \langle 1, f\rangle_\Psi^2 - \varepsilon = \mathsf{d}_\Psi(A)^2 - \varepsilon$$

$\square$

The remainder of the paper involves showing Theorem 2.1.4 using functional analysis techniques.

Let us further develop the Besicovitch seminorm. It is indeed a seminorm since Minkowski's inequality

$$\left(\sum_{n\in\Phi_N}|f(n)+h(n)|^2\right)^{1/2} \leq \left(\sum_{n\in\Phi_N}|f(n)|^2\right)^{1/2} + \left(\sum_{n\in\Phi_N}|h(n)|^2\right)^{1/2}$$

proves $\|f+h\|_\Phi \leq \|f\|_\Phi + \|h\|_\Phi$. Furthermore, we have the following identities:

Proposition 2.1.3. *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. If $\langle f, h\rangle_\Phi$ exists and $\|f\|_\Phi$ and $\|h\|_\Phi$ are finite, then $|\langle f, h\rangle_\Phi| \le \|f\|_\Phi \|h\|_\Phi$.

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.

Proof. 1. Let $N \in \mathbb{N}$ be such that there exists $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\Psi_n = \Phi_{\sigma(n)}$ for all $n \ge N$ and $\sigma(N) < \sigma(N+1) < \dots$. Then

$$\|f\|_\Psi^2 = \lim_{N\to\infty} \sup_{n\ge N} \frac{1}{|\Psi_n|} \sum_{m\in\Psi_n} |f(m)|^2 = \lim_{N\to\infty} \sup_{n\ge N} \frac{1}{|\Phi_{\sigma(n)}|} \sum_{m\in\Phi_{\sigma(n)}} |f(m)|^2$$
$$= \lim_{N\to\infty} \sup_{n\in\sigma(\{N,N+1,\dots\})} \frac{1}{|\Phi_n|} \sum_{m\in\Phi_n} |f(m)|^2 \le \lim_{N\to\infty} \sup_{n\ge\sigma(N)} \frac{1}{|\Phi_n|} \sum_{m\in\Phi_n} |f(m)|^2 = \|f\|_\Phi^2$$

since $\sigma(\{N, N+1, \dots\}) \subseteq \{\sigma(N), \sigma(N)+1, \dots\}$.

2. By the triangle inequality and normal Cauchy-Schwarz, we have

$$|\langle f, h\rangle_\Phi| \le \lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n)\overline{h(n)}| \le \lim_{N\to\infty} \frac{1}{|\Phi_N|} \left(\sum_{n\in\Phi_N} |f(n)|^2\right)^{1/2} \left(\sum_{n\in\Phi_N} |h(n)|^2\right)^{1/2}$$
$$= \lim_{N\to\infty} \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n)|^2\right)^{1/2} \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |h(n)|^2\right)^{1/2} = \|f\|_\Phi \|g\|_\Phi$$

3. In general, the following result is true: if $\{a_n\}_{n\in\mathbb{N}}$ is a sequence that is bounded above, then there exists a subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ where $n_1 < n_2 < \dots$ converging to $\ell := \limsup_{n\to\infty} a_n \in \mathbb{R}$. Indeed, pick $n_1 \in \mathbb{N}$ so that $|a_{n_1} - \ell| < 1$, and pick $n_{k+1} > n_k$ so that $|a_{n_{k+1}} - \ell| < \frac{1}{k+1}$. Then any subsequence of $\{a_{n_k}\}_{k\in\mathbb{N}}$ must also converge to ℓ. Here, $a_N = \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n)|^2\right)^{1/2}$, $\ell = \|f\|_\Phi$, and we define $\Psi_k = \Phi_{n_k}$.

4. Choose $K \in \mathbb{N}$ so that $N \ge K$ implies

$$\left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n)|^2\right)^{1/2} \le \|f\|_\Phi + 1 \quad \text{and} \quad \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |h(n)|^2\right)^{1/2} \le \|h\|_\Phi + 1$$

Then by Cauchy-Schwarz, we have

$$\left|\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} f(n)\overline{h(n)}\right| \le \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n)|^2\right)^{1/2} \left(\frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |h(n)|^2\right)^{1/2}$$
$$< (\|f\|_\Phi + 1)(\|h\|_\Phi + 1) < \infty$$

Therefore,

$$\left\{\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} f(n)\overline{h(n)} : N \geq K\right\} \subseteq \mathbb{C}$$

is bounded and hence contains a convergent subsequence by Bolzano-Weierstrass.

□

Now, we will see how the Besicovitch seminorm interacts with $\mathsf{L}^2(\mathbb{N},\Phi)$. By Minkowski's inequality, $\mathsf{L}^2(\mathbb{N},\Phi)$ is a vector space; however, it is not a Hilbert space since the limit defining $\langle f,h\rangle_\Phi$ may not even exist for all $f,h \in \mathsf{L}^2(\mathbb{N},\Phi)$. Not all is lost, since we can just use a few technical results to work around this fact. Call a sequence $j \mapsto f_j : \mathbb{N} \to \mathbb{C}$ *Cauchy* with respect to $\|\cdot\|_\Phi$ if for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $j,k \geq N \implies \|f_k - f_j\|_\Phi \leq \varepsilon$. Then we have the following completeness lemma:

Lemma 2.1.1. *If $j \mapsto f_j$ is 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 the interval $[a,b]$, then so does f.*

Proof. We will define Ψ in such a way so that

$$f(n) = \begin{cases} 0 & n \notin \bigcup_{K\geq 1}\Psi_K \\ f_M(n) & M = \min\{K \in \mathbb{N} : n \in \Psi_K\}\end{cases}$$

satisfies the theorem statement. In order to properly choose Ψ, define $\Xi_K := \Psi_K \setminus \bigcup_{k=1}^{K-1}\Psi_k$ and $\zeta_K := \Psi_K \setminus \Xi_K$. Then using $|x+y|^2 \leq 2|x|^2 + 2|y|^2$, the following bounds hold for $K \geq j$:

$$\begin{aligned}
\frac{1}{2}\sum_{n\in\Psi_K}|f_j(n) - f(n)^2| &\leq \sum_{n\in\Psi_K}|f_j(n) - f_K(n)|^2 + \sum_{n\in\Psi_K}|f_K(n) - f(n)|^2 \\
&= \sum_{n\in\Psi_K}|f_j(n) - f_K(n)|^2 + \sum_{n\in\zeta_K}|f_K(n) - f(n)|^2 \\
&= \sum_{n\in\Psi_K}|f_j(n) - f_K(n)|^2 + \sum_{i=1}^{K-1}\sum_{n\in\Xi_i}|f_K(n) - f_i(n)|^2 \\
&\leq |\Psi_K|\max\{|f_j(n) - f_K(n)|^2 : n \in \Psi_K\} \\
&\quad + K\max\left\{\sum_{n\in\Xi_i}|f_K(n) - f_i(n)|^2 : 1 \leq i \leq K-1\right\}
\end{aligned}$$

It suffices to choose Ψ carefully so that $n \in \Psi_K \implies |f_j(n) - f_K(n)|^2 < \frac{2}{j}$ and

$$\max\left\{\sum_{n\in\Xi_i}|f_K(n) - f_i(n)|^2 : 1 \leq i \leq K-1\right\} \leq \frac{|\Psi_K|}{K^2}$$

Suppose that $|f - f_{n_k}|_\Psi \to 0$ for some subsequence $f_{n_1}, f_{n_2}, \ldots$ of $\{f_j\}_{j\in\mathbb{N}}$. Then for all $\varepsilon > 0$, there exists $K \in \mathbb{N}$ such that $k \geq K \implies \|f_{n_k} - f\|_\Psi \leq \frac{\varepsilon}{2}$ from convergence and $j \geq n_k \implies \|f_j - f_{n_k}\| \leq \frac{\varepsilon}{2}$. By the triangle inequality, we find $\|f_j - f\|_\Phi \leq \varepsilon$ for all $j \geq \max\{K, n_k\}$, which means we are allowed to pass to subsequences of $\{f_j\}_{j\in\mathbb{N}}$ in this proof. Hence, assume without loss of generality that $k \geq j \implies \|f_k - f_j\|_\Phi^2 \leq \frac{1}{j}$, or else pass to such a subsequence. In particular, $\|f_j\|_\Phi^2 \leq (\|f_1\|_\Phi + 1)^2 =: C$, so we can find yet another subsequence $N_1 < N_2 < \dots$ for which

$$N \geq N_k \implies \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f_j(n) - f_k(n)|^2 \leq \frac{2}{j} \quad \text{and} \quad \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f_j(n)|^2 \leq 2C$$

By further refining $N_1, N_2, \dots$, we may further assume that

$$\max\left\{ \sum_{n\in\Phi_{N_i}} |f_K(n) - f_i(n)|^2 : 1 \leq i \leq K-1 \right\} \leq \frac{|\Phi_{N_K}|}{K^2}$$

which implies the desired bound with $\Psi_K = \Phi_{N_K}$. Moreover, given the way we defined f, if f_j takes values in $[a, b]$, of course this f takes values in $[a, b]$. □

Finally, one last important algebraic identity for the Besicovitch seminorm is Bessel's inequality.

Theorem 2.1.5. *Given is a sequence $\{u_j\}_{j\in\mathbb{N}} \subseteq \mathsf{L}^2(\mathbb{N}, \Phi)$ 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 $j \neq k$. 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*

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

Proof. Fix $J \in \mathbb{N}$, and define

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

Then

$$\begin{aligned}
0 &\leq \left[u - \sum_{j=1}^{J} u_j [u, u_j]_N, u - \sum_{k=1}^{J} u_k [u, u_k]_N \right]_N \\
&= [u, u]_N - \sum_{j=1}^{J} [u, u_j]_N \, [u_j, u]_N - \sum_{k=1}^{J} \overline{[u, u_k]_N} \, [u, u_k]_N + \sum_{j=1}^{J} \sum_{k=1}^{K} [u, u_j]_N \overline{[u, u_k]_N} [u_j, u_k]_N \\
&= [u, u]_N - 2 \sum_{j=1}^{J} |[u, u_j]_N|^2 + \sum_{j=1}^{J} \sum_{k=1}^{K} [u, u_j] \overline{[u, u_k]_N} [u_j, u_k]_N \\
&= [u, u]_N - \sum_{j=1}^{J} (2 - [u_j, u_j]_N) |[u, u_j]_N|^2 + \sum_{\substack{1\leq j,k\leq J \\ j\neq k}} [u, u_j] \overline{[u, u_k]_N} [u_j, u_k]_N \\
&\to \|u\|_\Phi^2 - \sum_{j=1}^{J} |\langle u, u_j\rangle_\Phi|^2 \quad \text{as } N \to \infty
\end{aligned}$$

We conclude by sending $J \to \infty$. □

2.2 Functional analysis proof of Theorem 2.1.4

The key insight in proving Theorem 2.1.4 is the idea to decompose 1_A into pseudo-random and structured components. We will show that we can decompse 1_A as a sum of $f_{\mathsf{wm}} + f_{\mathsf{c}}$, where f_{wm} is the *weak mixing* (pseudo-random) component and f_{c} is the *compact* (structured) component. This decomposition is stable under shifts by $m \in \mathbb{N}$, i.e., $\mathsf{R}^m f_{\mathsf{wm}} + \mathsf{R}^m f_{\mathsf{c}}$ is the decomposition of $\mathsf{R}^m 1_A = 1_{A-m}$. This allows us to replace

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

However, this decomposition is not stable under shifts by ultrafilters, so to proceed we will need another decomposition of 1_A into $f_{\mathsf{anti}} + f_{\mathsf{Bes}}$, where the structured component f_{Bes} is a *Besicovitch almost-periodic* function, and the pseudo-random component f_{anti} is orthogonal to $e^{2\pi i n\theta}$ for all $\theta \in [0, 1)$. Therefore, we are left with three terms:

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

We prove that the first term is zero by pseudo-randomness, the second term is positive by the shifting properties of f_{Bes}, and the third term is non-negative using a very delicate argument. Finally, the curious reader might wonder why the first decomposition is necessary, and why not just applying the second decomposition to the first term? The answer to this is that the pseudo-randomness guaranteed by weak mixing function is a slightly stronger form of pseudo-randomness than in the other decomposition.

2.2.1 Decomposition 1: $f = f_{\mathsf{Bes}} + f_{\mathsf{anti}}$

A *trigonometric polynomial* is any function $a : \mathbb{N} \to \mathbb{C}$ of the form

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

for some $c_1, \ldots, c_J \in \mathbb{C}$ and some *frequencies* $0 \le \theta_1, \ldots, \theta_J < 1$. Then given a Følner sequence Φ, we define the space of *Besicovitch almost periodic* functions, written $\mathsf{Bes}(\mathbb{N}, \Phi)$, to be the set of all functions $f : \mathbb{N} \to \mathbb{C}$ such that for all $\varepsilon > 0$ there exists a trigonometric polynomial a with $\|f - a\|_\Phi < \varepsilon$. Note that $\mathsf{Bes}(\mathbb{N}, \Phi) \subseteq \mathsf{L}^2(\mathbb{N}, \Phi)$. We also define $\mathsf{Bes}(\mathbb{N}, \Phi)^\perp$ as the set of all $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ 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 $\theta \in [0, 1)$.

The first $\mathsf{L}^2(\mathbb{N}, \Phi)$ decomposition of great importance is the following.

Theorem 2.2.1. *For every Følner sequence Φ on $\mathbb{N}$ and $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$, there is a subsequence Ψ of Φ, $f_{\mathsf{Bes}} \in \mathsf{Bes}(\mathbb{N}, \Phi)$, and $f_{\mathsf{anti}} \in \mathsf{Bes}(\mathbb{N}, \Phi)^\perp$ such that $f = f_{\mathsf{Bes}} + f_{\mathsf{anti}}$. Moreover,*

$$\|f - f_{\mathsf{Bes}}\|_\Psi = \inf\{\|f - g\|_\Psi : g \in \mathsf{Bes}(\mathbb{N}, \Psi)\}$$

and if f takes values in $[a, b]$, then so does f_{Bes}.

This theorem seems almost trivial, as it suggests taking f_{Bes} to be the orthogonal projection of f onto $\mathsf{Bes}(\mathbb{N}, \Psi)$. However, the complication comes from the fact that $\mathsf{L}^2(\mathbb{N}, \Phi)$ is not quite a Hilbert space, which means the proof of Theorem 2.2.1 is just a bunch of technical workarounds.

First, we start with making clear the definition of "projection" in the context of Følner sequences. We call U a *projection family* if it assigns to each Følner sequence Φ a collection $U(\Phi)$ of $\mathsf{L}^2(\mathbb{N}, \Phi)$ functions with the following properties:

1. $U(\Phi)$ is a vector subspace of $\mathsf{L}^2(\mathbb{N}, \Phi)$.
2. $U(\Phi)$ contains the constant functions and is closed under pointwise complex conjugation.
3. $u, v \in U(\Phi) \implies \langle u, v\rangle_\Phi$ exists.
4. $u, v \in U(\Phi)$ and u, v are real-valued, then $(n \mapsto \max\{u(n), v(n)\}) \in U(\Phi)$.
5. $U(\Phi)$ is closed with respect to the topology on $\mathsf{L}^2(\mathbb{N}, \Phi)$ induced by the seminorm $\|\cdot\|_\Phi$.
6. If Ψ agrees with a subsequence of Φ then $U(\Phi) \subseteq U(\Psi)$.

As suggested by Theorem 2.2.1, we will need to show that $\Phi \mapsto \mathsf{Bes}(\mathbb{N}, \Phi)$ is a projection family.

Theorem 2.2.2. $\Phi \mapsto \mathsf{Bes}(\mathbb{N}, \Phi)$ *is a projection family.*

Proof. We verify the axioms.

1. Let $\alpha, \beta \in \mathbb{C}$, $f, h \in \mathsf{Bes}(\mathbb{N}, \Phi)$, and $\varepsilon > 0$. There exist trigonometric polynomials $a, b : \mathbb{N} \to \mathbb{C}$ such that $\|f - a\|_\Phi < \frac{\varepsilon}{2|\alpha|}$ and $\|h - b\|_\Phi < \frac{\varepsilon}{2|\beta|}$. Thus, $\alpha a + \beta b$ is a trigonometric polynomial such that
$$\|(\alpha f + \beta h) - (\alpha a + \beta b)\|_\Phi \le |\alpha|\|f - a\|_\Phi + |\beta|\|g - b\|_\Phi < \varepsilon$$
Hence, $\alpha f + \beta h \in \mathsf{Bes}(\mathbb{N}, \Phi)$.
2. The constants are trigonometric polynomials, and if $a : \mathbb{N} \to \mathbb{C}$ is a trigonometric polynomial and $f \in \mathsf{Bes}(\mathbb{N}, \Phi)$, then
$$\begin{aligned}\|f - a\|_\Phi &= \left(\limsup_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |f(n) - a(n)|^2\right)^{1/2} \\ &= \left(\limsup_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |\overline{f(n)} - \overline{a(n)}|^2\right)^{1/2} = \|\overline{f} - \overline{a}\|_\Phi\end{aligned}$$
where $\overline{a}$ is also a triogonometric polynomial, so $\overline{f} \in \mathsf{Bes}(\mathbb{N}, \Phi)$.
3. First, we show that $\langle a, b\rangle_\Phi$ exists for all trigonometric polynomials $a, b : \mathbb{N} \to \mathbb{C}$. It suffices to show that
$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} e^{2\pi i n\theta} = 0$$

for all $0 < \theta < 1$. This follows from

$$\begin{aligned}
\left|\frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} e^{2\pi i n\theta}\right| &= \frac{1}{|\Phi_N|}\frac{1}{|1-e^{2\pi i\theta}|}\left|\sum_{n\in\Phi_N} e^{2\pi i n\theta} - \sum_{n\in\Phi_N} e^{2\pi i(n+1)\theta}\right| \\
&= \frac{1}{|\Phi_N|}\frac{1}{|1-e^{2\pi i\theta}|}\left|\sum_{n\in\Phi_N} e^{2\pi i n\theta} - \sum_{n\in\Phi_N+1} e^{2\pi i n\theta}\right| \\
&\leq \frac{1}{|\Phi_N|}\frac{1}{|1-e^{2\pi i\theta}|}\sum_{n\in\Phi_N\triangle(\Phi_N+1)} 1 \\
&= \frac{|\Phi_N\triangle(\Phi_N+1)|}{|\Phi_N|}\frac{1}{|1-e^{2\pi i\theta}|} \to 0 \quad \text{as } N\to\infty
\end{aligned}$$

Now we show that if $u, v \in \mathsf{Bes}(\mathbb{N}, \Phi)$ and $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ are two convergent sequences of trigonometric polynomials, then

$$\lim_{n\to\infty}\langle a_n, v\rangle_\Phi = \left\langle \lim_{n\to\infty} a_n, v\right\rangle_\Phi \quad \text{and} \quad \lim_{n\to\infty}\langle u, b_n\rangle_\Phi = \left\langle u, \lim_{n\to\infty} b_n\right\rangle_\Phi$$

It suffices to just prove the first one, since the second one is the symmetric identity. Using the inequality

$$\begin{aligned}
&\limsup_{N\to\infty}\left|\frac{1}{|\Phi_N|}\sum_{m\in\Phi_N}(A(m)-a_n(m))\overline{v(m)}\right| \\
\leq &\|A-a_n\|_\Phi \sup\left\{\left(\frac{1}{|\Phi_N|}\sum_{m\in\Phi_N}|v(m)|^2\right)^{1/2} : N\in\mathbb{N}\right\}
\end{aligned}$$

we obtain continuity of $\langle\cdot,\cdot\rangle_\Phi$ in the first variable. Suppose v is itself a trigonometric polynomial. Then if $u = \lim_{n\to\infty} a_n$, we find $n \mapsto a_n$ is a Cauchy sequence with respect to Φ, and hence $\langle a_n, v\rangle_\Phi$ is Cauchy by the Cauchy-Schwarz ineqaulity. Hence, $\langle u, v\rangle_\Phi$ converges in this case. Playing the same trick with the second coordinate, we prove the claim.

4. Given $u, v \in \mathsf{Bes}(\mathbb{N}, \Phi)$, we need to show that the function $n \mapsto \max\{u(n), v(n)\} = \frac{u(n)+v(n)+|u(n)-v(n)|}{2}$ is in $\mathsf{Bes}(\mathbb{N}, \Phi)$. By linearity, it suffices to show that $w \in \mathsf{Bes}(\mathbb{N}, \Phi) \implies |w| \in \mathsf{Bes}(\mathbb{N}, \Phi)$. The idea is to approximate $z \mapsto |z|$ with a polynomial using Stone-Weierstrass and applying that to a trigonometric polynomial approximating w. Let $\varepsilon > 0$ and $a : \mathbb{N} \to \mathbb{C}$ be a trigonometric polynomial such that $\|w - a\|_\Phi < \varepsilon/2$, so $\||w| - |a|\|_\Phi < \varepsilon/2$. Note that $|a|$ is bounded, so by Stone-Weierstrass there exists a polynomial $b \in \mathbb{C}[z]$ such that $|z| \leq \sup_{n\in\mathbb{N}}|a(n)| \implies |b(z) - |z|| < \varepsilon/2$. Finally, we know that $n \mapsto b(a(n))$ is a trigonometric polynomial and $\||w| - b \circ a\|_\Phi \leq \varepsilon$, so $|w| \in \mathsf{Bes}(\mathbb{N}, \Phi)$.

5. This follows from the fact that $\mathsf{Bes}(\mathbb{N}, \Phi)$ is the closure in $\mathsf{L}^2(\mathbb{N}, \Phi)$ of the space of trigonometric polynomials with respect to $\|\cdot\|_\Phi$.

6. If Ψ eventually agrees with Φ, and $f \in \mathsf{Bes}(\mathbb{N}, \Phi)$, then since $\|f - a\|_\Psi \leq \|f - a\|_\Phi$, $f \in \mathsf{Bes}(\mathbb{N}, \Psi)$.

□

The purpose of projection families is the following theorem.

Theorem 2.2.3. *Given a projection family U and a Følner sequence Φ, for every $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ there exists a subsequence Ψ of Φ and $f_U \in U(\Psi)$ such that:*

1. *$f - f_U \in U(\Psi)^\perp$,*
2. *$\|f - f_U\|_\Psi = \inf\{\|f - g\|_\Psi : g \in U(\Psi)\}$, and*
3. *if f takes values in an interval $[a, b]$, then so does f_U.*

Theorem 2.2.3 follows naturally from the suggestion of projecting f onto $\mathsf{Bes}(\mathbb{N}, \Phi)$. In order to prove it, we will need the following technical result.

Theorem 2.2.4. *Given a projection family U and a Følner sequence Φ, for every $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ there exists a subsequence Ψ of Φ such that $u \in U(\Psi) \implies \langle f, u\rangle_\Psi$.*

Proof of Theorem 2.2.3 assuming Theorem 2.2.4. Let Ψ be a subsequence of Φ guaranteed by Theorem 2.2.4 such that $\langle f, u\rangle_\Psi$ exists for all $u \in U(\Psi)$. We will find a sequence $\{u_k\}_{k\in\mathbb{N}} \subseteq U(\Psi)$ with $\|f - u_k\|_\Psi^2$ converging to $\delta := \inf\{\|f - g\|_\Psi^2 : g \in U(\Psi)\}$, and then we will extract a limit point of $\{u_k\}_{k\in\mathbb{N}}$ using Lemma 2.1.1. Select $\{u_k\}_{k\in\mathbb{N}}$ so that $\|f - u_k\|_\Psi^2 < \delta + \frac{1}{k}$. The sequence $\{u_k\}_{k\in\mathbb{N}}$ is Cauchy since

$$\|u_k - u_j\|_\Psi^2 = 2\|f - u_k\|_\Psi^2 + 2\|f - u_j\|_\Psi^2 - 4\left\|f - \frac{u_j + u_k}{2}\right\|_\Psi^2 \leq 2\left(\delta + \frac{1}{k}\right) + 2\left(\delta + \frac{1}{j}\right) - 4\delta = \frac{2}{k} + \frac{2}{j}$$

Therefore, by Lemma 2.1.1, there exists some $f_U \in \mathsf{L}^2(\mathbb{N}, \Psi)$ such that $\lim_{k\to\infty}\|f_U - u_k\|_\Psi = 0$. Since $U(\Psi)$ is closed, $f_U \in U(\Psi)$, and by Minkowski's inequality we find that $\|f - f_U\|_\Psi^2 = \delta$ exact.

It remains to show that $h := f - f_U \in U(\Psi)^\perp$. Let $0 \leq \theta < 1$ and $u(n) = \overline{e^{2\pi i n\theta}}$, so $\|u\|_\Psi = 1$. We need to show that $\langle h, u\rangle_\Psi = 0$. Indeed, let $I := \langle h, u\rangle_\Psi$, which exists by the choice of Ψ. Then we have

$$\begin{aligned}\|h - Iu\|_\Psi^2 &= \lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Psi_N} \left[|h(N)|^2 - 2\,\mathrm{Re}(h(n)\overline{Iu(n)}) + |Iu(n)|^2\right] \\ &\leq \|h\|_\Psi^2 - |I|^2(2 - |u|_\Psi^2) = \|h\|_\Psi^2 - |I|^2 = \delta - |I|^2\end{aligned}$$

On the other hand, $h - Iu = f - (f_U + Iu)$ where $f_U + Iu \in U(\Psi)$, so $\|h - Iu\|_\Psi^2 \geq \delta$. Hence, $I = 0$.

If f takes values in the interval $[a, b]$, repeat the proof but with the sequence $\{u'_k\}_{k\in\mathbb{N}}$ where

$$u'_k(n) = \begin{cases} a & \mathrm{Re}(u_k(n)) < a \\ b & \mathrm{Re}(u_k(n)) > b \\ \mathrm{Re}(u_k(n)) & \text{otherwise} \end{cases}$$

□

Proof of Theorem 2.2.4. We prove the statement by a greedy construction similar to the Gram-Schmidt algorithm for finding an orthogonal basis of a vector space. We construct a sequence of Følner sequences $\Phi^{(0)} \supseteq \Phi^{(1)} \supseteq \dots$ and elements $u_0, u_1, \dots$ such that $u_0, \dots, u_{k-1} \in U(\Phi^{(k-1)})$ and $\langle f, u_i \rangle_{\Phi^{(k-1)}}$ exists for all $0 \le i \le k-1$. Start with $u_0 = 0$ and $\Phi^{(0)} = 0$, and construct up to u_{k-1} and $\Phi^{(k-1)}$. At this moment, we have one of two cases:

Case 1: For all Følner subsequences Φ' of $\Phi^{(k-1)}$ and $u \in U(\Phi')$ with $\langle u, u_i \rangle_{\Phi'} = 0$ for all $0 \le i \le k-1$, we have $\|u\|_{\Phi'} = 0$. In this case, we claim that the construction is already complete. Let $\Psi = \Phi^{(k-1)}$ and let $u \in U(\Psi)$. Then

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

for all $0 \le j \le k-1$. Hence,

$$\left\| u - \sum_{i=0}^{k-1} u_i \right\|_\Psi = 0$$

so

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

converges as $N \to \infty$.

Case 2: Otherwise. Then there exists a Følner subsequence Φ' of $\Phi^{(k-1)}$ and $u \in U(\Phi')$ with $\langle u, u_i \rangle_{\Phi'} = 0$ for all $0 \le i \le k-1$ but $\|u\|_{\Phi'} \ne 0$. By scaling, we can also impose the condition that $\|u\|_{\Phi'} = 1$, in fact. Let δ_k be the supremum of $|\langle f, u \rangle_{\Phi'}|$ over all possible us that satisfy these properties for some Φ', and choose $\Phi^{(k)}$ and u_k so that $|\langle f, u_k \rangle_{\Phi^{(k)}}| > \delta_k - \frac{1}{k}$.

If this algorithm ever enters case 1, we're already done. Therefore, assume we constantly loop in case 2. Define the sequence $\Psi_N := \Phi_N^{(N)}$, which is a Følner sequence since it is a subsequence of $\Phi^{(1)}$. Let $u \in U(\Psi)$. We claim that $\langle f, u \rangle_\Psi$ exists and is equal to $\sum_{i=1}^\infty \langle f, u_i \rangle_\Psi \overline{\langle u, u_i \rangle_\Psi}$. This is well-defined since $\langle u, u_i \rangle_\Psi$ exists since $u, u_i \in U(\Psi)$, $\langle f, u_i \rangle_\Psi$ exists by the construction of Ψ, and the series is absolutely convergent by Theorem 2.1.5.

As before, write

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

By absolute convergence, we know that the second term tends to 0 as $k \to \infty$. Furthermore, we can bound δ_k by noting that

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

again by Theorem 2.1.5 and hence $\delta_k \to 0$ as $k \to \infty$ too. We conclude that $\langle f, u \rangle_\Psi$ exists and is equal to $\sum_{i=1}^{\infty} \langle f, u_i \rangle_\Psi \overline{\langle u, u_i \rangle_\Psi}$. □

2.2.2 Decomposition 2: $f = f_{\mathsf{wm}} + f_{\mathsf{c}}$

The next decomposition we will use concerns the so-called Jacobs-de Leeuw-Glicksberg splitting. In particular, we will obtain a discrete version of this splitting theorem that will allow us to decompose 1_A in $\mathsf{L}^2(\mathbb{N}, \Phi)$ into *compact* and *weak mixing* components.

Let $(\mathscr{H}, \|\cdot\|_{\mathscr{H}})$ be a Hilbert space and let U be an isometry (and hence a *linear* isometry since $\mathscr{H}$ is a Hilbert space). We call an element $x \in \mathscr{H}$ *compact* if $\{\mathsf{U}^n x : n \in \mathbb{N}\}$ is a pre-compact subset of $(\mathscr{H}, \|\cdot\|_{\mathscr{H}})$, or equivalently for all $\varepsilon > 0$ there exists $K \in \mathbb{N}$ such that

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

for all $m \in \mathbb{N}$. We call an element $x \in \mathscr{H}$ *weak mixing* if, for all $\varepsilon > 0$ and $y \in \mathscr{H}$, the set $\{n \in \mathbb{N} : |\langle \mathsf{U}^n x, y \rangle| \geq \varepsilon\}$ has zero density with respect to every Følner sequence on $\mathbb{N}$. The set of compact and weak mixing elements is denoted by $\mathscr{H}_{\mathsf{c}}$ and $\mathscr{H}_{\mathsf{wm}}$.

Let us consider some analytical properties of $\mathscr{H}_{\mathsf{c}}$ and $\mathscr{H}_{\mathsf{wm}}$.

Theorem 2.2.5. *$\mathscr{H}_{\mathsf{c}}$ and $\mathscr{H}_{\mathsf{wm}}$ are closed and U-invariant subspaces of $\mathscr{H}$.*

Proof. We break this theorem into the 6 claims.

- $\mathscr{H}_{\mathsf{c}}$ is a subspace: Let α_1, α_2 be scalars and $x_1, x_2 \in \mathscr{H}_{\mathsf{c}}$. Note that

  $$\{\mathsf{U}^n x_1 : n \in \mathbb{N}\} \times \{\mathsf{U}^n x_2 : n \in \mathbb{N}\}$$

  is precompact by Tychonoff's theorem. Furthermore, $f(z_1, z_2) = \alpha_1 z_1 + \alpha_2 z_2$ is continuous so

  $$\{\mathsf{U}^n(\alpha_1 x_1 + \alpha_2 x_2) : n \in \mathbb{N}\} = f(\{\mathsf{U}^n x_1 : n \in \mathbb{N}\} \times \{\mathsf{U}^n x_2 : n \in \mathbb{N}\})$$

  is precompact since the continuous image of a precompact set is precompact. Therefore, $\alpha_1 x_1 + \alpha_2 x_2 \in \mathscr{H}_{\mathsf{c}}$.

- $\mathscr{H}_{\mathsf{c}}$ is closed: Let $x \in \mathscr{H}$ be a limit point of $\mathscr{H}_{\mathsf{c}}$. Given $\varepsilon > 0$, we know there exists $x' \in \mathscr{H}_{\mathsf{c}}$ such that $\|x - x'\|_{\mathscr{H}} < \frac{\varepsilon}{4}$ and $K \in \mathbb{N}$ such that

  $$\min\{\|\mathsf{U}^m x' - \mathsf{U}^k x'\|_{\mathscr{H}} : 1 \leq k \leq K\} \leq \frac{\varepsilon}{2}$$

  for all $m \in \mathbb{N}$. Then we have

  $$\begin{aligned} \|\mathsf{U}^m x - \mathsf{U}^k x\|_{\mathscr{H}} &= \|\mathsf{U}^m x' - \mathsf{U}^k x' + \mathsf{U}^m(x - x') - \mathsf{U}^k(x - x')\|_{\mathscr{H}} \\ &\leq \|\mathsf{U}^m x' - \mathsf{U}^k x'\|_{\mathscr{H}} + \|\mathsf{U}^m(x - x')\|_{\mathscr{H}} + \|\mathsf{U}^k(x - x')\|_{\mathscr{H}} \\ &= \|\mathsf{U}^m x' - \mathsf{U}^k x'\|_{\mathscr{H}} + 2\|x - x'\|_{\mathscr{H}} \leq \|\mathsf{U}^m x' - \mathsf{U}^k x'\|_{\mathscr{H}} + \frac{\varepsilon}{2} \end{aligned}$$

so

$$\min\{\|\mathsf{U}^m x - \mathsf{U}^k x\|_{\mathscr{H}} : 1 \le k \le K\} \le \min\left\{\|\mathsf{U}^m x' - \mathsf{U}^k x'\|_{\mathscr{H}} + \frac{\varepsilon}{2} : 1 \le k \le K\right\} \le \varepsilon$$

for all $m \in \mathbb{N}$. Hence, $x \in \mathscr{H}_{\mathsf{c}}$.

- $\mathscr{H}_{\mathsf{c}}$ is U-invariant: Given $x \in \mathscr{H}_{\mathsf{c}}$, we can use the identity

$$\|\mathsf{U}^m(\mathsf{U}x) - \mathsf{U}^k(\mathsf{U}x)\|_{\mathscr{H}} = \|\mathsf{U}(\mathsf{U}^m x) - \mathsf{U}(\mathsf{U}^k x)\|_{\mathscr{H}} = \|\mathsf{U}^m x - \mathsf{U}^k x\|_{\mathscr{H}}$$

- $\mathscr{H}_{\mathsf{wm}}$ is a subspace: Let α_1, α_2 be scalars and $x_1, x_2 \in \mathscr{H}_{\mathsf{wm}}$. Given $\varepsilon > 0$ and $y \in \mathscr{H}$, we have

$$\begin{aligned}\{n \in \mathbb{N} : |\langle \mathsf{U}^n(\alpha_1 x_1 + \alpha_2 x_2), y\rangle| \ge \varepsilon\} &= \{n \in \mathbb{N} : |\alpha_1\langle \mathsf{U}^n x_1, y\rangle + \alpha_2\langle \mathsf{U}^n x_2, y\rangle| \ge \varepsilon\} \\ &\subseteq \left\{n \in \mathbb{N} : |\langle \mathsf{U}^n x_1, y\rangle| \ge \frac{\varepsilon}{2|\alpha_1|} \vee |\langle \mathsf{U}^n x_2, y\rangle| \ge \frac{\varepsilon}{2|\alpha_2|}\right\} \\ &= \left\{n \in \mathbb{N} : |\langle \mathsf{U}^n x_1, y\rangle| \ge \frac{\varepsilon}{2|\alpha_1|}\right\} \\ &\cup \left\{n \in \mathbb{N} : |\langle \mathsf{U}^n x_2, y\rangle| \ge \frac{\varepsilon}{2|\alpha_2|}\right\}\end{aligned}$$

where, for any Følner sequence, the right-hand side is the union of two density-zero sets. Therefore, the left-hand side has density zero with respect to all Følner sequence, so $\alpha_1 x_1 + \alpha_2 x_2 \in \mathscr{H}_{\mathsf{wm}}$.

- $\mathscr{H}_{\mathsf{wm}}$ is closed: Let $x \in \mathscr{H}$ be a limit point of $\mathscr{H}_{\mathsf{wm}}$. Given $\varepsilon > 0$ and $y \in \mathscr{H}$, we can find an $x' \in \mathscr{H}_{\mathsf{wm}}$ such that $\|x - x'\|_{\mathscr{H}}\|y\|_{\mathscr{H}} < \frac{\varepsilon}{2}$. Then

$$|\langle \mathsf{U}^n(x - x'), y\rangle| \le \|\mathsf{U}^n(x - x')\|_{\mathscr{H}}\|y\|_{\mathscr{H}} = \|x - x'\|_{\mathscr{H}}\|y\|_{\mathscr{H}} < \frac{\varepsilon}{2}$$

by Cauchy-Schwarz, so

$$\begin{aligned}\{n \in \mathbb{N} : |\langle \mathsf{U}^n x, y\rangle| \ge \varepsilon\} &= \{n \in \mathbb{N} : |\langle \mathsf{U}^n x', y\rangle + \langle \mathsf{U}^n(x - x'), y\rangle| \ge \varepsilon\} \\ &\subseteq \{n \in \mathbb{N} : |\langle \mathsf{U}^n x', y\rangle| + |\langle \mathsf{U}^n(x - x'), y\rangle| \ge \varepsilon\} \\ &\subseteq \left\{n \in \mathbb{N} : |\langle \mathsf{U}^n x', y\rangle| \ge \frac{\varepsilon}{2}\right\}\end{aligned}$$

which has density zero with respect to all Følner sequences. Hence, $x \in \mathscr{H}_{\mathsf{wm}}$.

- $\mathscr{H}_{\mathsf{wm}}$ is U-invariant: We have

$$\{n \in \mathbb{N} : |\langle \mathsf{U}^n(\mathsf{U}x), y\rangle| \ge \varepsilon\} = \{n \in \mathbb{N} : |\langle \mathsf{U}^{n+1}x, y\rangle| \ge \varepsilon\} \subseteq \{n \in \mathbb{N} : |\langle \mathsf{U}^n x, y\rangle| \ge \varepsilon\}$$

□

With this in mind, then the Jacobs-de Leeuw-Glicksberg splitting can be stated as follows.

Theorem 2.2.6. *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 order to use this result, we will need to extend the notions of compact and weak mixing functions to $\mathsf{L}^2(\mathbb{N}, \Phi)$.

Definition 1. A function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is called *compact* along Φ if, for every $\varepsilon > 0$, one can find $K \in \mathbb{N}$ such that

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

Definition 2. A function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is called *weak mixing* along Φ if, for every bounded function $h : \mathbb{N} \to \mathbb{C}$ 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

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

for all $\varepsilon > 0$.

We seek the following splitting result.

Theorem 2.2.7. *For every $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$, there exists 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 $a \le f_{\mathsf{c}} \le b$.*

We will employ the “there’s-nothing-to-do" strategy of corresponding the continuous and discrete definitions of compact/weak mixing, and directly utilizing the Jacobs-de Leeuw-Glicksberg splitting theorem. We will need to draw the bridge between the following analogues:

- $\langle F, H\rangle_\mu = \int_X F\overline{H}\ d\mu$
- $\|F\|_\mu^2 = \int_X |F|^2\ d\mu$
- $\|\mathsf{U}^m F - \mathsf{U}^k F\|_\mu$
- $\{n \in \mathbb{N} : |\langle \mathsf{U}^n F, H\rangle|_\mu \ge \varepsilon\}$ zero density

- $\langle f, h\rangle_\Phi = \lim_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Phi_N} f(n)\overline{g(n)}$
- $\|f\|_\Phi^2 = \limsup_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n\in\Phi_N} |f(n)|^2$
- $\|\mathsf{R}^m f - \mathsf{R}^k f\|_\Phi$
- $\overline{\mathsf{d}}_\Phi(\{n \in \mathbb{N} : |\langle \mathsf{R}^n f, h\rangle_\Phi| > \varepsilon\}) = 0$

Therefore, we want to find a compact metric space X and a probability measure μ on X such that it is easy to go between $\mathsf{L}^2(X, \mu)$ and $\mathsf{L}^2(\mathbb{N}, \Phi)$. Naturally, μ will be chosen to be the weak accumulation point of the measures

$$\mu_N := \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} \delta_{a_n}$$

for a properly chosen sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq X$ so that

$$\|F\|_\mu^2 = \limsup_{N\to\infty} \|F\|_{\mu_N}^2 = \limsup_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} \int_X |F|^2\ d\delta_{a_n} = \limsup_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} |F(a_n)|^2$$

Now for the aforementioned technical lemmas.

Lemma 2.2.1. *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. Let f, Ψ, and h satisfy the lemma statement, and let $\varepsilon > 0$. We have

$$|\langle \mathsf{R}^n f, h\rangle_\Psi| \leq \|\mathsf{R}^n f\|_\Psi \|h\|_\Psi = \|f\|_\Psi \|h\|_\Psi =: M$$

Additionally, for N sufficiently large,

$$\frac{1}{|\Psi_N|}\left|\left\{n \in \Psi_N : |\langle \mathsf{R}^n f, h\rangle_\Psi| > \frac{\varepsilon}{2}\right\}\right| < \frac{\varepsilon}{2M}$$

hence

$$\begin{aligned}
\frac{1}{|\Psi_N|}\sum_{n\in\Psi_N}|\langle \mathsf{R}^n f, h\rangle_\Psi| &= \frac{1}{|\Psi_N|}\sum_{\substack{n\in\Psi_N\\ |\langle \mathsf{R}^n f, h\rangle_\Psi|\leq\varepsilon/2}}|\langle \mathsf{R}^n f, h\rangle_\Psi| + \frac{1}{|\Psi_N|}\sum_{\substack{n\in\Psi_N\\ |\langle \mathsf{R}^n f, h\rangle_\Psi|>\varepsilon/2}}|\langle \mathsf{R}^n f, h\rangle_\Psi| \\
&\leq \frac{1}{|\Psi_N|}\sum_{\substack{n\in\Psi_N\\ |\langle \mathsf{R}^n f, h\rangle_\Psi|\leq\varepsilon/2}}\frac{\varepsilon}{2} + \frac{1}{|\Psi_N|}\sum_{\substack{n\in\Psi_N\\ |\langle \mathsf{R}^n f, h\rangle_\Psi|>\varepsilon/2}} M \\
&\leq \frac{\varepsilon}{2}\cdot\frac{|\Psi_N|}{|\Psi_N|} + M\cdot\frac{1}{|\Psi_N|}\left|\left\{n \in \Psi_N : |\langle \mathsf{R}^n f, h\rangle_\Psi| > \frac{\varepsilon}{2}\right\}\right| = \varepsilon
\end{aligned}$$

Sending $\varepsilon \to 0$, we find

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

as $N \to \infty$, as desired. ☐

Lemma 2.2.2. *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. Let Ψ be a subsequence for which $\|f\|_\Xi = \|f\|_\Phi$ for every subsequence Ξ of Ψ. Refine Ψ so that $\langle f, h\rangle_\Psi$ exists, and further so that $\langle \mathsf{R}^n f, \mathsf{R}^m h\rangle_\Psi$ exist for all $n, m \in \mathbb{N}$. It suffices to show that $\langle f, h\rangle_\Psi = 0$ for all $n, m \in \mathbb{N}$.

Let $\varepsilon > 0$, and pick $K \in \mathbb{N}$ such that $\|h\|_\Psi \|\mathsf{R}^m f - \mathsf{R}^k f\|_\Phi < \varepsilon$. We have

$$\begin{aligned}
|\langle f, h\rangle_\Psi| = |\langle \mathsf{R}^m f, \mathsf{R}^m h\rangle_\Psi| &\leq |\langle \mathsf{R}^m f - \mathsf{R}^k f, \mathsf{R}^m h\rangle_\Psi| + |\langle \mathsf{R}^k f, \mathsf{R}^m h\rangle_\Psi| \\
&\leq \|\mathsf{R}^m f - \mathsf{R}^k f\|_\Psi \|\mathsf{R}^m h\|_\Psi + \sum_{k=1}^{K}|\langle \mathsf{R}^k f, \mathsf{R}^m h\rangle_\Psi| \\
&= \|\mathsf{R}^m f - \mathsf{R}^k f\|_\Phi \|h\|_\Psi + \sum_{k=1}^{K}\limsup_{N\to\infty}\frac{1}{|\Psi_N|}\sum_{m\in\Psi_N}|\langle \mathsf{R}^k f, \mathsf{R}^m h\rangle_\Psi| \\
&= \|\mathsf{R}^m f - \mathsf{R}^k f\|_\Phi \|h\|_\Psi < \varepsilon
\end{aligned}$$

☐

For the remaining lemmas, we will need more language from ergodic theory. Given a compact space X, its Borel probability measure μ, and $\mathsf{T} : X \to X$ a measurable map that preserves μ, we call the triple (X, μ, T) a *measure preserving system*.

Lemma 2.2.3. *Let (X, μ, T) be a measure preserving system. For the isometry $\mathsf{U}f = f \circ \mathsf{T}$ of the Hilbert space $\mathsf{L}^2(X, \mu)$ the constant functions are compact, $|\varphi|$ is compact whenever φ is, and both $\min\{\varphi, \psi\}$ and $\max\{\varphi, \psi\}$ are compact whenever φ, ψ are compact and real-valued. Moreover, if $a \leq \varphi \leq b$ for some $a \leq b$, then $a \leq \varphi_{\mathsf{c}} \leq b$.*

Proof. Since the constant functions are fixed points of U, they are clearly compact. If φ is compact, then $|\varphi|$ is too since

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

The next part follows from

$$\min\{\varphi, \psi\} = \frac{\varphi + \psi - |\varphi - \psi|}{2}, \quad \max\{\varphi, \psi\} = \frac{\varphi + \psi + |\varphi - \psi|}{2}$$

And finally, using the fact that φ_{c} is the unique element of $\mathscr{H}_{\mathsf{c}}$ such that

$$\|\varphi - \varphi_{\mathsf{c}}\|_{\mathscr{H}} = \inf\{\|\varphi - \varphi'\|_{\mathscr{H}} : \varphi' \in \mathscr{H}_{\mathsf{c}}\}$$

we note that φ_{c} must be real-valued by comparison to $\mathrm{Re}(\varphi_{\mathsf{c}}) \in \mathscr{H}_{\mathsf{c}}$, and $a \leq \varphi_{\mathsf{c}} \leq b$ by comparison $\max\{\varphi_{\mathsf{c}}, a\}$ and $\min\{\varphi_{\mathsf{c}}, b\}$. □

Finally, the bridge between the discrete and continuous version of the Jacobs-de Leeuw-Glicksberg splitting. Recall that a dynamical system (X, S) is called *topologically transitive* if for any pair U, V of non-empty open subsets of X, there exists some $k \in \mathbb{N}$ such that $\mathsf{S}^k(U) \cap V \neq \varnothing$. We will show that every bounded sequence can be identified as the evaluation of some continuous function along the orbit of some point in some transitive topological system.

Lemma 2.2.4. *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 $\mathsf{S} : X \to X$, 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 $a_i(n) = F_i(\mathsf{S}^n x)$ for all $n \in \mathbb{N}$ and $i \in J$.*

Proof. For each $i \in J$, let $D_i \subseteq \mathbb{C}$ be a compact set containing the image of a_i, which exists since a_i is bounded. Let

$$Y := \prod_{i \in J} D_i^{\mathbb{N}}$$

which is compact by Tychonoff's theorem and corresponds to all sequences $y : J \times \mathbb{N} \to \mathbb{C}$ such that $y(i, n) \in D_i$ for all $n \in \mathbb{N}$ and $i \in J$. For $y \in Y$, define the action

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

which is continuous. Then the lemma statement is satisfied with $x(i, n) := a_i(n)$, X the orbit closure of x under S, and $F_i(y) := y(i, 0)$. □

Finally, we can proceed with the proof of Theorem 2.2.7.

Proof of Theorem 2.2.7. First assume $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ is bounded. Use Lemma 2.2.4 to find X, S, a continuous function $F : X \to \mathbb{C}$, and a point $x \in X$ with a dense orbit under S such that $F(\mathsf{S}^n(x)) = f(n)$ for all $n \in \mathbb{N}$. Using [Gla03] Theorem A.4, we can find a subsequence Ψ of Φ such that the measures

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

converge weakly to an S-invariant Borel probability measure μ on X. This is particularly advantageous (and motivated) due to the identity

$$\|H\|_\mu^2 = \int_X |H|^2 \, d\mu = \limsup_{N\to\infty} \int_X |H|^2 \, d\mu_N = \limsup_{N\to\infty} \frac{1}{|\Psi_N|} \sum_{n \in \Psi_N} |H(\mathsf{S}^n x)|^2 = \|n \mapsto H(\mathsf{S}^n x)\|_\Psi^2$$

for all $H \in \mathsf{C}(X)$. This defines our system (X, μ, S), along with the isometry U, $H \mapsto H \circ \mathsf{S}$. This allows us to use the continuous Jacobs-de Leeuw-Glicksberg splitting theorem to recover $F = F_{\mathsf{c}} + F_{\mathsf{wm}}$.

We need to use this decomposition to recover a subsequence Ψ of Φ and $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N}, \Psi)$ such that f_{c} and $f_{\mathsf{wm}} := f - f_{\mathsf{c}}$ are compact and weak mixing along Ψ, respectively. It is tempting to choose $f_{\mathsf{c}}(n) := F_{\mathsf{c}}(\mathsf{S}^n x)$, but it is not guaranteed that F_{c} is continuous, which will be important for later calculations. Hence, consider an approximation $\{H_j\}_{j\in\mathbb{N}} \subseteq \mathsf{C}(X)$ of F_{c} so that $\|F_{\mathsf{c}} - H_j\|_\mu < \frac{1}{j}$, let $h_j(n) = H_j(\mathsf{S}^n x)$ for all $n \in \mathbb{N}$, and extract f_{c} (with $f_{\mathsf{wm}} = f - f_{\mathsf{c}}$) as a limit of the (Cauchy) sequence $\{h_j\}_{j\in\mathbb{N}}$ via Lemma 2.1.1.

First we show f_{c} is compact along Ψ. Let $\varepsilon > 0$, and let $K \in \mathbb{N}$ be such that

$$\min\{\|\mathsf{S}^m F_{\mathsf{c}} - \mathsf{S}^k F_{\mathsf{c}}\|_\mu : 1 \le k \le K\} < \frac{\varepsilon}{2}$$

and let $j > \frac{8}{\varepsilon}$ so that $\|H_j - F_{\mathsf{c}}\|_\mu < \frac{\varepsilon}{8}$. Therefore,

$$\|\mathsf{R}^m f_{\mathsf{c}} - \mathsf{R}^k f_{\mathsf{c}}\|_\Psi \le \|\mathsf{R}^m h_j - \mathsf{R}^k h_j\|_\Psi + \frac{\varepsilon}{4} = \|\mathsf{S}^m H_j - \mathsf{S}^k H_j\|_\mu + \frac{\varepsilon}{4} \le \|\mathsf{S}^m F_{\mathsf{c}} - \mathsf{S}^k F_{\mathsf{c}}\|_\mu + \frac{\varepsilon}{2}$$

where the result follows from taking the minimum among $1 \le k \le K$. Additionally, if f takes values in $[a, b]$, then F does too, so F_{c} does too, so H_j can be chosen so that they do too, so h_j does too, and hence f_{c} does too.

Now we show f_{wm} is weak mixing along Ψ. Let $h : \mathbb{N} \to \mathbb{C}$ be bounded, and assume without loss of generality that $\langle \mathsf{R}^n f, h\rangle_\Psi$ exists for all $n \in \mathbb{N}$. We need a way to upper bound $\langle \mathsf{R}^n f, h\rangle_\Psi$ in terms of $\langle \mathsf{S}^n F, H\rangle_\mu$ so that $\{n \in \mathbb{N} : |\langle \mathsf{R}^n f, h\rangle_\Psi| > \varepsilon\}$ is the subset of a zero density set. Apply Lemma 2.2.4 again to get another $\tilde{X}$, $\tilde{\mathsf{S}} : \tilde{X} \to \tilde{X}$, $\tilde{F} \in \mathsf{C}(\tilde{X})$, and $\tilde{x} \in \tilde{X}$ with a dense orbit under $\tilde{\mathsf{S}}$ so that $h(n) = \tilde{F}(\tilde{\mathsf{S}}^n(\tilde{x}))$. Let $Z \subseteq X \times \tilde{X}$ be the orbit closure of $(x, \tilde{x})$ under $\mathsf{S} \times \tilde{\mathsf{S}}$. Pass to a subsequence Ψ' such that

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

coverges weakly to an invariant probability measure ν on Z. Note that $F_{\mathsf{wm}} \otimes 1 \in \mathsf{L}^2(Z, \nu)$ is weak mixing since

$$\varphi \in \mathsf{C}(X), \psi \in \mathsf{C}(\tilde{X}) \implies |\langle F_{\mathsf{wm}} \otimes 1, \varphi \otimes \psi\rangle_\nu| \le |\langle F_{\mathsf{wm}}, \varphi\rangle_\mu| \cdot \sup_{z \in \tilde{X}} |\psi(z)|$$

by Cauchy-Schwarz. Now let $\varepsilon > 0$, and pick $j \in \mathbb{N}$ large enough so that $\|h_j - f_{\mathsf{c}}\|_{\Psi'} \leq \frac{\varepsilon}{4}$.

$$\begin{aligned}
|\langle \mathsf{R}^m f_{\mathsf{wm}}, h\rangle|_\Psi = |\langle \mathsf{R}^m (f - f_{\mathsf{c}}), h\rangle|_\Psi &\leq |\langle \mathsf{R}^m f_{\mathsf{wm}}, h\rangle|_\Psi + \frac{\varepsilon}{4} \\
&= \left| \lim_{N\to\infty} \frac{1}{|\Psi'_N|} \sum_{n\in\Psi'_N} (f - h_j)(n+m)\overline{h(n)} \right| + \frac{\varepsilon}{4} \\
&= \left| \lim_{N\to\infty} \sum_{n\in\Psi'_N} \int_Z (\mathsf{S}\times\tilde{\mathsf{S}})^m ((F - H_j)\otimes 1)\overline{(1\otimes\tilde{F})}\, d\delta_{(\mathsf{S}\times\tilde{\mathsf{S}})^n(x,\tilde{x})} \right| + \frac{\varepsilon}{4} \\
&= \left| \int_Z (\mathsf{S}\times\tilde{\mathsf{S}})^m ((F - H_j)\otimes 1)\overline{(1\otimes\tilde{F})}\, d\nu \right| + \frac{\varepsilon}{4} \\
&\leq \left| \int_Z (\mathsf{S}\times\tilde{\mathsf{S}})^m ((F - H_j)\otimes 1)\overline{(1\otimes\tilde{F})}\, d\nu \right| + \frac{\varepsilon}{2}
\end{aligned}$$

Finally, conclude that f_{wm} is mixing via

$$\{n\in\mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}}, h\rangle_\Psi| > \varepsilon\} \subseteq \left\{ n\in\mathbb{N} : \left| \int_Z (\mathsf{S}\times\tilde{\mathsf{S}})^m ((F - H_j)\otimes 1)\overline{(1\otimes\tilde{F})}\, d\nu \right| > \frac{\varepsilon}{2} \right\}$$

Now, assume $f \in \mathsf{L}^2(\mathbb{N},\Phi)$ is arbitrary, and pick a sequence $\{f_j\}_{j\in\mathbb{N}}$ of bounded functions converging to f with respect to Φ. Construct a sequence $\{\Psi^{(j)}\}_{j\in\mathbb{N}}$ of nested Følner sequences, starting with $\Psi^{(0)} = \Phi$, so that $f_j = f_{j,\mathsf{c}} + f_{j,\mathsf{wm}}$ where $f_{j,\mathsf{c}}$ and $f_{j,\mathsf{wm}}$ are compact and weak mixing along $\Psi^{(j)}$, respectively. We claim that $\{f_{j,\mathsf{c}}\}_{j\in\mathbb{N}}$ is Cauchy with respect to Ψ. Indeed,

$$\begin{aligned}
\|f_j - f_\ell\|_\Psi^2 &= \|f_{j,\mathsf{c}} + f_{j,\mathsf{wm}} - f_{\ell,\mathsf{c}} - f_{\ell,\mathsf{wm}}\|_\Psi^2 \\
&= \|f_{j,\mathsf{c}} - f_{\ell,\mathsf{c}}\|_\Psi^2 + \|f_{j,\mathsf{wm}} - f_{\ell,\mathsf{wm}}\|_\Psi^2 + \sum_{i_1,i_2\in\{j,\ell\}} \langle f_{i_1,\mathsf{c}}, f_{i_2,\mathsf{wm}}\rangle_\Psi \\
&= \|f_{j,\mathsf{c}} - f_{\ell,\mathsf{c}}\|_\Psi^2 + \|f_{j,\mathsf{wm}} - f_{\ell,\mathsf{wm}}\|_\Psi^2 \geq \|f_{j,\mathsf{c}} - f_{\ell,\mathsf{c}}\|_\Psi^2
\end{aligned}$$

so it inherits Cauchyness from the Cauchyness of $\{f_j\}_{j\in\mathbb{N}}$ with respect to Φ. Refining Ψ and applying Lemma 2.1.1, we get a compact function $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N},\Psi)$ such that $\|f_{j,\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \to 0$ as $j\to\infty$. With $f_{\mathsf{wm}} = f - f_{\mathsf{c}}$, we find $\|f_{\mathsf{wm}} - f_{j,\mathsf{wm}}\|_\Psi \to 0$ as $j\to\infty$, so f_{wm} is weak mixing. □

2.2.3 Reduction of Theorem 2.1.4

Recall the statement of Theorem 2.1.4:

Theorem 2.1.4. *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. Then for every $\varepsilon > 0$, there exists a Følner 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*

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

First, we reduce this theorem a bit.

Theorem 2.2.8. *Fix $\varepsilon > 0$ 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, nonnegative, 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. $\{n \in \mathbb{N} : \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi < \varepsilon/3\} \in \mathsf{p}$,

U3. $\|\mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}} - f_{\mathsf{Bes}}\|_\Psi < \varepsilon/3$, *and*

U4. $\langle f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq 0$.

Proof of Theorem 2.2.8 $\implies$ *Theorem 2.1.4.* First, employ the entire work of Sections 2.2.1 and 2.2.2 to write $f = f_{\mathsf{Bes}} + f_{\mathsf{anti}} = f_{\mathsf{c}} + f_{\mathsf{wm}}$ for some subsequence Ψ of Φ. As previously mentioned, we can in general decompose

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

Therefore it suffices to find a non-principal ultrafilter $\mathsf{p} \in \beta\mathbb{N}$ to put this all together. Namely, we seek the inequalities

$$\langle \mathsf{R}^n f_{\mathsf{wm}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi = 0, \quad \langle \mathsf{R}^n f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f\rangle_\Psi \geq -\frac{\varepsilon}{3}, \text{ and } \quad \langle \mathsf{R}^n f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq \langle 1, f\rangle_\Psi^2 - \frac{2\varepsilon}{3}$$

as $n \to \mathsf{p}$.

To verify the hypotheses of Theorem 2.2.8, we need f_{Bes} to be bounded and nonnegative (which follows from f being bounded and nonnegative), $f_{\mathsf{anti}} = f - f_{\mathsf{Bes}}$ to be bounded and real-valued (obvious), and f_{c} to be bounded, nonnegative, and compact along Ψ. Therefore, using $\varepsilon/\|f\|_\Phi$, we can recover a desirable ultrafilter p satisfying U1–U4. We now can bound the three terms.

1. Given $\delta > 0$, we have

$$\{n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}}, \mathsf{R}^{\mathsf{p}}\rangle_\Psi| < \delta\} \notin \mathsf{p} \implies \{n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}}, \mathsf{R}^{\mathsf{p}}\rangle_\Psi| \geq \delta\} \in \mathsf{p}$$

where the latter has zero density, contradicting U1. Therefore, $\{n \in \mathbb{N} : |\langle \mathsf{R}^n f_{\mathsf{wm}}, \mathsf{R}^{\mathsf{p}}\rangle_\Psi| < \delta\} \in \mathsf{p}$. Sending $\delta \to 0$, we have

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

2. We have

$$|\langle \mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi| \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f_{\mathsf{anti}}\|_\Psi \leq \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_\Psi \|f\|_\Psi < \frac{\varepsilon}{3}$$

by Cauchy-Schwarz and U2. Therefore, by U4,

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq \langle f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi - \frac{\varepsilon}{3} \geq -\frac{\varepsilon}{3}$$

3. We have

$$\begin{aligned}\langle f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}}\rangle_\Psi &\overset{CS+U3}{\geq} \langle f_{\mathsf{c}}, f_{\mathsf{Bes}}\rangle_\Psi - \frac{\varepsilon}{3} = \|f_{\mathsf{Bes}}\|_\Psi^2 + \langle f_{\mathsf{c}} - f_{\mathsf{Bes}}, f_{\mathsf{Bes}}\rangle_\Psi - \frac{\varepsilon}{3} \\ &= \|f_{\mathsf{Bes}}\|_\Psi^2 + \langle f_{\mathsf{anti}} - f_{\mathsf{wm}}, f_{\mathsf{Bes}}\rangle_\Psi - \frac{\varepsilon}{3} \\ &\overset{CS}{\geq} \langle 1, f_{\mathsf{Bes}}\rangle_\Psi^2 + \langle f_{\mathsf{anti}} - f_{\mathsf{wm}}, f_{\mathsf{Bes}}\rangle_\Psi - \frac{\varepsilon}{3} = \langle 1, f\rangle_\Psi^2 - \frac{\varepsilon}{3}\end{aligned}$$

Therefore, by U2 we have

$$\lim_{n\to\mathsf{p}} \langle \mathsf{R}^n f_{\mathsf{c}}, \mathsf{R}^{\mathsf{p}} f_{\mathsf{anti}}\rangle_\Psi \geq \langle 1, f\rangle_\Psi^2 - \frac{2\varepsilon}{3}$$

□

2.2.4 Proof of Theorem 2.2.8

Since non-principal ultrafilters are generally hard to explicitly construct (especially without the axiom of choice), most reasoning about non-principal ultrafilters comes from the topological and measure theoretical properties of $\beta\mathbb{N}$. In particular, our approach is to find a set of ultrafilters of strictly positive measure such that almost all of them satisfy U1–U4, and this works because the set of principal ultrafilters has measure zero.

Recall that $\mathcal{M}(\Phi)$ is 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

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

We say a property holds Φ *almost everywhere* if the set of ultrafilters with that property has full measure with respect to every $\mu \in \mathcal{M}(\Phi)$.

Lemma 2.2.5. *Let Φ be a Følner sequence on $\mathbb{N}$. Then Φ almost every $\mathsf{p} \in \beta\mathbb{N}$ satisfies $\overline{\mathsf{d}}_\Phi(E) > 0$ for every $E \in \mathsf{p}$.*

Proof. Let $\mu \in \mathcal{M}(\Phi)$, and define

$$\mathsf{Ess}(\Phi) := \{\mathsf{p} \in \beta\mathbb{N} : \overline{\mathsf{d}}_\Phi(E) > 0 \ \forall\ E \in \mathsf{p}\} = \bigcap_{\substack{E \subseteq \mathbb{N} \\ \overline{\mathsf{d}}_\Phi(E)=0}} \mathsf{cl}(\mathbb{N} \setminus E) = \beta\mathbb{N} \setminus \bigcup_{\substack{E \subseteq \mathbb{N} \\ \overline{\mathsf{d}}_\Phi(E)=0}} \mathsf{cl}(E)$$

which is closed and hence Borel. Since $\bigcup_{\substack{E \subseteq \mathbb{N} \\ \overline{\mathsf{d}}_\Phi(E)=0}} \mathsf{cl}(E)$ is compact and also an open covering of itself, let $\{\mathsf{cl}(E_\alpha)\}_{\alpha \in I}$ be a finite subcover. Then note that $\mu(\mathsf{cl}(E_\alpha)) \leq \overline{\mathsf{d}}_\Phi(E_\alpha) = 0$, which allows us to conclude that $\mu(\mathsf{Ess}(\Phi)) = 1$. □

In order to attain our goals, we will work with Bohr and Bohr$_0$ sets, a class of sets that are relatively “large" and have desirable topological properties, such as being closed under intersections and playing well with the spaces we have already studied.

Definition 3. 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. If U contains the identity element of K, it is called a *Bohr$_0$ set.*

Lemma 2.2.6. *If A and B are Bohr$_0$ sets, then so is $A \cap B$.*

Proof. If $A = a^{-1}(U)$ for $a : \mathbb{N} \to K$ and $B = b^{-1}(V)$ for $b : \mathbb{N} \to L$, then $A \cap B = c^{-1}(U \times V)$, where $c : \mathbb{N} \to K \times L,\ n \mapsto (a(n), b(n))$ is a homomorphism. □

Lemma 2.2.7. *If $B \subseteq \mathbb{N}$ is a non-empty Bohr set, then for every Følner sequence Φ its indicator function 1_B is in $\mathsf{Bes}(\mathbb{N}, \Phi)$ and $\mathsf{d}_\Phi(B) > 0$.*

The key idea is using strict positivity of the Haar measure on non-empty sets.

Proof. Suppose $B = a^{-1}(U)$ where $a : \mathbb{N} \to K$. Assume without loss of generality that $a(\mathbb{N})$ has a dense image else we can replace K with $\overline{a(\mathbb{N})}$. Then we have

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

where

$$\mu_N = \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} \delta_{a(n)}$$

Letting μ be a weak limit point of $\{\mu_N\}_{N\in\mathbb{N}}$, we claim that μ is the Haar measure on K. Since μ is a weak limit point, it must be invariant under $a(\mathbb{N})$. Also, it's easy to see that $\overline{a(\mathbb{N})}$ is a subgroup of K, so μ is Haar. We therefore have

$$\mathsf{d}_\Phi(B) = \lim_{N\to\infty} \mu_N(U) = \mu(U) > 0$$

by [Gla03] Theorem A.5.

Finally, note that the set of linear combinations of $\overline{a(\mathbb{N})}$ is dense in $\mathsf{L}^2(K, m)$, so for all $\varepsilon > 0$ we can find some f such that $\|f - 1_U\|_\mu < \varepsilon$. By continuity, $\|f \circ a - 1_B\|_\Phi = \|f - 1_U\|_\mu < \varepsilon$, where $f \circ a$ is a trigonometric polynomial. We conclude that $1_B \in \mathsf{Bes}(\mathbb{N}, \Phi)$. ☐

Lemma 2.2.8. *Given $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ that is compact along Φ and $\varepsilon > 0$, the set $\{n \in \mathbb{N} : \|\mathsf{R}^n f - f\|_\Phi < \varepsilon\}$ contains a Bohr$_0$ set B_c.*

Proof. Let $g(n) := \|\mathsf{R}^{|n|} f - f\|_\Phi$ for $n \in \mathbb{Z}$. Then $\Omega := \overline{\{\mathsf{R}^k g : k \in \mathbb{Z}\}}$ is compact since f is compact along Φ. By defining $(\mathsf{R}^n g) \star (\mathsf{R}^k g) := \mathsf{R}^{k+n} g$ for $n, k \in \mathbb{Z}$ and extending to all of Ω by continuity, we make Ω into a compact topological group. Let $U_\varepsilon := \{\varphi : \mathbb{Z} \to [0, \infty) : \varphi(0) < \varepsilon\}$. Furthermore, $a(n) := \mathsf{R}^n g$ is a homomorphism from $\mathbb{N}$ to $(\Omega, \star)$, and

$$\{n \in \mathbb{N} : \|\mathsf{R}^n f - f\|_\Phi < \varepsilon\} = \{n \in \mathbb{N} : a(n) \in U_\varepsilon\} \supseteq \{n \in \mathbb{N} : a(n) \in U_\varepsilon\}$$

for all $0 < \eta < \varepsilon$. Picking $\eta > 0$ such that ∂U_η has measure 0, $B_\mathsf{c} := a^{-1}(U_\eta)$ satisfies the problem statement. ☐

Lemma 2.2.9. *Suppose $f \in \mathsf{Bes}(\Psi)^\perp$ is real-valued and bounded. Then for every non-empty Bohr set $B \subseteq \mathbb{N}$ and every bounded function $h : \mathbb{N} \to \mathbb{R}$, the set*

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

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

Proof. Let $\mu \in \mathcal{M}(\Psi)$. By Lemma 2.2.7, d_Ψ exists and is positive, so $\mu(\mathsf{cl}(B)) = \mathsf{d}_\Psi(B) > 0$, which allows us to define a new probability measure $\mu_B(\Omega) := \mu(\Omega \cap \mathsf{cl}(B))/\mu(\mathsf{cl}(B))$.

We need to show that the map

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

is measurable, which comes from the fact that the map $\mathsf{p} \mapsto (\mathsf{R}^{\mathsf{p}} f)(n) = \lim_{m\to\mathsf{p}} f(n+m)$ is continuous. From here, we want to show that

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

or sufficiently (by Fatou's lemma),

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

However, the result follows from the bound

$$\begin{aligned}
\left| \int_{\beta\mathbb{N}} (\mathsf{R}^{\mathsf{p}} f)(n) \; d\mu_B(\mathsf{p}) \right| &= \frac{1}{\mu(\mathsf{cl}(B))} \left| \int_{\beta\mathbb{N}} 1_{\mathsf{cl}(B)}(\mathsf{p})(\mathsf{R}^{\mathsf{p}} f)(n) \; d\mu(\mathsf{p}) \right| \\
&\leq \limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} 1_B(m) f(n+m) \right| \\
&= \limsup_{N\to\infty} \left| \frac{1}{|\Psi_N|} \sum_{m\in\Psi_N} 1_{B+n}(m) f(m) \right| = 0
\end{aligned}$$

since $f \in \mathsf{Bes}(\mathbb{N}, \Psi)^{\perp}$ and $m \mapsto 1_{B+n}(m)$ is Besicovitch almost periodic along Ψ. □

Finally, we need one more ingredient to complete the proof of Theorem 2.2.8. The proof to this is long and is found in the next section.

Theorem 2.2.9. *Let Φ be a Følner sequence on $\mathbb{N}$ and let $f \in \mathsf{Bes}(\mathbb{N}, \Phi)$. For every $\varepsilon > 0$ there exists a Bohr$_0$ set B and a subsequence Ψ of Φ such that for Ψ almost every $\mathsf{p} \in \mathsf{cl}(B)$, $\|\mathsf{R}^{\mathsf{p}} f - f\|_{\Psi} < \varepsilon$.*

Proof of Theorem 2.2.8. Let $\varepsilon > 0$, Φ, $f_{\mathsf{Bes}} \in \mathsf{Bes}(\mathbb{N}, \Phi)$, $f_{\mathsf{anti}} \in \mathsf{Bes}(\mathbb{N}, \Phi)$, and $f_{\mathsf{c}} \in \mathsf{L}^2(\mathbb{N}, \Phi)$ be given as in the theorem statement.

From Lemma 2.2.8, obtain a Bohr$_0$ set $B_{\mathsf{c}} \subseteq \{n \in \mathbb{N} : \|\mathsf{R}^n f_{\mathsf{c}} - f_{\mathsf{c}}\|_{\Phi} < \varepsilon/3\}$. From Theorem 2.2.9, obtain a subsequence Ψ of Φ and a Bohr$_0$ set B_{Bes} such that $\|\mathsf{R}^{\mathsf{p}} f_{\mathsf{Bes}} - f_{\mathsf{Bes}}\|_{\Psi} < \varepsilon/3$ for Ψ almost every $\mathsf{p} \in \mathsf{cl}(B_{\mathsf{Bes}})$. Intersect these to get $B := B_{\mathsf{c}} \cap B_{\mathsf{Bes}}$, also Bohr$_0$. Finally, apply Lemma 2.2.9 to find

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

has positive measure with respect to all $\mu \in \mathcal{M}(\Phi)$. Finally, note that Ψ almost every $\mathsf{p} \in \mathsf{P}$ satisfies

U1. because Lemma 2.2.5,

U2. because $\mathsf{p} \in \mathsf{cl}(B_{\mathsf{c}})$,

U3. because $\mathsf{p} \in \mathsf{cl}(B_{\mathsf{Bes}})$, and

U4. because $\mathsf{p} \in \mathsf{P}$.

Therefore, we can find one such p non-principal. □

2.2.5 Proof of Theorem 2.2.9

TODO

2.3 Discussion

There are several components that worked magically well together in this proof and would generalize to a vast domain of Ramsey-type problems.

First, there's the ultrafilters, which have a beautiful topology and translate smoothly into a combinatorial interpretation. In this proof, their main combinatorial use is that if p is a non-principal ultrafilter, every element of p must necessarily be infinite. For finite Ramsey problems, it's often easy to create constructions using induction, but when we are doing infinite Ramsey-type problems, we need more structure about how these infinite sets interact, especially under intersections. Plus, the ultrafilters play especially well with measure theory, so when it comes time to actually find a non-principal ultrafilter we can reduce it to finding a set of ultrafilters with positive measure.

Then, there's the Følner sequences. A space like $\mathsf{L}^2(\mathbb{N}, \Phi)$ has a great Hilbert-like structure due to the Besicovitch seminorm and similar inner product. They play well with shifts, and the majority of the time they behave very similarly to a Hilbert space upon passing to the proper subsequences. It would be interesting to define a function space that is a weaker notion than Hilbert spaces but provides the same benefits as them, particularly the completeness result. For instance, it may be of interest to study a space with a sequence of norms and inner products, and determine the behavior of the space under the limiting behavior of the norms and inner products.

One of the biggest difficulties in the proof was converting from the combinatorial setting to the ultrafilter setting. This construction was made inductively by restricting our candidates for B and C. This allowed us to have more structure in how we selected our values for B and C.

Now let's re-examine the two decompositions used, why they were chosen, and how we proved them for $\mathsf{L}^2(\mathbb{N}, \Phi)$. The first one was decomposing $\mathsf{L}^2(\mathbb{N}, \Phi)$ into $\mathsf{Bes}(\mathbb{N}, \Phi)$ and $\mathsf{Bes}(\mathbb{N}, \Phi)^{\perp}$. $\mathsf{Bes}(\mathbb{N}, \Phi)$ is a very structured space that behaves well under linear shifts due to the periodic nature of complex exponentials. As with most orthogonal decompositions, our main goal was to project our function $f \in \mathsf{L}^2(\mathbb{N}, \Phi)$ onto $\mathsf{Bes}(\mathbb{N}, \Phi)$, however we cannot quite simply project our function because we do not have the right Hilbert structure. Therefore, we had to work around this shortcoming by using technical lemmas, but we have the intuition of what to prove and managed to make it work.

For the compact and weak mixing decomposition, we needed a decomposition that reacts better to shifts by ultrafilters in order to bound the quantity in Theorem 2.1.4. Conveniently, there is a general splitting technique for Hilbert spaces (the Jacobs-de Leeuw-Glicksberg) which we could actually steal by embedding our ultrafilters into a compact metric space along with an isometry (corresponding to R^n), along with the proper measure that aligns with the norm on $\mathsf{L}^2(\mathbb{N}, \Phi)$. From here, we borrow some ideas from dynamical systems to complete the decomposition.

Finally, after picking and proving the right two decompositions, we were able to bound the cross terms to get the desired bound.

Chapter 3

Extending Moreira's results

3.1 Problem statement

We chose to approach Question 6.3 in [MRR19].

Conjecture 2. Suppose $A \subseteq \mathbb{N}$ has positive upper density. Then there exists infinite sets $B, C, D \in \mathbb{N}$ such that $B + C + D$ is contained in A.

Conjecture 3. Suppose $A \subseteq \mathbb{N}$ has positive upper density. Then for all $k \in \mathbb{N}$, there exist infinite sets $B_1, \ldots, B_k \in \mathbb{N}$ such that $B_1 + \cdots + B_k$ is contained in A.

Let's just examine Conjecture 2 for now. The very first no-work approaches one could pose right after proving the Erdos sumset conjecture would be: given A with positive upper density and B, C infinite such that $B + C \subseteq A$, (1) does either of B, C necessarily have positive upper density? If so, we could find infinite D, E such that $D + E$ is contained within that set just by applying the theorem one more time. (2) Is there a smarter construction (as in Section 2.1.1) that we could make to guarantee this density condition? (3) Is it possible that the exact same construction works, but by choosing p more carefully to guarantee the density condition? (4) Can we somehow embed the three subsets into the problem statement?

A completely different approach would be (5) by trying to use a method more general than ultrafilters that could accomodate more than 2 sets? (6) Could we find a completely different construction that would still admit ultrafilters?

We examine each of these proposals below.

1. This is not true. For instance, if $A = \mathbb{N}$, then any B, C satisfy $B + C \subseteq A$, so they could be any infinite sets and need not have positive upper density. It was worth a try though.

2. Recall the current construction, which first finds a set $L \subseteq \mathbb{N}$ of target values of B and a sequence $\{e_n\}_{n \in \mathbb{N}}$ of target values of C. We're first and foremost going to need to have L and $\{e_n\}_{n \in \mathbb{N}}$, bottom line. In order to get our bound on $\{e_n\}_{n \in \mathbb{N}}$, we would need to replace

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

with something where we can control the gaps like

$$\bigcap_{\ell \in F} (A - \ell) \cap \{m \in \mathbb{N} : \mathsf{d}_\Phi((A - m) \cap L) > \varepsilon\} \text{ is piecewise syndetic}$$

which is a reasonable choice since piecewise syndetic sets play well with ultrafilters. We would also need to bound our indexing sequence σ, which we got by defining $f_N := \frac{1}{N}\sum_{i=1}^{n} 1_{B_n}$, $f = \limsup_{N\to\infty} f_N$ and proving $\int_X f \, d\mu > 0$, then finding a point where $f(x) > 0$. We would need to choose f_N in such a way so that $f(x) > 0$ implies that there exists an enumeration of that value that has positive upper density. Furthermore, we would need to bound the density of L, which was constructed via $L := A - \mathsf{p}$. Nonprincipal ultrafilters necessarily contain only infinite sets, so is there a way to ensure they contain sets of positive upper density?

3. We ended up finding quite a large number of set of ultrafilters that satisfied Theorem 2.2.8. In fact, we found a set of ultrafilters of strictly positive measure that satisfy Theorem 2.2.8. Is it possible that we could impose more conditions on p so that the construction works as-is? All we are looking for is a way to find a bound that comes with the sets B, C that we constructed.

4. Moreira did prove the Erdos sumset conjecture for countable amenable groups, so we are working with something more general than just subsets of $\mathbb{N}$. Is there a way we can embed our setting into some different countable amenable group A' so that a sumset $B' + C'$ in that group corresponds to (or at least must contain) a sumset $B + C + D \subseteq A$? Often times embeddings might allow us to get a proof to the more general problem very cheaply.

5. This is the least plausible approach in my opinion, though certainly worth mentioning. Most of the paradigms in our analysis are binary. Ultrafilters rely on the binary $A \in \mathsf{p} \iff A^c \notin \mathsf{p}$. Inner products are a binary operation that take two functions as arguments. One possible three-function analog to an inner product is perhaps a convolution, perhaps combining f and g via a convolution such as

$$\int_0^t \int_0^{\tau_1} f(\tau_1)g(\tau_2)h(t - \tau_1 - \tau_2)\, d\tau_2\, d\tau_1$$

   though it doesn't make too much sense because usually in inner products there is a conjugate attached to the second term.

   However, this approach isn't *completely* unreasonable. Ramsey's theorem was originally stated as coloring the edges (*pairs* of vertices) in a graph, but it was generalized to higher order hypergraphs. Another domain to consider perhaps is 3-colorings.

6. When constructing inductively, we would need the following three (instead of two) criteria:

$$a_{n+1} \in \bigcap_{\substack{i\le n\\ j\le n}} (A - b_i - c_j), \quad b_{n+1} \in \bigcap_{\substack{i\le n+1\\ j\le n}} (A - a_i - c_j), \quad \text{and } c_{n+1} \in \bigcap_{\substack{i\le n+1\\ j\le n+1}} (A - a_i - b_j)$$

Applying the techniques in the Moreira paper, we will likely need to begin with the right ultrafilter reformulation.

Bibliography

[Gla03] Eli Glasner. *Ergodic theory via joinings*. Number 101. American Mathematical Soc., 2003.

[Mor17] Joel Moreira. Monochromatic sums and products in âĎŢ. *Annals of Mathematics*, pages 1069–1090, 2017.

[MRR19] Joel Moreira, Florian K Richter, and Donald Robertson. A proof of a sumset conjecture of erdos. *Annals of Mathematics*, 189(2):605–652, 2019.