Characteristic Factors for Multiple Recurrence and Combinatorial Applications

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By
Donald Robertson, MPhys
Graduate Program in Mathematics
The Ohio State University
2015

Dissertation Committee:
Vitaly Bergelson, Advisor
Neil Falkner
Alexander Leibman
ABSTRACT

Since Furstenberg’s proof of Szemerédi’s theorem in the 1970s a mutually beneficial relationship has blossomed between ergodic theory and density Ramsey theory. In this thesis we contribute to this relationship by using the method of characteristic factors to prove some extensions and generalizations of Szemerédi’s theorem, building on recent work by many authors. Specifically, in Chapter 3 we prove a two-sided version of the Furstenberg correspondence theorem for amenable groups, in Chapter 4 we describe characteristic factors for correlations that generalize Szemerédi’s theorem to amenable groups, in Chapter 5 we enlarge the collections of polynomials for which the multidimensional polynomial Szemerédi theorem ins known to hold, and in Chapter 7 we prove that every large subset of an amenable group contains a two-sided finite products set.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Vitaly Bergelson for his commitment to his students and his infectious enthusiasm for mathematics. His tutelage has in no small part contributed to the production of this thesis and to my development as a mathematician.

I would like to thank committee members Neil Falkner and Alexander Leibman for their patience and their participation in several beneficial discussions.

Many thanks to co-authors Cory Christopherson and Pavel Zorin-Kranich, and fellow students Moy, John, Bill, Younghwan, AJ, Joel, Daniel, Florian, Marc and Runlin, with whom I enjoyed working and from whom I learned a great deal.

Finally, heartfelt gratitude to Yanyan Zhang for her support and understanding.
VITA

June 2004 .................................. Portree High School

June 2008 .................................. MPhys Mathematics and Theoretical Physics
                                         University of St Andrews

June 2008 to Present ....................... Graduate Teaching Associate
                                         Department of Mathematics
                                         Ohio State University

PUBLICATIONS

V. Bergelson and D. Robertson. “Polynomial multiple recurrence over rings of integers”. In: Ergodic Theory and Dynamical Systems FirstView (Mar. 2015), pp. 1–25. issn: 1469-4417

FIELDS OF STUDY

Major Field: Mathematics
Specialization: Ergodic Theory
CONTENTS

Abstract ................................................................. ii

Acknowledgements ..................................................... iii

Vita ................................................................. iv

List of Figures ......................................................... vii

1 Introduction ......................................................... 1
   1.1 Combinatorial problems ...................................... 1
   1.2 From intersections to correlations .......................... 3
   1.3 Characteristic factors ....................................... 5

2 Preliminaries ....................................................... 7
   2.1 Probability spaces and group actions ...................... 7
   2.2 Uniformly continuous functions ............................. 9
   2.3 Amenable groups ............................................. 10
   2.4 Disintegrations .............................................. 14
   2.5 Joinings .................................................... 16
   2.6 Ultrafilters .................................................. 17
   2.7 Structured sets in groups .................................. 26
   2.8 Finite IP sets .............................................. 31

3 Correspondence Principles ........................................ 34
   3.1 A correspondence principle for amenable groups .......... 34
   3.2 A correspondence principle for discrete amenable groups . 37

4 Characteristic Factors ............................................ 39
   4.1 Borel-Hilbert bundles ...................................... 39
   4.2 Almost periodicity and weak mixing ........................ 42
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>The dichotomy</td>
<td>43</td>
</tr>
<tr>
<td>4.4</td>
<td>Characteristic factors for amenable averages</td>
<td>48</td>
</tr>
<tr>
<td>4.5</td>
<td>Lifting positivity</td>
<td>51</td>
</tr>
<tr>
<td>4.6</td>
<td>Recurrence results</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>Polynomial Recurrence</td>
<td>58</td>
</tr>
<tr>
<td>5.1</td>
<td>Polynomial Szemerédi theorems</td>
<td>58</td>
</tr>
<tr>
<td>5.2</td>
<td>Single recurrence for polynomials over rings of integers</td>
<td>62</td>
</tr>
<tr>
<td>5.3</td>
<td>Examples of intersective polynomials</td>
<td>65</td>
</tr>
<tr>
<td>5.4</td>
<td>Gowers-Host-Kra norms for commuting actions</td>
<td>66</td>
</tr>
<tr>
<td>5.5</td>
<td>Characteristic factors for some polynomial averages</td>
<td>70</td>
</tr>
<tr>
<td>5.6</td>
<td>Multiple recurrence for polynomials over rings of integers</td>
<td>74</td>
</tr>
<tr>
<td>5.7</td>
<td>Single recurrence for polynomials over countable fields</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>Satedness</td>
<td>83</td>
</tr>
<tr>
<td>6.1</td>
<td>Relative independence</td>
<td>83</td>
</tr>
<tr>
<td>6.2</td>
<td>Sated systems and sated extensions</td>
<td>85</td>
</tr>
<tr>
<td>6.3</td>
<td>Characteristic factors via satedness</td>
<td>89</td>
</tr>
<tr>
<td>6.4</td>
<td>Approximate satedness</td>
<td>91</td>
</tr>
<tr>
<td>6.5</td>
<td>Ergodicity and satedness</td>
<td>92</td>
</tr>
<tr>
<td>7</td>
<td>Mixing Groups</td>
<td>95</td>
</tr>
<tr>
<td>7.1</td>
<td>Weakly mixing groups</td>
<td>95</td>
</tr>
<tr>
<td>7.2</td>
<td>Characteristic factors</td>
<td>96</td>
</tr>
<tr>
<td>7.3</td>
<td>Finite products sets</td>
<td>100</td>
</tr>
<tr>
<td>7.4</td>
<td>Limits along minimal idempotent ultrafilters</td>
<td>102</td>
</tr>
</tbody>
</table>

Bibliography ............................................. 106
LIST OF FIGURES

4.1 Characteristic factors for commuting actions of amenable groups 49
6.1 Functorial sub-$\sigma$-algebra 85
6.2 Satedness 86
6.3 Existence of sated extensions 88
CHAPTER 1

INTRODUCTION

This thesis is concerned with proving results in density combinatorics using ergodic theory. In this introductory chapter we give some examples of problems in density combinatorics, describe the connection with ergodic theory by outlining Furstenberg’s ergodic proof of Szemerédi’s theorem, and state some of the combinatorial results we will prove in this thesis.

1.1 Combinatorial problems

The types of combinatorial problems we are interested in come from density Ramsey theory. The schema for problems in this area of combinatorics is as follows: given a set with some structure, does any large enough subset have some of the same structure? In order to make such questions precise, one needs to decide what one means by structure and how the size of a subset is to be determined. The prototypical example of such a problem is Szemerédi’s theorem, in which one considers collections of positive integers and structure refers to arithmetic progressions. By an arithmetic progression with \( k \) steps one means a set of the form

\[
\{a, a + d, \ldots, a + kd\}
\]

for some \( a, d \in \mathbb{N} \). Precisely, Szemerédi’s theorem is as follows.

**Theorem 1.1.2** (Szemerédi theorem [Sze75]). For every \( k \in \mathbb{N} \) and every \( \delta > 0 \) there is \( N \in \mathbb{N} \) such that any subset \( A \) of \( \{1, \ldots, n\} \) with \( |A| \geq \delta n \) and \( n \geq N \) contains an arithmetic progression with \( k \) steps.

This result was conjectured by Erdős and Turán [ET36] in 1936. The \( k = 3 \) case was first proved in 1952 by Roth [Rot52], and the case \( k = 4 \) was settled by Szemerédi [Sze69] in 1969. In 1977, Furstenberg [Fur77] gave a proof of Szemerédi’s theorem using ergodic theory, establishing a connection between dynamics and combinatorics that continues to benefit both subjects. We outline Furstenberg’s proof of Szemerédi’s theorem in the following
sections as it serves as the basis for many applications of ergodic theory to combinatorics, including those we will pursue in this thesis.

Szemerédi’s theorem has been generalized in many ways. One such generalization, which was a triumph of ergodic theory over combinatorics at the time, was Furstenberg and Katznelson’s multidimensional version.

**Theorem 1.1.3** (Multidimensional Szemerédi theorem [FK78]). For every $k, d \in \mathbb{N}$ and every $\delta > 0$ there is $N \in \mathbb{N}$ such that any set $A \subset \{1, \ldots, n\}^d$ with $n \geq N$ and $|A| \geq n^d \delta$ contains a set of the form

$$\{a, \ldots, ka\}^d + b$$

for some $a \in \mathbb{N}$ and some $b \in \mathbb{N}^d$.

The multidimensional Szemerédi theorem can be thought of as a version of Szemerédi’s theorem in the group $\mathbb{Z}^d$, and this raises the question of whether an analogous result holds in other groups. One can naturally phrase an analogue of Theorem 1.1.3 for an arbitrary countable amenable group; such a result was proved in the case analogous to $k = 2$ by Bergelson, McCutcheon and Zhang [BMZ97], and in general recently by Austin [Aus13]. In Chapter 4 we consider the extent to which Furstenberg’s first ergodic proof of Szemerédi’s theorem can be adapted to the setting of amenable groups.

One could also ask whether the step size $d$ in the progressions in Szemerédi’s theorem can have additional arithmetic properties. For example, can one guarantee a step size that is a prime number, or a perfect square? These questions are not obvious even for $k = 1$. In [Sár78] Sárközy proved that one can always choose $d$ to be of the form $p - 1$ for some prime $p$. That one can always choose $d$ to be a square when $k = 1$ was proved independently by Furstenberg [Fur81b] and Sárközy [Sár78].

More generally, and somewhat later, Bergelson and Leibman [BL96] gave a polynomial version of Furstenberg and Katznelson’s multidimensional Szemerédi theorem, a special case of which we now state.

**Theorem 1.1.4** (Polynomial Szemerédi Theorem [BL96]). For every $k, d \in \mathbb{N}$, every choice of polynomials $p_{1,1}, \ldots, p_{k,d}$ over $\mathbb{Z}$ with zero constant term, and every $\delta > 0$, there is $N \in \mathbb{N}$ such that any set $A \subset \{1, \ldots, n\}^d$ with $n \geq N$ and $|A| \geq n^d \delta$ contains a set of the form

$$\{p_{1,1}(a), \ldots, p_{k,1}(a)\} \times \cdots \times \{p_{1,d}(a), \ldots, p_{k,d}(a)\} + b$$

for some $a \in \mathbb{N}$ and some $b \in \mathbb{Z}^d$.

It is not only the polynomials with zero constant that verify Theorem 1.1.4. For example, when $k = 1$ and $d = 1$, the conclusion of Theorem 1.1.4 holds with $p_{1,1}(X) = X^2 - 1$, and with $p_{1,1}(X) = (X^2 - 13)(X^2 - 17)(X^2 - 221)$. The question of describing precisely those
families of polynomials for which Theorem 1.1.4 holds remains open in general. For the case \( d = 1 \) a complete description was given by Bergelson, Leibman and Lesigne [BLL08]. In Chapter 5 we will extend the description of those families of polynomials for which Theorem 1.1.4 is known to hold in by considering polynomials over algebraic number fields (see Theorem 5.1.5).

We now take the liberty of describing another result in density combinatorics – the density Hales-Jewett theorem – which is concerned not with progressions in sets of integers but with maps between sets of integers. Given \( n \in \mathbb{N} \), write \([n]\) for the set \( \{1, \ldots, n\} \). Fixing \( k \in \mathbb{N} \), for any \( d \in \mathbb{N} \) one can consider the set \([k]^{[d]}\). Its elements are called \textbf{words} and written as strings of length \( d \) without punctuation. The structures on \([k]^{[d]}\) that the density Hales-Jewett theorem is concerned with are the combinatorial lines, which one can define as follows.

\textbf{Definition 1.1.5.} Fix \( k, d \in \mathbb{N} \). A \textbf{combinatorial line} in \([k]^{[d]}\) is specified by the following data:

- a partition \( W_0 \sqcup W_1 \) of \([d]\) with \( W_1 \neq \emptyset \);
- a map \( c : W_0 \rightarrow [k] \);

and consists of all words \( w \in [k]^{[d]} \) that agree with \( c \) on \( W_0 \) and are constant on \( W_1 \).

For example, if \( k = 3 \) and \( d = 4 \) then

\[ \{12132, 22232, 32332\} \]

is a combinatorial line in \([3]^{[5]}\) with \( W_0 = \{2, 4, 5\} \) and \( W_1 = \{1, 3\} \).

\textbf{Theorem 1.1.6 (Density Hales-Jewett [FK91]).} For every \( k \in \mathbb{N} \) and every \( \delta > 0 \) there is \( D \in \mathbb{N} \) such that any set \( A \subset [k]^{[d]} \) with \( d \geq D \) and \( |A| \geq k^d \delta \) contains a combinatorial line.

The density Hales-Jewett theorem is quite powerful: not only does it imply Szemerédi’s theorem, but it also implies its own higher dimensional analogues, and therefore the multidimensional Szemerédi theorem as well. Proofs of these statements can be found in [McC99].

\section*{1.2 From intersections to correlations}

The conclusion of Szemerédi’s theorem (Theorem 1.1.2) is that \( A \) contains a set of the form (1.1.1), which is the same as saying that the intersection

\[ A \cap (A - d) \cap \cdots \cap (A - kd) \quad (1.2.1) \]

is non-empty for some \( d \in \mathbb{N} \). Loosely speaking, the conclusion has been phrased dynamically: in terms of a set \( A \) and its intersection with certain of its translates. In [Fur77]
Furstenberg showed that the following result implies Szemerédi’s theorem by introducing a formal bridge between density combinatorics and ergodic theory.

**Theorem 1.2.2.** Let $\mu$ be a probability measure on a measurable space $(X, \mathcal{B})$. For any invertible, measurable, measure-preserving transformation $T : X \to X$, any $B \in \mathcal{B}$ with $\mu(B) > 0$ and any $k \in \mathbb{N}$, one has

$$\mu(B \cap T^{-n}B \cap \cdots \cap T^{-kn}B) > 0$$

(1.2.3)

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

The following argument justifies the implication stated above.

**Proposition 1.2.4.** Theorem 1.2.2 implies Theorem 1.1.2.

*Proof.* Suppose that Theorem 1.1.2 is false. Then we can find $k \in \mathbb{N}$, $\delta > 0$ and a sequence $i \mapsto A_i$ of sets $A_i \subset \{1, \ldots, n_i\}$ with $n_i \to \infty$ such that $|A_i| \geq \delta n_i$ for all $i$ yet none of the $A_i$ contains an arithmetic progression of length $k$.

Write $f_i$ for the function $Z \to \{0, 1\}$ defined to take the value 1 on $A_i$ and 0 elsewhere. Define a sequence $i \mapsto \mu_i$ of Borel measures on $\{0, 1\}^Z$ by

$$\mu_i = \frac{\delta f_1 + \cdots + \delta f_i}{i}$$

where $\delta_x$ is the point mass at $x \in \{0, 1\}^Z$, and let $\mu$ be a subsequential limit of this sequence in the weak* topology. Put $B = \{x \in \{0, 1\}^Z : x(0) = 1\}$. The condition $|A_i| \geq \delta n_i$ implies that $\mu(B) > 0$. Since the map $T : X \to X$ given by $(Tx)(n) = x(n + 1)$ is measurable and measure-preserving, we have (1.2.3) by Theorem 1.2.2. The definition of $\mu$ implies that some $A_i$ contains an arithmetic progression with $k$ steps, yielding the desired contradiction. \qed

In fact, the preceding argument essentially shows that, for any set $A \subset \mathbb{Z}$, we can find an invertible, measurable, measure-preserving transformation $T$ on a probability space $(X, \mathcal{B}, \mu)$ and a set $B \in \mathcal{B}$ with

$$\mu(B) = \limsup_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{N}$$

such that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |(A - n_1) \cap \cdots \cap (A - n_k) \cap \{1, \ldots, N\}| \geq \mu(T^{-n_1}B \cap \cdots \cap T^{-n_k}B)$$

(1.2.5)

for all $n_1, \ldots, n_k \in \mathbb{Z}$. Relations such as (1.2.5) between combinatorics and correlations have become known as Furstenberg correspondence principles. In Chapter 3 we will develop a
two-sided Furstenberg correspondence principle for amenable groups based on the work in [BF09].

Fix an invertible, measurable, measure-preserving transformation $T$ on a probability space $(X, \mathcal{B}, \mu)$ and $B \in \mathcal{B}$ with $\mu(B) > 0$. To prove Theorem 1.2.2 it suffices to show that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(B \cap T^{-n}B \cap \cdots \cap T^{-kn}B) > 0$$

(1.2.6)

and to do so one proceeds by identifying a sub-$\sigma$-algebra $\mathcal{C}_k$ of $\mathcal{B}$ that characterizes the behavior of the above limit. In the next section we outline how such a sub-$\sigma$-algebra can be constructed.

1.3 Characteristic factors

Fix $k \in \mathbb{N}$ and fix an invertible, measurable, measure-preserving transformation $T$ on a probability space $(X, \mathcal{B}, \mu)$ and $B \in \mathcal{B}$ with $\mu(B) > 0$. A sub-$\sigma$-algebra $\mathcal{C}_k$ is a characteristic factor for the average (1.2.6) if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int f_0 \cdot T^n f_1 \cdots T^{kn} f_k \, d\mu = 0$$

for all $f_0, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ such that $\mathbb{E}(f_i | \mathcal{C}_k) = 0$ for some $0 \leq i \leq k$. Here $\mathbb{E}(f_i | \mathcal{C}_k)$ is the orthogonal projection of $f_i$ on the closed subspace $L^2(X, \mathcal{C}_k, \mu)$ of $L^2(X, \mathcal{B}, \mu)$. Writing

$$1_B = 1_B - \mathbb{E}(1_B | \mathcal{C}_k) + \mathbb{E}(1_B | \mathcal{C}_k)$$

one sees that it suffices to prove (1.2.6) under the assumption that $B \in \mathcal{C}_k$. The goal is to determine a sub-$\sigma$-algebra $\mathcal{C}_k$ with properties that make it both characteristic and useful for proving positivity.

The sub-$\sigma$-algebras $\mathcal{C}_k$ are determined inductively. It is easy to see that one can take $\mathcal{C}_1$ to consist of those sets $C \in \mathcal{B}$ that are invariant in the sense that $\mu(C \triangle (T^n)^{-1}C) = 0$ for all $n \in \mathbb{Z}$. Indeed, the mean ergodic theorem (see Theorem 2.3.6 below) implies that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} T^n 1_B = \mathbb{E}(1_B | \mathcal{C}_1)$$

in $L^2(X, \mathcal{B}, \mu)$.

The induction step relies on the van der Corput trick (see Theorem 2.3.7 below), an application of which implies that

$$\limsup_{N \to \infty} \left\| \frac{1}{N} \sum_{n=1}^{N} T^n f_1 \cdot T^{2n} f_2 \cdots T^{kn} f_k \right\|$$

$$\leq \min \left\{ \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left\| \mathbb{E}(f_i \cdot T^{in} f_i | \mathcal{C}_{k-1}) \right\|^2 : 1 \leq i \leq k \right\}$$
for any \( f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu) \) and any \( k \geq 2 \). Fixing \( 1 \leq i \leq k \), the averages above are controlled by realizing them as averages for the transformation \( T^i \times T^i \) on the probability space

\[(X \times X, \mathcal{B} \otimes \mathcal{B}, \nu)\]

where \( \nu \) is the \( T^i \times T^i \) invariant probability measure defined by

\[
\int \phi \otimes \psi \, d\nu = \int \mathbb{E}(\phi \mid \mathcal{C}_{k-1}) \cdot \mathbb{E}(\psi \mid \mathcal{C}_{k-1}) \, d\mu
\]

for all \( \phi, \psi \in L^\infty(X, \mathcal{B}, \mu) \). Indeed

\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \| \mathbb{E}(f_i \cdot T^{in} f_i \mid \mathcal{G}_{k-1}) \|^2
\]

\[
= \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int (f_i \otimes f_i) \cdot (T^i \times T^i)^n (f_i \otimes f_i) \, d\nu
\]

\[
= \int (f_i \otimes f_i) \cdot \mathbb{E}(f_i \otimes f_i \mid \mathcal{I}) \, d\nu
\]

where \( \mathcal{I} \) is the \( \sigma \)-algebra of \( T^i \times T^i \) invariant sets. The next step is to relate \( T^i \times T^i \) invariant functions in \( L^2(X^2, \mathcal{B}^2, \nu) \) to a space of functions in \( L^2(X, \mathcal{B}, \mu) \). One does this by considering the sub-\( \sigma \)-algebras \( \mathcal{D} \) of \( \mathcal{B} \) such that

\[
\mathbb{E}(\phi \otimes \phi \mid \mathcal{I}) = \mathbb{E}(\mathbb{E}(\phi \mid \mathcal{D}) \otimes \mathbb{E}(\phi \mid \mathcal{D})) \mid \mathcal{I} \quad \text{(1.3.1)}
\]

for all \( \phi \in L^\infty(X, \mathcal{B}, \mu) \). As we will see in Chapter 4, there is a sub-\( \sigma \)-algebra \( \mathcal{G}_k \) satisfying (1.3.1) that, relative to \( \mathcal{G}_{k-1} \) has a rigid structure that can be precisely described (see Definition 4.2.1).
In this chapter we recall facts from measure theory and group theory that we will use later in the thesis, and give some background on ultrafilters and notions from ergodic Ramsey theory. Most of the material is quite standard. For more details see [EW11], [Gla03], [HS12], [McC99] and the in-line references given below.

2.1 Probability spaces and group actions

By a probability space we mean a tuple \((X, \mathcal{B}, \mu)\) where \(X\) is any set, \(\mathcal{B}\) is a \(\sigma\)-algebra of subsets of \(X\), and \(\mu\) is a measure defined on \(\mathcal{B}\) such that \(\mu(X) = 1\). Note that the last requirement precludes \(X\) from being the empty set.

Given a sub-\(\sigma\)-algebra \(\mathcal{C}\) of \(\mathcal{B}\) and \(f\) in \(L^2(X, \mathcal{B}, \mu)\) the conditional expectation of \(f\) on \(\mathcal{C}\), denoted \(E(f|\mathcal{C})\), is the orthogonal projection of \(f\) onto the closed subspace \(L^2(X, \mathcal{C}, \mu)\) of \(L^2(X, \mathcal{B}, \mu)\).

Let \(G\) be a Hausdorff topological group. By a measure-preserving action of \(G\) on a probability space \((X, \mathcal{B}, \mu)\) we mean a measurable map \(G \times X \to X\) such each of the induced maps \(T^g : X \to X\) is measure-preserving and satisfies \(T^g(T^h x) = T^{gh} x\) for all \(g, h \in G\). Two measure-preserving actions \(T\) and \(S\) of \(G\) are said to commute if \(T^g S^h = S^h T^g\) for all \(g, h\) in \(G\), and if they do \((g, x) \mapsto T^g S^h x\) is also a measure-preserving action of \(G\) on \((X, \mathcal{B}, \mu)\).

We will only consider continuous, measure-preserving actions of locally compact, Hausdorff topological groups on compact, Hausdorff spaces. Precisely, by a \(G\) system we mean a tuple \((X, \mathcal{B}, \mu, T)\) consisting of compact, Hausdorff space \(X\) equipped with a Baire probability measure \(\mu\) and a continuous action \(T : G \times X \to X\) that is measure-preserving. We often write \(X\) for \((X, \mathcal{B}, \mu, T)\) and \(L^p(X)\) for the corresponding real Banach space \(L^p(X, \mathcal{B}, \mu)\). Given a system \(X\) each \(T^g\) induces a unitary operator on \(L^2(X)\) given by \((T^g f)(x) = f(T^g x)\). It is immediate that \(T^g(T^h f) = T^{gh} f\) for all \(g, h\) in \(G\) and all \(f \in L^2(X)\). A system \(X = (X, \mathcal{B}, \mu, T)\) is ergodic if any set \(B \in \mathcal{B}\) satisfying \(\mu(B \Delta (T^g)^{-1} B) = 0\) for all \(g \in G\)
has either zero measure or full measure. This is equivalent to the absence of non-constant, invariant functions in $L^2(X)$. When $G$ is countable, ergodicity is equivalent to the statement that any $B \in \mathcal{B}$ satisfying $B = (T^g)^{-1}B$ for all $g \in G$ has either zero measure or full measure.

Let $X = (X, \mathcal{B}, \mu, T)$ be a system. A sub-$\sigma$-algebra $\mathcal{C}$ of $\mathcal{B}$ is $T$ invariant if, for every $g \in G$ and every $C \in \mathcal{C}$ one can find $D \in \mathcal{C}$ such that $\mu(D \triangle (T^g)^{-1}C) = 0$. When this is the case each $T^g$ commutes with the conditional expectation on $\mathcal{C}$, which is to say that $E(T^g f | \mathcal{C}) = T^g E(f | \mathcal{C})$ for all $f \in L^2(X)$ and all $g \in G$.

Given $G$ systems $X = (X, \mathcal{B}, \mu, T)$ and $Y = (Y, \mathcal{D}, \nu, S)$ we say that $X$ is an extension of $Y$, or that $Y$ is a factor of $X$, if there is a continuous, measure-preserving map $\pi : X \to Y$ satisfying $\pi \circ T^g = S^g \circ \pi$ for all $g \in G$. Such a map is called a factor map and we sometimes write $\pi : X \to Y$ to denote this.

Given a continuous action $T$ of a Hausdorff topological group $G$ on a compact, Hausdorff space $X$, we can consider the compact, convex set $\mathcal{M}_T(X)$ of Baire probability measures on $X$ that are $T$ invariant. We will need the fact that the extreme points of $\mathcal{M}_T(X)$ are the ergodic measures.

**Lemma 2.1.1.** Let $T$ be a continuous action of a Hausdorff topological group $G$ on a compact, Hausdorff space $X$. Then the extreme points of $\mathcal{M}_T(X)$ are precisely the ergodic measures.

**Proof.** Let $\mu$ be an extreme point. If it is not ergodic then there is a set $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and $\mu(B \triangle (T^g)^{-1}B) = 0$ for all $g \in G$. But then $\mu$ is a convex combination of the $T$ invariant probability measures $\mu | B$ and $\mu | (X \setminus B)$.

Suppose now that $\mu$ is not an extreme point. Then $\mu = t \eta + (1-t) \lambda$ for some $0 < t < 1$ and some $\eta, \lambda \in \mathcal{M}_T(X)$. This implies $\eta \ll \mu$, so $\eta = f \mu$ for some $f \in L^1(X, \mu)$. Invariance implies that $f$ is $T$ invariant, and therefore constant. Thus $\mu = \eta$, a contradiction.

The following result lets us write any measure in $\mathcal{M}_T(X)$ as a convex combination of the extreme points in a certain weak sense.

**Theorem 2.1.2** (Choquet-Bishop-de Leeuw [Phe01, Page 17]). Let $E$ be a compact, convex subset of a locally convex space. For any $\mu \in E$ there is a Baire probability measure $\eta$ on $E$ that represents $\mu$ in the sense that $\phi(\mu) = \int \phi \, d\eta$.
for every affine, continuous function $\phi : E \to \mathbb{R}$, and such that $\eta$ vanishes on every Baire subset of $E$ disjoint from the extreme points of $E$.

At the above level of generality it is not possible to strengthen the sense in which $\eta$ is supported on the extreme points of $E$. This is because the set of extreme points of a compact, convex subset of a locally convex space need not be Baire measurable.

### 2.2 Uniformly continuous functions

In the next section we will define amenability and describe some basic results pertaining to measure preserving actions of amenable groups. Before doing so we collect in this section relevant facts about uniformly continuous functions, on which the definition of amenability lies, and which will play an important role in Chapter 3. Given a group $G$ and any $g \in G$ write $L_g$ and $R_g$ for the maps $G \to G$ defined by $L_g(x) = gx$ and $R_g(x) = xg$ respectively. Given a function $f$ defined on $G$, put $(L_gf)(x) = f(gx)$ and $(R_gf)(x) = f(xg)$ for all $g, x \in G$.

**Definition 2.2.1.** Let $G$ be a Hausdorff topological group. A function $f : G \to \mathbb{C}$ is left uniformly continuous if, for every $\varepsilon > 0$, one can find an open neighbourhood $V$ of the identity in $G$ such that $|f(vx) - f(x)| < \varepsilon$ for all $v \in V$ and all $x \in G$, and right uniformly continuous if, for every $\varepsilon > 0$, one can find an open neighbourhood $V$ of the identity in $G$ such that $|f(xv) - f(x)| < \varepsilon$ for all $v \in V$ and all $x \in G$. A function $f : G \to \mathbb{C}$ is uniformly continuous if it is both left uniformly continuous and right uniformly continuous.

It is straightforward to check that the set $\text{LUC}(G)$ of bounded, left uniformly continuous functions on $G$ forms a $C^*$ algebra under pointwise operations and the topology of uniform convergence. Similarly, bounded right uniformly continuous function form a $C^*$ algebra $\text{RUC}(G)$. Write $\text{UC}(G)$ for the $C^*$ algebra $\text{RUC}(G) \cap \text{LUC}(G)$.

**Lemma 2.2.2.** Let $G$ be a Hausdorff topological group. If $f \in \text{LUC}(G)$ then $L_gf \in \text{LUC}(G)$ and $R_gf \in \text{LUC}(G)$ for all $g \in G$.

**Proof.** Fix $\varepsilon > 0$ and let $V$ be an open neighbourhood of the identity in $G$ such that $|f(vx) - f(x)| < \varepsilon$ for all $v \in V$ and all $x \in G$. We have

$$|(L_gf)(vx) - (L_gf)(x)| = |f(gvg^{-1}gx) - (gx)|$$

and

$$|(R_gf)(vx) - (R_gf)(x)| = |f(vxg) - (xg)|$$

so $L_gf$ is left uniformly continuous via the neighbourhoods $g^{-1}Vg$ and $R_gf$ is left uniformly continuous via the neighbourhoods $V$. \qed
One can similarly show that $\text{RUC}(G)$, and therefore $\text{UC}(G)$, is invariant under $L$ and $R$ as well.

The preceding lemma proves that $G$ permutes $\text{LUC}(G)$ via left and right multiplication. Next, we prove that these permutations give rise to continuous actions of $G$.

**Lemma 2.2.3.** Let $G$ be a Hausdorff topological group. Then the map

$$G \times G \times \text{LUC}(G) \to \text{LUC}(G)$$

defined by $(g_1, g_2, \phi) \mapsto L_{g_1}R_{g_2}\phi$ is continuous.

**Proof.** Fix $f \in \text{LUC}(G)$ and $\varepsilon > 0$. We need to show that the set

$$\{(g_1, g_2, \phi) \in G \times G \times \text{LUC}(G) : \|L_{g_1}R_{g_2}\phi - f\| < \varepsilon\}$$

is open, so fix a point $(g_1, g_2, \phi)$ therein. Put $\delta = \varepsilon - \|L_{g_1}R_{g_2}\phi - f\|$. There is an open neighbourhood $U$ of $\phi$ in $\text{LUC}(G)$ such that $\|\psi - \phi\| < \delta/3$ for all $\psi \in U$, there is an open neighbourhood $V_1$ of $g_1$ in $G$ such that $\|L_{h_1}\phi - L_{g_1}\phi\| < \delta/3$ for all $h_1 \in V_1$, and there is an open neighbourhood $V_2$ of $g_2$ in $G$ such that $\|R_{h_2}\phi - R_{g_2}\phi\| < \delta/3$ for all $h_2 \in V_2$. Thus if $(h_1, h_2, \psi) \in V_1 \times V_2 \times U$ then

$$\|L_{h_1}R_{h_2}\psi - f\| \leq \|L_{h_1}R_{h_2}\psi - L_{h_1}L_{h_2}\phi\| + \|L_{h_1}R_{h_2}\phi - L_{g_1}R_{g_2}\phi\| + \|L_{g_1}R_{g_2}\phi - f\| < \frac{\delta}{3} + \|L_{h_1}R_{h_2}\phi - L_{h_1}R_{g_2}\phi\| + \|L_{h_1}R_{g_2}\phi - L_{g_1}R_{g_2}\phi\| + \|L_{g_1}R_{g_2}\phi - f\|$$

as desired. \qed

### 2.3 Amenable groups

Much of ergodic theory and density combinatorics functions at the level of amenable groups, which is the point of view taken in this thesis. We take the following as our definition of amenability.

**Definition 2.3.1.** A locally compact, Hausdorff topological group $G$ is **amenable** if there is a left invariant mean on $\text{LUC}(G)$, by which one means a non-negative, linear functional $M : \text{LUC}(G) \to \mathbb{C}$ with $M(1) = 1 = \|M\| \|f\|$ such that $M(L_gf) = M(f)$ for all $g \in G$ and all $f \in \text{LUC}(G)$.

The definition does not require $G$ to be locally compact, but we take the liberty of requiring this as we only consider locally compact groups.

**Lemma 2.3.2.** If $G$ is amenable there is a mean on $\text{UC}(G)$ that is left and right invariant.
Proof. Since $G$ is amenable there is a left-invariant mean $M_L$ on $LUC(G)$. One can easily check, using the map $x \mapsto x^{-1}$ that there is a right-invariant mean $M_R$ on $RUC(G)$. For each $f \in UC(G)$ define a function $\phi_f : G \to \mathbb{C}$ by $\phi_f(g) = M_L(R_g f)$.

We claim that, for every $f \in UC(G)$, the function $\phi_f$ belongs to $RUC(G)$. It is certainly bounded. To see that it is right uniformly continuous, note that

$$|\phi_f(xv) - \phi_f(x)| = |M_L(R_{xv}f - R_x f)| \leq \|M_L\| \cdot \|R_{xv}f - R_x f\|
$$

because $M_L$ is bounded and $R_x$ is isometric for the uniform norm. Since $f$ is right uniformly continuous, the right-hand side above can be made arbitrarily small provided $v$ is close enough to 1, so $\phi_f$ is indeed right uniformly continuous.

Define a mean $M$ on $UC(G)$ by $M(f) = M_R(\phi_f)$ for all $f \in UC(G)$. We claim that $M$ is left and right invariant. Since $L$ and $R$ commute we have $\phi_{L_g f} = \phi_f$ and

$$M(L_g f) = M_R(\phi_{L_g f}) = M_R(\phi_f) = M(f)
$$

for all $g \in G$, so $M$ is left invariant. Also

$$M(R_g f) = M_R(\phi_{R_g f}) = M_R(R_g \phi_f) = M_R(\phi_f) = M(f)
$$

for all $g \in G$ so $M$ also right invariant as desired. \qed

There are many equivalent definitions of amenability. We recall some of these equivalent definitions in the following theorem. Write $L_{1, +}^1(G, m) = \{ f \in L^1(G, m) : f \geq 0, \|f\|_1 = 1\}$.

**Theorem 2.3.3.** For any locally compact, Hausdorff topological group $G$ with a left Haar measure $m$, the following properties are equivalent.

1. There is a left-invariant mean on $LUC(G)$.

2. There is a left-invariant mean on $L^\infty(G, m)$.

3. Every continuous, affine action of $G$ on a non-empty compact, convex subset of a locally convex topological vector space has a fixed point.

4. There exists a net $\iota \mapsto \Phi_\iota$ of compact, positive-measure subsets of $G$ such that

$$\frac{m(\Phi_\iota \triangle g \Phi_\iota)}{m(\Phi_\iota)} \to 0 \quad (2.3.4)
$$

uniformly on compact subsets of $G$ as $\iota \to \infty$.

5. There exists a net $\iota \mapsto \phi_\iota$ in $L_{1, +}^1(G, m)$ such that

$$\|\phi_\iota - L_g \phi_\iota\|_1 \to 0 \quad (2.3.5)
$$

uniformly on compact subset of $G$ as $\iota \to \infty$. 11
6. Every continuous action of $G$ on a non-empty compact, Hausdorff space has an invariant Baire probability measure.

7. Every continuous, minimal, strongly proximal action of $G$ on a compact, Hausdorff space is trivial.

Proof. The equivalence of the first five properties can be found in [BHV08, Appendix G], and that the last two properties are equivalent is the content of [Fur03, Proposition 7]. One can easily show that the third property implies the sixth, and, as described in Chapter 3, that the sixth property implies the first follows from the Gelfand-Naimark theorem applied to the C$^*$ algebra $\text{LUC}(G)$.

Any net $\iota \mapsto \Phi_\iota$ of compact, positive measure subsets of $G$ satisfying (2.3.4) uniformly on compact subsets of $G$ is called a left Følner net in $G$, and any net $\iota \mapsto \phi_\iota$ in $L^1(G, m)$ satisfying (2.3.4) is called a left Reiter net in $G$.

As a consequence of [Pat88, Theorem 4.16], when $G$ is amenable and $\sigma$-compact, one can always find a left Følner sequence in $G$, by which one means a left Følner net indexed by the directed set $\mathbb{N}$.

We now turn to basic facts about measure-preserving actions of amenable groups. Although one could attempt to describe the results below using any of the equivalent definitions of amenability listed in Definition 2.3.1, we will follow the traditional approach, doing so using left Følner nets. To facilitate this recall some some notation.

Given a left Følner net $\Phi$ in an amenable group $G$ and a measurable map $u : G \to \mathbb{C}$ write

$$\lim_{g \to \Phi} u(g) = v$$

for any $v \in \mathbb{C}$ if

$$\frac{1}{m(\Phi_\iota)} \int_{\Phi_\iota} u(g) \, dm(g) \to v$$

as $\iota \to \infty$. Similarly, write

$$\frac{1}{m(\Phi_\iota)} \int_{\Phi_\iota} |u(g) - v| \, dm(g) \to 0$$

as $\iota \to \infty$. When $G = \mathbb{Z}$ and $\Phi$ is the net $\iota \mapsto \{1, \ldots, \iota\}$ it is immediate that the above notions of convergence are just Cesàro convergence and strong Cesàro convergence. Given an amenable group $G$ with a left Haar measure $m$, a left Følner net $\Phi$ in $G$, and a $G$ system $X$, we can consider for any $f, \xi \in L^2(X)$ and any $\iota$ the integral

$$\int \int 1_{\Phi_\iota}(g) \cdot f(T(g, x)) \cdot \xi(x) \, d(\mu \otimes m)(x, g) = \int \int _{\Phi} \xi \cdot T^g f \, d\mu \, dm(g)$$
since the map $T : G \times X \to X$ is measurable. Given $h \in L^2(X)$, write
\[
\lim_{g \to \Phi} T^g f = h
\]
if one has
\[
\frac{1}{m(\Phi)} \int \int \xi \cdot T^g f \, d\mu \, dm(g) \to \int \xi \cdot h \, d\mu
\]
for all $\xi \in L^2(X)$.

The first important result in the ergodic theory of amenable groups is the mean ergodic theorem, which was originally proved for $\mathbb{R}$ actions by von Neumann [Neu32].

**Theorem 2.3.6** (Mean ergodic theorem). Let $G$ be an amenable group with a left Haar measure $m$ and a left Følner net $\Phi$. For any $G$ system $X$ one has
\[
\lim_{g \to \Phi} T^g f = \mathbb{E}(f \mid \mathcal{I} X)
\]
for every $f \in L^2(X)$.

**Proof.** If $f = \phi - T^h \phi$ for some $h \in G$ and some $\phi \in L^2(X)$ then
\[
\frac{1}{m(\Phi)} \int \int \xi \cdot T^g f \, d\mu \, dm(g)
\]
for all $\xi \in L^2(X)$ so
\[
\frac{1}{m(\Phi)} \int \int \xi \cdot T^g f \, d\mu \, dm(g) \leq \frac{m(\Phi \triangle \Phi)}{m(\Phi)} \|f\|_2 \|\xi\|_2 \to 0
\]
because $\Phi$ is a left Følner net. The orthogonal complement in $L^2(X)$ of the span of such $f$ is the space of invariant functions.

A crucial tool in the study of multiple recurrence for actions of amenable groups is the van der Corput trick, which began life as a tool in uniform distribution: it was used by Weyl to prove his result [Wey16] about equidistribution of certain polynomials mod 1. The version we present here is based on [BMZ97, Lemma 4.2].

**Theorem 2.3.7** (van der Corput trick). Let $G$ be an amenable group. For any Hilbert space $\mathcal{H}$ and any bounded, measurable function $u : G \to \mathcal{H}$ we have
\[
\lim_{\kappa} \sup_{\kappa} \left\| \frac{1}{m(\Phi)} \int u(g) \, dm(g) \right\|
\]
\[
\leq \lim_{\kappa} \sup_{\kappa} \frac{1}{m(\Phi)^2} \int \int \left( \lim_{\kappa} \sup_{\kappa} \frac{1}{m(\Phi)} \int (u(gh), u(gl)) \, dm(g) \right) \, dm(h) \, dm(l)
\]

13
for any left Følner sequence in $G$.

**Proof.** By using the Følner property one can show that it suffices to consider

$$\left\| \frac{1}{m(\Phi_N)} \int \frac{1}{m(\Phi_i)} \int u(gh) \, dm(g) \, dm(h) \right\|$$

which is bounded by

$$\frac{1}{m(\Phi_i)} \int \frac{1}{m(\Phi_N)^2} \int \int \langle u(gh), u(gl) \rangle \, dm(h) \, dm(l) \, dm(g)$$

as a consequence of Fubini’s theorem and the Cauchy-Schwarz inequality. \hfill $\square$

We will make use of the following version of the van der Corput trick in Chapter 5.

**Proposition 2.3.8** ([Lei05a, Lemma 4(i)]). Let $G$ be a discrete, countable, abelian group and let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$. Fix a bounded map $u : G \to \mathcal{H}$. Then

$$\limsup_{N \to \infty} \left\| \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} u(g) \right\|^2 \leq \frac{1}{|\Phi_H|^2} \sum_{h,l \in \Phi_H} \limsup_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \langle u(g + h), u(g + l) \rangle$$

for any Følner sequence $\Phi$ in $G$ and any $H$ in $\mathbb{N}$.

### 2.4 Disintegrations

Let $(X, \mathcal{B})$ be a measurable space. A sub-$\sigma$-algebra $\mathcal{C}$ of $\mathcal{B}$ is **countably generated** if there is a sequence $n \mapsto C_n$ of sets in $\mathcal{C}$ such that $\sigma(\bigcup \{C_i : i \in \mathbb{N}\}) = \mathcal{C}$. Given such a $\sigma$-algebra, to any point $x \in X$ we can associate the **atom**

$$[x]_{\mathcal{C}} = \bigcap \{C_i : x \in C_i\} \cap \bigcap \{X \setminus C_i : x \notin C_i\}$$

and declare $x \sim_{\mathcal{C}} y$ if and only if $[x]_{\mathcal{C}} = [y]_{\mathcal{C}}$. This relation is in fact an equivalence relation, and thus $X$ is partitioned by the atoms of $\mathcal{C}$. Moreover, $[x]_{\mathcal{C}}$ is the smallest set containing $x$ that belongs to $\mathcal{C}$ in the sense that

$$[x]_{\mathcal{C}} = \bigcap \{C \in \mathcal{C} : x \in C\}$$

for every $x \in X$, because any point that can be separated from $x$ by a set in $\mathcal{C}$ can be separated from $x$ by one of the $C_i$.

**Example 2.4.1.**

1. The $\sigma$-algebra $\mathcal{B}$ of Borel sets on $[0, 1]$ is countably generated and $[x]_{\mathcal{B}} = \{x\}$ for all $x \in [0, 1]$. 

14
2. Write $\pi$ for the projection $[0,1]^2 \to [0,1]$ on the first coordinate. The $\sigma$-algebra $\mathcal{B}_1 = \pi^{-1}\mathcal{B}$ on $[0,1] \times [0,1]$ is countably generated and its atoms are the fibres of $\pi$.

3. (See [EW11, Exercise 6.1.2].) Fix $\alpha, \beta \in \mathbb{R}$ with $\{1, \alpha, \beta\}$ linearly independent over $\mathbb{Q}$, and consider the action $T^n(x, y) \mapsto (x + n\alpha, y + n\beta)$ of $\mathbb{Z}$ on the torus $T^2 = \mathbb{R}^2 / \mathbb{Z}^2$, which is ergodic. The $\sigma$-algebra $\mathcal{D}$ of Borel sets $B \subset T^2$ such that $T^{-1}B = B$ is not countably generated. Indeed, suppose $\mathcal{D}$ is generated by a sequence $D_1, D_2, \ldots$ of Borel sets. Since $T$ is ergodic with respect to Lebesgue measure, each of the sets $D_i$ has either zero measure or full measure. Consider the set
\[
\bigcap \{D_i : \mu(D_i) = 1\} \cap \bigcap \{T^2 \setminus D_i : \mu(D_i) = 0\}
\]
which has full measure, and is the atom associated to any point it contains. Fix such a point $x$. Since $[x]_{\mathcal{D}}$ is the smallest set in $\mathcal{D}$ that contains $x$, and the orbit $\{T^n x : n \in \mathbb{Z}\}$ is certainly in $\mathcal{D}$, it must be that $[x]_{\mathcal{D}}$ is the orbit of $x$. But the orbit has measure zero since it is countable, giving the desired contradiction.

As the third example above shows, not all sub-$\sigma$-algebras of countably generated $\sigma$-algebras are countably generated. However, we do have the following result. Given a probability space $(X, \mathcal{B}, \mu)$ and sub-$\sigma$-algebras $\mathcal{C}$ and $\mathcal{D}$ of $\mathcal{B}$, write $\mathcal{C} \sim \mathcal{D}$ if for any $C \in \mathcal{C}$ one can find $D \in \mathcal{D}$ such that $\mu(C \triangle D) = 0$, and vice versa.

**Proposition 2.4.2** ([EW11, Lemma 5.17]). Let $(X, \mathcal{B}, \mu)$ be a compact, metric probability space. To every sub-$\sigma$-algebra $\mathcal{C}$ of $\mathcal{B}$ we can associate a countably generated sub-$\sigma$-algebra $\mathcal{D} \subset \mathcal{C}$ such that $\mathcal{C} \sim \mathcal{D}$.

Given a compact metric space $X$, a Borel probability measure $\mu$ on $X$, and a sub-$\sigma$-algebra $\mathcal{C}$ of the Borel sets, one can find a family $x \mapsto \mu_x$ of measures indexed by $X$, a convex combination of which equals $\mu$. Moreover, when $\mathcal{C}$ is countably generated, the measure $\mu_x$ lives on the atom $[x]_{\mathcal{C}}$. The following result makes this precise.

**Theorem 2.4.3** ([EW11, Theorem 5.14]). Let $X$ be a compact metric space and let $\mathcal{C}$ be a sub-$\sigma$-algebra of the Borel sets. There is a set $C \in \mathcal{C}$ with $\mu(C) = 1$ and a Borel probability measure $\mu_x$ on $X$ for every $x \in C$ with the following properties.

1. For any Borel measurable function $f : X \to \mathbb{R}$ that is $\mu$ integrable the map
\[
x \mapsto \int f \, d\mu_x
\]
is defined on a full-measure set in $\mathcal{C}$ and measurable with respect to $\mathcal{C}$. Moreover, this property uniquely determines the family $\mu_x$ for almost every $x \in X$. 

15
2. For any Borel measurable function \( f \) that is integrable

\[
\mathbb{E}(f|\mathcal{C})(x) = \int f \, d\mu_x
\]

almost everywhere.

3. If \( \mathcal{C} \) is countably generated then \( \mu_x([x]_{\mathcal{C}}) = 1 \) for every \( x \in C \) and \( \mu_x = \mu_y \) whenever \( [x]_{\mathcal{C}} = [y]_{\mathcal{C}} \).

Care is taken to speak of a function \( f : X \to \mathbb{R} \) rather than an equivalence class of functions in \( L^1(X, \mathscr{B}, \mu) \) because the measures \( \mu_x \) may well be singular with respect to \( \mu \). Although each integrable function \( f : X \to \mathbb{R} \) defines an equivalence class in the space \( L^1(X, \mathscr{B}, \mu_x) \) for almost every \( x \), changing \( f \) on a set of \( \mu \) measure zero may not preserve all of these classes. Write \( \langle \cdot, \cdot \rangle_x \) for the inner product on \( L^2(X, \mathscr{B}, \mu_x) \) and \( \| \cdot \|_x \) for the corresponding norm.

2.5 Joinings

We now recall some basic facts about joinings. Let \((X_i, \mathscr{B}_i, \mu_i), 1 \leq i \leq k\) be probability spaces and let \( \pi_i \) be the projection from \( X_1 \times \cdots \times X_k \) to \( X_i \). We say that a probability measure \( \nu \) on \((X_1 \times \cdots \times X_k, \mathscr{B}_1 \otimes \cdots \otimes \mathscr{B}_k)\) is a standard measure if 

\[
\nu(\pi_i^{-1}(B_i)) = \mu_i(B_i)
\]

for all \( B_i \in \mathscr{B}_i \) and all \( 1 \leq i \leq k \). A sequence \( \nu_n \) of standard measures is said to converge to a standard measure \( \nu \) if

\[
\nu_n(B_1 \times \cdots \times B_k) \to \nu(B_1 \times \cdots \times B_k)
\]

for all \( B_i \) in \( \mathscr{B}_i \). A joining of systems \( X_1, \ldots, X_k \) is any system \( X = (X, \mathscr{B}, \nu, T) \) where \( X = X_1 \times \cdots \times X_k, \mathscr{B} = \mathscr{B}_1 \otimes \cdots \otimes \mathscr{B}_k, T^q = T^q_1 \times \cdots \times T^q_k \) and \( \nu \) is a standard measure that is \( T \)-invariant. Given a factor \( Y_i \) of \( X_i \) for each \( i \), we can consider the system \( Y = (Y, \mathscr{D}, \eta, T) \) made from \( X \) by projecting \( \nu \) onto the product \( (Y, \mathscr{D}) \) of the underlying measurable spaces \( (Y_i, \mathscr{D}_i) \). Call a joining \( X \) of the \( X_i \) a conditional product joining relative to the factors \( Y_i \) if

\[
\int f_1 \otimes \cdots \otimes f_k \, d\nu = \int \mathbb{E}(f_1|Y_1) \otimes \cdots \otimes \mathbb{E}(f_k|Y_k) \, d\eta
\]

for all \( f_i \) in \( L^\infty(X_i) \). Here \( f_1 \otimes \cdots \otimes f_k \) denotes the function mapping \((x_1, \ldots, x_k)\) to \( f_1(x_1) \cdots f_k(x_k) \). We can re-write (2.5.1) as

\[
\nu = \int \mu_{i,y_1} \otimes \cdots \otimes \mu_{i,y_k} \, d\eta(y_1, \ldots, y_k)
\]

if \( \mu_{i,y_i} \) is the almost-surely defined disintegration of \( \mu_i \) over \( Y_i \).
Fix an amenable group $G$ with a left Haar measure $m$. Let $X = (X, \mathcal{B}, \mu, T)$ be a $G^k$ system and write $T_1, \ldots, T_k$ for the coordinate $G$ actions. Fixing a left Følner net $\Phi$ in $G$, one can define the measure $\nu_k$ on $(X^{k+1}, \mathcal{B}^{k+1})$ by

$$\int f_1 \otimes \cdots \otimes f_{k+1} \, d\nu_k = \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f_{k+1} \cdot \prod_{i=1}^k T_i^g \cdots T_1^g f_i \, d\mu \, dm(g) \quad (2.5.3)$$

for any $f_1, \ldots, f_{k+1}$ in $L^\infty(X, \mathcal{B}, \mu)$. The existence of the limit is justified by [ZK13, Theorem 1.3]. Using the fact that $\Phi$ is a Følner net, one can show that $\nu_k$ is invariant. Thus the measure $\nu_k$ yields a joining of the systems

$$(X, \mathcal{B}, \mu, T_k \cdots T_1), \ldots, (X, \mathcal{B}, \mu, I)$$

called the Furstenberg joining of the actions $T_1, \ldots, T_k$.

Given two systems $X_1 = (X_1, \mathcal{B}_1, \mu_1, T_1)$ and $X_2 = (X_2, \mathcal{B}_2, \mu_2, T_2)$ having a common factor $Y = (Y, \mathcal{D}, \mu, T)$ via factor maps $\pi_1$ and $\pi_2$ respectively, one can form their relatively independent joining over $Y$, which is the system

$$X_1 \times_Y X_2 = (X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \nu, T_1 \times T_2)$$

where $\nu$ is the measure defined by

$$\int f_1 \otimes f_2 \, d\nu = \int \mathbb{E}(f_1 \mid \pi_1^{-1} \mathcal{D}) \cdot \mathbb{E}(f_2 \mid \pi_2^{-1} \mathcal{D}) \, d\mu$$

for all $f_1$ in $L^\infty(X_1)$ and all $f_2$ in $L^\infty(X_2)$. The set

$$\{(x_1, x_2) : \pi_1 x_1 = \pi_2 x_2\}$$

has full measure with respect to $\nu$ so $Y$ is a factor of $X_1 \times_Y X_2$ in an unambiguous way.

One can similarly form the relatively independent self-joining $X \times_{\mathcal{D}} X$ of a system $X$ over a $T$ invariant sub-$\sigma$-algebra $\mathcal{D}$. If $\mu_x$ is an almost-surely defined disintegration of $\mu$ over $\mathcal{D}$ then

$$\nu = \int \mu_x \otimes \mu_x \, d\mu(x) \quad (2.5.4)$$

is the disintegration of $\nu$ over $\mathcal{D}$.

### 2.6 Ultrafilters

We recall here the basic facts about ultrafilters needed to define idempotent ultrafilters and minimal idempotent ultrafilters, which we will use later as a replacement for Følner sequences on countable groups that are not amenable. The material in this section is standard, and can all be found in [HS12].
Definition 2.6.1. A filter on a set $X$ is a non-empty collection $\mathcal{F}$ of subsets of $X$ with the following properties:

1. $\emptyset \notin \mathcal{F}$;
2. if $A \in \mathcal{F}$ and $A \subset B \subset X$ then $B \in \mathcal{F}$;
3. if $A, B \in \mathcal{F}$ then $A \setminus B \in \mathcal{F}$.

Example 2.6.2. For any infinite set $X$ the collection $\mathcal{F}$ of co-finite subsets of $X$ is a filter, called the Fréchet filter. (It is necessary to assume $X$ is infinite as otherwise $\emptyset$ is cofinite.)

Example 2.6.3. Let $X$ be a set and let $\mu$ be a finitely additive measure on the power set of $X$. Then the collection $\{ A \subset X : \mu(A) = 1 \}$ is a filter on $X$.

Example 2.6.4. Let $X$ be a non-empty set and fix $x \in X$. Then $\mathcal{F} = \{ A \subset X : x \in A \}$ is a filter.

Example 2.6.5. When $X = \emptyset$ the only collections of subset of $X$ are $\emptyset$ and $\{ \emptyset \}$. Neither is a filter.

Lemma 2.6.6. A non-empty collection of non-empty subsets of $X$ has the finite intersection property if and only if there is a filter on $X$ that contains all sets in $\mathcal{D}$.

Proof. If $\mathcal{D}$ is such a collection then

$$\mathcal{F} = \{ A \subset X : A \supset D_1 \cap \cdots \cap D_n \text{ for some } D_1, \ldots, D_n \in \mathcal{D} \}$$

is a filter containing all sets in $\mathcal{D}$, and the other direction follows from the defining properties of a filter.

Any finite cover of a non-empty set $X$ can be used to make a filter larger.

Lemma 2.6.7. Let $X$ be a non-empty set and let $\mathcal{F}$ be a filter on $X$. For any finite cover $E_1 \cup \cdots \cup E_n$ of $X$ there is $1 \leq i \leq n$ such that

$$\{ A \subset X : A \supset E_i \cap F \text{ for some } F \in \mathcal{F} \}$$

(2.6.8)

is a filter on $X$.

Proof. We first claim that there is $1 \leq i \leq n$ such that $F \cap E_i \neq \emptyset$ for all $F \in \mathcal{F}$. Indeed if not then for each $1 \leq i \leq n$ we can find $F_i \in \mathcal{F}$ such that $F_i \cap E_i = \emptyset$. But $F_1 \cap \cdots \cap F_n \in \mathcal{F}$ so it is non-empty and therefore $F_1 \cap \cdots \cap F_n \in \mathcal{F} \cap E_j$ is non-empty for some $1 \leq j \leq n$, which is a contradiction.
Fix $1 \leq i \leq n$ such that $F \cap E_i \neq \emptyset$ for all $F \in \mathcal{F}$. We need to show that (2.6.8) is a filter on $X$. Certainly it satisfies conditions (1) and (2). Fix $A_1, A_2$ in (2.6.8) and choose $F_1, F_2 \in \mathcal{F}$ such that $A_1 \supset F_1 \cap E_i$ and $A_2 \supset F_2 \cap E_i$. Then $A_1 \cap A_2 \supset F_1 \cap F_2 \cap E_i$ belongs to (2.6.8) because $F_1 \cap F_2$ belongs to $\mathcal{F}$.

**Definition 2.6.9.** Let $X$ be a non-empty set. An **ultrafilter** on $X$ is a filter $p$ on $X$ with the additional property

4. if $X = A \cup B$ then $A \in p$ or $B \in p$.

**Example 2.6.10.** Let $X$ be a non-empty set and fix $x \in X$. Then

$$p_x = \{A \subset X : x \in A\} \quad (2.6.11)$$

is easily seen to be an ultrafilter on $X$. Such ultrafilters are said to be **principal**.

Ultrafilters can be pushed forward using maps.

**Lemma 2.6.12.** Let $p$ be an ultrafilter on a set $X$ and let $f : X \to Y$ be a map. Then $f[p] = \{B \subset Y : f^{-1}(B) \in p\}$ is an ultrafilter on $Y$.

**Proof.** Certainly $f[p]$ is non-empty and contains $Y$. The rest of the properties are easy to check. \hfill $\square$

**Lemma 2.6.13.** Let $p$ be an ultrafilter on a non-empty set $X$. For any finite partition $E_1 \cup \cdots \cup E_n$ of $X$ there is $1 \leq i \leq n$ such that $E_i \in p$.

**Proof.** When $i = 1$ the result is trivial. Let $1 \leq j \leq n$ be minimal with the property that $E_1 \cup \cdots \cup E_j$ belongs to $p$. If $j = 1$ we are done. Otherwise $E_1 \cup \cdots \cup E_{j-1} \notin p$ so $E_j \cup \cdots \cup E_n \in p$. It follows that

$$E_j = (E_1 \cup \cdots \cup E_j) \cap (E_j \cup \cdots \cup E_n) \in p$$

as desired. \hfill $\square$

The above lemma allows us to prove that every non-principle ultrafilter contains the Fréchet filter.

**Lemma 2.6.14.** Let $p$ be a non-principle ultrafilter on a set $X$. Then $p$ contains the Fréchet filter.

**Proof.** Let $A$ be a cofinite subset of $X$ and let $b_1, \ldots, b_n$ be an enumeration of the complement. Writing $X = A \cup \{b_1\} \cup \cdots \cup \{b_n\}$ we see that $A \in p$ as otherwise $p$ is principle. \hfill $\square$
Theorem 2.6.15. Let $X$ be a non-empty set. For any filter $\mathcal{F}$ on $X$ there is an ultrafilter $p$ on $X$ such that $\mathcal{F} \subset p$.

Proof. Let $\Omega$ be the set of filters on $X$ that contain all sets in $\mathcal{F}$. It is partially ordered by containment. It is straightforward to check that any chain in $\Omega$ has a maximal element: the union of all filters in the chain. By Zorn’s lemma there is a maximal filter $p$ in $\Omega$. We claim that $p$ is an ultrafilter on $X$. Certainly $p$ is a filter. It remains to prove that $p$ satisfies (2.6.9). Suppose $X = E_1 \cup E_2$ and that neither $E_1$ nor $E_2$ belong to $p$. Lemma 2.6.7 furnishes $1 \leq i \leq 2$ and a filter $\mathcal{G}$ containing $p$ and $E_i$. But then $\mathcal{G} \not\subset p$, contradicting maximality. \qed

Theorem 2.6.15 proves the existence of ultrafilters other than the principal ultrafilters described in Example 2.6.10. The proof is non-constructive as it relies on the axiom of choice (in the form of Zorn’s lemma). In fact, the existence of non-principal ultrafilters is equivalent to the Boolean prime ideal theorem, which is weaker than the axiom of choice.

A filter on a set $X$ gives a notion of convergence for maps from $X$ to topological spaces. Throughout, by a neighborhood of a point $y$ in a topological space $Y$ we can a subset $U$ of $Y$ having $y$ in its interior.

Definition 2.6.16. Let $\mathcal{F}$ be a filter on a set $X$. Let $Y$ be a topological space and let $f : X \to Y$ be a map. For any $y \in Y$ write

$$\lim_{x \to \mathcal{F}} f(x) = y$$

if $f^{-1}U \in \mathcal{F}$ for every neighborhood $U$ of $y$.

Example 2.6.17. Let $X$ be a non-empty set and let $f : X \to Y$ be a map into a topological space. Then

$$\lim_{z \to p} f(z) = f(x)$$

for every $x \in X$.

Since non-principle ultrafilters contains the Fréchet filter, convergence along a non-principal ultrafilter extends the usual notion of limit.

Lemma 2.6.18. Let $p$ be a non-principle ultrafilter on a set $X$ and let $f : X \to Y$ be a map into a topological space. If $\lim_{x \to \infty} f(x) = y$ then $\lim_{x \to p} f(x) = y$.

Proof. Since $p$ is non-principle it contains the Fréchet filter by Lemma 2.6.14. \qed

When $Y$ is Hausdorff limits along filters are unique, if they exist.
Lemma 2.6.19. Let $\mathcal{F}$ be a filter on a set $X$. Let $Y$ be a Hausdorff topological space and let $f : X \to Y$ be a map. If

$$\lim_{x \to \mathcal{F}} f(x) = y_1 \text{ and } \lim_{x \to \mathcal{F}} f(x) = y_2,$$

then $y_1 = y_2$.

Proof. Suppose that $y_1$ and $y_2$ are distinct. Since $Y$ is Hausdorff we can find disjoint open subsets $U_1$ and $U_2$ of $Y$ containing $y_1$ and $y_2$ respectively. Thus the disjoint sets $f^{-1}(U_1)$ and $f^{-1}(U_2)$ belong to $\mathcal{F}$, which is a contradiction. $\square$

Limits along ultrafilters exist in any compact topological space.

Lemma 2.6.20. Let $p$ be an ultrafilter on a set $X$. Let $Y$ be a compact topological space and let $f : X \to Y$ be a map. Then

$$\lim_{x \to p} f(x) = y \quad (2.6.21)$$

for some $y \in Y$.

Proof. Suppose there is no $y \in Y$ for which (2.6.21) exists. Then for every $y \in Y$ there is an open neighborhood $U_y$ of $y$ such that $f^{-1}(U_y) \notin p$. By compactness there are finitely many points $y_1, \ldots, y_n$ in $Y$ such that the corresponding neighborhoods $U_{y_1}, \ldots, U_{y_n}$ cover $Y$. The sets $f^{-1}(U_{y_1}), \ldots, f^{-1}(U_{y_n})$ therefore cover $X$, so one of them must belong to $p$, a contradiction. $\square$

The set of ultrafilters on a non-empty set can be given a compact Hausdorff topology. This is done by defining, for any subset $A$ of $X$, its closure $\text{cl}(A)$ in $\beta X$ and showing that the collection of all such closures generates a topology of open sets on $\beta X$.

Definition 2.6.22. Let $X$ be a non-empty set. Denote by $\beta X$ the set of ultrafilters on $X$. Define $\text{cl}(A) = \{ p \in \beta X : A \in p \}$ for every $A \subset X$.

It is immediate that $\text{cl}(A \cap B) = \text{cl}(A) \cap \text{cl}(B)$ for any subsets $A,B$ of a non-empty set $X$. Thus there is a unique topology on $\beta X$ for which the collection $\{ \text{cl}(A) : A \subset X \}$ forms a base. Unless otherwise stated, we always assume $\beta X$ is equipped with this topology.

Proposition 2.6.23. For any non-empty set $X$ the space $\beta X$ is Hausdorff and compact.

Proof. To see that $\beta X$ is Hausdorff, fix distinct ultrafilters $p$ and $q$ on $X$. There is $A \subset X$ such that $A \in p$ and $A \notin q$. Thus $\text{cl}(A)$ and $\text{cl}(X \setminus A)$ are disjoint open sets containing $p$ and $q$ respectively.

Now suppose that $\beta X$ is not compact. This means that there is an infinite open cover of $\beta X$ having no finite sub cover. Without loss of generality we may assume that the cover
consists of sets from the base of the topology. Let \( C = \{ \text{cl}(A_i) : i \in I \} \) be such a cover. Consider the collection \( D = \{ X \setminus A_i : i \in I \} \). Either \( D \) has the finite intersection property or it does not.

If \( D \) has the finite intersection property then there is an ultrafilter \( p \) on \( X \) that contains \( D \) by Lemma 2.6.6 and Theorem 2.6.15. But such an ultrafilter is contained in none of the \( \text{cl}(A_i) \), a contradiction.

If \( D \) does not have the finite intersection property then one can find \( i_1, \ldots, i_k \) in \( I \) such that \( X \setminus A_i \cap \cdots \cap X \setminus A_{i_k} = \emptyset \). But then \( A_{i_1} \cup \cdots \cup A_{i_k} = X \) so \( \text{cl}(A_{i_1}) \cup \cdots \cup \text{cl}(A_{i_k}) = \beta X \), which is also a contradiction.

Our next goal is to show that \( \beta X \) is the Stone-Čech compactification of \( X \) considered as a discrete space via the continuous embedding \( x \mapsto p_x \) defined by (2.6.11). We do so by showing that \( \beta X \) has the necessary universal property.

**Proposition 2.6.24.** Let \( X \) be a non-empty set and let \( f : X \to H \) be a map into a compact Hausdorff space. Then there is a unique continuous map \( \beta f : \beta X \to H \) such that \( (\beta f)(p_x) = f(x) \) for all \( x \in X \).

**Proof.** For each \( p \) in \( \beta X \) there is a unique point \( y \in Y \) such that (2.6.21) holds by Lemmas 2.6.19 and 2.6.20. This defines a map \( \beta f : \beta X \to Y \) by

\[
(\beta f)(p) = \lim_{x \to p} f(x)
\]

for all \( p \) in \( \beta X \). We have \( (\beta f)(p_x) = f(x) \) for all \( x \in X \) by Example 2.6.17.

It remains to show that \( \beta f : \beta X \to Y \) is continuous. To this end, fix an open subset \( U \) of \( Y \). We have

\[
(\beta f)^{-1}(U) = \{ p \in \beta X : \lim_{x \to p} f(x) \in U \} = \{ p \in \beta X : \{ x \in X : f(x) \in U \} \in p \} = \text{cl}(f^{-1}(U))
\]

so \( (\beta f)^{-1}(U) \) is open in \( \beta X \) as desired. \( \square \)

**Example 2.6.25.** Let \( X \) be a non-empty set. Then

\[
\lim_{x \to p} p_x = p
\]

for every \( p \) in \( \beta X \). Indeed, for any \( A \subset X \) with \( p \in \text{cl}(A) \) we have

\[
\{ x \in X : p_x \in \text{cl}(A) \} = A \in p
\]

as desired.
We now turn to ultrafilters on semigroups and the induced binary operation on the set of ultrafilters on a semigroup. Let $X$ be a semigroup. Fix each $x \in X$ define maps $L_x : X \to X$ and $R_x : X \to X$ by $L_x(y) = xy$ and $R_x(y) = yx$ respectively for every $y \in X$.

We can use $L_x$ and $R_x$ to permute $\beta X$ via the push-forward defined in Lemma 2.6.12.

Fix $q$ in $\beta X$ and consider the map $X \mapsto \beta X$ defined by $x \mapsto L_x[q]$ for every $x \in X$. By Proposition 2.6.24 we can extend this to a continuous map $p \mapsto L_p[q]$ defined on $\beta X$ upon setting

$$L_p[q] = \lim_{x \to p} L_x[q]$$

for every $p \in \beta X$. We claim that

$$L_p[q] = \{A \subset X : \{x \in X : x^{-1}A \in q\} \in p\}$$

for every $p \in \beta X$. Indeed the right-hand side is an ultrafilter and for any $A$ therein we have $\{x \in X : L_x[q] \in \text{cl}(A)\} \in p$ by definition.

Fixing again $q$ in $\beta X$, one can similarly consider the map $X \mapsto \beta X$ defined by $x \mapsto R_x[q]$ for every $x \in X$. Its extension to a continuous map $p \mapsto R_p[q]$ is given by

$$R_p[q] = \{A \subset X : \{x \in X : Ax^{-1} \in q\} \in p\}$$

for all $p \in \beta X$.

Thus we define operations $\ltimes$ and $\rtimes$ on $\beta X$ by

$$p \ltimes q = \{A \subset X : \{x \in X : x^{-1}A \in q\} \in p\} = L_p[q]$$
$$q \rtimes p = \{A \subset X : \{x \in X : Ax^{-1} \in q\} \in p\} = R_p[q]$$

for any two ultrafilters $p$ and $q$ in $\beta X$. These operations are associative. Indeed

$$(p \ltimes q) \ltimes r = \{A \subset X : \{x \in X : x^{-1}A \in r\} \in p \ltimes q\}$$
$$= \{A \subset X : \{y \in X : y^{-1}\{x \in X : x^{-1}A \in r\} \in q\} \in p\}$$
$$= \{A \subset X : \{y \in X : \{x \in X : x^{-1}(y^{-1}A) \in r\} \in q\} \in p\}$$
$$= \{A \subset X : \{y \in X : y^{-1}A \in q \rtimes r\} \in p\} = p \ltimes (q \rtimes r)$$

and a similar calculation holds for $\rtimes$.

**Example 2.6.26.** For any $x, y \in X$ we have $p_x \ltimes p_y = p_{xy} = p_x \rtimes p_y$.

**Lemma 2.6.27.** Let $X$ be a commutative semigroup. Then $p \ltimes q = q \rtimes p$ for all $p, q$ in $\beta X$.

Fix a semigroup $G$. We now consider how the multiplications $L$ and $R$ on $\beta G$ interact with left and right actions of $G$ respectively. Let $T : G \times X \to X$ be a left action of $G$ on a
compact Hausdorff topological space $X$ via continuous maps. Thus the induced continuous maps $T^x : X \to X$ satisfy $T^g \circ T^h = T^{gh}$ for all $g, h \in X$. Fix $p$ in $\beta G$. Define

$$T^p x = \lim_{g \to p} T^g x$$

for each $x \in X$. Note that $T^p$ need not be continuous, as the following example shows.

**Example 2.6.28.** Consider the action of the multiplicative semigroup $\mathbb{N}$ on $[0, 1]$ defined by $T^nx = x^n$. Then

$$T^p x = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$$

for every $0 \leq x \leq 1$ and every non-principle ultrafilter $p$ on $\mathbb{N}$ by Lemma 2.6.18.

We now show that $p \mapsto T^p$ induces an action of $\beta G$ on $X$ with respect to $\preceq$.

**Proposition 2.6.29.** Let $T : G \times X \to X$ be a left action of a semigroup $G$ on a compact Hausdorff space $X$ via continuous maps. Then $T^p \cdot q = T^p \circ T^q$ for all $p, q$ in $\beta G$.

**Proof.** Fix $x \in X$. Put $z = T^p(T^q x)$. Let $U$ be an arbitrary neighborhood of $z$. We have

$$p \ni \{h \in G : T^h(T^q x) \in U\}$$

$$= \{h \in G : T^q x \in (T^h)^{-1}U\}$$

$$= \{h \in G : \{g \in G : T^g x \in (T^h)^{-1}U\} \in q\}$$

$$= \{h \in G : \{g \in G : T^{hg} x \in U\} \in q\}$$

$$= \{h \in G : h^{-1}\{g \in G : T^g x \in U\} \in q\}$$

so $\{g \in G : T^g x \in U\} \in p \cdot q$. Since $U$ was an arbitrary neighborhood of $z$ we see that $T^{p \cdot q} = z$ as desired. \qed

The version of Proposition 2.6.29 for right actions is obtained mutatis mutandis.

**Proposition 2.6.30.** Let $T : G \times X \to X$ be a right action of a semigroup $G$ on a compact Hausdorff space $X$ via continuous maps. Then $T^p \cdot q = T^q \circ T^p$ for all $p, q$ in $\beta G$.

Of crucial importance for applications to ergodic theory are the idempotent ultrafilters, whose definition and existence we now turn to.

**Definition 2.6.31.** Let $p$ be an ultrafilter on a semigroup $G$. We say that $p$ is **left idempotent** if $p \cdot p = p$ and **right idempotent** if $p \cdot p = p$.

Note that when $G$ is abelian there is no distinction between left idempotents and right idempotents. The existence of idempotent ultrafilters follows from a standard argument that we now present.
**Theorem 2.6.32** ([Ell58, Lemma 1]). Let $G$ be a semigroup. Then there is $q$ in $\beta G$ such that $q \ltimes q = q$.

**Proof.** First we note that $\beta G$ has a minimal, non-empty compact sub-semigroup. Indeed, note that the intersection of a chain of non-empty, compact sub-semigroups is non-empty by compactness, so Zorn’s furnishes a minimal such sub-semigroup $S$.

Fix $q$ in $S$. Consider $S \ltimes q = \{ p \ltimes q : p \in S \}$. It is a non-empty sub-semigroup of $S$ that is compact because the map $p \mapsto p \ltimes q$ is continuous. Therefore $S \ltimes q = S$ by minimality.

Consider now the collection $\{ p \in S : p \ltimes q = q \}$, which is non-empty because $S \ltimes q = S$. It is also a compact sub-semigroup of $S$ so by minimality it equals $S$ and therefore contains $q$ as desired. \qed

We conclude this section by describing minimal ultrafilters. Note that any minimal left ideal $S$ for $\ltimes$ is automatically compact because $S = \beta G \ltimes p$ for any $p \in S$.

**Definition 2.6.33.** A left idempotent ultrafilter is a **minimal left idempotent** if it belongs to a minimal left ideal in $\beta G$, and a right idempotent ultrafilter is **minimal right idempotent** if it belongs to a minimal right ideal in $\beta G$.

As in the proof of Theorem 2.6.32, every non-zero left ideal contains a minimal left ideal, so Ellis’s lemma [Ell58, Lemma 1] implies that every minimal left ideal contains a left idempotent ultrafilter. The following lemma tells us that sets in left minimal idempotent ultrafilters have a certain largeness property. Recall that $S \subseteq G$ is **left syndetic** if there is a finite subset $F$ of $G$ for which $SF^{-1} = G$.

**Lemma 2.6.34** (cf. [Ber03, Theorem 2.4]). Let $G$ be a semigroup and let $p$ be a minimal left idempotent ultrafilter on $G$. For any $A \in p$ the set $AA^{-1} = \{ gh^{-1} : g, h \in A \}$ is left syndetic.

**Proof.** Fix $A \in p$. Let $X$ be a minimal left ideal containing $p$. Since $\beta G \ltimes p$ is a left ideal contained in $X$ it must be equal to $X$, so continuity of the map $q \mapsto p \ltimes q$ implies $X$ is compact. Consider the action $T$ of $G$ on $X$ defined by $T^g(p) = \delta_g \ltimes p$. This action is continuous. The set $U := \mathrm{cl}(A) \cap X$ is open in $X$ and contains $p$. We claim that the collection $\{(T^g)^{-1}U : g \in G\}$ covers $X$. Indeed, if not then the complement $V$ of the union of this collection is a closed, non-empty, $T$ invariant subset of $X$. This implies that $V$ is a left ideal, because any $q$ in $V$ satisfies

$$\beta G \ltimes q = \mathrm{cl}(G) \ltimes q = \mathrm{cl}(G \ltimes q) \subset \mathrm{cl}(V) = V$$

by continuity. By minimality $V = X$, which is a contradiction.
We have shown that the sets \((T^g)^{-1}U\) cover \(X\). By compactness we can extract a finite sub-cover \((T^{g_1})^{-1}U, \ldots, (T^{g_n})^{-1}U\). Thus for every \(g \in G\) there is some \(1 \leq i \leq n\) for which \(T^{g}p \in (T^{g_i})^{-1}U\). We can rewrite this as \(\delta_{g,g} \preceq p \in U\), which is the same as \((g_i g)^{-1}A \in p\).

So for every \(g \in G\) there is \(1 \leq i \leq n\) such that \((g_i g)^{-1}A \in p\). Putting \(F = \{g_1, \ldots, g_n\}\), this is the same as saying \(\{g \in G : g^{-1}A \in p\}\) is left syndetic. Since \(A \in p\) the larger set \(\{g \in G : A \cap g^{-1}A \in p\}\) is also left syndetic. But \(g \in AA^{-1}\) if and only if \(A \cap g^{-1}A\) is non-empty, so \(AA^{-1} \supset \{g \in G : A \cap g^{-1}A \in p\}\) is left syndetic, as desired.

We conclude this section by stating a version of the van der Corput trick for ultrafilters.

**Proposition 2.6.35.** Let \(G\) be a group and let \(\mathcal{H}\) be a Hilbert space. For any sequence \(u : G \to \mathcal{H}\) that is norm-bounded and any idempotent ultrafilter \(p\) on \(G\), if

\[
\left| \lim_{h \to p} \lim_{g \to p} \langle u(hg), u(g) \rangle \right| < \varepsilon
\]

then \(\| \lim_{g \to p} u(g) \|^2 \leq \varepsilon\), the latter limit being taken in the weak topology on \(\mathcal{H}\).

**Proof.** See [Sch07, Lemma 4].

### 2.7 Structured sets in groups

In this section we recall some relations between different types of structured sets in groups. We begin by defining sets of positive density.

**Definition 2.7.1.** Given a left Følner net \(\Phi\) in an amenable group \(G\) with a left Haar measure \(m\) and a measurable subset \(E\) of \(G\), define the **upper density** of \(E\) with respect to \(\Phi\) by

\[
\overline{d}_\Phi(E) = \limsup_i \frac{m(E \cap \Phi_i)}{m(\Phi_i)}
\]

the **lower density** of \(E\) with respect to \(\Phi\) by

\[
\underline{d}_\Phi(E) = \liminf_i \frac{m(E \cap \Phi_i)}{m(\Phi_i)}
\]

and, if the corresponding limit exists, the **density** of \(E\) with respect to \(\Phi\) as the common value \(d_\Phi(E)\) above.

Finite-index subgroups have the expected density with respect to every left Følner net.

**Lemma 2.7.2.** Let \(G\) be an amenable group and let \(H \subset G\) be a finite index subgroup that is also a measurable subset of \(G\). Then

\[
d_\Phi(H) = \frac{1}{[G : H]}
\]

for all left Følner nets \(\Phi\) in \(G\).
Proof. Let \( g_1, \ldots, g_k \) be left coset representatives for \( H \). We have
\[
\lim_{i} \frac{m(\Phi_i \cap H)}{m(\Phi_i)} - \frac{m(\Phi_i \cap gH)}{m(\Phi_i)} = 0
\]
for any \( g \in G \) so
\[
1 = \limsup_{i} \frac{m(\Phi_i \cap g_1H)}{m(\Phi_i)} + \cdots + \frac{m(\Phi_i \cap g_kH)}{m(\Phi_i)} = k \limsup_{i} \frac{m(H \cap \Phi_i)}{m(\Phi_i)}
\]
with the same holding for the limit inferior.

We next related positive density to syndeticity.

Definition 2.7.3. A subset \( S \) of a topological group \( G \) is left syndetic if there is a compact set \( K \subset G \) such that \( K^{-1}S = G \), and left thick if for every compact set \( K \subset G \) there is \( g \in G \) such that \( Kg \subset S \). Write \( S \) for the collection of left syndetic subsets of \( G \) and \( T \) for the collection of left thick subsets of \( G \). Declare that a subset of \( G \) belongs to \( S^* \) if and only if its intersection with every left syndetic set is non-empty, and that a subset of \( G \) belongs to \( T^* \) if and only if its intersection with every left thick set is non-empty.

Although it is immaterial for groups, we write \( K^{-1} \) in the definition of left syndetic so that the results below apply without modification to topological semigroups as well.

Lemma 2.7.4. For any topological group \( G \) one has \( S^* = T \) and \( T^* = S \).

Proof. First, let \( S \) be syndetic in \( G \) and let \( T \) be thick in \( G \). By syndeticity there exists a compact set \( K \subset G \) such that \( K^{-1}S = G \). Since \( T \) is thick there exists \( g \in G \) such that \( Kg \subset S \). We can find \( k \in K \) such that \( kg \in S \) because \( K^{-1}S = G \). Certainly \( kg \in T \), so the intersection of any left syndetic set with any left thick set is non-empty. This implies \( T \subset S^* \) and \( S \subset T^* \).

We now prove that if \( G = P \cup Q \) then either \( P \) is thick or \( Q \) is syndetic. If \( P \) is not thick then there exists a compact set \( K \) such that for each \( g \in G \) we have \( Kg \cap Q \neq \emptyset \). Thus for any \( g \in G \) we can find \( k \in K \) and \( q \in Q \) such that \( kg = q \). This implies that \( K^{-1}Q = G \) as desired.

Now, if \( P \) does not belong to \( T \) then its complement is syndetic so \( P \) does not belong to \( S^* \). Similarly, if \( Q \) does not belong to \( S \) then its complement is thick so \( Q \) does not belong to \( T^* \). Thus \( T \supset S^* \) and \( S \supset T^* \).

Lemma 2.7.5. Let \( G \) be an amenable group and let \( E \subset G \) be measurable. If \( \tilde{d}_\Phi(E) > 0 \) for every left Følner net \( \Phi \) in \( G \) then \( E \) is left syndetic.
Proof. Fix a left Følner net \( \Phi \) in \( G \). If \( E \) is not left syndetic then for every compact subset \( K \) of \( G \) we can find \( g_K \in G \setminus K^{-1}E \), which is to say \( Kg_K \cap E = \emptyset \). In particular, for every \( i \) we can find \( g_i \) such that \( \Phi_i g_i \cap E = \emptyset \). Define a net \( \Psi \) by \( \Psi_i = \Phi_i g_i \). We have

\[
\frac{m(\Psi_i \triangle g_i \Psi_i)}{m(\Psi_i)} = \frac{m(\Phi_i \triangle g_i \Phi_i \triangle (g_i))}{m(\Phi_i \triangle (g_i))}
\]

where \( \triangle \) is the modular function, so \( \Psi \) is also a left Følner net. By construction \( d(\Psi_i(E)) = 0 \) as desired.

We will need the following results about averages in discrete, countable, amenable groups later.

**Lemma 2.7.6.** Let \( G \) be a discrete, countable, amenable group. If \( \phi : G \to [0, \infty) \) has the property that

\[
\liminf_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \phi(g) > 0 \tag{2.7.7}
\]

for every left Følner sequence \( \Phi \) in \( G \), then there is some \( c > 0 \) such that

\[
\liminf_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \phi(g) \geq c
\]

for every left Følner sequence \( \Phi \) in \( G \).

Proof. If not then for every \( k \in \mathbb{N} \) there is left a Følner sequence \( N \mapsto \Phi_{k,N} \) such that

\[
0 \leq \liminf_{N \to \infty} \frac{1}{|\Phi_{k,N}|} \sum_{g \in \Phi_{k,N}} \phi(g) < \frac{1}{k}
\]

and defining \( \Phi_N = \Phi_{k_N,N} \) with \( k_N \to \infty \) sufficiently quickly gives a Følner sequence \( \Phi \) for which (2.7.7) does not hold.

**Lemma 2.7.8.** Let \( G \) be a discrete, countable amenable group. If \( \phi : G \to [0, \infty) \) is bounded and (2.7.7) holds for every left Følner sequence \( \Phi \) in \( G \), then there is a constant \( c > 0 \) such that \( \{u \in G : \phi(u) \geq c\} \) is left syndetic.

Proof. Choose \( c \) as in the conclusion of Lemma 2.7.6. We claim that

\[
A = \{u \in G : \phi(u) \geq c/2\}
\]

is left syndetic. If not then \( \overline{d}_\Phi(A) = 0 \) for some left Følner sequence \( \Phi \) by Lemma 2.7.5. But

\[
c \leq \limsup_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in A \cap \Phi_N} \phi(g) + \limsup_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N \cap (G \setminus A)} \phi(g) \leq c/2
\]

makes this impossible.
The following result will also be useful later.

**Lemma 2.7.9.** Let \( g \mapsto \phi(g) \) be a sequence from a countable, discrete, amenable group \( G \) to a normed vector space \( (X, \| \cdot \|) \) and let \( \Phi \) be a left Følner sequence in \( G \). If
\[
\lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \| \phi(g) - x \| = 0
\]
then \( d_\mathbb{F}(\{ g \in G : \| \phi(g) - x \| \geq \varepsilon \}) = 0 \) for every \( \varepsilon > 0 \).

**Definition 2.7.10.** Let \( F = \{ \alpha \subset \mathbb{N} : 0 < |\alpha| < \infty \} \). For any sequence \( n \mapsto \phi_n \) in a topological group \( G \) and any \( \alpha = \{ k_1 < \cdots < k_m \} \) in \( F \), define \( I_\alpha(\phi) = \phi_{k_1} \cdots \phi_{k_m} \) and \( D_\alpha(\phi) = \phi_{k_m} \cdots \phi_{k_1} \). Then \( FPI(\phi) = \{ I_\alpha(\phi) : \alpha \in F \} \) is the **increasing finite products set** or **increasing IP set** determined by the sequence \( \phi \) and \( FPD(\phi) = \{ D_\alpha(\phi) : \alpha \in F \} \) is the **decreasing finite products set** or **decreasing IP set** determined by the sequence \( \phi \). Put \( I_\emptyset(\phi) = D_\emptyset(\phi) = 1 \). Lastly, define by
\[
FP(\phi) = \{ I_\alpha(\phi)D_\beta(\phi) : \alpha, \beta \in F \cup \\{ \emptyset \}, \alpha \cap \beta = \emptyset, \alpha \cup \beta \neq \emptyset \}
\]
the **two-sided finite products set** determined by the sequence \( \phi \).

When \( G \) is abelian there is no distinction between increasing and decreasing IP sets, and we speak simply of IP sets.

The term IP was introduced in [FW78], the initials standing for “idempotence” or “infinite-dimensional parallelopiped”. Note that \( FP(\phi) \) contains both \( FPI(\phi) \) and \( FPD(\phi) \) because we can take either \( \beta \) or \( \alpha \) to be empty.

**Definition 2.7.11.** A subset of a group \( G \) is \( FPI^* \) if its intersection with every increasing finite products set is non-empty, and \( FPD^* \) if its intersection with every decreasing finite products set is non-empty. When \( G \) is abelian there is no distinction between \( FPI^* \) sets and \( FPD^* \) sets, and we accordingly speak simply of IP* sets.

**Lemma 2.7.12.** Every \( FPI^* \) subset of a group \( G \) is left syndetic.

**Proof.** We show that every set \( S \subset G \) that is not left syndetic contains an \( FPI \) set. From the proof of Lemma 2.7.4 we know that \( S \) is left thick. Thus for any finite set \( F \subset G \) we can find \( g_F \in G \) such that \( Fg_F \subset S \).

We proceed by inductively constructing an \( FPI \) set in \( S \). To begin, let \( \phi_1 \in G \) be such that \( \{1\} \phi_1 \subset S \). For the induction step, assume that we have found \( \phi_1, \ldots, \phi_n \in G \) such that \( F = \{ I_\alpha : \emptyset \neq \alpha \subset [n] \} \subset S \) and choose \( \phi_{n+1} \in G \) such that \( F\phi_{n+1} \subset S \).
The following notion of structure mixes products sets and density.

**Definition 2.7.13.** The **left upper Banach density** of a measurable subset $S$ of an amenable group $G$ is defined by

$$\overline{d}(S) = \sup \left\{ \overline{d}_\Phi(S) : \Phi \text{ a left Følner net in } G \right\}$$

and $S \subset G$ is AIP$^*$ (with A standing for “almost”) if it is of the form $A \setminus B$ where $A$ is an FPI$^*$ subset of $G$ and $\overline{d}(B) = 0$. We say that $S \subset G$ is AIP$^*_{\,+}$ if it is a shift of an AIP$^*$ set.

**Lemma 2.7.14.** Every AIP$^*_+\,+\,$ subset of a countable, discrete, abelian group $G$ is syndetic.

**Proof.** Every IP$^*$ subset of $G$ is syndetic by Lemma 2.7.12. Let $A \subset G$ be IP$^*_+\,$ and let $B \subset G$ have zero upper Banach density. Shifts of syndetic sets are themselves syndetic so $A$ is syndetic by the above argument, and therefore has positive upper density with respect to every Følner sequence. Now $\overline{d}_\Phi(B) = 0$ for every Følner sequence, so $\overline{d}_\Phi(A \setminus B) > 0$ for every Følner sequence. It now follows from Lemma 2.7.5 that $A \setminus B$ is syndetic. 

**Definition 2.7.15.** A subset of a group $G$ is a **left central** or C set if it belongs to some minimal left idempotent ultrafilter, and a **left central**$^*$ or C$^*$ set if it belongs to every minimal left idempotent ultrafilter. A subset of an abelian group is a C$^*_+$ if it is a shift of a C$^*$ set.

**Lemma 2.7.16.** Let $G$ be a group and let $A \subset G$ be left central$^*$ set. Then $A$ is left syndetic.

**Proof.** Suppose $A$ is not left syndetic. Then its complement $B$ is left thick (see Lemma 2.7.4) meaning that for every finite subset $F$ of $G$ there is some $h \in G$ such that $Fh \subset B$. This implies that $\mathcal{F} = \{ g^{-1}B : g \in G \}$ has the finite intersection property. By Lemma 2.6.6 and Theorem 2.6.15 there is an ultrafilter $p$ on $G$ that contains all sets from $\mathcal{F}$. The collection $I$ of ultrafilters that contain all sets from $\mathcal{F}$ is a closed subset of $\beta G$. Moreover, it is a left-ideal. Indeed, if $q \in \beta G$ and $p \in I$ then it is immediate that $g^{-1}B \in q \uplus p$ for all $g \in G$. Since any left ideal contains a minimal left idempotent ultrafilter, there is a minimal idempotent in $I$. This implies that $B$ is left central, so $A$ is not left central$^*$. 

**Example 2.7.17.** Following the proof of [Ber03, Theorem 2.20] one can construct a $C^*_+$ subset of $\mathbb{Z}^m$ that is not syndetic. Therefore, in order to produce a syndetic set that is not AIP$^*_+$, it suffices to show that every AIP$^*$ subset of $\mathbb{Z}^m$ is a $C^*$ set. Let $S$ be an AIP$^*$ set and write $S = A \setminus B$ where $A$ is IP$^*$ and $\overline{d}(B) = 0$. Certainly $A$ is C$^*$. But every central set has positive upper Banach density by [Ber03, Theorem 2.4(iii)], so $A \setminus B$ remains C$^*$. 

30
Lemma 2.7.18. Let $G$ be a group and let $T$ be a distal action of $G$ on a compact metric space $(X,d)$ by continuous maps. Then

$$\lim_{g \to p} T^g x = x$$

(2.7.19)

for every $x \in X$ and every idempotent ultrafilter $p$ on $G$.

Proof. Fix $x \in X$ and an idempotent ultrafilter $p$ in $\beta G$. We have

$$\lim_{g \to p} T^g (\lim_{h \to p} T^h x) = \lim_{g \to p} \lim_{h \to p} T^{gh} x = \lim_{g \to p} T^g x =: y$$

because $p \ast p = p$ so $x$ and $y$ are proximal. By distality they must be equal. \qed

Corollary 2.7.20. Let $G$ be a group and let $T$ be a distal action of $G$ on a compact metric space $(X,d)$. For every $x \in X$ and every neighborhood $U$ of $x$ the set \{ $g \in G : T^g x \in U$ \} is IP$^\ast$.

Proof. Fix $x \in X$ and let $U$ be a neighborhood of $x$. Since $T$ is distal we have \{ $g \in G : T^g x \in U$ \} \in $p$ for every idempotent ultrafilter $p$ on $G$. But any set that belongs to every idempotent ultrafilter is IP$^\ast$ (see [HS12] for details). \qed

2.8 Finite IP sets

Let $\mathcal{F}$ be the collection of all finite, non-empty subsets of $\mathbb{N}$. Write $\alpha < \beta$ for elements of $\mathcal{F}$ if $\max \alpha < \min \beta$. A subset of $\mathcal{F}$ is an FU set if it contains a sequence $\alpha_1 < \alpha_2 < \cdots$ from $\mathcal{F}$ and all finite unions of sets from the sequence. Write $\mathcal{F}_r$ for all finite, non-empty subsets of $\{1, \ldots, r\}$. A subset of $\mathcal{F}_r$ (or of $\mathcal{F}$) is an FU$^s$ set if it contains sets $\alpha_1 < \cdots < \alpha_s$ from $\mathcal{F}_r$ (or from $\mathcal{F}$) and all finite unions. For any IP$^s_r$ set $A \supset \text{FS}(x_1, \ldots, x_r)$ in an abelian group $G$ there is a map $\mathcal{F}_r \to G$ given by $\alpha \mapsto \sum \{x_i : i \in \alpha\}$, and for any IP set in $G$ there is a map $\mathcal{F} \to G$ defined similarly.

Furstenberg and Katznelson [FK85] showed that any IP$^s_r$ set $A$ in $\mathbb{Z}$ satisfies

$$\liminf_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{N} \geq \frac{1}{2^r - 1}$$

so for any $r \in \mathbb{N}$ one can construct an IP$^s$ set that is not IP$^s_r$. The set $k\mathbb{N}$, with $k$ large enough, is one such example. As the following example shows, by removing well-spread IP$^s_r$ sets from $\mathbb{Z}$, it is possible to construct a set that is IP$^s$ but never IP$^s_r$.

Example 2.8.1. Let $A_r$ be the IP$^s_r$ set with generators $x_1 = \cdots = x_r = 2^{2^r}$ so that $A_r = \{i \cdot 2^{2^r} : 1 \leq i \leq r\}$. Let $A$ be the union of all the $A_r$. We claim that $A$ cannot contain an IP set, from which it follows that $\mathbb{N}\setminus A$ is IP$^s$. Since $A$ contains IP$^s_r$ sets for arbitrarily large $r$ we also have that $\mathbb{N}\setminus A$ is not IP$^s_r$ for any $r$. 

31
Suppose that \( x_n \) is a sequence generating an \( \mathcal{IP} \) set in \( A \). If one can find \( x_i \in A_r \) and \( x_j \in A_s \) with \( r < s \) then \( x_j + x_i \) does not belong to \( A \) because the gaps in \( A_s \) are larger than the largest element in \( A_r \). On the other hand, if all \( x_i \) belong to the same \( A_r \) then some combination of them is not in \( A \) because the gap between \( A_r \) and \( A_{r+1} \) is too large.

A family \( \mathcal{I} \) of subsets of \( G \) is said to have the \textbf{Ramsey property} if \( S_1 \cup S_2 \) belonging to \( \mathcal{I} \) always implies that at least one of \( S_1 \) or \( S_2 \) contains a member of \( \mathcal{I} \). It follows from the reformulation of Hindman’s theorem [Hin74], stated below, that the collection of all \( \mathcal{IP} \) subsets of a group \( G \) has the Ramsey property.

**Theorem 2.8.2** ([Hin74, Corollary 3.3]). For any coloring of \( \mathcal{F} \) one can find \( \alpha_1 < \alpha_2 < \cdots \) in \( \mathcal{F} \) such that the collection of all finite unions of the sets \( \alpha_i \) is monochromatic.

It follows from Hindman’s theorem that the intersection of an \( \mathcal{IP}^* \) set with an \( \mathcal{IP} \) set contains an \( \mathcal{IP} \) set, and that the intersection of two \( \mathcal{IP}^* \) sets is therefore \( \mathcal{IP}^* \). The collection of all \( \mathcal{IP}_r \) sets does not have the Ramsey property; nevertheless, there is a suitable replacement that allows one to deduce properties of \( \mathcal{IP}_r^* \) sets similar to those just described for \( \mathcal{IP}^* \) sets.

**Proposition 2.8.3.** For any \( s \) and \( k \) in \( \mathbb{N} \) there is an \( r \) such that any \( k \)-coloring of any \( \mathcal{IP}_r \) set yields a monochromatic \( \mathcal{IP}_s \) set.

**Proof.** Suppose to the contrary that one can find \( s \) and \( k \) in \( \mathbb{N} \) such that, for any \( r \) there is a \( k \)-coloring of an \( \mathcal{IP}_r \) set \( A_r \) having no monochromatic \( \mathcal{IP}_s \) subset. This coloring of \( A_r \) gives rise to a coloring \( c_r \) of \( \mathcal{F}_r \) via the canonical map \( \mathcal{F}_r \rightarrow A_r \). That no \( A_r \) contains a monochromatic \( \mathcal{IP}_s \) set implies that no \( \mathcal{F}_r \) contains a monochromatic \( \mathcal{FU}_s \) set. We now use Hindman’s theorem to reach a contradiction.

Let \( \alpha_i \) be an enumeration of \( \mathcal{F}_r \). We construct a coloring \( c : \mathcal{F}_r \rightarrow \{1, \ldots, k\} \) by induction on \( i \). To begin note that \( \alpha_1 \in \mathcal{F}_r \), whenever \( r > \max \alpha_1 \) so we can find a strictly increasing sequence \( r(1,n) \) in \( \mathbb{N} \) such that \( c_{r(1,n)}(\alpha_1) \) takes the same value for all \( n \). Put \( c(\alpha_1) = c_{r(1,n)}(\alpha_1) \). Now, assuming that we have found a strictly increasing sequence \( r(i,n) \) such that, for each \( 1 \leq j \leq i \) the color \( c_{r(i,n)}(\alpha_j) \) is constant in \( n \) and equal to \( c(\alpha_j) \), choose a strictly increasing subsequence \( r(i+1,n) \) of \( r(i,n) \) such that \( c_{r(i+1,n)}(\alpha_{i+1}) \) is constant and let this value be \( c(\alpha_{i+1}) \). The colors of \( \alpha_1, \ldots, \alpha_i \) are unchanged and the induction argument is concluded.

By Hindman’s theorem we can find \( \beta_1 < \cdots < \beta_s \) in \( \mathcal{F} \) such that \( B = \mathcal{FU}(\beta_1, \ldots, \beta_s) \) is monochromatic, meaning \( c \) is constant on \( B \). Choose \( i \) such that \( B \subset \{\alpha_1, \ldots, \alpha_i\} \) and then choose \( n \) so large that \( r(i,n) > \max \beta_s \). It follows that \( B \subset \mathcal{F}_{r(i,n)} \) is monochromatic because \( c_{r(i,n)}(\beta) = c(\beta) \) for all \( \beta \in B \). Thus \( \mathcal{F}_{r(i,n)} \) contains a monochromatic \( \mathcal{FU}_s \) set, which is a contradiction. \( \square \)
With this version of partition regularity for $\text{IP}_r$ sets we can deduce some facts about $\text{IP}_r^*$ sets.

**Proposition 2.8.4.** Given any $s \in \mathbb{N}$ there is some $r \in \mathbb{N}$ such that any $\text{IP}_r^*$ set intersects any $\text{IP}_r$ set in an $\text{IP}_s$ set.

*Proof.* Let $A$ be an $\text{IP}_r^*$ set and choose by the previous proposition some $r$ such that any two-coloring of an $\text{IP}_r$ set yields a monochromatic $\text{IP}_s$ set. Let $B$ be an $\text{IP}_r$ set. One of $B \cap A$ and $B \setminus A$ contains an $\text{IP}_s$ set. It cannot be $B \setminus A$ because $A$ is $\text{IP}_r^*$ and disjoint from it. Thus $A \setminus B$ contains an $\text{IP}_s$ set as desired. \hfill \Box

**Proposition 2.8.5.** Given any $r, s \in \mathbb{N}$ there is some $\alpha(r, s) \in \mathbb{N}$ such that if $A$ is $\text{IP}_r^*$ and $B$ is $\text{IP}_s^*$ then $A \setminus B$ is $\text{IP}_{\alpha(r, s)}^*$.

*Proof.* Let $A$ be $\text{IP}_r^*$ and let $B$ be $\text{IP}_s^*$ with $r \geq s$. Choose $q$ so large that $A \cap C$ contains an $\text{IP}_r$ set whenever $C$ is an $\text{IP}_q$ set. This is possible by the previous result. Since $A \cap C$ contains an $\text{IP}_r$ set and $r \geq s$ the set $(A \cap C) \cap B$ must be non-empty. Since $C$ was arbitrary $A \cap B$ is an $\text{IP}_q^*$ set. Put $\alpha(r, s) = q$. \hfill \Box
CHAPTER 3
CORRESPONDENCE PRINCIPLES

As we have described in Section 1.2, Furstenberg translated Szemerédi’s theorem into an ergodic-theoretic statement via a correspondence principle in [Fur77]. A correspondence principle for discrete, countable, amenable groups was proved by Bergelson, McCutcheon and Zhang in [BMZ97], and this was generalized to locally compact, second countable groups by Bergelson and Furstenberg in [BF09]. In this chapter we describe in Section 3.1 a two-sided version of the correspondence principle from [BF09] for amenable groups, and give in Section 3.2 a more explicit proof of the version for countable, amenable groups. Our correspondence principle is similar to the one in [BCRZK14].

3.1 A correspondence principle for amenable groups

It turns out that it is not possible to find a correspondence principle that applies to arbitrary Borel sets of positive upper density, even in $\mathbb{R}$. Indeed, setting $\Phi_N = [0, N]$ in $\mathbb{R}$, [BBB94, Theorem D] exhibits for all but countably many $\alpha > 0$ a subset $A$ of $\mathbb{R}$ with $d_{\Phi}(A) = 1/2$ such that $d_{\Phi}((A - n\alpha) \cap A) = 0$ for all $n \in \mathbb{N}$. On the other hand, whenever $\alpha > 0$ is irrational the sequence $n\alpha t$ is uniformly distributed mod 1 for any real, non-zero $t$ (this was first proved in [Csi30]) so the set $\{n\alpha : n \in \mathbb{N}\}$ is a set of measurable recurrence, meaning that for any $\mathbb{R}$ system $X$ and any $B \in \mathcal{B}$ with $\mu(B) > 0$ one can find $n$ such that $\mu(B \cap T^{-n\alpha} B) > 0$. Thus there is no analogue of (1.2.5) for arbitrary Borel subset $A$ of $\mathbb{R}$.

An appropriate one-sided correspondence principle for general locally compact, second countable groups was given in [BF09, Theorem 1.1] by restricting one’s attention to neighbourhoods of measurable sets having positive upper density. We describe a two-sided version adequate for our needs, using means instead of Følner sequences.

As is [BF09], the correspondence principle will rest on an application of the Gelfand-Naimark theorem (see [GRS64]) applied to a $C^*$ algebra. Here, we use the $C^*$ algebra $UC(G)$ of bounded, uniformly continuous functions on $G$ under pointwise operations and the topology of uniform convergence.
By the Gelfand-Naimark theorem there is a compact, Hausdorff space $X$ such that $\text{UC}(G)$ and $\text{C}(X)$ are isomorphic as $C^*$ algebras. In fact, one can take $X$ to be the space of $C^*$ morphisms $\text{UC}(G) \to \mathbb{C}$ equipped with the weak* topology. Define actions $T_1$ and $T_2$ of $G$ on $X$ by

$$(T_g^1 x)(f) = x(L_g f) \quad (T_g^2 x)(f) = x(R_{g^{-1}} f)$$

for all $g \in G$, all $x \in X$ and all $f \in \text{UC}(G)$. Since $T_1$ and $T_2$ commute they define an action $T$ of $G^2$ on $X$ by $T^{(g_1, g_2)} x = T_{g_1}^1 T_{g_2}^2 x$. The corresponding map

$$T : G \times G \times X \to X$$

is continuous by Lemma 2.2.3. Moreover, invariant means on $\text{UC}(G)$ are in bijective correspondence with $T$ invariant measures on $X$. Indeed, given a mean $M$ on $\text{UC}(G)$, one can define a Baire measure $\mu_M$ on $X$ via

$$\int f \, d\mu_M = M(\hat{f})$$

where $\hat{f}$ is the uniformly continuous function on $G$ corresponding to $f \in \text{C}(X)$, and given a Baire measure $\mu$ on $X$ one can define a mean $M_\mu$ on $\text{UC}(G)$ by

$$M_\mu(f) = \int \hat{f} \, d\mu$$

(3.1.1)

where $\hat{f}$ is the continuous function on $X$ corresponding to $f \in \text{UC}(G)$.

We will use the above construction to produce a correspondence principle for sets with a certain largeness property that we now describe.

**Definition 3.1.2.** Let $G$ be an amenable group with a left Haar measure $m$. A subset $S$ of $G$ is **substantial** if one can find $f \in \text{UC}(G)$ with $1_S \geq f \geq 0$ and a mean $M$ on $\text{UC}(G)$ that is left and right invariant such that $M(f) > 0$.

Note that our definition of substantial differs from the one given in [BF09] in that it is two-sided. One can create substantial sets from sets with positive density.

**Example 3.1.3.** Let $G$ be an amenable group with a left Haar measure $m$ and let $E \subset G$ have positive upper density with respect to a two-sided mean $M$ on $L^\infty(G, m)$. For any open neighbourhood $U$ of 1 in $G$ the set $UEU$ is substantial.

**Proof.** Let $V$ be a symmetric, open neighbourhood of 1 in $G$ such that $VV \subset U$. Fix a continuous function $\phi : G \to [0, 1]$ with support in $V$ and $\phi(1) = 1$. The function $f = \phi \ast 1_E \ast \tilde{\phi}$ is uniformly continuous and $0 \leq f \leq 1_{UEU}$ by construction.

We now turn to the correspondence principle itself.

35
Theorem 3.1.4. Let $G$ be an amenable group and let $f \geq 0$ be uniformly continuous. Assume there is a two-sided invariant mean $M$ on $\text{UC}(G)$ such that $M(f) > 0$. Then one can find a $G^2$ system $X = (X, \mathcal{B}, \mu, T)$, and a continuous function $\xi : X \to [0, \infty)$ such that $M(f) = \|\xi\|_1$ and
\[
M(L_{g_1} R_{h_1}^{-1} f \cdots L_{g_k} R_{h_k}^{-1} f) = \int T_1^{g_1} T_2^{h_1} \cdots T_1^{g_k} T_2^{h_k} \xi d\mu
\]
(3.1.5)
for all $g_1, \ldots, g_k, h_1, \ldots, h_k \in G$.

Proof. Fix a two-sided invariant mean $M$ on $\text{UC}(G)$ with $M(f) > 0$. Let $X$ be the Gelfand spectrum of $\text{UC}(G)$ and let $\xi \in C(X)$ be the continuous function corresponding to $f \in \text{UC}(G)$. We have seen above that left and right multiplication of $G$ on $\text{UC}(G)$ induce actions $T_1$ and $T_2$ of $G$ on $X$ that together form a continuous $G^2$ action $T$ on $X$. Let $\mu$ be the Baire probability measure on $X$ corresponding to the mean $M$. By invariance of $M$ the measure $\nu$ is $T$ invariant. It follows from the isomorphism $\text{UC}(G) \cong C(X)$ of $C^*$ algebras that (3.1.5) holds for this $\mu$.

By varying the mean in (3.1.5) we can assume that the system obtained is ergodic. That one can associate an ergodic action with $E$ was first proved in [BHK05] using ideas from [Fur81a].

Theorem 3.1.6. Let $G$ be an amenable group and let $f \geq 0$ be uniformly continuous. Assume there is a two-sided invariant mean $M$ on $\text{UC}(G)$ such that $M(f) > 0$. Then one can find an ergodic $G^2$ system $X_{\text{erg}} = (X, \mathcal{B}, \nu, T)$, a continuous function $\xi : X \to [0, \infty)$ such that $M(f) = \|\xi\|_1$, and a two-sided invariant mean $M_{\text{erg}}$ on $\text{UC}(G)$ such that
\[
M_{\text{erg}}(L_{g_1} R_{h_1}^{-1} f \cdots L_{g_k} R_{h_k}^{-1} f) = \int T_1^{g_1} T_2^{h_1} \cdots T_1^{g_k} T_2^{h_k} \xi d\nu
\]
(3.1.7)
for all $g_1, \ldots, g_k, h_1, \ldots, h_k \in G$.

Proof. Let $X = (X, \mathcal{B}, \mu, T)$ be as in the proof of Theorem 3.1.4. We claim that there is a probability measure $\nu$ on $(X, \mathcal{B})$ that is ergodic for the $G^2$ action $T$ such that $\int \xi d\nu \geq \int \xi d\mu$. Indeed, consider the set
\[
\Lambda = \left\{ \lambda \in \mathcal{M}(X) : \int \xi d\lambda \geq \int \xi d\mu \right\}
\]
which is a Baire subset of $\mathcal{M}_T(X)$. By the Choquet-Bishop-de Leeuw theorem (Theorem 2.1.2) there is a measure $\eta$ on $\mathcal{M}_T(X)$ such that
\[
\phi(\mu) = \int \phi d\eta
\]
for every continuous, affine function $\phi : \mathcal{M}_T(X) \to \mathbb{R}$. Moreover, $\eta$ vanishes on the Baire sets of $\mathcal{M}_T(X)$ that are disjoint from the extreme points of $\mathcal{M}_T(X)$. If $\Lambda$ is disjoint from the extreme points of $\mathcal{M}_T(X)$ then

$$\int \xi \, d\mu > \int \xi \, d\nu \, d\eta(\nu)$$

contradicting the above representation, so there must exist an extreme point $\nu$ of $\mathcal{M}_T(X)$ with $\int \xi \, d\mu \leq \int \xi \, d\nu$. The measure $\nu$ is ergodic by Lemma 2.1.1. Since there is a mean $M_{\text{erg}}$ on UC($G$) that agrees with $\nu$ in the sense that (3.1.1) holds for all $f \in \text{UC}(G)$, we have (3.1.7) with this mean. \qed

### 3.2 A correspondence principle for discrete amenable groups

For countable, amenable groups the construction in the previous section can be made explicit. We present here a left-sided version; two-sided versions can also be produced (see [BCRZK14, Theorem 3.2]).

**Theorem 3.2.1.** Let $G$ be a countable, amenable group and let $S \subset G$ has positive upper density with respect to a left Følner sequence $\Phi$ in $G$. There is a compact, metric $G$ system $X = (X, \mathcal{B}, \mu, T)$ and $B \in \mathcal{B}$ with $\mu(B) = \overline{d}_\Phi(S)$ such that

$$\overline{d}_\Phi(g_1^{-1}S \cap \cdots \cap g_n^{-1}S) \geq \mu((Tg_1)^{-1}B \cap \cdots \cap (Tg_n)^{-1}B)$$

(3.2.2)

for all $g_1, \ldots, g_k$ in $G$.

**Proof.** Put $X = \{0, 1\}^G$. Define a $G$ action on $X$ by $(T^g x)(h) = x(hg)$. Write $T$ for the induced action $(T^g \phi)(x) = \phi(T^g x)$ on continuous functions $\phi : X \to \mathbb{R}$. Let $B = \{x \in X : x(1) = 1\}$. By passing to a subsequence of $\Phi$ we may assume that the sequence $N \mapsto \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \delta_{T^g 1_S}$ converges in the weak* topology to a measure $\mu = \overline{d}_\Phi(S)$. We claim that $\mu$ is $T$ invariant. Indeed

$$\int T^h \phi \, d\mu = \lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \int T^h \phi \, d\delta_{T^g 1_S}$$

$$= \lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} (T^h \phi)(T^g 1_S) = \lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} \phi(T^{hg} 1_S) = \int \phi \, d\mu$$

for any $\phi \in \text{C}(X)$ and any $g \in G$ because $\Phi$ is a left Følner sequence. Lastly, note that

$$\mu((Tg_1)^{-1}B \cap \cdots \cap (Tg_n)^{-1}B)$$

$$= \lim \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} |(Tg_1)^{-1}B \cap \cdots \cap (Tg_n)^{-1}B \cap \Phi_N|$$

$$\leq \overline{d}_\Phi(g_1^{-1}S \cap \cdots \cap g_n^{-1}S)$$

37
for all $g_1, \ldots, g_n \in G$ because
\[ \delta_{T^{g_1}}((T^{g_1})^{-1}B \cap \cdots \cap (T^{g_l})^{-1}B) = 1 \]
if and only if $g \in g_1^{-1}S \cap \cdots \cap g_l^{-1}S \neq \emptyset$.

As in the previous section, by varying the Følner sequence in the left-hand side of (3.2.2) one can assume that the system $X$ is ergodic.

**Theorem 3.2.3.** Let $G$ be a countable, amenable group and let $S \subset G$ has positive upper density with respect to a left Følner sequence $\Phi$ in $G$. There is an ergodic, compact, metric $G$ system $X = (X, \mathcal{B}, \nu, T)$, $\mathcal{B} \in \mathcal{B}$ with $\mu(B) = \overline{d}_\Phi(S)$, and a left Følner sequence $\Psi$ in $G$ such that
\[ \overline{d}_\Psi(g_1^{-1}S \cap \cdots \cap g_l^{-1}S) \geq \nu((T^{g_1})^{-1}B \cap \cdots \cap (T^{g_k})^{-1}B) \] (3.2.4)
for all $g_1, \ldots, g_k$ in $G$.

**Proof.** Let $(X, \mathcal{B}, \mu, T)$ and $B$ be as in Theorem 3.2.1. Fix an ergodic, $T$ invariant measure $\nu$ on $(X, \mathcal{B})$ satisfying $\nu(B) \geq \mu(B)$. By repeatedly applying the mean ergodic theorem and passing to a subsequence of $\Phi$ we can find $x \in X$ such that
\[ \lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f(T^g x) \]
exists for every $f \in C(X)$. Fix a sequence $n \mapsto f_n$ in $C(X)$ with dense image. By passing to a further subsequence of $\Phi$ if necessary we may suppose that
\[ \left| \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f_k(T^g x) - \int f \, d\nu \right| < \frac{1}{N} \]
for all $N \in \mathbb{N}$ and all $1 \leq k \leq N$. One can find a sequence $N \mapsto g_N$ in $G$ such that
\[ \left| \frac{1}{|\Phi_N|} \sum_{g \in \Phi_N} f_k(T^g(T^{g_N}1_S)) - \int f \, d\nu \right| < \frac{1}{N} \]
holds for all $N \in \mathbb{N}$ and all $1 \leq k \leq N$. Thus $1_S$ is generic for $\nu$ along the left Følner sequence $N \mapsto \Phi_N g_N$. Applying this to the indicator functions of the cylinder sets in $\{0,1\}^N$ gives (3.2.4).
CHAPTER 4

CHARACTERISTIC FACTORS

In this chapter we describe recent work [Rob14] on characteristic factors for (2.5.3) when $G$ is second countable using the classical dichotomy between weak-mixing and almost periodicity over a factor. We begin this chapter with a preliminary section on Borel-Hilbert bundles before proving in Section 4.2 the aforementioned dichotomy. We describe our characteristic factors in Section 4.6. In the final two sections of this chapter we use this description to prove a special case of multiple recurrence for (2.5.3).

The dichotomy between almost periodicity and weak mixing over a factor for actions of $\mathbb{Z}$ was proved by Furstenberg [Fur77] in his original proof of Szemerédi’s theorem, and we follow broadly the same approach, tempered with the van der Corput trick (Theorem 2.6.35) and some techniques from [BM07]. Throughout this chapter we only consider second countable groups and systems on compact metric spaces.

4.1 Borel-Hilbert bundles

In this section we describe how to associate Borel-Hilbert bundles with extensions and relatively independent joinings. We begin by recalling the definition of a Borel-Hilbert bundle. For details, see [Dix81], [Gla03] and [Wil07].

Let $(Y, \mathcal{D}, \mu)$ be a probability space and let $\mathcal{H} = \{H_y : y \in Y\}$ be a collection of separable, real Hilbert spaces. Write $\langle \cdot, \cdot \rangle_y$ for the inner product on $H_y$ and $\|\cdot\|_y$ for the norm. From $Y$ and $\mathcal{H}$ we can form the total space $Y \ast \mathcal{H} = \{(y, h) : y \in Y, h \in H_y\}$ which comes with a projection $\pi : Y \ast \mathcal{H} \to Y$. The spaces $H_y$ are called the fibers of the total space. A section of $Y \ast \mathcal{H}$ is any map $f : Y \to Y \ast \mathcal{H}$ such that $\pi \circ f$ is the identity. The image of a point $y$ under a section $f$ is a point in $Y \ast \mathcal{H}$ which we will write as $(y, f_y)$. Thus $f_y$ belongs to $H_y$. To any section $f$ we can associate the map $\tilde{f} : Y \ast \mathcal{H} \to \mathbb{R}$ defined by $\tilde{f}(y, h) = \langle f_y, h \rangle_y$. A Borel-Hilbert bundle is a Hilbert bundle $Y \ast \mathcal{H}$ equipped with a $\sigma$-algebra of subsets of $Y \ast \mathcal{H}$ for which:

(i) the projection $Y \ast \mathcal{H} \to Y$ is measurable;
(ii) there is a sequence $f_n$ of sections such that:

a) the maps $\tilde{f}_n$ are measurable;

b) for each $n, m$ the map $Y \to \mathbb{R}$ given by $y \mapsto \langle f_n(y), f_m(y) \rangle_y$ is measurable;

c) the functions $\tilde{f}_n$ and $\pi$ separate points on $Y \ast \mathcal{H}$.

Write $B(Y \ast \mathcal{H})$ for the set of all measurable sections. From these data one can form the Hilbert space $L^2(Y \ast \mathcal{H}, \mu)$ from the set

$$L^2(Y \ast \mathcal{H}, \mu) = \left\{ f \in B(Y \ast \mathcal{H}) : \int \|f_y\|^2 \, d\lambda(y) < \infty \right\}$$

of square-integrable sections by identifying sections that agree almost surely. In this sense, the Hilbert space $L^2(Y \ast \mathcal{H}, \mu)$ is the direct integral of the Hilbert space $\mathcal{H}_y$.

Fix now a locally compact, second countable group $G$ and a compact, metric $G$ system $X = (X, \mathcal{B}, \mu, T)$. We describe how to associate to any $T$ invariant sub-$\sigma$-algebra $\mathcal{D}$ of $\mathcal{B}$ a Borel Hilbert bundle. Fix a countable, dense subgroup $\Gamma$ of $G$, an almost-surely defined disintegration $\mu_x$ of $\mu$ over $\mathcal{D}$ and a countable sub-algebra $A = \{A_1, A_2, \ldots\}$ of $\mathcal{B}$ that generates $\mathcal{D}$ and has the property that $(T^\gamma)^{-1}A$ is $A$ for every $A \in A$ and every $\gamma \in \Gamma$. For each $n, m$ the function $x \mapsto (1_{A_n}, 1_{A_m})_x$ is defined on a full-measure subset of $X$ and is measurable there. Let $X_0$ be a $\Gamma$ invariant, full-measure subset of $X$ on which $\mu_x$ and all of the functions $x \mapsto (1_{A_n}, 1_{A_m})_x$ are defined and on which $T^\gamma \mu_x = \mu_{T^\gamma x}$ for all $\gamma \in \Gamma$. Put $\mathcal{H}_x = L^2(X, \mathcal{B}, \mu_x)$ when $x \in X_0$ and put $\mathcal{H}_x = \{0\}$ otherwise. Each $\mathcal{H}_x$ is separable because $\mathcal{B}$ is countably generated. Let $\mathcal{H}$ be the collection $\{\mathcal{H}_x : x \in X\}$. Define a sequence $f_n$ of sections by taking $f_{n,x} = 1_{A_n}$ when $x \in X_0$ and $f_{n,x} = 0$ otherwise. Equip $X \ast \mathcal{H}$ with the smallest $\sigma$-algebra of subsets for which $\pi$ and the maps $\tilde{f}_n$ are measurable. It is immediate from the construction that this $\sigma$-algebra makes $X \ast \mathcal{H}$ into a $\sigma$-Hilbert bundle. Moreover, a section $f : X \to X \ast \mathcal{H}$ is measurable with respect to this $\sigma$-algebra if and only if $x \mapsto \langle f_x, f_{n,x} \rangle_x$ is measurable for each $n$. We call $X \ast \mathcal{H}$ the Borel-Hilbert bundle corresponding to the sub-$\sigma$-algebra $\mathcal{D}$. The Hilbert space $L^2(X \ast \mathcal{H}, \mu)$ is isomorphic to $L^2(X, \mathcal{B}, \mu)$. Thus to any $\phi$ in $L^2(X)$ we can associate an almost-surely defined, square-integrable section $x \mapsto \phi_x$ and vice versa.

We now recall how $\Gamma$ acts on sections of $X \ast \mathcal{H}$. Fix $\gamma \in \Gamma$. Since $T^\gamma \mu_x = \mu_{T^\gamma x}$ whenever $x \in X_0$ the map $T^\gamma_x : \mathcal{H}_{T^\gamma x} \to \mathcal{H}_x$ given by $(T^\gamma_x f)(z) = f(T^\gamma z)$ is well-defined and unitary. Define $T^\gamma_x : \mathcal{H}_{T^\gamma x} \to \mathcal{H}_x$ to be the zero map when $x \notin X_0$. The family of maps $\{T^\gamma_x : x \in X\}$ induces a map $T^\gamma$ on sections of $X \ast \mathcal{H}$ such that $(T^\gamma f)_x = T^\gamma_x f_{T^\gamma x}$. If $f$ is a measurable section then so is $T^\gamma f$. Also $T^\gamma(T^\eta f) = T^{\gamma \eta} f$ for all $\gamma, \eta \in \Gamma$.

It remains to relate the Borel-Hilbert bundle associated with a relatively independent self-joining $X \times_\mathcal{D} X$ over $\mathcal{D}$ to the Hilbert bundle associated with $X$ and $\mathcal{D}$. Let $\mathcal{A}$ be a countable, $\Gamma$ invariant algebras that generate $\mathcal{B}$. The countable algebra generated by
\{A_1 \times A_2 : A_1, A_2 \in \mathcal{A}\} is \(\Gamma\) invariant and generates \(\mathcal{B} \otimes \mathcal{B}\). We can thus simultaneously form the Borel-Hilbert bundles \(X \ast \mathfrak{H}\) and \(X \ast \mathfrak{K}\) corresponding to \(\mathcal{D}\) as a sub-\(\sigma\)-algebra of \(\mathcal{B}\) and of \(\mathcal{B} \otimes \mathcal{B}\) respectively. From (2.5.4) we see that \(\mathfrak{K}_x = \mathfrak{H}_x \otimes \mathfrak{H}_x\) for \(\mu\) almost every \(x\). Thus for any section \(H\) of \(X \ast \mathfrak{K}\) and almost every \(x\) the corresponding member \(H_x\) of \(\mathfrak{K}_x\) induces a compact operator \(H_x : \mathfrak{H}_x \to \mathfrak{H}_x\) defined by

\[
(\phi \ast H_x)(x_2) = \int \phi_1(x_1) \cdot H(x_1, x_2) \, d\mu_x(x_1) \tag{4.1.1}
\]

for any \(\phi \in \mathfrak{H}_x\). This family of operators induces a map taking almost-surely defined sections of \(X \ast \mathfrak{H}\) to almost-surely defined sections of \(X \ast \mathfrak{H}\). If \(H\) is a measurable section of \(X \ast \mathfrak{K}\) the induced map preserves measurability of sections because

\[
x \mapsto \langle 1_{A_n} \ast H, 1_{A_m} \rangle_x = \langle H, 1_{A_n} \otimes 1_{A_m} \rangle_x
\]

is measurable for all \(n, m \in \mathbb{N}\). However, the induced map need not preserve square-integrability. It may happen that \(f\) is a square-integrable section of \(Y \ast \mathfrak{H}\) and \(x \mapsto f_x \ast H_x\) is not a square-integrable section of \(Y \ast \mathfrak{H}\). As the following proposition shows, we avoid this problem when the norms \(\|H_x\|_x\) are bounded almost-surely. In this case we write \(f \ast H\) for the element of \(L^2(X)\) corresponding to the square integrable section \(x \mapsto f_x \ast H_x\) of \(X \ast \mathfrak{H}\).

**Proposition 4.1.2.** Let \(H\) be a section of \(X \ast \mathfrak{H}\). If the norms \(\|H_x\|_x\) are essentially bounded and \(f\) is a square-integrable section of the bundle \(Y \ast \mathfrak{H}\) then \(x \mapsto f_x \ast H_x\) is a square-integrable section of \(Y \ast \mathfrak{H}\).

**Proof.** See Section F.3 in [Wil07]. \(\square\)

Given a measurable section \(H\) of \(X \ast \mathfrak{K}\) we can spectrally decompose the compact operator \(H_x : \mathfrak{K}_x \to \mathfrak{K}_x\) for almost every \(x \in X\). The following theorem, due to Furstenberg and Katznelson, shows that this spectral decomposition is measurable.

**Theorem 4.1.3** (3.7 in [FK91]). Let \(X = (X, \mathcal{D}, \mu, T)\) be a system and let \(\mathcal{D}\) be a sub-\(\sigma\)-algebra of \(\mathcal{B}\). Form the corresponding Borel-Hilbert bundle \(X \ast \mathfrak{H}\). Let \(H_x\) be a measurable family of positive-definite, self-adjoint, compact operators on \(\mathfrak{K}_x\). Let \(\lambda_n(x)\) be a decreasing enumeration of the positive eigenvalues of \(H_x\), counting multiplicities. There is a sequence \(\Psi_n\) of square integrable sections of \(X \ast \mathfrak{H}\) such that \(\Psi_{n,x} \ast H_x = \lambda_n(x)\Psi_{n,x}\) whenever \(\lambda_n(x)\) is defined, \(\Psi_{n,y} = 0\) otherwise, and \(\{\Psi_{n,x} : n \in \mathbb{N}\} \setminus \{0\}\) is orthonormal in almost every fiber.
4.2 Almost periodicity and weak mixing

Fix a second countable, amenable group $G$ and a countable, dense subgroup $\Gamma$ of $G$. Let $X = (X, \mathcal{B}, \mu, T)$ be a system and let $\mathcal{D}$ be a $T$ invariant sub-$\sigma$-algebra of $\mathcal{B}$. As in the previous section, fix a disintegration $\mu_x$ of $\mu$ over $\mathcal{D}$ and form the corresponding Borel-Hilbert bundle $X \ast \mathcal{D}$.

**Definition 4.2.1.** We say that $f$ in $L^2(X)$ is **almost-periodic** over $\mathcal{D}$ if for every $\varepsilon > 0$ one can find a finite set $\Xi \subset L^\infty(X)$ and $E \subset X$ with $\mu(E) > 1 - \varepsilon$ such that

$$\min\{|T^\gamma f - \xi|_x : \xi \in \Xi\} \leq \varepsilon$$

for each $\gamma \in \Gamma$ and almost every $x \in E$.

The closure in $L^2(X)$ of the set of functions that are almost-periodic over $\mathcal{D}$, which we denote $\mathcal{A}(X|\mathcal{D})$, forms a closed subspace of $L^2(X)$ that contains the constant functions. Moreover, if $f$ is almost-periodic then so is $|f|$. Thus condition (c) of [FK91, Lemma 3.1] is satisfied and there exists a sub-$\sigma$-algebra $\mathcal{C}$ of $\mathcal{B}$ such that $\mathcal{A}(X|\mathcal{D}) = L^2(X, \mathcal{C}, \mu)$. This lets us approximate any $f$ in $\mathcal{A}(X|\mathcal{D})$ arbitrarily well by a function in $L^\infty(X, \mathcal{C}, \mu)$ that is almost-periodic over $\mathcal{D}$ as follows: truncate $f$ at a high level and then re-define $f$ to be zero on certain fibers of the factor map as in the proof of Theorem 9.1 in [FKO82]. Since $\mathcal{A}(X|\mathcal{D})$ is closed and invariant under $T^\gamma$ for each $\gamma \in \Gamma$, it is also $T$ invariant. Thus $\mathcal{C}$ is $T$ invariant. When $\mathcal{D}$ is the trivial factor, write $\mathcal{A}(X)$ for $\mathcal{A}(X|\mathcal{D})$.

**Definition 4.2.3.** The **Kronecker factor** of a $G$ system $X$ is the sub-$\sigma$-algebra corresponding to $\mathcal{A}(X)$.

An element $f \in L^2(X)$ is measurable with respect to the Kronecker factor if and only if the orbit $\{T^g f : g \in G\}$ is totally bounded in $L^2(X)$.

**Definition 4.2.4.** We say that $X$ is **compact** over $\mathcal{D}$ if $\mathcal{A}(X|\mathcal{D}) = L^2(X)$ and **weak-mixing** over $\mathcal{D}$ if $\mathcal{A}(X|\mathcal{D}) = L^2(X, \mathcal{D}, \mu)$.

Almost periodic functions generalize the notion of having a totally bounded orbit. There is also a generalization of the notion of a finite-dimensional sub-representation due to Furstenberg [Fur77], the definition of which we now recall.

**Definition 4.2.5.** A function $f$ in $L^2(X)$ is an **eigenfunction** over $\mathcal{D}$ if the closed subspace $\mathcal{M}$ of $L^2(X)$ spanned by the orbit of $f$ is a finite-rank $L^\infty(X, \mathcal{D}, \mu)$ module.

Being a finite-rank $L^\infty(X, \mathcal{D}, \mu)$ module means that we can find $\phi_1, \ldots, \phi_d$ in $L^2(X)$ such that

$$\{\alpha^1 \phi_1 + \cdots + \alpha^d \phi_d : \alpha^1, \ldots, \alpha^d \in L^\infty(\mathcal{D})\}$$
is dense in \( \mathcal{M} \).

Denote by \( \mathcal{E}(X|\mathcal{D}) \) the closed subspace of \( L^2(X) \) spanned by the eigenfunctions over \( \mathcal{D} \). When \( Y \) is the trivial factor, write \( \mathcal{E}(X) \) for \( \mathcal{E}(X|\mathcal{D}) \).

We end this section with some examples of extensions with certain properties.

**Example 4.2.6.** Fix an amenable group \( G \). Let \( X_1 = (X_1, \mathcal{B}_1, \mu_1, T_1) \) be an almost periodic \( G \) system and let \( X_2 = (X_2, \mathcal{B}_2, \mu_2, T_2) \) be a weakly mixing \( G \) system. Then \( X_1 \times X_2 \to X_1 \) is weakly mixing extension and \( X_1 \times X_2 \to X_2 \) is a compact extension. This follows from the fact that functions of the form \( f_1 \otimes f_2 \) with \( f_i \in L^\infty(X_i) \) are dense in \( L^2(X_1 \times X_2) \), and that fact the conditional expectation on \( X_i \) is given by integration against \( \mu_{3-i} \).

**Example 4.2.7.** Let \( \phi : \mathbb{T} \to \mathbb{T} \) be measurable and let \( \alpha \) be irrational. Let \( X = (\mathbb{T}^2, \mathcal{B}, \mu) \) be the \( \mathbb{Z} \) system with \( \mathcal{B} \) the Borel \( \sigma \)-algebra, \( \mu \) the Lebesgue measure and \( T(x, y) = (x + \alpha, y + \phi(x)) \). Projection on the first coordinate gives a factor map \( \pi : X \to Y \) where \( Y = (Y, \mathcal{D}, \mu, S) \) is the \( \mathbb{Z} \) system with Lebesgue measure \( \mu \) and \( S(x) = x + \alpha \). The conditional expectation on \( \pi^{-1}\mathcal{D} \) is given by

\[
E(f|\pi^{-1}X)(x) = \int f(x, y) \, d\mu(y)
\]

for all \( f \) in \( L^2(X) \). Write \( \exp(x) \) for \( e^{2\pi ix} \). Fix \( p, q \in \mathbb{Z} \) with \( q \neq 0 \). We show that \( f(x, y) = \exp(px + qy) \) is not weakly mixing for \( \pi : X \to Y \). From

\[
T^n(x, y) = \left( x + n\alpha, y + \sum_{k=0}^{n-1} \phi(x + k\alpha) \right)
\]

we get

\[
(T^n\phi)(x, y) = \exp(pm\alpha) \exp(px) \exp(qy) \exp \left( q \sum_{k=0}^{n-1} \phi(x + k\alpha) \right)
\]

so \( |E(\phi \cdot T^n\phi|\pi^{-1}Y)|^2 = 1 \) almost surely.

For example, when \( \phi(x) = x \), we have the skew product \( T(x, y) = (x + \alpha, y + x) \). In this case the extension \( \pi : X \to Y \) is easily seen to be almost periodic.

### 4.3 The dichotomy

Fix again a system \( X = (X, \mathcal{B}, \mu, T) \) and a \( T \) invariant sub-\( \sigma \)-algebra \( \mathcal{D} \). Let \( \mathcal{C} \) be the sub-\( \sigma \)-algebra corresponding to \( \mathcal{A}(X|\mathcal{D}) \). Our first goal in this section is to relate \( \mathcal{A}(X|\mathcal{D}) \) and to \( \mathcal{A}(X \times_\mathcal{D} X|\mathcal{D}) \) by showing that any \( H \) in \( L^\infty(X \times_\mathcal{D} X) \) that is almost-periodic over \( \mathcal{D} \) satisfies

\[
\langle H, f_1 \otimes f_2 \rangle = \langle H, E(f_1|\mathcal{C}) \otimes E(f_2|\mathcal{C}) \rangle \tag{4.3.1}
\]

for any \( f_1, f_2 \) in \( L^\infty(X) \). This is similar to Proposition 4.4.4 in [McC99].
Proposition 4.3.2. For any $H$ in $L^\infty(X \times \varphi X)$ almost-periodic over $\mathcal{D}$ and any $f$ in $L^2(X)$ the element $f \ast H$ of $L^2(X)$ is almost-periodic over $\mathcal{D}$.

Proof. It suffices to prove this when $f$ is in $L^\infty(X)$. Fix $\varepsilon > 0$. We have to find a finite subset $\Xi$ of $L^\infty(X)$ and a subset $E$ of $X$ with $\mu(E) > 1 - \varepsilon$ such that

$$\min\{\|T^\gamma(f \ast H) - \xi\|_x : \xi \in \Xi\} \leq \varepsilon \tag{4.3.3}$$

for each $\gamma \in \Gamma$ and almost every $x \in E$. Almost-periodicity of $H$ over $\mathcal{D}$ implies the existence of a finite subset $\Psi$ of $L^\infty(X \times \varphi X)$ and a subset $F_1$ of $X$ with $\mu(F_1) > 1 - \varepsilon/16$ such that

$$\min\{\|(T^\gamma \times T^\gamma)H - \psi\|_x : \psi \in \Psi\} < \varepsilon/16$$

for all $\gamma \in \Gamma$ and almost all $x \in F_1$. Write $\Psi = \\{\psi_1, \ldots, \psi_k\}$. Let $\gamma_n$ be an enumeration of $\Gamma$. For each $1 \leq i \leq k$ and almost every $x \in F_1$ we have a compact operator $\psi_{i,x}$ on $L^2(X)$. Thus for each $1 \leq i \leq k$ and almost every $x$ we can find a positive integer $M_i(x)$ such that

$$\{(T^{\gamma_n}f_{\Gamma^\gamma_n,x}) \ast \psi_{i,x} : 1 \leq n \leq M_i(x)\}$$

is $\varepsilon/16$-dense in $\{(T^\gamma f) \ast \psi_{i,x} : \gamma \in \Gamma\}$. Each of the functions $M_i$ is measurable. Thus we can find $N$ in $\mathbb{N}$ so large that $F_2 = M_1^{-1}[1, N] \cap \cdots \cap M_k^{-1}[1, N]$ has measure at least $1 - \varepsilon/16$. Put

$$\Xi = \{(T^{\gamma_n}f) \ast \psi_i : 1 \leq i \leq k, 1 \leq n \leq N\}$$

and $E = F_1 \cap F_2$. Fix $\gamma$ in $\Gamma$ and almost any $x$ in $E$. We can choose $i$ such that

$$\|(T^\gamma \times T^\gamma)H - \psi_i\|_x \leq \varepsilon/16$$

and then guarantee

$$\|(T^\gamma f_{\Gamma^\gamma_n,x}) \ast \psi_{i,x} - (T^{\gamma_n}f_{\Gamma^\gamma_n,x}) \ast \psi_{i,x}\|_x \leq \varepsilon/16$$

holds for some $1 \leq n \leq N$. From

$$(T^\gamma(f \ast H))_x(x_2) = \int f(x_1)H(x_1, T^\gamma x_2) \, d\mu_1,\Gamma^\gamma x(x_1)$$

$$= \int f(T^\gamma x_1)H(T^\gamma x_1, T^\gamma x_2) \, d\mu_{1,x}(x_1)$$

we have

$$\|T^\gamma(f \ast H) - (T^{\gamma_n}f) \ast \psi_i\|_x$$

$$\leq \|T^\gamma(f \ast H) - (T^\gamma f) \ast \psi_i\|_x + \|(T^\gamma f) \ast \psi_i - (T^{\gamma_n}f) \ast \psi_i\|_x$$

$$\leq \|(T^\gamma f) \ast (T^\gamma \times T^\gamma)H - (T^\gamma f) \ast \psi_i\|_x + \|(T^\gamma f) \ast \psi_i - (T^{\gamma_n}f) \ast \psi_i\|_x$$

so (4.3.3) holds as desired. \qed
Proposition 4.3.4. For any $H$ in $L^\infty(X \times \mathcal{D} X)$ that is almost-periodic over $\mathcal{D}$ and any $f_1, f_2$ in $L^2(X)$ we have (4.3.1).

Proof. We have

$$\langle H, f_1 \otimes f_2 \rangle = \langle H, (f_1 - \mathbb{E}(f_1|\mathcal{C}) + \mathbb{E}(f_1|\mathcal{C})) \otimes f_2 \rangle$$

and a similar equality holds for $f_2$ so it suffices to prove that $\langle H, f_1 \otimes f_2 \rangle$ is zero when either $f_1$ or $f_2$ is orthogonal to $\mathcal{C}$. The two cases are similar. In the latter we have

$$\langle H, f_1 \otimes f_2 \rangle = \iint f_1(x_1)H(x_1, x_2)f_2(x_2)\,d(\mu_x \otimes \mu_x)(x_1, x_2)\,d\mu(y)$$

$$= \iint (f_1 * H)(x_2)f_2(x_2)\,d\mu_x(x_2)\,d\mu(x) = \langle f_1 * H, f_2 \rangle$$

which is zero by Proposition 4.3.2.

Corollary 4.3.5. Let $\mathcal{I}$ be the sub-$\sigma$-algebra of $T \times T$ invariant sets in $X \times \mathcal{D} X$. For any $f_1, f_2$ in $L^\infty(X)$ with either $f_1$ or $f_2$ orthogonal to $\mathcal{A}(X|\mathcal{D})$ we have $\mathbb{E}(f_1 \otimes f_2|\mathcal{I}) = 0$ in $L^2(X \times \mathcal{D} X)$.

Proof. For any $T \times T$ invariant function $H$ in $L^2(X \times \mathcal{D} X)$ we have

$$\langle \mathbb{E}(f_1 \otimes f_2|\mathcal{I}), H \rangle = \langle f_1 \otimes f_2, H \rangle = \langle \mathbb{E}(f_1|\mathcal{C}) \otimes \mathbb{E}(f_2|\mathcal{C}), H \rangle = 0$$

by (4.3.1), because invariant functions are certainly almost-periodic.

Theorem 4.3.6. $\mathcal{A}(X \times \mathcal{D} X|\mathcal{D}) = \mathcal{A}(X|\mathcal{D}) \otimes \mathcal{A}(X|\mathcal{D})$.

Proof. It is straightforward to check that if $f_1$ and $f_2$ in $L^\infty(X)$ are both almost-periodic over $\mathcal{D}$ then $f_1 \otimes f_2$ in $L^2(X \times \mathcal{D} X)$ is almost-periodic over $\mathcal{D}$. On the other hand, if $H$ belongs to $\mathcal{A}(X \times \mathcal{D} X|\mathcal{D})$ and is orthogonal to $\mathcal{A}(X|\mathcal{D}) \otimes \mathcal{A}(X|\mathcal{D})$ then by Proposition 4.3.4 we have $\langle H, f_1 \otimes f_2 \rangle = 0$ for all $f_1, f_2$ in $L^\infty(X)$ so $H = 0$.

We now turn to a description of the orthogonal complement of $\mathcal{A}(X|\mathcal{D})$ in $L^2(X)$. Fix a left Følner sequence $\Phi$ in $G$.

Definition 4.3.7. A function $f$ in $L^2(X)$ is weakly mixing over $\mathcal{D}$ if

$$\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int \|\mathbb{E}(\phi \cdot T^g f|\mathcal{D})\|^2 \,d\mu \,dm(g) = 0$$

for every $\phi$ in $L^\infty(X)$.

The set $\mathcal{W}(X|\mathcal{D})$ of weakly mixing functions is a closed, $T$ invariant subspace of $L^2(X)$. Proposition 4.3.2 lets us prove the following result.

Theorem 4.3.8. We have $L^2(X) = \mathcal{A}(X|\mathcal{D}) \oplus \mathcal{W}(X|\mathcal{D})$. 

45
Proof. First we show that if \( f \) in \( L^2(X) \) is orthogonal to \( A(X|D) \) then \( f \) belongs to \( W(X|D) \). Fix \( \phi \) in \( L^\infty(X) \). Let \( \mathcal{I} \) denote the sub-\( \sigma \)-algebra of \( T \times T \) invariant sets and put \( H = \mathbb{E}(\phi \otimes \phi|\mathcal{I}) \) in \( L^2(X \times \mathcal{I}, X) \). We have
\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int |\mathbb{E}(\phi \cdot T^g f|D)|^2 \, d\mu \, dm(g)
\]
\[
= \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int (\phi \otimes \phi) \cdot (T \times T)^g (f \otimes f) \, d\nu \, dm(g)
\]
\[
= \int \mathbb{E}(\phi \otimes \phi|\mathcal{I}) \cdot (f \otimes f) \, d\nu = \langle H, f \otimes f \rangle = \langle f \ast H, f \rangle = 0
\]
by the mean ergodic theorem and Proposition 4.3.2. Since \( \phi \) was arbitrary, \( f \) is weakly mixing over \( D \).

Now we show that \( A(X|D) \) and \( W(X|D) \) are orthogonal. Fix \( f \) in \( W(X|D) \). It suffices to prove that \( f \) is orthogonal to any \( \phi \) in \( L^\infty(X) \) that is almost-periodic over \( D \). Fix \( \varepsilon > 0 \). Since \( \phi \) is almost periodic we can find a subset \( E \) of \( X \) with \( \mu(E) > 1 - \varepsilon \) and a finite subset \( \Xi = \{\xi_1, \ldots, \xi_k\} \) of \( L^2(X) \) such that (4.2.2) holds for all \( \gamma \in \Gamma \) and all \( x \in E \). Fix \( g \in G \). Since \( \Gamma \) is dense in \( G \) we can find some \( \gamma \) in \( \Gamma \) such that \( \|T^\gamma g - \gamma^\gamma g\|_X < \varepsilon \) for all \( x \) in a subset \( E_g \) of \( X \) with \( \mu(E_g) > 1 - \varepsilon \). For each \( x \) in \( E \) choose \( 1 \leq i(x) \leq k \) so that \( \|T^\gamma g - \xi_{i(x)}\|_X < \varepsilon \). Put \( F = E \cap E_g \). Cauchy-Schwarz gives
\[
\int T^\gamma f \cdot T^g f \, d\mu_x \leq \left| \int \xi_{i(x)} \cdot T^g f \, d\mu_x \right| + 2\varepsilon \|T^g f\|_x
\]
\[
\leq \sum_{i=1}^k |\mathbb{E}(\xi_i \cdot T^g f|D)(x)| + 2\varepsilon \|T^g f\|_x
\]
for any \( x \in F \). Combining this with
\[
\int T^\gamma f \cdot T^g f \, d\mu_x \leq \int |T^g f| \, d\mu_x \|\phi\|_\infty \leq \|T^g f\|_x \|\phi\|_\infty
\]
which holds (in particular) for almost-every \( x \notin F \) we get
\[
|\langle \phi, f \rangle| \leq \sum_{i=1}^k \int |\mathbb{E}(\xi_i \cdot T^g f|D)| \, d\mu + 2\varepsilon \|T^g f\| + \int 1_{Y \setminus F}(x) \cdot \|T^g f\|_x \, d\mu(x) \|\phi\|_\infty
\]
\[
\leq \sum_{i=1}^k \int |\mathbb{E}(\xi_i \cdot T^g f|D)| \, d\mu + 2\varepsilon \|f\| + \sqrt{2\varepsilon} \|f\| \cdot \|\phi\|_\infty
\]
by integrating, applying Cauchy-Schwarz, and noting that \( \mu(X \setminus F) \leq 2\varepsilon \). Finally, averaging over the Følner sequence \( \Phi \) and applying Cauchy-Schwarz once more gives
\[
|\langle \phi, f \rangle| \leq 2\varepsilon \|f\| + \sqrt{2\varepsilon} \|f\| \|\phi\|_\infty + \sum_{i=1}^k \left( \frac{1}{m(\Phi_N)} \int \int |\mathbb{E}(\xi_i \cdot T^g f|D)|^2 \, d\mu \, dm(g) \right)^{1/2}
\]
which, upon using the fact that \( f \) is weakly-mixing and noting that \( \varepsilon \) was arbitrary, gives
\[ \langle \phi, f \rangle = 0. \]
Since the definition of $\mathcal{A}(X|\mathcal{D})$ does not depend on the choice of a left Følner sequence $\Phi$, the above proposition implies that $\mathcal{W}(X|\mathcal{D})$ is also independent of $\Phi$. Moreover, since what it means to be a weakly mixing function does not depend on a choice of disintegration, neither does the space $\mathcal{A}(X|\mathcal{D})$, and $\mathcal{A}(X|\mathcal{D}) = \mathcal{A}(X|\mathcal{E})$ whenever $\mathcal{D}$ and $\mathcal{E}$ are equal up to sets of measure zero.

We conclude this section by using Theorem 4.3.8 to show that $\mathcal{A}(X|\mathcal{D}) = \mathcal{E}(X|\mathcal{D})$

**Theorem 4.3.11.** $L^2(X) = \mathcal{E}(X|\mathcal{D}) \oplus \mathcal{W}(X|\mathcal{D})$.

**Proof.** Using the fact that the orbit of an eigenfunction is contained in a $L^\infty(X,\mathcal{D},\mu)$ module having finite rank, one can show that every eigenfunction over $\mathcal{D}$ is almost-periodic over $\mathcal{D}$. Thus $\mathcal{E}(X|\mathcal{D}) \subset \mathcal{W}(X|\mathcal{D})^\perp$ by Theorem 4.3.8.

It remains to prove that $\mathcal{E}(X|\mathcal{D})^\perp \subset \mathcal{W}(X|\mathcal{D})$. Fix $f$ in $L^2(X)$ orthogonal to $\mathcal{E}(X|\mathcal{D})$. For any $\phi$ in $L^\infty(X)$ we have

$$\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int |E(\phi \cdot T^g f|\mathcal{D})|^2 \, d\mu \, dm(g) = \langle f \cdot H, f \rangle$$

as in the proof of Theorem 4.3.8, where $H = E(\phi \otimes \phi|\mathcal{D})$. Thus it suffices to prove that $f \cdot H$ is in $\mathcal{E}(X|\mathcal{D})$. Let $\Psi_n$ and $\lambda_n$ be as in Theorem 4.1.3. Since $\{\Psi_{n,x} : n \in \mathbb{N}\}$ spans the image of $H_x$ for almost-every $x$, it suffices to prove that each $\Psi_n$ is in $\mathcal{E}(X|\mathcal{D})$. To this end, fix $n$ in $\mathbb{N}$ and denote by $\theta(x)$ the multiplicity of the eigenvalue $\lambda_n(x)$ if $\lambda_n(x)$ is defined, and put $\theta(x) = 0$ otherwise. Each of the functions $\lambda_m$ is measurable, so $\theta$ is too. For each $k \in \mathbb{N}$ let $\Omega_k = \theta^{-1}(k)$.

We will show that the orbit of $1_{\Omega_k}(x)\Psi_n$ is a finite-rank $L^\infty(X,\mathcal{D},\mu)$ module. Fix $\gamma$ in $\Gamma$. We have $T^\gamma H_{T^\gamma x} = H_x T^\gamma$ for almost-every $x$ because $H$ is $T \times T$ invariant. Since $T^\gamma$ is unitary on almost every fiber, the operators $H_x$ and $H_{T^\gamma x}$ have the same spectrum, so each of the functions $\lambda_m$ is $T^\gamma$ invariant. This implies $\Omega_k$ is $T^\gamma$-invariant. Also

$$(H \ast T^\gamma \Phi_n)_x = (T^\gamma (H \ast \Phi_n))_x = T^\gamma (H_{T^\gamma x} \ast \Phi_{n,T^\gamma x}) = \lambda_n(x)(T^\gamma \Psi_n)_x$$

so in almost every fiber, the dimension the $\Gamma$-orbit of the square-integrable section $x \mapsto 1_{\Omega_k}(x)\Psi_{n,x}$ of $X \ast \mathcal{D}$ is bounded by $k$. Thus $x \mapsto 1_{\Omega_k}(x)\Psi_{n,x}$ corresponds to an eigenfunction. Summing over $k$ proves that $\Psi_n$ is in $\mathcal{E}(X|\mathcal{D})$ as desired. \qed

Combining Theorems 4.3.8 and 4.3.11 yields the following result, a basic version of which will be used later.

**Corollary 4.3.12.** $\mathcal{A}(X|\mathcal{D}) = \mathcal{E}(X|\mathcal{D})$. 47
4.4 Characteristic factors for amenable averages

In this section we define the characteristic factors $\mathcal{C}_{k,i}$ associated to commuting, measurable actions $T_1, \ldots, T_k$ and prove that

$$\lim_{N \to \infty} \frac{1}{\mu(\Phi_N)} \int_{\Phi_N} \prod_{i=1}^{k} T_{g_i}^j f_i - \prod_{i=1}^{k} T_{g_i}^j \mathbb{E}(f_i|\mathcal{C}_{k,i}) \, dm(g) = 0$$

(4.4.1)

in $L^2(X, \mathcal{B}, \mu)$ for any $f_1, \ldots, f_k$ in $L^\infty(X, \mathcal{B}, \mu)$.

To define the $\mathcal{C}_{k,i}$ fix $k$ in $\mathbb{N}$ and let $T_1, \ldots, T_k$ be commuting, measurable actions of $G$ on a probability space $(X, \mathcal{B}, \mu)$. Let $\mathcal{C}_{1,1}$ be the sub-$\sigma$-algebra of $T_1$ invariant sets. It is invariant under all of the actions $T_2, \ldots, T_k$ because they each commute with $T_1$. Suppose by induction that for some $1 \leq l \leq k-1$ we have defined sub-$\sigma$-algebras $\mathcal{C}_{l,1}, \ldots, \mathcal{C}_{l,l}$ such that

1. for each $1 \leq j \leq l$, $\mathcal{C}_{l,j}$ is $T_l \cdots T_j$ invariant;
2. for each $1 \leq j \leq l$ and every $l+1 \leq i \leq k$, $\mathcal{C}_{l,j}$ is $T_i$ invariant.

For each $1 \leq j \leq l$ form the $G$ system $X_j = (X, \mathcal{B}, \mu, T_{l+1} \cdots T_j)$ and let $\mathcal{C}_{l+1,j}$ be the sub-$\sigma$-algebra of $\mathcal{B}$ corresponding to $\mathcal{A}(X_j|\mathcal{C}_{l,j})$. It is invariant under $T_{l+1} \cdots T_j$ because it consists of $T_{l+1} \cdots T_j$ almost periodic functions, and (if $l < k-1$) it is $T_l$ invariant for all $l+2 \leq i \leq k$ because the actions commute. To define $\mathcal{C}_{l+1,l+1}$ form the $G$ system $X_{l+1} = (X, \mathcal{B}, \mu, T_{l+1})$ and let $\mathcal{C}_{l+1,l+1}$ be the sub-$\sigma$-algebra corresponding to $\mathcal{A}(X_{l+1}|\mathcal{C}_{l,1} \lor \cdots \lor \mathcal{C}_{l,l})$. It is $T_{l+1}$ invariant because it consists of the $T_{l+1}$ almost-periodic functions over $Y_{l+1}$, and (if $l < k-1$) it is $T_l$ invariant for all $l+2 \leq i \leq k$ because the actions commute. This concludes the inductive construction. Figure 4.1 shows how the $\mathcal{C}_{k,i}$ are related for $k \leq 4$. The remainder of this section constitutes a proof of the following theorem.

**Theorem 4.4.2.** Let $X$ be a $G^k$ system Then

$$\lim_{N \to \infty} \frac{1}{\mu(\Phi_N)} \int_{\Phi_N} \prod_{i=1}^{k} T_{g_i}^j f_i - \prod_{i=1}^{k} T_{g_i}^j \mathbb{E}(f_i|\mathcal{C}_{k,i}) \, dm(g) = 0$$

(4.4.2)

for any $f_i$ in $L^\infty(X)$.

Since the limit

$$\lim_{N \to \infty} \frac{1}{\mu(\Phi_N)} \int_{\Phi_N} \prod_{i=1}^{k} T_{g_i}^j f_i \, dm(g)$$

48
is known to exist (see [ZK13]) it suffices to prove that
\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f_{k+1} \cdot \prod_{i=1}^{k} T_i^\theta \cdot \int_{\Phi_N} f_i \ d\mu \ dm(g)
\]
\[
= \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f_{k+1} \cdot \prod_{i=1}^{k} T_i^\theta \cdot \int_{\Phi_N} f_i \mathbb{E}(f_i|\mathcal{C}_{k,i}) \ d\mu \ dm(g)
\] (4.4.3)
for any \( f_1, \ldots, f_{k+1} \) in \( L^\infty(X, \mathcal{B}, \mu) \). We will prove (4.4.3) by induction on \( k \). The case \( k = 1 \) follows from the mean ergodic theorem: we have
\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int T_1^\theta f_1 \cdot f_2 \ d\mu \ dm(g) = \int \mathbb{E}(f_1|\mathcal{C}_{1,1}) \cdot f_2 \ d\mu
\]
by Theorem 2.3.6, which can be re-written as
\[
\int f_1 \otimes f_2 \ d\nu_1 = \int \mathbb{E}(f_1|\mathcal{C}_{1,1}) \otimes f_2 \ d\nu_1
\]
where \( \nu_1 \) is the Furstenberg joining for the action \( T_1 \). For the inductive step we need the following application of the van der Corput trick, which is a version of Lemma 4.7 in [Aus10].

**Theorem 4.4.4.** Let \( T_1, \ldots, T_k \) be commuting, measurable actions of \( G \) on a compact, metric probability space \( (X, \mathcal{B}, \mu) \). Let \( \nu_{k-1} \) be the Furstenberg joining of the actions \( T_1, \ldots, T_{k-1} \) and let \( \nu_k \) be the Furstenberg joining of the actions \( T_1, \ldots, T_k \). Suppose we have sub-\( \sigma \)-algebras \( \mathcal{E}_1, \ldots, \mathcal{E}_{k-1} \) with each \( \mathcal{E}_i \) invariant under \( T_{k-1} \cdots T_i \) and \( T_k \) such that
\[
\int f_1 \otimes \cdots \otimes f_k \ d\nu_{k-1} = \int \mathbb{E}(f_1|\mathcal{E}_1) \otimes \cdots \otimes \mathbb{E}(f_{k-1}|\mathcal{E}_{k-1}) \otimes f_k \ d\nu_{k-1}
\]
for all \( f_1, \ldots, f_k \) in \( L^\infty(X, \mathcal{B}, \mu) \). Put \( \mathcal{E}_k = \mathcal{E}_1 \lor \cdots \lor \mathcal{E}_{k-1} \). Then
\[
\int f_1 \otimes \cdots \otimes f_{k+1} \ d\nu_k = \int \mathbb{E}(f_1|\mathcal{F}_1) \otimes \cdots \otimes \mathbb{E}(f_k|\mathcal{F}_k) \otimes f_{k+1} \ d\nu_k
\] (4.4.5)

\[
\begin{array}{cccc}
\mathcal{C}_{4,1} & \mathcal{C}_{4,2} & \mathcal{C}_{4,3} & \mathcal{C}_{4,4} \\
T_4 T_3 T_2 T_1 & T_4 T_3 T_2 & T_4 T_3 & T_4 \\
\mathcal{C}_{3,1} & \mathcal{C}_{3,2} & \mathcal{C}_{3,3} & \mathcal{C}_{3} \\
T_3 T_2 T_1 & T_3 T_2 & T_3 & \mathcal{C}_{2} \\
\mathcal{C}_{2,1} & \mathcal{C}_{2,2} & \mathcal{C}_{2} \\
T_2 T_1 & T_2 & \mathcal{C}_{1} \\
\mathcal{C}_{1,1} & \mathcal{C}_{1} \\
\end{array}
\]

Figure 4.1: The sub-\( \sigma \)-algebras \( \mathcal{C}_{k,i} \) for \( k \leq 4 \). A line indicates that the upper \( \sigma \)-algebra corresponds to the functions almost-periodic for the labeled action over the lower \( \sigma \)-algebra.
for all \( f_1, \ldots, f_{k+1} \) in \( L^\infty(X, \mathcal{B}, \mu) \) where, for each \( 1 \leq i \leq k \) the sub-\( \sigma \)-algebra \( \mathcal{F}_i \) corresponds to the functions that are \( T_k \cdots T_1 \) almost-periodic over \( \mathcal{E}_i \).

**Proof.** Fix \( f_1, \ldots, f_{k+1} \) in \( L^\infty(X, \mathcal{B}, \mu) \) with \( \|f_i\|_\infty \leq 1 \) for all \( 1 \leq i \leq k + 1 \). Since \( \mathcal{E}_i \) is \( T_{k-1} \cdots T_1 \) invariant and contained in \( \mathcal{E}_k \) for each \( 1 \leq i \leq k - 1 \), we have

\[
T_{k-1}^\theta \cdots T_i^\theta \mathbb{E}(f_i|\mathcal{E}_i) = \mathbb{E}(T_{k-1}^\theta \cdots T_i^\theta \mathbb{E}(f_i|\mathcal{E}_i)|\mathcal{E}_k)
\]

for each \( 1 \leq i \leq k - 1 \). Thus we can re-write our assumption as

\[
\int f_1 \otimes \cdots \otimes f_k \, d\nu_{k-1} = \int \mathbb{E}(f_1|\mathcal{E}_1) \otimes \cdots \otimes \mathbb{E}(f_k|\mathcal{E}_k) \, d\nu_{k-1}
\]

using (2.5.3). We proceed by applying the van der Corput trick to the sequence

\[
u(g) = \prod_{i=1}^k T_{k-1}^\theta \cdots T_i^\theta f_i
\]

in \( L^2(X, \mathcal{B}, \mu) \). From (2.5.3) and (4.4.6) we see that

\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \langle u(hg), u(lg) \rangle \, dm(g) = \int \bigotimes_{i=1}^k (T_{k-1}^h \cdots T_i^h f_i \cdot T_{k-1}^l \cdots T_i^l f_i) \, d\nu_{k-1}
\]

\[
= \int \bigotimes_{i=1}^k \mathbb{E}(T_{k-1}^h \cdots T_i^h f_i \cdot T_{k-1}^l \cdots T_i^l f_i|\mathcal{E}_i) \, d\nu_{k-1}
\]

for any \( h, l \in G \). Using (2.5.3) once more yields

\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int \langle u(hg), u(lg) \rangle \, dm(g) \leq \| \mathbb{E}(T_{k-1}^h \cdots T_i^h f_i \cdot T_{k-1}^l \cdots T_i^l f_i|\mathcal{E}_i) \|
\]

for each \( 1 \leq i \leq k \), the norm taken in \( L^2(X, \mathcal{B}, \mu) \). Put \( X_i = (X, \mathcal{B}, \mu, T_k \cdots T_1) \). Let \( \mathcal{I}_i \) be the sub-\( \sigma \)-algebra of \( T_k \cdots T_i \times T_k \cdots T_i \) invariant sets in the relatively independent joining \( X_i \times_{\mathcal{E}_i} X_i \). We have

\[
\limsup_{H \to \infty} \frac{1}{m(\Phi_H)^2} \int \int \| \mathbb{E}(T_{k-1}^h \cdots T_i^h f_i \cdot T_{k-1}^l \cdots T_i^l f_i|\mathcal{E}_i) \| \, dm(h) \, dm(l)
\]

\[
\leq \lim_{H \to \infty} \left\| \frac{1}{m(\Phi_H)} \int (T_k \cdots T_i \times T_k \cdots T_i)^h (f_i \otimes f_i) \, dm(h) \right\| = \| \mathbb{E}(f_i \otimes f_i|\mathcal{I}_i) \|
\]

in \( L^2(X_i \times_{\mathcal{E}_i} X_i) \) by Cauchy-Schwarz and the mean ergodic theorem. By Corollary 4.3.5 the conditional expectation \( \mathbb{E}(f_i \otimes f_i|\mathcal{I}_i) \) will be zero if \( f_i \) is orthogonal to \( \mathcal{A}(X_i|\mathcal{E}_i) \). Since \( 1 \leq i \leq k \) was arbitrary, (4.4.5) follows from the van der Corput trick.

\( \square \)
Taking $e_i = e_{k-1,i}$ in the preceding theorem proves (4.4.3) and concludes the proof of Theorem 4.4.2. We conclude this section with another application of the van der Corput trick that is sometimes useful.

**Theorem 4.4.7.** Let $T_1, \ldots, T_k$ be commuting, measurable actions of $G$ on a separated, countably generated probability space $(X, \mathcal{B}, \mu)$. Let $\nu_{k-1}$ be the Furstenberg joining of the actions $T_1, \ldots, T_{k-1}$ and let $\nu_k$ be the Furstenberg joining of the actions $T_1, \ldots, T_k$. Let $\mathcal{I}_k$ denote the sub-$\sigma$-algebra of $\mathcal{B}^k$ consisting of

$$T_k T_{k-1} \cdots T_1 \times \cdots \times T_k T_{k-1} \times T_k$$

invariant sets. If $E(\mu | \mathcal{I}_k) = 0$ in $L^2(X^k, \mathcal{B}^k, \nu_{k-1})$ for some $f_1, \ldots, f_k$ in $L^\infty(X^k, \mathcal{B}^k, \mu)$ then

$$\int f_1 \otimes \cdots \otimes f_k \otimes f_{k+1} \, d\nu_k = 0$$

for all $f_{k+1}$ in $L^2(X, \mathcal{B}, \mu)$.

**Proof.** Fix $f_1, \ldots, f_k$ in $L^\infty(X, \mathcal{B}, \mu)$ satisfying $E(f_1 \otimes \cdots \otimes f_k | \mathcal{I}_k) = 0$. Applying the van der Corput trick as in Theorem 4.4.4 gives

$$\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int_{\Phi_N} \langle u(hg), u(lg) \rangle \, dm(g) = \int \left( T_{k}^{h} \cdots T_{i}^{h} f_{i} \cdot T_{l}^{k} \cdots T_{i}^{l} f_{i} \right) \, d\nu_{k-1}$$

for any $h, l \in G$. From this we get

$$\lim_{H \to \infty} \frac{1}{m(\Phi_H)^2} \int \int \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int_{\Phi_N} \langle u(hg), u(lg) \rangle \, dm(g) \, dm(h) \, dm(l)$$

$$= \lim_{H \to \infty} \frac{1}{m(\Phi_H)} \int \left( T_{k}^{h} \cdots T_{i}^{h} f_{i} \, dm(h) \right)^2 = \int E(f_1 \otimes \cdots \otimes f_k | \mathcal{I}_k)^2 \, d\nu_{k-1}$$

where the norm is determined by $\nu_{k-1}$ and the last equality follows from Theorem 2.3.6. The conclusion follows from the van der Corput trick and the fact that strong convergence implies weak convergence.

\[\square\]

### 4.5 Lifting positivity

In this section we prove a technical result, based on Theorem 9.1 in [FKO82], that allows us to lift multiple recurrence from one level of Figure 4.1 to the next, provided the sub-$\sigma$-algebras in the lower level are all equal. In the next section we will use this to prove some multiple recurrence results.
Theorem 4.5.1. Let \( T_1, \ldots, T_k \) be commuting, measurable actions of \( G \) on a separated, countably-generated probability space \((X, \mathcal{B}, \mu)\). Let \( \mathcal{D} \) be a sub-\( \sigma \)-algebra that is \( T_k \cdots T_i \) invariant for all \( 1 \leq i \leq k \). Suppose that
\[
\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f \prod_{i=1}^{k} T_{k}^g \cdots T_{i}^g \mathbb{E}(f|\mathcal{D}) \, d\mu \, dm(g) > 0
\]
for any \( f > 0 \) in \( L^\infty(X, \mathcal{B}, \mu) \). For each \( 1 \leq i \leq k \), let \( E_i \) be a sub-\( \sigma \)-algebra of \( \mathcal{B} \) that is \( T_k \cdots T_i \) invariant, and suppose that \( E_i \to \mathcal{D} \) is \( T_k \cdots T_i \) almost-periodic. Then
\[
\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f \cdot \prod_{i=1}^{k} T_{k}^g \cdots T_{i}^g \mathbb{E}(f|E_i) \, d\mu \, dm(g) > 0
\]
for any \( f > 0 \) in \( L^\infty(X, \mathcal{D}, \mu) \).

Proof. It suffices to prove the theorem when \( f = 1_B \) for some set \( B \in \mathcal{B} \) having positive measure. Let \( \mu_x \) be the disintegration of \( \mu \) over \( \mathcal{D} \). For each \( 1 \leq i \leq k \) write \( f_i \) for \( \mathbb{E}(1_B|E_i) \).

From
\[
\mu(B \cap \{f_i = 0\}) = \int f_i \cdot 1_{\{f_i = 0\}} \, d\mu = 0
\]
it follows that \( f_i \) is positive on almost all of \( B \). Thus we can find a set \( D_1 \) in \( \mathcal{D} \) with positive measure and some \( \alpha > 0 \) such that
\[
\int f \cdot f_1 \cdots f_k \, d\mu_x > \alpha
\]
for all \( x \) in \( D_1 \). Fix \( \varepsilon = \alpha/4k \).

For any \( x \in X \) and any non-empty subset \( F \) of \( \Gamma \) define
\[
\mathcal{L}(x, F) = \{ (T_k^a \cdots T_i^a f_1, \ldots, T_k^a f_k) : a \in F \} \subset L^2(X, \mathcal{B}, \mu_x)^k
\]
and equip it with the max norm coming from \( \| \cdot \|_x \) on the constituents.

Claim. There is a subset \( D_2 \) of \( D_1 \) with positive measure such that \( \mathcal{L}(x, \Gamma) \) is totally bounded for each \( x \in D_2 \).

Proof. For each \( j \) in \( \mathbb{N} \) put \( \varepsilon_j = \mu(D_1)/2^{j+k+1} \). Since each \( f_i \) is \( T_k \cdots T_i \) almost-periodic over \( \mathcal{D} \) one can find finite subsets \( \Xi_j^1 \) of \( L^\infty(X, \mathcal{B}, \mu) \) and subsets \( E_j^1 \) with measure at least \( 1 - \varepsilon_j \) such that for each \( \gamma \) in \( \Gamma \) we have
\[
\min\{\|T_k^\gamma \cdots T_i^\gamma f_i - \xi\|_x : \xi \in \Xi_j^1 \} < \varepsilon_j
\]
for every \( x \) in \( E_j^1 \). Put
\[
D_2 = D_1 \setminus \bigcup_{j=1}^{\infty} E_j^1 \cup \cdots \cup E_j^k
\]
and note that \( \mu(D_2) \geq \mu(D_1)/2 \).
We will be interested in separated subsets of $\mathcal{L}(x, \Gamma)$ so define
\[
\text{Sep}(F, t) := \bigcap_{a \neq b} \bigcap_{k} \bigcup_{a \in F, b \in F} \{ x \in X : \|T_k^a \cdots T_i^a f_i - T_k^b \cdots T_i^b f_i\|_x > t \}
\]
for any finite, non-empty subset $F$ of $\Gamma$ and any positive $t$. It belongs to $\mathcal{D}$ and when $F$ is a singleton it is all of $X$. The fact that $\mathcal{L}(x, \Gamma)$ is totally bounded whenever $x \in D_2$ implies that there is a bound on the cardinality of the finite sets $F$ for which $x$ belongs to $\text{Sep}(F, \varepsilon)$. Thus the $\mathcal{D}$ measurable sets
\[
Q(F) := \text{Sep}(F, \varepsilon) \setminus \bigcup \{ \text{Sep}(E, \varepsilon) : E \subset \Gamma \text{ with } |F| < |E| < \infty \}
\]
cover almost all of $D_2$ as $F$ runs through the finite subsets of $\Gamma$ and we can fix a finite, non-empty subset $F$ of $\Gamma$ such that $Q(F) \cap D_2$ has positive measure. For each $x$ in $Q(F) \cap D_2$ we can find some $n \in \mathbb{N}$ with the property that $x \in \text{Sep}(F, \varepsilon + 1/n)$ because of the strict inequalities and finite number of conditions in the definition of $\text{Sep}(F, \varepsilon)$. Thus we can find some $\eta > 0$ with the property that $Q(F) \cap \text{Sep}(F, \varepsilon + \eta) \cap D_2$ has positive measure. Define a function $\Psi$ by
\[
\Psi : Q(F) \cap \text{Sep}(F, \varepsilon + \eta) \cap D_2 \to [0, 2]^{F \times F \times \{1, \ldots, k\}}
\]
\[
\Psi(x) : (a, b, i) \mapsto \|T_k^a \cdots T_i^a f_i - T_k^b \cdots T_i^b f_i\|_x
\]
and partition $[0, 2]^{F \times F \times \{1, \ldots, k\}}$ into cubes of side length $\eta/2$. Since $\Psi$ is measurable we can find a cell $D$ in the pull-back partition that has positive measure. Now $D$ belongs to $\mathcal{D}$, so by hypothesis
\[
\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int \prod_{i=1}^k T_k^a \cdots T_i^a 1_D \cdot 1_D \, d\mu \, dm(g) > 0
\]
and thus there exists $\zeta > 0$ and a subset $\Delta$ of $G$ with positive lower density such that
\[
\int \prod_{i=1}^k T_k^a \cdots T_i^a 1_D \cdot 1_D \, d\mu > \zeta
\]
for any $g$ in $\Delta$.

**Claim.** For any $g \in \Delta$ there is a subset $E_g$ of $(T_k^g \cdots T_i^g)^{-1} D \cap \cdots \cap (T_k^g)^{-1} D \cap D$ with measure at least $\zeta/2$ such that for any $x \in E_g$ one can find $b \in F$ satisfying $\|T_k^b \cdots T_i^b f_i - f_i\|_x < 2\varepsilon$ for every $1 \leq i \leq k$.

**Proof.** Fix $g \in \Delta$. Since $\Gamma$ is dense and each $T_i^g$ is unitary we can find $\gamma$ in $\Gamma$ such that
\[
\|T_k^a \cdots T_i^a f_i - T_k^a \cdots T_i^a \gamma f_i\|^2 \leq \min\{\eta^2 \zeta / 2^{2k+2} |F|, \varepsilon^2 \}
\]
(4.5.3)
for all $a \in F$ and all $1 \leq i \leq k$. It follows from Chebyshev’s inequality that there is a subset $E_g$ of $(T_k^g \cdots T_1^g)^{-1} D \cap \cdots \cap (T_k^g)^{-1} D \cap D$ with $\mu(E_g) \geq \zeta/2$ such that
\[
\|T_k^g \cdots T_i^g f_i - T_k^\alpha \cdots T_i^\alpha f_i\|_x \leq \eta/4
\]
(4.5.4)
for all $a \in F$, all $1 \leq i \leq k$, and all $x \in E_g$.

If $1 \in F^\gamma$ then the claim follows immediately from (4.5.3), so assume otherwise. In this case the subset $F^\gamma \cup \{1\}$ of the subgroup $\Gamma$ has cardinality strictly larger than $F$ so $x$ does not belong to $\text{Sep}(F^\gamma \cup \{1\}, \varepsilon)$. Thus we can find $\alpha \neq \beta$ in $F^\gamma \cup \{1\}$ such that
\[
|T_k^\alpha \cdots T_i^\alpha f_i - T_k^\beta \cdots T_i^\beta f_i|_x \leq \varepsilon
\]
(4.5.5)
for all $1 \leq i \leq k$, and the proof will be concluded if we can show that one of $\alpha$ or $\beta$ must be 1. Fix $a \neq b$ in $F$. (If $|F| = 1$ then one of $\alpha$ or $\beta$ must be 1.) That $x$ belongs to $\text{Sep}(F, \varepsilon + \eta)$ tells us
\[
\|T_k^a \cdots T_i^a f_i - T_k^b \cdots T_i^b f_i\|_x > \varepsilon + \eta
\]
holds for some $1 \leq i \leq k$. Since $T_k^g \cdots T_i^g x$ belongs to $D$ we must have
\[
\|T_k^a \cdots T_i^a f_i - T_k^b \cdots T_i^b f_i\|_x \geq \varepsilon + \eta/2
\]
because the function $x \mapsto \Psi(x)(a, b, i)$ takes values in an interval of length at most $\eta/2$. Now $a \neq b$ in $F$ were arbitrary so, combined with (4.5.4), this forces one of $\alpha$ or $\beta$ to be 1 as otherwise (4.5.5) is contradicted. \hfill \square

We can now finish the proof. Fix $g \in \Delta$ and let $E_g$ be as in the claim. For any $x \in E_g$ we can find some $b \in F$ such that $\|T_k^b \cdots T_i^b f_i - f_i\|_x \leq 2\varepsilon$ for all $1 \leq i \leq k$. Thus
\[
\int f \cdot \prod_{i=1}^k T_k^b \cdots T_i^b f_i \, d\mu_x \geq \alpha - 2k\varepsilon = \frac{\alpha}{2}
\]
for any $x$ in the subset $E_g$ of $D$. Summing over $b \in F$ on the left hand side weakens the inequality and removes the dependence of $b$ on $x$ and $g$. This allows us to integrate over $E_g$, obtaining
\[
\sum_{b \in F} \int f \cdot \prod_{i=1}^k T_k^b \cdots T_i^b f_i \, d\mu \geq \frac{\zeta \alpha}{4}
\]
which, after averaging over $\Delta$ using the Følner sequence $\Phi$, gives
\[
\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int f \cdot \prod_{i=1}^k T_i^g \cdots T_i^g f_i \, d\mu \geq \liminf_{N \to \infty} \frac{m(\Delta \cap \Phi_N)}{m(\Phi_N)} \cdot \frac{\zeta \alpha}{4|F|}
\]
concluding the proof. \hfill \square
4.6 Recurrence results

Bergelson conjectured in [Ber96] that for any $G^k$ system $X$ we have

$$\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int 1_B \cdot \prod_{i=1}^{k} T_k \cdots T_i \cdot T_i^g 1_B \, d\mu \, dm(g) > 0 \quad (4.6.1)$$

for every $B$ in $\mathcal{B}$ with positive measure. In this section we verify this conjecture when $k = 2$ without additional assumptions, and when $k = 3$ assuming $T_1, T_2$ and $T_2T_1$ are ergodic. The $k = 2$ case was previously obtained for countable, amenable groups in [BMZ97], and the conjecture was verified for arbitrary $k$ by Austin [Aus13].

**Theorem 4.6.2.** Let $T_1, T_2$ be commuting, measurable actions of $G$ on a separated, countably generated probability space $(X, \mathcal{B}, \mu)$. Then

$$\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f \cdot T_2^g T_1^g f \cdot T_2^g f \, d\mu \, dm(g) > 0 \quad (4.6.3)$$

for any $f > 0$ in $L^\infty(X, \mathcal{B}, \mu)$.

**Proof.** By Theorem 4.4.2 it suffices to prove that

$$\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int \int f \cdot T_2^g T_1^g \mathcal{E}(f|\mathcal{E}_{2,1}) \cdot T_2^g \mathcal{E}(f|\mathcal{E}_{2,2}) \, d\mu \, dm(g) > 0 \quad (4.6.4)$$

for all $f > 0$ in $L^\infty(X, \mathcal{B}, \mu)$. If $f$ is of the form $1_B$ for some $B \in \mathcal{E}_1$ with $\mu(B) > 0$ then $T_2T_1 f : T_2 f = T_2 f$ so in this case (4.6.3) follows from the mean ergodic theorem. Thus we have (4.6.3) whenever $f$ is $\mathcal{E}_1$ measurable. Applying Theorem 4.5.1 with $\mathcal{D} = \mathcal{E}_1$, $\mathcal{E}_1 = \mathcal{E}_{2,1}$ and $\mathcal{E}_2 = \mathcal{E}_{2,2}$ yields (4.6.4). $\square$

When $k = 3$ we cannot use Theorem 4.5.1 to prove (4.6.1) because the sub-$\sigma$-algebras $\mathcal{E}_{2,1}$ and $\mathcal{E}_{2,2}$ need not agree and because the behavior of $T_3$ with respect to the extensions $\mathcal{E}_{2,i} \to \mathcal{E}_{1,1}$ is unknown. However, if $T_1$ is ergodic then $\mathcal{E}_{1,1}$ is trivial and the sub-$\sigma$-algebras $\mathcal{E}_{2,1}$ and $\mathcal{E}_{2,2}$ consist of functions that are almost-periodic for $T_2T_1$ and $T_2$ respectively over the trivial factor. We will prove below that if $T_2$ and $T_2T_1$ are ergodic any function almost-periodic for $T_2$ or $T_2T_1$ over the trivial factor is necessarily almost periodic for $T_3$ over the trivial factor. This leads to a description of characteristic factors that allow us, under the aforementioned ergodicity assumptions, to prove Bergelson’s conjecture when $k = 3$.

Given a system $X$, recall that $f$ in $L^2(X)$ is an eigenfunction of $X$ if its $T$-orbit is contained in a $T$-invariant, finite-dimensional subspace of $L^2(X)$. In other words $f$ is an eigenfunction of $T$ if its orbit is contained in a finite-dimensional sub-representation of $L^2(X)$. Denote by $\mathcal{E}(X)$ or $\mathcal{E}(T)$ the closure of the subspace of $L^2(X)$ spanned by the eigenfunctions of $X$. 55
Proposition 4.6.5. Let $S_1$ and $S_2$ be commuting actions of $G$ on a separated, countably generated probability space $(X, \mathcal{B}, \mu)$. If $S_2$ is ergodic then $\mathcal{E}(S_2) \subset \mathcal{E}(S_1)$.

Proof. Let $f$ be an eigenfunction of $S_2$ and let $\mathcal{M}$ be an $S_2$-invariant, finite-dimensional subspace of $L^2(X, \mathcal{B}, \mu)$ containing the orbit of $f$. Without loss of generality, we can assume $\mathcal{M}$ is irreducible. Let $\mathcal{N}$ be the closed subspace of $L^2(X, \mathcal{B}, \mu)$ spanned by the sub-representations of $G$ on $L^2(X, \mathcal{B}, \mu)$ induced by $S_2$ that are equivalent to $\mathcal{M}$. By Proposition 1.4 in [BR88] the multiplicity of $\mathcal{M}$ in $L^2(X, \mathcal{B}, \mu)$ is bounded by its dimension, so $\mathcal{N}$ is finite-dimensional. Fix $g \in G$ and put $\mathcal{M}_g = \{ S_1^g f : f \in \mathcal{M} \}$. Since $S_1$ and $S_2$ commute the representations of $G$ on $\mathcal{M}$ and $\mathcal{M}_g$ determined by $S_2$ are equivalent. Thus $\mathcal{M}_g \subset \mathcal{N}$. This implies $S_1^g f \in \mathcal{N}$ for all $g \in G$, so $f$ is contained in $\mathcal{E}(S_1)$. \hfill \Box

Proposition 4.6.6. Let $S_1$ and $S_2$ be commuting actions of $G$ on a separated, countably generated probability space $(X, \mathcal{B}, \mu)$. If $S_2$ is ergodic then $\mathcal{E}(S_2) \subset \mathcal{E}(S_2S_1)$.

Proof. Let $f$ be an eigenfunction of $S_2$. Form $\mathcal{M}$ and $\mathcal{N}$ as in the proof of Proposition 4.6.5. Fix $g \in G$. Put $\mathcal{M}_g = S_2^g S_1^g \mathcal{M}$. We have $\mathcal{M}_g = S_1^g \mathcal{M}$, which is equivalent to $\mathcal{M}$ and therefore contained in $\mathcal{N}$, as desired. \hfill \Box

We can now give a proof of Bergelson’s conjecture when $k = 3$ and the actions $T_1, T_2$ and $T_2T_1$ are all ergodic.

Theorem 4.6.7. Let $T_1, T_2, T_3$ be commuting, measurable actions of $G$ on a separated, countably generated probability space $(X, \mathcal{B}, \mu)$. Suppose that the actions $T_1, T_2$ and $T_2T_1$ are ergodic. Then

$$\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int_{\Phi_N} \int f \cdot T_3^g T_2^g T_1^g f \cdot T_3^g T_2^g f \cdot T_3^g f \, d\mu \, d\mu(g) > 0$$

(4.6.8)

for any $f > 0$ in $L^\infty(X, \mathcal{B}, \mu)$.

Proof. Ergodicity of $T_1$ means $\mathcal{C}_{1,1}$ is trivial, so $\mathcal{C}_{2,1}$ and $\mathcal{C}_{2,2}$ correspond to the functions that are $T_2T_1$ and $T_2$ almost-periodic over the trivial factor respectively. Let $\mathcal{D}$ be the sub-$\sigma$-algebra of $\mathcal{B}$ corresponding to the functions that are almost-periodic for $T_3$ over the trivial factor. Combining Corollary 4.3.12 with Propositions 4.6.5 and 4.6.6 gives $\mathcal{C}_{2,2} \subset \mathcal{C}_{2,1} \subset \mathcal{D}$. This implies any $\mathcal{C}_{2,1}$ measurable function $f$ is almost-periodic for both $T_2T_1$ and $T_3$, so $f \in \mathcal{E}(T_3T_2T_1)$.

We begin by showing that

$$\liminf_{N \to \infty} \frac{1}{m(\Phi_N)} \int_{\Phi_N} \int T_3^g T_2^g T_1^g f \cdot T_3^g T_2^g f \cdot T_3^g f \cdot f \, d\mu \, d\mu(g) > 0$$

56
whenever \( f = 1_B \) is \( \mathcal{C}_{2,1} \) measurable and \( \mu(B) > 0 \). Put \( \varepsilon = \mu(B)^2/6 \). Since \( \mathcal{C}_{2,1} \subset \mathcal{D} \) the set \( \Omega_1 \) of \( g \) in \( G \) for which \( \| T_3^g 1_B - 1_B \| \leq \varepsilon \) and \( \| T_3^g T_2^g T_1^g 1_B - 1_B \| \leq \varepsilon \) is a measurable \( \text{IP}^* \) subset of \( G \). By the argument on page 50 of [Ber96] the set \( \Omega_2 \) consisting of those \( g \) for which \( \mu(B) \geq \mu(B)^2 - \varepsilon \) is \( \text{IP}^* \). It is also measurable, so the intersection \( \Omega = \Omega_1 \cap \Omega_2 \) is measurable and \( \text{IP}^* \). We have

\[
\lim_{N \to \infty} \frac{1}{m(\Phi_N)} \left( \int T_3^g T_2^g T_1^g 1_B \cdot T_3^g T_2^g 1_B \cdot T_3^g 1_B \cdot 1_B \, d\mu \, dm(g) \right)
\geq \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \left( \int 1_{\Omega}(g) \left( \int T_3^g T_2^g T_1^g 1_B \cdot 1_B \, d\mu - 2\varepsilon \right) \, dm(g) \right)
\geq \lim_{N \to \infty} \frac{1}{m(\Phi_N)} \int 1_{\Omega}(g) \, dm(g) \cdot \left( \mu(B)^2 - 3\varepsilon \right) = \frac{d_{\Phi}(\Omega)\mu(B)^2}{2} > 0
\]

because every \( \text{IP}^* \) set has positive lower density.

The fact that \( \mathcal{C}_{2,2} \subset \mathcal{C}_{2,1} \) implies

\[
\int f_1 \otimes f_2 \otimes f_3 \, d\nu_2 = \int \mathbb{E}(f_1|\mathcal{C}_{2,1}) \otimes \mathbb{E}(f_2|\mathcal{C}_{2,1}) \otimes \mathbb{E}(f_3|\mathcal{C}_{2,1}) \, d\nu_2
\]

for all \( f_1, f_2, f_3 \) in \( L^\infty(X, \mathcal{B}, \mu) \) where \( \nu_2 \) is the Furstenberg joining for the actions \( T_1 \) and \( T_2 \). Applying Theorem 4.4.4 with \( \mathcal{D}, \mathcal{C}_i = \mathcal{C}_{2,1} \) gives sub-\( \sigma \)-algebras \( \mathcal{E}_{3,i} \) that are characteristic for (4.6.8). Moreover \( \mathcal{E}_{3,i} \to \mathcal{C}_{2,1} \) is compact for \( T_3 \cdot \cdots \cdot T_i \). Finally, using Theorem 4.5.1 with \( k = 3, \mathcal{D} = \mathcal{C}_{2,1} \) and \( \mathcal{E}_i = \mathcal{E}_{3,i} \) yields (4.6.8). \( \square \)
CHAPTER 5

POLYNOMIAL RECURRENCE

In this chapter we describe recent work [BR15] on an extension of the polynomial Szemerédi theorem of Bergelson and Leibman [BL96]. Specifically, we describe how the classification by Bergelson, Leibman and Lesigne [BLL08] of those polynomials to which Bergelson and Leibman’s polynomial Szemerédi theorem (see Theorem 1.1.4) applies can be extended to certain multidimensional situations by considering polynomials over algebraic number fields. We also discuss recent work [BR14] on single recurrence for polynomials over countable fields that strengthening recent results of McCutcheon and Windsor [MW14].

5.1 Polynomial Szemerédi theorems

The polynomial ergodic Szemerédi theorem in [BL96] implies, in particular, that

$$R = \{ n \in \mathbb{Z} : \mu(B \cap T_{p_1}(n)B \cap \cdots \cap T_{p_k}(n)B) > 0 \}$$

(5.1.1)

has positive lower density, meaning that

$$\liminf_{N \to \infty} \frac{|R \cap \{1, \ldots, N\}|}{N} > 0,$$

for any $p_1, \ldots, p_k \in \mathbb{Z}[x]$ each having zero constant term. It was shown in [BM96] that (5.1.1) is syndetic under the same assumptions, and the later work [BM00] implies it is large in the stronger sense (see Definition 2.7.11) of being $IP^*$.

The task of determining precisely which families $p_1, \ldots, p_k$ of polynomials have the property that (5.1.1) is syndetic was undertaken in [BLL08]. There it was shown polynomials $p_1, \ldots, p_k$ have the property that (5.1.1) is syndetic whenever $T$ is an action of $\mathbb{Z}$ on $(X, \mathcal{B}, \mu)$ and $\mu(B) > 0$ if and only if the polynomials are jointly intersective, which means that for any finite index subgroup $\Lambda$ of $\mathbb{Z}$, one can find $\zeta$ in $\mathbb{Z}$ such that $\{p_1(\zeta), \ldots, p_k(\zeta)\} \subset \Lambda$.

The polynomial ergodic Szemerédi theorem in [BL96] actually implies the following multi-dimensional result: for any action $T$ of $\mathbb{Z}^m$ on a probability space $(X, \mathcal{B}, \mu)$ and any
B with \( \mu(B) > 0 \) the set
\[
\{ n \in \mathbb{Z}^d : \mu(B \cap T^{p_1(n)}B \cap \cdots \cap T^{p_k(n)}B) > 0 \} \tag{5.1.2}
\]
has positive lower density for any polynomial mappings \( p_1, \ldots, p_k : \mathbb{Z}^d \rightarrow \mathbb{Z}^m \) each having zero constant term. In (5.1.2) and below we write \( T^{p_i(u)} \) for \( T_1^{p_{i,1}(u)} \cdots T_m^{p_{i,m}(u)} \) when \( p_i = (p_{i,1}, \ldots, p_{i,m}) \). As in the \( m = 1 \) case above, [BM00] implies that (5.1.2) is IP\(^*\). There is no known characterization of those polynomial mappings \( p_1, \ldots, p_k \) for which (5.1.2) is non-empty. By considering finite systems, one can show that joint intersectivity (defined below in general) is a necessary condition; it is conjectured in [BLL08] that it is also sufficient.

Since [Fur77], the sizes of sets such as (5.1.1) have been studied by considering the limiting behavior of averages such as
\[
\frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} \mu(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) \tag{5.1.3}
\]
where \( N \mapsto \Phi_N \) is some Følner sequence in \( \mathbb{Z} \). In [BLL08] the works of Host and Kra [HK05] and Ziegler [Zie07] on characteristic factors are combined with [Lei05a] to prove that the limiting behavior of the average (5.1.3) can be approximated arbitrarily well by replacing \((X, \mathcal{B}, \mu)\) with quotients \( G/\Gamma \) of certain nilpotent Lie groups by a cocompact subgroup on which \( \mathbb{Z} \) acts via \( T(g\Gamma) = ag\Gamma \) for some \( a \in G \). Upon passing to this more tractable setting, it is shown in [BLL08] that (5.1.3) is positive in the limit as \( N \rightarrow \infty \) when \( p_1, \ldots, p_k \) are jointly intersective.

It is not possible to proceed like this when studying (5.1.2) because there is currently no general version of the work of Host and Kra [HK05] and Ziegler [Zie07] for actions of \( \mathbb{Z}^m \). In this chapter we enlarge the class of polynomial mappings \( p_1, \ldots, p_k \) for which (5.1.2) is known to be non-empty by working with polynomials over rings of integers of algebraic number fields. As we will see, this is a setting where it is possible to reduce to the case of commuting translations on homogeneous spaces of nilpotent Lie groups, which will allow us to show that (5.1.2) is large. Our techniques also allow us to improve upon the main result in [BLL08] by strengthening the largeness property of the set (5.1.1). To describe our results we recall some definitions.

**Definition 5.1.4.** Let \( R \) be a commutative ring with identity. Polynomials \( p_1, \ldots, p_k \) in \( R[x_1, \ldots, x_d] \) are said to be **jointly intersective** if, for any finite index subgroup \( \Lambda \) of \( R \), one can find \( \zeta \) in \( \mathbb{R}^d \) such that \( \{ p_1(\zeta), \ldots, p_k(\zeta) \} \subset \Lambda \). When \( d = 1 \) we say that \( p_1 \) is **intersective**.

See Section 5.3 for some examples of intersective polynomials.

We can now state our main result. Given an algebraic number field \( L \), write \( \mathcal{O}_L \) for its ring of integers.
Theorem 5.1.5. Let $L$ be an algebraic number field and let $p_1,\ldots,p_k$ be jointly intersective polynomials in $O_L[x_1,\ldots,x_d]$. For any ergodic action $T$ of the additive group of $O_L$ on a compact metric probability space $(X,\mathcal{B},\mu)$ and any $B \in \mathcal{B}$ with $\mu(B) > 0$ there is $c > 0$ such that
\[
\{u \in O_L^d : \mu(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) \geq c\}
\] (5.1.6)
is $AIP^*_+$. In particular, taking $L = \mathbb{Q}$ shows that (5.1.1) is an $AIP^*_+$ subset of $\mathbb{Z}$. As was shown in Example 2.7.17, being $AIP^*_+$ is a stronger property than being syndetic, so Theorem 5.1.5 constitutes a strengthening of [BL96, Theorem 1.1].

The following version of the Furstenberg correspondence principle (which follows from Theorem 3.2.3) allows us to use Theorem 5.1.5 to find polynomial configurations in large subsets of $O_L$.

Theorem 5.1.7. For any $E \subset O_L$ there is an ergodic action $T$ of $O_L$ on a compact metric probability space $(X,\mathcal{B},\mu)$ and $B \in \mathcal{B}$ with $\mu(B) = \overline{d}(E)$ such that
\[
\overline{d}((E - u_1) \cap \cdots \cap (E - u_k)) \geq \mu(T^{u_1}B \cap \cdots \cap T^{u_k}B)
\] (5.1.8)
for every $u_1,\ldots,u_k$ in $O_L$.

Combining Theorems 5.1.5 and 5.1.7 gives the following combinatorial result.

Theorem 5.1.9. Let $L$ be an algebraic number field and let $E \subset O_L$ have positive upper Banach density. For any jointly intersective polynomials $p_1,\ldots,p_k$ in $O_L[x_1,\ldots,x_d]$ there is a constant $c > 0$ such that the set
\[
\{u \in O_L^d : \overline{d}(E \cap (E - p_1(u)) \cap \cdots \cap (E - p_k(u))) \geq c\}
\] (5.1.10)
is $AIP^*_+$.

Whenever $O_L$ is finitely partitioned, one of the partitions has positive upper Banach density. As a result, Theorem 5.1.9 yields new examples of the polynomial van der Waerden theorem, extending [BLL08, Theorem 1.5].

Corollary 5.1.11. Let $L$ be an algebraic number field. For any finite partition $E_1 \cup \cdots \cup E_k$ of $O_L$ there is $1 \leq i \leq k$ such that, for any jointly intersective polynomials $p_1,\ldots,p_k$ in $O_L[x_1,\ldots,x_d]$ the set (5.1.10) is $AIP^*_+$. So far, such polynomial van der Waerden results have only been proved via multiple recurrence of measure-preserving dynamical systems. It would be interesting to have a proof that only used topological dynamics, or a purely combinatorial proof.
Upon fixing a basis $e_1, \ldots, e_m$ for $O_L$ as a $Z$ module, defining actions $T_1, \ldots, T_m$ of $Z$ by $T_i^n = T^{ne_i}$, and writing

$$p_i(u) = p_{i,1}(u)e_1 + \cdots + p_{i,m}(u)e_m$$

(5.1.12)

for some polynomials $p_{i,j}$ in $Z[x_1, \ldots, x_{dm}]$, we see that Theorem 5.1.5 implies

$$\left\{ u \in Z^{md} : \int 1_B \prod_{i=1}^k T_1^{p_{i,1}(u)} \cdots T_m^{p_{i,m}(u)} 1_B \, d\mu > 0 \right\}$$

is AIP$^*_+$, extending [BL96, Theorem A] to certain families of intersective polynomials. Indeed, if for some polynomials $p_{i,j}$ from $Z[x_1, \ldots, x_d]$, one can find an algebraic number field $L$, jointly intersective polynomials $p_1, \ldots, p_k$ in $O_L[x_1, \ldots, x_d]$, and a basis $e_1, \ldots, e_m$ for $O_L$ over $Z$ such that (5.1.12) holds, then the polynomial mappings

$$(p_{1,1}, \ldots, p_{1,m}), \ldots, (p_{k,1}, \ldots, p_{k,m}) : Z^d \to Z^m$$

are good for recurrence.

It would be interesting to know whether (5.1.6) is AIP$^*_+$ without the ergodicity assumption. We show that it is syndetic.

**Theorem 5.1.13.** Let $L$ be an algebraic number field and let $p_1, \ldots, p_k$ be jointly intersective polynomials in $O_L[x_1, \ldots, x_d]$. For any action $T$ of the additive group of $O_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$ and any $B \in \mathcal{B}$ with $\mu(B) > 0$ there is $c > 0$ such that

$$\left\{ u \in O_L^d : \mu(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) \geq c \right\}$$

(5.1.14)

is syndetic.

Our proof of Theorem 5.1.5 consists of two main steps. First we show, by combining Leibman’s polynomial convergence result [Lei05a] with Griesmer’s description [Gri09] of characteristic factors for certain actions of $Z^m$, that upon restricting our attention to a very large subset of $O_L^d$ – one whose complement has zero upper Banach density – it suffices to consider (5.1.6) when $(X, \mathcal{B}, \mu)$ has the structure of a nilrotation, the definition of which we now recall.

**Definition 5.1.15.** By a nilmanifold we mean a homogeneous space $G/\Gamma$ where $G$ is a nilpotent Lie group and $\Gamma$ is a discrete, cocompact subgroup of $G$. A nilrotation is an action $T$ of $Z^m$ on a nilmanifold $G/\Gamma$ of the form $T^u(g\Gamma) = \phi(u)g\Gamma$ for some homomorphism $\phi : Z^m \to G$. The nilpotency degree of a nilrotation is the minimal length of a shortest central series for $G$. 
The second step in the proof of Theorem 5.1.5 is to use results from [BLL08] about polynomial orbits of nilrotations to show that, within the very large subset of $O_L$ mentioned above, we can achieve the desired multiple recurrence.

It is natural to ask how large the intersection in (5.1.6) can be. When $k = 1$ we show it is as large as can be expected, extending results in [Fur81b], [Sár78] and [KMF78].

**Theorem 5.1.16.** Let $L$ be an algebraic number field and let $p \in O_L[x_1, \ldots, x_d]$ be an intersective polynomial. For any action $T$ of the additive group of $O_L$ on a probability space $(X, \mathcal{B}, \mu)$ and any $B$ in $\mathcal{B}$ the set

$$\{ u \in O_L^d : \mu(B \cap T^{p(u)}B) > \mu(B)^2 - \varepsilon \}$$  \hspace{1cm} (5.1.17)

is AIP$^*$ for any $\varepsilon > 0$.

When $p$ has zero constant term one can use [BFM96, Theorem 1.8] to show that (5.1.17) is IP$^*$. It follows immediately that (5.1.17) is IP$^*_+$ when $p$ has a zero in $O_L^d$, but it is unknown whether (5.1.17) is IP$^*_+$ if one only assumes $p$ is intersective, even in the case $L = \mathbb{Q}$. More generally, one could ask whether a version of Theorem 5.1.16 holds for a given intersective polynomial $p$ over an arbitrary integral domain $R$. Under the additional assumption that $p$ has zero constant term it was shown in [BLM05] that $\{ u \in R : \mu(B \cap T^{p(u)}B) > 0 \}$ has positive density with respect to some Følner sequence in $R$, but whether this set is syndetic is unknown. We cannot proceed as in the proof of Theorem 5.1.16, or apply [BFM96, Theorem 1.8], at such a level of generality due to complications that arise when the additive group of the ring is not finitely generated. However, if the ring is a countable field then Theorem 5.7.2 below gives a version of Theorem 5.1.9 in that setting.

### 5.2 Single recurrence for polynomials over rings of integers

In this section we prove Theorem 5.1.16, which relies on the following lemmas.

**Lemma 5.2.1.** Let $L$ be an algebraic number field. If $p \in O_L[x_1, \ldots, x_d]$ and the induced map $O_L^d \to O_L$ is a non-zero homomorphism of abelian groups then $p(O_L^d)$ is a finite-index subgroup of $O_L$.

**Proof.** Write $p(x_1, \ldots, x_d) = a_1 x_1 + \cdots + a_d x_d$. Certainly the image of $p$ is a subgroup of $O_L$. Since some $a_i$ is non-zero, $p(O_L^d)$ contains the ideal generated by $a_i$, which is non-zero. But every non-zero ideal in the ring of integers of an algebraic number field has finite index (see [Jan96, Section I.8]).

**Lemma 5.2.2.** Let $R$ be a commutative ring and let $G$ be an abelian, compact, Hausdorff topological group. Fix an additive homomorphism $\psi : R \to G$. For any $k \in \mathbb{N}$, any
polynomial \( p \in \mathbb{R}[x_1, \ldots, x_k] \) with \( p(0) = 0 \), and any idempotent ultrafilter \( p \) on the additive group of \( \mathbb{R}^k \) we have \( \lim_{r \to p} \psi(p(r)) = 0 \).

**Proof.** The proof is by induction on the degree of \( p \). When \( p \) has degree 1 the map \( r \mapsto \psi(p(r)) \) is an additive homomorphism so we have

\[
\lim_{r \to p} \psi(p(r)) = \lim_{r \to p} \psi(p(r + s)) = \lim_{r \to p} \psi(p(r)) + \psi(p(s)) = 2 \lim_{r \to p} \psi(p(r))
\]

by idempotence so the limit in question is zero.

For the induction step, write \( \psi(p(r + s)) = \psi(p(r)) + \psi(p(s)) + \psi(q(r, s)) \) for some polynomial \( q \) with twice as many indeterminates as \( p \) and zero constant. By induction we have

\[
\lim_{r \to p} \lim_{s \to p} \psi(q(r, s)) = 0
\]

so we again have (5.2.3) and the limit in question is zero. \( \square \)

**Lemma 5.2.4.** Let \( G \) be an abelian group and let \( H \) be a finite index subgroup. If \( T \) is an action of \( G \) on a probability space \( (X, \mathcal{B}, \mu) \) and \( f \in L^2(X, \mathcal{B}, \mu) \) is invariant under \( T|H \) then \( f \) is a finite sum of eigenfunctions of \( T \).

**Proof.** Let \( g_1, \ldots, g_n \) be coset representatives for \( H \) with \( g_1 = 0 \). Writing any \( g \in G \) as \( h + g_i \) for some \( i \) and some \( h \in H \), we see that \( T^g f = T^{g_i} f \). Thus the subspace \( K \) of \( L^2(X, \mathcal{B}, \mu) \) spanned by \( \{f, \ldots, T^{g_n} f\} \) is \( T \)-invariant. The unitary representation of \( G \) on \( K \) decomposes as a direct sum of one-dimensional representations because \( G \) is abelian. In particular \( f \) is a sum of eigenfunctions. \( \square \)

**Proof of Theorem 5.1.16.** Let \( T \) be an action of the additive group of \( \mathcal{O}_L \) on a probability space \( (X, \mathcal{B}, \mu) \). Fix \( B \in \mathcal{B} \) and \( \varepsilon > 0 \). Let \( P \) be the orthogonal projection in \( L^2(X, \mathcal{B}, \mu) \) onto the closed subspace \( \mathcal{K}_c \) spanned by the eigenfunctions of \( T \). Put \( f = 1_B - P1_B \).

We begin by proving that

\[
\lim_{u \to \Phi} |\langle \phi, T^{p(u)} f \rangle|^2 = 0
\]

for every Følner sequence \( \Phi \) in \( \mathcal{O}^d_L \) and every \( \phi \) that is orthogonal to \( \mathcal{K}_c \) and satisfies \( \| \phi \| \leq 1 \). Since \( \mathcal{K}_c \) is \( T \)-invariant we can assume \( p(0) = 0 \). In terms of the product system we have

\[
\limsup_{N \to \infty} \left( \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} |\langle \phi, T^{p(u)} f \rangle|^2 \right)^2
\]

\[
\leq \limsup_{N \to \infty} \left( \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} (T \times T)^{p(u)} (f \otimes f) \right)^2
\]

\[
\leq \frac{1}{|\Phi_H|^2} \sum_{h,l \in \Phi_H} \limsup_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{u \in \Phi} |\langle f, T^{p(u+h)-p(u+h)} f \rangle|^2
\]

63
for every $H \in \mathbb{N}$ by Cauchy-Schwarz and an application of the van der Corput trick (see Proposition 2.3.8) to the sequence $g(u) = (T \times T)^{p(u)}(f \otimes f)$. For all but a density zero set of $(h, l) \in \mathcal{O}_L^d$ the polynomial $u \mapsto p(u + l) - p(u + h)$ is non-constant and has degree smaller than that of $p$. Taking the lim sup as $H \to \infty$ above, it therefore suffices by an induction argument to prove (5.2.5) when $p$ has degree 1. Thus we may assume $p$ is an additive homomorphism $\mathcal{O}_L^d \to \mathcal{O}_L$. Lemma 5.2.1 implies $R := p(\mathcal{O}_L^d)$ is a finite index subgroup. Applying the mean ergodic theorem to the product system we see that the limit

$$\lim_{u \to \Phi} (T \times T)^{p(u)}(f \otimes f)$$

is invariant under $(T \times T)|R$. By Lemma 5.2.4 the limit is a sum of eigenfunctions of $T \times T$. Since the eigenfunctions of $T \times T$ are spanned by functions of the form $\phi_1 \otimes \phi_2$ where $\phi_1$ and $\phi_2$ are eigenfunctions of $T$, we see that (5.2.5) is zero as desired.

Since $\Phi$ was an arbitrary Følner sequence, Lemma 2.7.9 implies that

$$\{u \in \mathcal{O}_L^d : |\langle 1_B, T^{p(u)}1_B \rangle - \langle 1_B, T^{p(u)}P1_B \rangle| \geq \varepsilon \}$$

has zero upper Banach density.

Let $f_1, \ldots, f_r$ be eigenfunctions of $T$ with eigenvalues $\chi_1, \ldots, \chi_r$ such that $\|f_1 + \cdots + f_r - P1_B\| \leq \varepsilon$. Define a map $\psi : \mathcal{O}_L^d \to \mathbb{T}^r$ by $\psi(u) = (\chi_1(u), \ldots, \chi_r(u))$ for all $u \in \mathcal{O}_L^d$. Let $e_1, \ldots, e_m$ be a basis for $\mathcal{O}_L$ as a $\mathbb{Z}$-module and write

$$p(u) = p_1(u)e_1 + \cdots + p_m(u)e_m$$

for polynomials $p_1, \ldots, p_m$ in $\mathbb{Z}[x_1, \ldots, x_{dm}]$. We claim that $p_1, \ldots, p_k$ are jointly intersective. Indeed, let $\Lambda = \mathbb{Z}\lambda$ be a finite index subgroup of $\mathbb{Z}$. Since $p$ is intersective we have $p(\zeta) \in (\lambda)$ for some $\zeta$ in $\mathcal{O}_L^d$, and this implies $\{p_1(\zeta), \ldots, p_k(\zeta)\} \subset \Lambda$ as desired. Writing

$$\psi(p(u)) = p_1(u)(\chi_1(e_1), \ldots, \chi_r(e_1)) + \cdots + p_m(u)(\chi_1(e_m), \ldots, \chi_r(e_m))$$

we can apply [BLL08, Proposition 3.6] to obtain $w$ in $\mathcal{O}_L^d$ for which $|\chi_i(p(w))| < \varepsilon/k$ for all $1 \leq i \leq k$. The polynomial $q(u) = p(u + w) - p(w)$ has zero constant term. Thus

$$\lim_{u \to p} T\psi(q(u)) = 0$$

for any idempotent ultrafilter $p$ on $\mathcal{O}_L^d$ by Lemma 5.2.2. Combining this with how $w$ was chosen, Corollary 2.7.20 implies

$$\{u \in \mathcal{O}_L^d : \langle 1_B, T^{p(u+w)}P1_B \rangle \geq \mu(B)^2 - \varepsilon \}$$

$$\supset \{u \in \mathcal{O}_L^d : \langle 1_B, T^{p(u+w)-p(w)}P1_B \rangle \geq \mu(B)^2 - 4\varepsilon \}$$

is $\text{IP}^*$. Thus the set

$$\{u \in \mathcal{O}_L^d : \langle 1_B, T^{p(u)}P1_B \rangle \geq \mu(B)^2 - \varepsilon \}$$

is $\text{IP}^*_+$ and (5.1.16) is $\text{AIP}^*_+$ as desired. 

\[ \square \]
5.3 Examples of intersective polynomials

We now turn to some examples. Since every non-zero ideal in $\mathcal{O}_L$ has finite index, polynomials $p_1, \ldots, p_k$ in $\mathcal{O}_L[x_1, \ldots, x_d]$ are jointly intersective if and only if, for any non-zero ideal $I$ in $\mathcal{O}_L$, one can find $\zeta$ in $\mathcal{O}_L^d$ such that $\{p_1(\zeta), \ldots, p_k(\zeta)\} \subset I$. It was shown in [BLL08, Proposition 6.1] that when $L = \mathbb{Q}$, polynomials $p_1, \ldots, p_k \in \mathbb{Z}[x]$ are jointly intersective if and only if there is an intersective polynomial $p \in \mathbb{Z}[x]$ such that $p|p_i$ for all $1 \leq i \leq k$. The same proof works for intersective polynomials of one variable over $\mathcal{O}_L$.

**Lemma 5.3.1.** Let $L$ be an algebraic number field and let $p_1, \ldots, p_k \in \mathcal{O}_L[x]$ be jointly intersective. Then there is an intersective polynomial $p \in \mathcal{O}_L[x]$ such that $p|p_i$ for all $1 \leq i \leq k$.

**Proof.** Let $p \in \mathcal{O}_L[x]$ be the greatest common divisor of $p_1, \ldots, p_k$ in $\mathcal{L}[x]$. Then one can find $h_1, \ldots, h_k \in \mathcal{L}[x]$ such that $h_1p_1 + \cdots + h_kp_k = p$. By clearing denominators we obtain $f_1p_1 + \cdots + f_kp_k = dp$ for polynomials $f_1, \ldots, f_k \in \mathcal{O}_L[x]$. Joint intersectivity of $p_1, \ldots, p_k$ now implies intersectivity of $dp$ and thus of $p$. \qed

**Example 5.3.2.** Let $K$ be an algebraic number field and fix $c \in \mathcal{O}_K$. Define $f$ in $\mathcal{O}_K[x]$ by $f(x) = x^2 + c$ for all $x \in \mathcal{O}_K$. We show that if $f$ is intersective then $f$ has a root in $\mathcal{O}_K$. The converse is immediate.

Suppose to the contrary that $f$ does not have a root in $\mathcal{O}_K$. Put $L = K(\sqrt{-c})$. Then $f$ is the minimal polynomial of $\sqrt{-c}$. Since $f$ is intersective it has a root modulo every prime ideal $\mathfrak{p}$ in $\mathcal{O}_K$. Thus $f$ is a product of two linear factors in the ring $\mathcal{O}_K/\mathfrak{p}[x]$. By Kummer’s theorem [Jan96, Page 37] this implies that $\mathfrak{p}\mathcal{O}_L$ is not prime and therefore factors in $\mathcal{O}_L$. This is a contradiction because one can always find prime ideals in $\mathcal{O}_K$ which remain prime when lifted to $\mathcal{O}_L$. Thus $f$ has a root in $\mathcal{O}_K$.

For a specific example, consider $f(x) = x^2 + 1$ over $\mathbb{Z}[i]$ and let $T_1, T_2$ be commuting, measure-preserving actions of $\mathbb{Z}$ on a probability space $(X, \mathcal{B}, \mu)$. Then $a + ib \mapsto T_1^aT_2^b$ is an action of $\mathbb{Z}[i] = \mathcal{O}_{\mathbb{Q}[i]}$ on $(X, \mathcal{B}, \mu)$. Theorem 5.1.16 tells us

$$\{u \in \mathbb{Z}[i] : \mu(B \cap T^p(u)B) \geq \mu(B)^2 - \varepsilon\}$$

is AIP* for any $B \in \mathcal{B}$ and any $\varepsilon > 0$. In terms of $\mathbb{Z}$-actions, we see that

$$\{(a, b) \in \mathbb{Z}^2 : \mu(B \cap T_1^{a^2-b^2}T_2^{2ab}B) \geq \mu(B)^2 - \varepsilon\}$$  \hspace{1cm} (5.3.3)

is AIP* for any $B \in \mathcal{B}$ and any $\varepsilon > 0$.

In this case we can actually say more. By replacing $b$ with $b + 1$ in (5.3.3) we obtain

$$\{(a, b) \in \mathbb{Z}^2 : \mu(B \cap T_1^{a^2-b^2-2b}T_2^{2ab}B) \geq \mu(B)^2 - \varepsilon\}$$

and this set is IP* by [BM00]. Thus (5.3.3) is IP*. 

65
Note that any non-constant, monic polynomial can be made intersective by passing to an extension in which it has a root. Our second example is of an intersective polynomial over $\mathbb{Z}[i]$ without a root. It is based on [BS66, Page 3], which is also discussed in [BLL08, Section 6].

Example 5.3.4. Write $L = \mathbb{Q}[i]$ and let $\alpha$ and $\beta$ be primes in $\mathcal{O}_L = \mathbb{Z}[i]$ distinct from $1+i$ such that $\alpha$ is a quadratic residue modulo $(\beta)$ and vice versa. Assume also that one of $\alpha$, $\beta$ or $\alpha \beta$ is a square modulo $(1+i)^5$. Then $f(x) = (x^2 - \alpha)(x^2 - \beta)(x^2 - \alpha \beta)$ in $\mathcal{O}_L[x]$ is intersective.

It suffices to prove that $f$ has a root modulo every non-zero ideal in $\mathcal{O}_L$. Since every non-zero, proper ideal in $\mathcal{O}_L$ factors a product of powers of prime ideals, the Chinese remainder theorem implies that it suffices to prove $f$ has a root modulo $p^n$ for every prime ideal $p$ in $\mathcal{O}_L$ and every $n \in \mathbb{N}$.

If $p = (z)$ for some prime $z$ distinct from $\alpha$, $\beta$ and $1+i$ then quadratic reciprocity in $\mathbb{Z}[i]$ implies that one of the factors of $f$ has a root modulo $p$. Since the root is non-zero in $\mathcal{O}_L/p$, Hensel’s lemma [Jan96, Page 105] implies that the same factor has a root modulo every power of $p$.

The same argument shows that $f$ has a root modulo $p^n$ when $p \in \{(\alpha), (\beta)\}$ by our assumption that $\alpha$ is a residue modulo $(\beta)$ and vice versa.

Lastly, if $p = (1+i)$ then one of the factors $h$ of $f$ has a root modulo $(1+i)^n$ for $n \leq 5$ by assumption. Suppose by induction that $h$ has a root $w$ modulo $(1+i)^n$ for some $n \geq 5$. If $(1+i)^{n+1}$ divides $h(w)$ then there is nothing to prove, so assume otherwise. We claim that $w + (1+i)^{n-2}$ is a root of $h$ modulo $(1+i)^{n+1}$. Since

$$h(w + (1+i)^{n-2}) = h(w) - iw(1+i)^n + (1+i)^{2n-4}$$

and $n \geq 5$, it suffices to prove that $(1+i)^{n+1}$ divides $h(w) - iw(1+i)^n$. Note that $1+i$ cannot divide $w$ because $\alpha$ and $\beta$ are primes distinct from $1+i$. Thus $h(w)$ and $-iw(1+i)^n$ are both divisible by $(1+i)^n$, but neither is divisible by $(1+i)^{n+1}$. Their sum is therefore divisible by $(1+i)^{n+1}$ as desired.

For example, one may take $\alpha = 7$ and $\beta = 5+2i$. Indeed $(3+5i)^2 = -16 + 30i$ is congruent to $5+2i$ modulo 7 and $(2+i)^2 = 3+4i$ is congruent to 7 modulo $(1+i)^5 = -4-4i$.

5.4 Gowers-Host-Kra norms for commuting actions

In this section we recall the construction of Gowers-Host-Kra seminorms for a $\mathbb{Z}^m$-system $X = (X, \mathcal{B}, \mu, T)$, which is totally analogous to the $m = 1$ case given in [HK05]. See [Gri09, Section 4.3.6] for more on these seminorms.
One defines inductively a sequence $X^{[k]}$ of $\mathbb{Z}^m$ systems as follows. Put $X^{[0]} = X$. Assuming that $X^{[k]} = (X^{[k]}, \mathcal{B}^{[k]}, \mu^{[k]}, T^{[k]})$ has been defined, put

$$X^{[k+1]} = X^{[k]} \times X^{[k]} \quad \mathcal{B}^{[k+1]} = \mathcal{B}^{[k]} \otimes \mathcal{B}^{[k]} \quad T^{[k+1]} = T^{[k]} \times T^{[k]}$$

and define $\mu^{[k+1]}$ to be the relatively independent self-joining of $\mu^{[k]}$ over the sub-$\sigma$-algebra $\mathcal{I}^{[k]} \subset \mathcal{B}^{[k]}$ of sets invariant under $T^{[k]}$. Thus for any $F_0, F_1$ in $L^\infty(X^{[k]})$ we have

$$\int F_0 \otimes F_1 \, d\mu^{[k+1]} = \int \mathbb{E}(F_0|\mathcal{I}_{[k]}) \cdot \mathbb{E}(F_1|\mathcal{I}_{[k]}) \, d\mu^{[k]}$$

for any Følner sequence $\Phi$ in $\mathbb{Z}^m$. For example

$$X^{[1]} = (X \times X, \mathcal{B} \otimes \mathcal{B}, T \times T, \mu \otimes \mathcal{I}_{[0]} \mu)$$

where $\mathcal{I}_{[0]}$ is the sub-$\sigma$-algebra of $T$-invariant sets. In particular $\mu^{[1]} = \mu \otimes \mu$ if $T$ is ergodic.

Given $f$ in $L^\infty(X)$ write $f^{[k]}$ for the function

$$f \otimes \cdots \otimes f = f \circ \pi_1 \cdots f \circ \pi_{2^k}$$

in $L^\infty(X^{[k]})$, where $\pi_1, \ldots, \pi_{2^k}$ are the coordinate projections $X^{[k]} \to X$. For each $k \geq 1$ the $k$th **Gowers-Host-Kra seminorm** $\| \cdot \|_k$ on $L^\infty(X)$ is defined by

$$\| f \|_k^2 = \int f^{[k]} \, d\mu^{[k]}$$

for all $f$ in $L^\infty(X)$, and $\| f \|_0 = \int f \, d\mu$. Note that

$$\| f \|_1^2 = \int f \otimes f \, d\mu^{[1]} = \int \mathbb{E}(f|\mathcal{I}_{[0]}) \cdot \mathbb{E}(f|\mathcal{I}_{[0]}) \, d\mu^{[0]}$$

for all $f$ in $L^\infty(X)$ so

$$\| f \|_0 \leq \| f \|_1$$

(5.4.1)

by Cauchy-Schwarz. When $k \geq 1$ we have

$$\| f \|_k^2 = \int \mathbb{E}(f^{[k-1]}|\mathcal{I}_{[k-1]}) \cdot \mathbb{E}(f^{[k-1]}|\mathcal{I}_{[k-1]}) \, d\mu^{[k-1]}$$

for all $f$ in $L^\infty(X)$. For any $k \geq 0$ and any Følner sequence $\Phi$ in $\mathbb{Z}^m$ we have

$$\lim_{u \to \Phi} \| f \cdot T^u f \|_k = \lim_{u \to \Phi} \int f^{[k]} \cdot T^u f^{[k]} \, d\mu^{[k]}$$

$$= \int \mathbb{E}(f^{[k]}|\mathcal{I}_{[k]}) \cdot \mathbb{E}(f^{[k]}|\mathcal{I}_{[k]}) \, d\mu^{[k]} = \| f \|_{k+1}^{2k+1}$$

(5.4.2)

for all $f$ in $L^\infty(X)$ by the mean ergodic theorem.

The key feature of the seminorms $\| \cdot \|_k$ is that, for ergodic $\mathbb{Z}^m$-systems their kernels are determined by $T$-invariant sub-$\sigma$-algebras $\mathcal{Z}_k$ of $\mathcal{B}$ that have a strong algebraic structure. This was proved for $m = 1$ by Host and Kra [HK05] and generalized to arbitrary $m$ by Griesmer as follows.
Theorem 5.4.3 ([Gri09]). Let $X = (X, \mathcal{B}, \mu, T)$ be an ergodic $\mathbb{Z}^m$-system. For each $k \in \mathbb{N}$ there is an invariant sub-$\sigma$-algebra $\mathcal{Z}_k$ of $\mathcal{B}$ with the property that $\|f\|_k = 0$ if and only if $\mathbb{E}(f|\mathcal{Z}_k) = 0$. Moreover, the factor corresponding to $\mathcal{Z}_k$ is an inverse limit of of a sequence of nilrotations of nilpotency degree at most $k$.

Proof. This is a combination of Lemma 4.4.3 and Theorem 4.10.1 in [Gri09]. 

We will use the following theorem to relate the Gowers-Host-Kra seminorms of an ergodic $\mathbb{Z}^m$-system $(X, \mathcal{B}, \mu, T)$ to those of the systems $(X^2, \mathcal{B}^2, T \times T, \mu_s)$ where $\mu_s$ is the ergodic decomposition of $T \times T$.

Theorem 5.4.4. Let $X = (X, \mathcal{B}, \mu, T)$ be an ergodic $\mathbb{Z}^m$ system with Kronecker factor $\mathbf{Z} = (\mathcal{Z}, \mathcal{X}, m, \mathcal{T})$. For each $s$ in $\mathbf{Z}$ define a measure $\mu_s$ on $(X \times X, \mathcal{B} \otimes \mathcal{B})$ by

$$\int f_1 \otimes f_2 \, d\mu_s = \int \mathbb{E}(f_1|\mathcal{Z})(z) \cdot \mathbb{E}(f_2|\mathcal{Z})(z - s) \, dm(z)$$

for all $f_1, f_2$ in $L^\infty(X)$. Then $\mu_s$ is the ergodic decomposition of $\mu \otimes \mu$.

Proof. The Kronecker factor $(X, \mathcal{X}, m)$ has the structure of a compact abelian group. Let $\alpha : \mathbb{Z}^m \to \mathbf{Z}$ be a homomorphism with dense image that determines $T$ on $(\mathbf{Z}, \mathcal{X}, m)$. Write $\pi$ for the factor map $X \to \mathbf{Z}$.

Write $X \times X$ for the system $(X^2, \mathcal{B} \otimes \mathcal{B}, \mu \otimes \mu, T \times T)$. If $F$ in $L^2(X \times X)$ is invariant then $F$ is $\pi^{-1}\mathcal{X} \otimes \pi^{-1}\mathcal{X}$ measurable. This is because any $T \times T$-invariant function can be approximated by linear combinations of products of eigenfunctions of $T$, which in turn is a consequence of Theorem 4.3.6 applied to the case $\mathcal{D} = \{\emptyset, X\}$. It follows that $F$ is of the form $\Psi \circ \pi$ for some $\Psi$ in $L^2(\mathbf{Z} \times \mathbf{Z})$. Thus we can write $\Psi$ as

$$\Psi = \sum_{i,j} c_{i,j} \chi_i \otimes \chi_j$$

where $\chi_i$ is an orthonormal basis of $L^2(\mathbf{Z})$ consisting of characters. Invariance of $\Psi$ gives

$$\Psi = (T \times T)^n\Psi = \sum_{i,j} c_{i,j} \chi_i(n \cdot \alpha) \chi_j(n \cdot \alpha) \chi_i \otimes \chi_j$$

(5.4.5)

for all $n$ in $\mathbb{Z}^d$. Thus $c_{i,j}(1 - \chi_i(n \cdot \alpha) \chi_j(n \cdot \alpha)) = 0$ for all $n$ in $\mathbb{Z}^d$ and all $i, j$. If $c_{i,j}$ is non-zero for some $i, j$ we have $\chi_i(n \cdot \alpha) \chi_j(n \cdot \alpha) = 1$ for all $n$ in $\mathbf{Z}$, and the character $\chi_i \chi_j$ takes the value 1 on the orbit of $\alpha$ so it is constant. Thus if $c_{i,j}$ is non-zero we have $\chi_i = \chi_j$, leading to the simplification

$$\Psi = \sum_i c_i \cdot \chi_i \otimes \chi_i$$

(5.4.6)

of (5.4.5). For any $i$ and any subset $U$ of $\mathbb{C}$ we have

$$(\chi_i \otimes \chi_i)^{-1}U = \{(z_1, z_2) : \chi_i(z_1 - z_2) \in U\} = \{(z_1, z_2) : z_1 - z_2 \in \chi_i^{-1}U\}$$
so $\chi_i \pi \otimes \chi_i \pi$ is measurable with respect to the sub-$\sigma$-algebra

$$\mathcal{I} = \sigma(\{(x_1, x_2) : \pi x_1 - \pi x_2 \in A \} : A \in \mathcal{F})$$

of $\mathcal{B} \otimes \mathcal{B}$. Since $F$ was an arbitrary invariant function in $L^2(\mathbb{X} \times \mathbb{X})$ and every set in $\mathcal{I}$ is invariant under $T \times T$, we have that $\mathcal{I}$ is the sub-$\sigma$-algebra of $T \times T$-invariant sets.

This suggests that for each $s \in \mathbb{Z}$ there is a measure on

$$\{(x_1, x_2) : \pi x_1 - \pi x_2 = s\}$$

that is ergodic for $T \times T$. To make this precise, fix $s \in \mathbb{Z}$ and let $m_s$ be the measure on $\mathbb{Z}^2$ obtained by pushing $m$ forward using the map $z \mapsto (z, z - s)$. Then, let $\mu_s$ be the measure on $(\mathbb{X}^2, \mathcal{B}^2)$ defined by

$$\int f_1 \otimes f_2 \, d\mu_s = \int \mathbb{E}(f_1|\mathbb{Z}) \otimes \mathbb{E}(f_2|\mathbb{Z}) \, dm_s$$

for all $f_1, f_2$ in $L^\infty(\mathbb{X}, \mathcal{B}, \mu)$. By definition of $\mu_s$ we have

$$\int f_1 \otimes f_2 \, d\mu_s = \int \mathbb{E}(f_1|\mathbb{Z})(z) \cdot \mathbb{E}(f_2|\mathbb{Z})(z - s) \, dm(z)$$

for all $f_1, f_2$ in $L^\infty(\mathbb{X}, \mathcal{B}, \mu)$. This proves $\mu_s$ depends measurably on $s$. It is immediate that each of the measures $\mu_s$ is $T \times T$-invariant. Moreover, our description of $\mathcal{I}$ implies that if $C$ is $T \times T$-invariant then $\mu_s(C)$ must be either 0 or 1, so each of the measures $\mu_s$ is ergodic. Lastly, note that

$$\int\int f_1 \otimes f_2 \, d\mu_s \, dm(s) = \int\int \mathbb{E}(f_1|\mathbb{Z})(z) \cdot \mathbb{E}(f_2|\mathbb{Z})(z - s) \, dm(z) \, dm(s)$$

by Fubini’s theorem, so $\mu_s$ is the ergodic decomposition of $\mu \otimes \mu$.

Write $\mu_s^{[k]}$ for $(\mu_s)^{[k]}$ and $\| \cdot \|_{s,k}$ for the $k$th Gowers-Host-Kra seminorm of the system $(\mathbb{X}^2, \mathcal{B}^2, T \times T, \mu_s)$.

**Proposition 5.4.7.** Let $T$ be an ergodic, measure-preserving action of $\mathbb{Z}^m$ on a compact metric probability space $(\mathbb{X}, \mathcal{B}, \mu)$ and let $\mu_s$ be the ergodic decomposition of $T \times T$. Then

$$\mu_s^{[k+1]} = \int \mu_s^{[k]} \, dm(s)$$

(5.4.8)

for every $k \geq 0$ and

$$\| f \|_{k+1}^{2^{k+1}} = \int \| f \otimes f \|_{s,k}^{2^k} \, dm(s)$$

for every $f$ in $L^\infty(\mathbb{X})$. 69
Proof. The proof is by induction on \( k \). When \( k = 0 \) we use ergodicity of \( \mu \) and Theorem 5.4.4 to obtain

\[
\|f\|_1^2 = \int f \otimes f \, d(\mu \otimes \mu) = \iint f \otimes f \, d\mu_s \, dm(s) = \int \|f \otimes f\|_{s,0} \, dm(s)
\]

for any \( f \) in \( L^\infty(X,\mathcal{B},\mu) \).

Suppose now that (5.4.8) holds for some \( k \geq 0 \). Fix a bounded, measurable function \( F : X^{[k+1]} \to \mathbb{R} \). Write \( \Phi_N = \{1, \ldots, N\}^m \). In this proof we will denote the measure with respect to which a conditional expectation is taken using a subscript.

The pointwise ergodic theorem for actions of \( \mathbb{Z}^m \) (see [DS58, VIII.6.9]) tells us that

\[
\lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} T^u_{[k+1]} F = E(F|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}}
\]

almost surely with respect to \( \mu^{[k+1]} \). It also implies that, for \( m \) almost every \( s \), we have

\[
\frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} T^u_{[k+1]} F \to E(F|\mathcal{I}_{[k+1]})_{\mu_s^{[k]}}
\]

almost surely with respect to \( \mu_s^{[k]} \). Thus (5.4.8) implies that for \( m \) almost every \( s \) we have

\[
E(f|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}} = E(f|\mathcal{I}_{[k+1]})_{\mu_s^{[k]}}
\]

on a set of full \( \mu_s^{[k]} \) measure. But then

\[
\int F_0 \otimes F_1 \, d\mu^{[k+2]} = \int E(F_0|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}} \cdot E(F_1|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}} \, d\mu^{[k+1]}
\]

\[
= \iiint E(F_0|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}} \cdot E(F_1|\mathcal{I}_{[k+1]})_{\mu^{[k+1]}} \, d\mu_s^{[k]} \, dm(s)
\]

\[
= \iiint E(F_0|\mathcal{I}_{[k+1]})_{\mu_s^{[k]}} \cdot E(F_1|\mathcal{I}_{[k+1]})_{\mu_s^{[k]}} \, d\mu_s^{[k]} \, dm(s)
\]

\[
= \int F_0 \otimes F_1 \, d\mu_s^{[k+1]} \, dm(s)
\]

for any bounded, measurable functions \( F_0, F_1 \) on \( X^{[k+1]} \) as desired. \( \square \)

5.5 Characteristic factors for some polynomial averages

In this section we describe characteristic factors for multiparameter correlations of the form

\[
\int f \cdot T^{p_1(u)} f \cdots T^{p_k(u)} f \, d\mu
\]  

(5.5.1)

where \( T \) is an ergodic action of \( \mathcal{O}_L \) on a compact metric probability space \( (X,\mathcal{B},\mu) \), the function \( f \) belongs to \( L^\infty(X,\mathcal{B},\mu) \) and \( p_1, \ldots, p_k \) are non-constant polynomials in
$\mathcal{O}_L[x_1, \ldots, x_d]$. A **characteristic factor** for (5.5.1) is a $T$ invariant sub-$\sigma$-algebra $C$ of $\mathcal{B}$ for which

$$\int f \cdot T^{p_1(u)} f \cdots T^{p_k(u)} f - \mathbb{E}(f|\mathcal{C}) \cdot T^{p_1(u)} \mathbb{E}(f|\mathcal{C}) \cdots T^{p_k(u)} \mathbb{E}(f|\mathcal{C}) \, d\mu \to 0$$

in $L^2(X, \mathcal{B}, \mu)$ for every $f \in L^\infty(X, \mathcal{B}, \mu)$ along some averaging scheme. We will be concerned with characteristic factors for convergence in density. Recall that polynomials $p_1, \ldots, p_k$ over a ring are said to be **essentially distinct** if $p_i - p_j$ is not constant for all $i \neq j$. Our main goal in this section is the following theorem.

**Theorem 5.5.2.** Let $L$ be an algebraic number field and fix $p_1, \ldots, p_k$ in $\mathcal{O}_L[x_1, \ldots, x_d]$ that are non-constant and essentially distinct. For any ergodic action $T$ of the additive group of $\mathcal{O}_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$ there is $r \in \mathbb{N}$ such that

$$\text{dlim}_{u \to \Phi} \int f \cdot T^{p_1(u)} f \cdots T^{p_k(u)} f - \mathbb{E}(f|\mathcal{Z}_r) \cdot T^{p_1(u)} \mathbb{E}(f|\mathcal{Z}_r) \cdots T^{p_k(u)} \mathbb{E}(f|\mathcal{Z}_r) \, d\mu = 0$$

for any Følner sequence $\Phi$ in $\mathcal{O}_L^d$ and any $f_1, \ldots, f_k$ in $L^\infty(X, \mathcal{B}, \mu)$.

The remainder of this section constitutes a proof of Theorem 5.5.2. Essentially, we follow Leibman's proof [Lei05a] of convergence of averages of the form (5.5.1) for $\mathbb{Z}$-actions to show that the limiting behavior of (5.5.1) along any Følner sequence is controlled by a certain Gowers-Host-Kra seminorm, and then apply Theorem 5.4.3. For this reason we prove only the results that require some modification for our setting. We then use Proposition 5.4.7 to obtain characteristic factors for convergence in density from those obtained for Cesàro convergence.

We begin with the following lemma.

**Lemma 5.5.3.** Let $p \in \mathcal{O}_L[x_1, \ldots, x_d]$ be a degree 1 polynomial with zero constant term. There is a constant $c \geq 0$ such that

$$\lim_{N \to \infty} \left\| \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} T^{p(u)} f \right\| \leq c \|f\|_2$$

(5.5.4)

for any $f$ in $L^\infty(X)$ and any Følner sequence $\Phi$ in $\mathcal{O}_L^d$.

**Proof.** Write $p(x_1, \ldots, x_d) = a_1 x_1 + \cdots + a_d x_d$ for some $a_i$ in $\mathcal{O}_L$, not all of which are zero. By the mean ergodic theorem we have

$$\lim_{N \to \infty} \left\| \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} T^{p(u)} f \right\|^2 = \|\mathbb{E}(f|\mathcal{Z}_r)\|^2$$

(5.5.5)
where $\mathcal{S}_a$ is the sub-$\sigma$-algebra of sets invariant under $T_a$ for all $a$ in the ideal $a$ generated by $\{a_1, \ldots, a_d\}$. By Lemma 5.2.1 the ideal $a$ is a finite-index subgroup. Thus

$$
\lim_{N \to \infty} \frac{[O_L : a]}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \| f \cdot T^u f \|_1 \geq \lim_{N \to \infty} \frac{1}{\vert \Phi_N \cap a \vert} \sum_{u \in \Phi_N \cap a} \| f \cdot T^u f \|_1 \geq \lim_{N \to \infty} \frac{1}{\vert \Phi_N \cap a \vert} \sum_{u \in \Phi_N \cap a} \| f \cdot T^u f \|_0 = |E(f | \mathcal{S}_a)|^2
$$

for any $f$ in $L^\infty(X)$ by Lemma 2.7.2, (5.4.1) and the mean ergodic theorem. Combining the above with (5.5.5) and Cauchy-Schwarz gives us

$$
\lim_{N \to \infty} \left\| \frac{1}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} T^{p(u)} f \right\|^2 \leq [O_L : a] \left( \lim_{N \to \infty} \frac{1}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \| f \cdot T^u f \|_1^2 \right)^{1/2}
$$

which, upon applying (5.4.2), yields (5.5.4) with $c^2 = [O_L : a]$.  

**Lemma 5.5.6.** Let $p \in O_L[x_1, \ldots, x_d]$ be a degree 1 polynomial with zero constant term. There is a constant $c \geq 0$ such that

$$
\lim_{N \to \infty} \frac{1}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \| f \cdot T^{p(u)} f \|_k^{2k} \leq c \| f \|_k^{2k+1}
$$

(5.5.7)

for every $f$ in $L^\infty(X)$, every Følner sequence $\Phi$ in $O_L^d$ and every $k$ in $\mathbb{N}$.

**Proof.** Write $p(x_1, \ldots, x_d) = a_1 x_1 + \cdots + a_d x_d$ for some $a_i$ in $O_L$ not all of which are zero, and let $a$ be the ideal in $O_L$ generated by $\{a_1, \ldots, a_d\}$. Let $\mathcal{S}_a$ be the sub-$\sigma$-algebra of $\mathcal{B}[k]$ consisting of sets that are invariant under $T_a[k]$ for all $a$ in $a$. For any Følner sequence $\Phi$ in $O_L^d$ and any $f$ in $L^\infty(X)$ we have

$$
\lim_{N \to \infty} \frac{1}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \| f \cdot T^{p(u)} f \|_k^{2k} = \lim_{N \to \infty} \frac{1}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \int f^{[k]}(T_a[k] T^{p(u)} f^{[k]} d\mu[k] \\
= \int E(f^{[k]} | \mathcal{S}_a)^2 d\mu[k] \\
\leq \lim_{N \to \infty} \frac{[O_L : a]}{\vert \Phi_N \vert} \sum_{u \in \Phi_N} \| f \cdot T^u f \|_k^{2k} = [O_L : a] \| f \|_k^{2k+1}
$$

by arguing as in Lemma 5.5.3. 

The next step is to obtain a version of Lemma 5.5.6 for multiple recurrence.
**Theorem 5.5.8.** Let \( p_1, \ldots, p_k \in \mathcal{O}[x_1, \ldots, x_d] \) be non-constant, essentially distinct linear polynomials with zero constant term. There is a constant \( c \geq 0 \) such that

\[
\limsup_{N \to \infty} \left\| \frac{1}{\Phi_N} \sum_{u \in \Phi_N} T^{p_i(u)} f_k \right\|_2 \leq c \| f_1 \|_\infty \cdots \| f_k \|_\infty
\]

for any \( f_1, \ldots, f_k \in L^\infty(X) \) and any Følner sequence \( \Phi \) in \( \mathcal{O}_L^d \).

**Proof.** The proof is by induction of \( k \). When \( k = 1 \) this is just Lemma 5.5.3. Put \( g(u) = T^{p_1(u)} f_1 \cdots T^{p_k(u)} f_k \) for each \( u \) in \( \mathcal{O}_L^d \) and note that in \( L^2(X) \) we have

\[
\langle g(u+h), g(u) \rangle = \int \prod_{i=1}^k T^{p_i(u)} (f_i \cdot T^{p_i(h)} f_i) \, d\mu
\]

\[
= \int f_k \cdot T^{p_k(h)} f_k \prod_{i=1}^{k-1} T^{p_i(u)-p_i(u)} (f_i \cdot T^{p_i(h)} f_i) \, d\mu
\]

so for any \( H \) in \( \mathbb{N} \) we have

\[
\limsup_{N \to \infty} \left\| \frac{1}{\Phi_N} \sum_{u \in \Phi_N} \prod_{i=1}^k T^{p_i(u)} f_i \right\|^2_2 \leq \frac{1}{\Phi_H^2} \sum_{h,i \in \Phi_H} \left\| f_i \right\|_\infty^2 \limsup_{N \to \infty} \left\| \frac{1}{\Phi_N} \sum_{u \in \Phi_N} \prod_{i=1}^{k-1} T^{p_i(u)-p_i(u)} (T^{p_i(h)} f_i \cdot T^{p_i(h)} f_i) \right\|
\]

\[
\leq \frac{1}{\Phi_H^2} \sum_{h,i \in \Phi_H} C \| f_1 \cdot T^{p_1(l-h)} f_1 \|_\infty \| f_2 \|_\infty^2 \cdots \| f_k \|_\infty^2
\]

by the van der Corput trick (Proposition 2.3.8) and induction. Applying Cauchy-Schwarz a number of times and then Lemma 5.5.6 gives the desired result. \( \square \)

Using a PET induction argument exactly as in [Lei05a], one can use Theorem 5.5.8 to obtain the following result, which gives characteristic factors for Cesàro averages.

**Theorem 5.5.9.** For any finite collection of non-constant, essentially distinct polynomials \( p_1, \ldots, p_k \) in \( \mathcal{O}[x_1, \ldots, x_d] \) there is \( r \) in \( \mathbb{N} \) such that for any Følner sequence \( \Phi \) in \( \mathcal{O}_L^d \), any action \( T \) of \( \mathcal{O}_L \) on a compact metric probability space \( (X, \mathcal{B}, \mu) \) and any \( f \) in \( L^\infty(X, \mathcal{B}, \mu) \) we have

\[
\lim_{u \to \Phi} \int f \cdot T^{p_1(u)} f \cdots T^{p_k(u)} f - \mathbb{E}(f | \mathcal{F}) \cdot T^{p_1(u)} \mathbb{E}(f | \mathcal{F}) \cdots T^{p_k(u)} \mathbb{E}(f | \mathcal{F}) = 0
\]

whenever \( \| f \|_r = 0 \).

The next step is to obtain a version of Theorem 5.5.9 for convergence in density. To do so we use product systems as in [BHK05]. Let \( p_1, \ldots, p_k \) be non-constant, essentially distinct polynomials in \( \mathcal{O}[x_1, \ldots, x_d] \) and let \( r \geq 1 \) be as in Theorem 5.5.9. Fix an ergodic
action $T$ of $O_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$ and let $\mu_s$ be the ergodic decomposition of $\mu \otimes \mu$. If $f$ in $L^\infty(X, \mathcal{B}, \mu)$ satisfies $\|f \otimes f\|_{s,r} = 0$ then
\[
\lim_{u \to \Phi} \int (f \otimes f) \cdot (T \times T)^{p_1(u)}(f \otimes f) \cdots (T \times T)^{p_k(u)}(f \otimes f) \, d\mu_s = 0 \tag{5.5.10}
\]
for any Følner sequence $\Phi$ in $O_L^d$. But from Proposition 5.4.7, if $\|f\|_{r+1} = 0$ then $\|f \otimes f\|_{s,r} = 0$ for almost every $s$, so (5.5.10) holds for almost every $s$. Integrating over $s$ concludes the proof of Theorem 5.5.2.

5.6 Multiple recurrence for polynomials over rings of integers

Let $T$ be an ergodic action of $O_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$. In the previous section we showed that, by neglecting a set of zero Banach density, it suffices to study the average (5.5.1) when $(X, \mathcal{B}, \mu)$ is an inverse limit of nilrotations. The goal of this section is to prove Theorem 5.1.5. We do so by exhibiting largeness of the set of multiple recurrence times for nilrotations.

**Theorem 5.6.1.** Let $L$ be an algebraic number field. For any jointly intersective polynomials $p_1, \ldots, p_k$ in $O_L[x_1, \ldots, x_d]$, any ergodic action $T$ of $O_L$ on a nilmanifold $(G/\Gamma, m)$ determined by a homomorphism $a : O_L \to G$, and any $B \subset G/\Gamma$ with $m(B) > 0$ there is $c > 0$ for which the set
\[
\left\{ u \in O_L^d : \int 1_B \cdot T^{p_1(u)}1_B \cdots T^{p_k(u)}1_B \, dm \geq c \right\} \tag{5.6.2}
\]
is AIP$^*$.  

**Proof.** Let $e_1, \ldots, e_m$ be a basis for $O_L$ thought of as a $\mathbb{Z}$-module. Using this basis we can identify $O_L^d$ with $\mathbb{Z}^{dm}$. For each $1 \leq i \leq k$ define polynomials $p_{i,1}, \ldots, p_{i,m} : \mathbb{Z}^{md} \to \mathbb{Z}$ by
\[
p_i(u) = p_{i,1}(u)e_1 + \cdots + p_{i,m}(u)e_m
\]
for each $u$ in $O_L^d$.

We claim that the polynomials $\{p_{i,j} : 1 \leq i \leq k, 1 \leq j \leq m\}$ are jointly intersective. Indeed, fix $\lambda$ in $\mathbb{Z} \setminus \{0\}$. Since $p$ is intersective there is $\zeta$ in $O_L^d$ such that $\{p_1(\zeta), \ldots, p_k(\zeta)\} \subset (\lambda)$. This means that, for each $i$, we can find $t_1, \ldots, t_m$ in $O_L$ such that
\[
p_{i,1}(\zeta)e_1 + \cdots + p_{i,m}(\zeta)e_m = \lambda(t_1e_1 + \cdots + t_me_m)
\]
from which it follows that $\lambda|p_{i,j}(\zeta)$.

Next, we show that (5.6.2) is syndetic following [BLL08]. Fix a nilpotent Lie group $G$ and a closed, cocompact subgroup $\Gamma$. Let $m$ be the $G$-invariant probability measure.
on the quotient \( X := G/\Gamma \). Fix \( B \subseteq X \) with \( m(B) > 0 \). Let \( a : \mathcal{O}_L \to G \) be a group homomorphism and let \( T \) be the induced action of \( \mathcal{O}_L \) on \( G/\Gamma \). Put \( a_i = a(e_i) \). Then
\[
a(p_i(u)) = a(p_i,1(u))e_1 + \cdots + a(p_i,m(u))e_m = a_{1}^{p_i,1(u)} \cdots a_{m}^{p_i,m(u)}
\]
for each \( 1 \leq i \leq k \) and every \( u \) in \( \mathbb{Z}^{dm} \). Define a polynomial sequence \( g : \mathbb{Z}^{dm} \to G^{k+1} \) by
\[
g(u) = (1, a_1^{p_1,1(u)} \cdots a_m^{p_1,m(u)}, \ldots, a_1^{p_k,1(u)} \cdots a_m^{p_k,m(u)})
\]
for all \( u \) in \( \mathbb{Z}^{dm} \). Let \( \Delta \) be the diagonal in \( X^{k+1} \) and let \( m_\Delta \) be the push-forward of \( m \) under the embedding of \( X \) in \( \Delta \). By [Lei05b], the closure
\[
Y = \bigcup \{g(u)\Delta : u \in \mathcal{O}_L^d\}
\]
is a finite union of sub-nilmanifolds of \( X^{k+1} \) and the sequence \( u \mapsto g(u)m_\Delta \) has an asymptotic distribution \( \mu \) in its orbit closure that is a convex combination of the Haar measures on the connected components of \( Y \). Thus we have
\[
\lim_{u \to \Phi} \int f_0 \cdot T^{p_1(u)}f_1 \cdots T^{p_k(u)}f_k \ dm
= \lim_{u \to \Phi} \int f_0 \otimes T^{p_1(u)}f_1 \otimes \cdots \otimes T^{p_k(u)}f_k \ dm_\Delta
= \lim_{u \to \Phi} \int f_0 \otimes f_1 \otimes \cdots \otimes f_k \ dm_{g(u)\Delta}
= \int f_0 \otimes f_1 \otimes \cdots \otimes f_k \ d\mu
\]
for any continuous functions \( f_0, f_1, \ldots, f_k : X \to \mathbb{R} \) and any Følner sequence \( \Phi \) in \( \mathbb{Z}^{dm} \). A density argument proves that the same is true for any \( f_0, f_1, \ldots, f_k \) in \( L^\infty(X) \). Thus for any \( B \) in \( \mathcal{B} \) we have
\[
\lim_{u \to \Phi} m(B \cap T^{-p_1(u)}B \cap \cdots \cap T^{-p_k(u)}B) = \mu(B^{k+1})
\]
for every Følner sequence \( \Phi \) in \( \mathbb{Z}^{dm} \). Following the argument on Page 376 of [BLL08] and applying [BLL08, Proposition 2.4] yields
\[
\lim_{u \to \Phi} \int 1_B \cdot T^{p_1(u)}1_B \cdots T^{p_k(u)}1_B \ dm > 0
\]
for every Følner sequence \( \Phi \) in \( \mathbb{Z}^{dm} \). By Lemma 2.7.6 there is some \( c > 0 \) such that
\[
\lim_{u \to \Phi} \int 1_B \cdot T^{p_1(u)}1_B \cdots T^{p_k(u)}1_B \ dm \geq c
\]
for every \( \Phi \). Thus
\[
\left\{ u \in \mathcal{O}_L^d : \int 1_B \cdot T^{p_1(u)}1_B \cdots T^{p_k(u)}1_B \ dm \geq \frac{c}{2} \right\}
\]
(5.6.3)
has positive density with respect to every Følner sequence and is therefore syndetic by Lemma 2.7.8.

It remains to prove (5.6.3) is \( \text{AIP}^*_+ \). Fix a continuous function \( f : X \to [0,1] \) with \( \|1_B - f\|_1 < c/(k+1) \). Define \( \varphi : \mathcal{O}^d_L \to \mathbb{R} \) by
\[
\varphi(u) = \int f \cdot T^{p_1(u)} f \cdots T^{p_k(u)} f \, dm
\]
for every \( u \in \mathcal{O}^d_L \). By [Lei15, Theorem 4.3] we can write \( \varphi \) as a sum of sequences \( \phi + \psi \) where \( \phi \) is a nilsequence and
\[
\lim_{u \to \Phi} \psi(u) = 0.
\]
for every Følner sequence. Thus there is a nilmanifold \( \tilde{X} = \hat{G}/\hat{\Gamma} \), a homomorphism \( b : \mathcal{O}^d_L \to \tilde{G} \), a continuous function \( h : \tilde{X} \to \mathbb{R} \) and some \( x \in \tilde{X} \) such that \( \phi(u) = h(b(u)x) \) for all \( u \in \mathcal{O}^d_L \). Combining the above, we obtain
\[
\left| \int 1_B \cdot T^{p_1(u)} 1_B \cdots T^{p_k(u)} 1_B \, dm - h(b(u)x) \right| \leq \frac{c}{8} + |\psi(u)|
\]
for every \( u \in \mathcal{O}^d_L \). The set \( \{ u \in \mathcal{O}^d_L : |\psi(u)| > c/8 \} \) has zero upper Banach density so syndeticity of (5.6.3) and Lemma 2.7.5 imply that \( h(b(w)x) \geq c/8 \) for some \( w \in \mathcal{O}^d_L \). The nilrotation \( b \) determines is distal by [Key66, Theorem 2.2], so
\[
\lim_{v \to p} h(b(v+w)x) = h(b(w)x)
\]
for every idempotent ultrafilter \( p \) in \( \beta \mathcal{O}^d_L \) by Lemma 2.7.18. It follows that
\[
\{ u \in \mathcal{O}^d_L : h(b(u)x) \geq c/8 \}
\]
is \( \text{IP}^*_+ \). Finally, (5.6.3) is \( \text{AIP}^*_+ \) as desired.

\[\Box\]

In order to deduce Theorem 5.1.5 from Theorem 5.6.1 we need the following preliminary result, based on [FKO82, Proposition 7.1].

**Proposition 5.6.5.** Fix a countable, commutative ring \( R \) and polynomials \( p_1, \ldots, p_l \) in \( R[x_1, \ldots, x_d] \). Let \( (X, \mathcal{B}, \mu) \) be a compact metric probability space and let \( T \) be an action of the additive group of \( R \) on \( (X, \mathcal{B}, \mu) \) by measurable, measure-preserving maps. Fix \( B \in \mathcal{B} \) with \( \mu(B) > 0 \). For any countably generated \( T \)-invariant sub-\( \sigma \)-algebra \( \mathcal{D} \subset \mathcal{B} \) and any \( D \in \mathcal{D} \) with \( \mu(B \triangle D) < \mu(B)/8l \) we can find \( E \in \mathcal{D} \) with \( \mu(E) > 0 \) such that
\[
\int T^{p_1(u)} 1_B \cdots T^{p_l(u)} 1_B \, d\mu \geq \frac{1}{2} \int T^{p_1(u)} 1_E \cdots T^{p_l(u)} 1_E \, d\mu
\]
for every \( u \in R \).
Proof. We have $\mu(D) \geq \mu(B) - \mu(D)/8l > 0$ because $|\mu(B) - \mu(D)| \leq \mu(B \triangle D)$. Let $x \mapsto \mu_x$ be a disintegration of $\mu$ over $\mathcal{D}$. Put

$$E = \{x \in D : \mu_x(B) > 1 - 1/2l\}$$

and note that

$$\mu(D \setminus B) = \iint 1_D 1_{X \setminus B} \, d\mu_x \, d\mu(x)$$
$$= \int 1_D(x) \mu_x(X \setminus B) \, d\mu(x)$$
$$\geq \int 1_D \triangle E(x) (1 - \mu_x(B)) \, d\mu(x) \geq \frac{\mu(D \setminus E)}{2l}$$

implies $\mu(D \setminus E) < \mu(B)/4$ as otherwise $\mu(B \triangle D) < \mu(B)/8l$ is contradicted. Thus $\mu(E) \geq \mu(B)/2$. Fix $u \in R$. If $x \in T^{-p_l(u)} E$ then $\mu_x(T^{-p_l(u)} B) > 1 - 1/2l$ because $\mathcal{D}$ is $T$-invariant. Thus if $x \in T^{-p_l(u)} E \cap \cdots \cap T^{-p_l(u)} E$ we have

$$\mu_x(T^{-p_l(u)} B \cap \cdots \cap T^{-p_l(u)} B) > \frac{1}{2}$$

and integrating over $T^{-p_l(u)} E \cap \cdots \cap T^{-p_l(u)} E$ gives (5.6.6). □

Here is the proof of Theorem 5.1.5.

Proof of Theorem 5.1.5. Let $T$ be an ergodic action of $\mathcal{O}_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$ and fix $B \in \mathcal{B}$ with $\mu(B) > 0$. Let $r$ be as in Theorem 5.5.2. Put $h = \mathbb{E}(1_B | \mathcal{Z}_r)$. We can assume that the polynomials $p_1, \ldots, p_k$ in $\mathcal{O}_L[x_1, \ldots, x_d]$ are distinct. Since distinct, jointly intersective polynomials are always essentially distinct, for every $\varepsilon > 0$ the set

$$\left\{ u \in \mathcal{O}_L^d : \left| \int 1_B \cdot T^{p_1(u)} 1_B \cdots T^{p_k(u)} 1_B \, d\mu - \int h \cdot T^{p_1(u)} h \cdots T^{p_k(u)} h \, d\mu \right| \geq \varepsilon \right\}$$

has zero upper Banach density by Theorem 5.5.2. Since $h$ is positive on $B$ we can find $C \in \mathcal{B}$ and $a > 0$ such that $a1_C \leq h$.

The factor corresponding to $\mathcal{Z}_r$ is an inverse limit of nilrotations by Theorem 5.4.3. Thus we can find a Borel subset $D$ of a nilrotation such that $\mu(C \triangle D) \leq \mu(C)/8(k + 1)$. Combining Proposition 5.6.5 with Theorem 5.6.1 implies there is some $c > 0$ such that

$$\left\{ u \in \mathcal{O}_L^d : \int h \cdot T^{p_1(u)} h \cdots T^{p_k(u)} h \, d\mu \geq c \right\}$$

is AIP$^*_+$. Picking $\varepsilon = c/2$ proves that (5.1.6) is also AIP$^*_+$ as desired. □

We conclude this section by giving a proof of Theorem 5.1.13.
Proof of Theorem 5.1.13. Let $T$ be an action of $\mathcal{O}_L$ on a compact metric probability space $(X, \mathcal{B}, \mu)$ and fix $B \in \mathcal{B}$ with $\mu(B) > 0$. Let $\mu_x$ be an ergodic decomposition for $\mu$. For almost every $x \in B$ we have $\mu_x(B) > 0$ so there is a constant $c_x > 0$ such that

$$R_x = \{ u \in \mathcal{O}_L^d : \mu_x(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) \geq c_x \}$$

is AIP$_r$ by Theorem 5.1.5 and therefore syndetic by Lemma 2.7.14. Thus for every Følner sequence $\Phi$ in $\mathcal{O}_L^d$ we have

$$\lim\inf_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} \mu_x(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) > 0$$

for almost every $x \in B$. Integrating over $B$ and applying Fatou’s lemma gives

$$\lim\inf_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} \int \mu_x(B \cap T^{p_1(u)}B \cap \cdots \cap T^{p_k(u)}B) \, d\mu > 0$$

so (5.1.14) is syndetic by Lemma 2.7.8.

5.7 Single recurrence for polynomials over countable fields

In this section we give a short proof of polynomial recurrence with large intersection for additive actions of finite-dimensional vector spaces over countable fields on probability spaces, strengthening and extending recent results from [MW14] regarding actions of fields having finite characteristic.

To state our result we recall some definitions from [BLM05]. Fix a countable field $F$. By a monomial we mean a mapping $F^n \to F$ of the form $(x_1, \ldots, x_n) \mapsto ax_1^{d_1} \cdots x_n^{d_n}$ for some $a \in F$ and integers $d_1, \ldots, d_n \geq 0$ not all zero. Let $V$ and $W$ be finite-dimensional vector spaces over $F$. A mapping $F^n \to W$ is a polynomial if it is a linear combination of vectors with monomial coefficients. A mapping $V \to W$ is a polynomial if, in terms of a basis of $V$ over $F$, it is a polynomial mapping $F^n \to W$. Note that whether a mapping is polynomial or not is independent of the basis chosen.

Definition 5.7.1. A subset $S$ of a countable, discrete, abelian group $G$ is said to be AIP$_r$ if it is of the form $A \setminus B$ where $A$ is IP$_r$ and $d(B) = 0$.

Here is our main result.

Theorem 5.7.2. Let $W$ be a finite-dimensional vector space over a countable field $F$ and let $T$ be an action of the additive group of $W$ on a probability space $(X, \mathcal{B}, \mu)$. For any polynomial $\phi : F^n \to W$, any $B \in \mathcal{B}$ and any $\varepsilon > 0$ the set

$$\{ u \in F^n : \mu(B \cap T^{\phi(u)}B) > \mu(B)^2 - \varepsilon \}$$

(5.7.3)

is AIP$_r$ for some $r \in \mathbb{N}$. 78
First we note that we may assume, by restricting our attention to the sub-$\sigma$-algebra generated by the orbit of $B$, that the probability space $(X, \mathcal{B}, \mu)$ is separable.

We begin with a corollary of the Hales-Jewett theorem, the density version of which was stated in Chapter 1 as Theorem 1.1.6. Write $\mathcal{P}A$ for the set of all subsets of a set $A$. Recall that, given $k, m \in \mathbb{N}$, a **combinatorial line** in $[k]^d$ is specified by a partition $W_0 \sqcup W_1$ of $[m]$ with $W_1 \neq \emptyset$ and a function $c : W_0 \to [k]$, and consists of all words $w \in [k]^d$ that agree with $c$ on $W_0$ and are constant on $W_1$. By an $r$-coloring of $[k]^d$ we mean a map $[k]^d \to [r]$. A subset of $[k]^d$ is **monochromatic** for a given coloring if it is contained in a single fibre of the coloring map. With these definitions we can state the Hales-Jewett theorem.

**Theorem 5.7.4** ([HJ63]). For every $k, r \in \mathbb{N}$ there is $\text{HJ}(k, r) \in \mathbb{N}$ such that for any $d \geq \text{HJ}(k, r)$ and any $r$-coloring of $[k]^d$ one can find a combinatorial line that is monochromatic.

**Corollary 5.7.5.** For any $k, r \in \mathbb{N}$ there is $\text{HJS}(k, r) \in \mathbb{N}$ such that any $r$-coloring

\[(\mathcal{P}\{1, \ldots, d\})^k \to \{1, \ldots, r\}\]  

with $d \geq \text{HJS}(k, r)$ contains a monochromatic configuration of the form

\[\{(\alpha_1 \cup \eta_1, \ldots, \alpha_k \cup \eta_k) : (\eta_1, \ldots, \eta_k) \in \{\emptyset, \gamma\}^k\}\]  

for some $\gamma, \alpha_1, \ldots, \alpha_k \subseteq \{1, \ldots, d\}$ with $\gamma$ non-empty and $\gamma \cap \alpha_i = \emptyset$ for each $1 \leq i \leq k$.

**Proof.** Let $\text{HJS}(k, r) = \text{HJ}(2^k, r)$. Fix $d \geq \text{HJS}(k, r)$ and an $r$-coloring (5.7.6). Define a map $\psi : [2^k]^d \to (\mathcal{P}[r])^k$ by declaring $\psi(w) = (\alpha_1, \ldots, \alpha_k)$ where $\alpha_i$ consists of those $j \in [d]$ for which the binary expansion of $w(j) - 1$ has a 1 in the $i$th position. Combinatorial lines in $[2^k]^d$ correspond via this map to configurations of the form (5.7.7) in $(\mathcal{P}[d])^k$. \hfill $\square$

We will use this corollary of the Hales-Jewett theorem to derive the following topological recurrence result. Given $n \in \mathbb{N}$ and a ring $R$, by a **monomial mapping** from $R^n$ to $R$ we mean any map of the form $(x_1, \ldots, x_n) \mapsto ax_1^{d_1} \cdots x_n^{d_n}$ for some $a \in R$ and some $d_1, \ldots, d_n \geq 0$ not all zero.

**Proposition 5.7.8** (cf [Ber10, Theorem 7.7]). Let $R$ be a commutative ring and let $T$ be an action of the additive group of $R$ on a compact metric space $(X, d)$ by isometries. For any monomial mapping $\phi : R^n \to R$, any $x \in X$ and any $\varepsilon > 0$ there is $r \in \mathbb{N}$ such that the set

\[\{u \in R^n : d(T^{\phi(u)}x, x) < \varepsilon\}\]

is $\text{IP}^*_r$.  

79
Proof. Write \( \phi(x_1, \ldots, x_n) = ax_1^{d_1} \cdots x_n^{d_n} \) for some \( a \in R \) and some \( d_i \geq 0 \) not all zero. Let \( d = d_1 + \cdots + d_n \). Put \( e_0 = 0 \) and \( e_i = d_1 + \cdots + d_i \) for each \( 1 \leq i \leq n \). Fix \( x \in X \) and \( \varepsilon > 0 \). Let \( V_1, \ldots, V_t \) be a cover of \( X \) by balls of radius \( \varepsilon/2^d \). Let \( r = HJS(d, t) \) be as in Corollary 5.7.5. Fix \( u_1, \ldots, u_t \) in \( R^n \). Given \( \alpha \subset \{1, \ldots, r\} \) write \( u_\alpha \) for \( \sum\{u_i : i \in \alpha\} \) and \( u_\alpha(i) \) for the \( i \)th coordinate of \( u_\alpha \). By choosing for each \( (\alpha_1, \ldots, \alpha_d) \in (P\{1, \ldots, r\})^d \) the minimal \( 1 \leq i \leq t \) such that

\[
T(au_{\alpha_1}(1) \cdots u_{\alpha_{i-1}}(1) \cdots u_{\alpha_{n-1}+1}(n) \cdots u_{\alpha_n}(n))x \in V_i
\]

we obtain via Theorem 5.7.5 sets \( \alpha_1, \ldots, \alpha_d, \gamma \subset \{1, \ldots, r\} \) with \( \gamma \) non-empty and disjoint from all \( \alpha_i \) which, combined with the expansion

\[
a_{\gamma}(1)^{d_1} \cdots u_{\gamma}(n)^{d_n} = a \prod_{k=1}^{n} \prod_{i=e_{k-1}+1}^{e_k} u_{\gamma}(k) + u_{\alpha_i}(k) - u_{\alpha_i}(k)
\]

and the fact that \( T \) is an isometry, yields \( d(T^{\phi(u)})x, x) < \varepsilon \) as desired. \( \square \)

As we now show, iterating the previous result yields a version for commuting actions of rings.

Corollary 5.7.9. Let \( R \) be a commutative ring and let \( T_1, \ldots, T_k \) be commuting actions of the additive group of \( R \) on a compact metric space \( (X,d) \) by isometries. For any monomial mappings \( \phi_1, \ldots, \phi_k : R^n \to R \), any \( x \in X \) and any \( \varepsilon > 0 \), there is \( r \in \mathbb{N} \) such that

\[
\{u \in R^n : d(T_1^{\phi_1(u)} \cdots T_k^{\phi_k(u)}x, x) < \varepsilon\}
\]

is IP*.

Proof. Fix \( 1 \leq i \leq k \). By applying Proposition 5.7.8 to the \( R \) action \( r \mapsto T_i^r \), we can find \( r_i \in \mathbb{N} \) such that

\[
Z_i = \{u \in R^n : d(T_i^{\phi_i(u)}x, x) < \varepsilon/k\}
\]

is IP*\( r_i \). By Proposition 2.8.5, the intersection \( Z_1 \cap \cdots \cap Z_k \) is IP* for some \( r \in \mathbb{N} \). Since the \( T_i \) are isometries, it follows that (5.7.10) is IP* as desired. \( \square \)

Combining the preceding lemma with the following facts from [BLM05] will lead to a proof of Theorem 5.7.2. Let \( \phi : V \to W \) be a polynomial and let \( T \) be an action of \( W \) on a probability space \( (X, \mathcal{B}, \mu) \). Assume that \( \phi V \) spans \( W \). As in [BLM05], say that \( f \) in \( L^2(X, \mathcal{B}, \mu) \) is weakly mixing for \( (T, \phi) \) if \( \text{dlim}_v(T^{\phi(v)}f, g) = 0 \) for all \( g \) in \( L^2(X, \mathcal{B}, \mu) \), where \( \text{dlim} \) denotes convergence with respect to the filter of sets whose complements have zero upper Banach density. This is the same as strong Cesàro convergence along every Følner sequence in \( V \). Call \( f \in L^2(X, \mathcal{B}, \mu) \) compact for \( T \) if \( \{T^v f : v \in V\} \) is pre-compact in the norm topology. Denote by \( \mathcal{H}_{wm}(T, \phi) \) the closed subspace of \( L^2(X, \mathcal{B}, \mu) \) spanned
by functions that are weakly mixing for \((T, \phi)\), and let \(\mathcal{H}_c(T)\) be the closed subspace of \(L^2(X, \mathcal{B}, \mu)\) spanned by functions compact for \(T\). We have \(L^2(X, \mathcal{B}, \mu) = \mathcal{H}_c(T) \oplus \mathcal{H}_{wm}(T, \phi)\) by [BLM05, Theorem 3.17].

**Proof of Theorem 5.7.2.** Write \(\phi = \phi_1 w_1 + \cdots + \phi_k w_k\) where the \(\phi_i\) are monomials \(F^n \to F\) and the \(w_i\) belong to \(V\). Fix \(B\) in \(\mathcal{B}\) and \(\varepsilon > 0\). Let \(f = P1_B\) be the orthogonal projection of \(1_B\) on \(\mathcal{H}_c(T)\). Let \(\Omega\) be the orbit closure of \(f\) in the norm topology under \(T\). Since \(f\) is compact, \(\Omega\) is a compact metric space. Applying Lemma 5.7.9 to the \(F\) actions \(x \mapsto T^xw_i\) and monomials \(\phi_i\) for \(1 \leq i \leq k\), we see that

\[\{ u \in F^n : \|f - T^{\phi(u)}f\| < \varepsilon/2 \}\]

is \(IP^*_r\). We have

\[\langle T^{\phi(u)}1_B, 1_B \rangle = \langle T^{\phi(u)}f, 1_B \rangle + \langle T^{\phi(u)}(1_B - f), 1_B \rangle\]

so the set

\[\{ u \in F^n : \langle T^{\phi(u)}1_B, 1_B \rangle \geq \langle f, 1_B \rangle - \varepsilon/2 + \langle T^{\phi(u)}(1_B - f), 1_B \rangle\]

is \(IP^*_r\). Since \(1_B - f\) is weakly mixing for \((T, \phi)\) the set

\[\{ u \in F^n : \langle T^{\phi(u)}1_B, 1_B \rangle \geq \langle f, 1_B \rangle - \varepsilon \}\]

is \(AIP^*_r\). Thus (5.7.3) is \(AIP^*_r\) by

\[\langle f, 1_B \rangle = \langle P1_B, P1_B \rangle \langle 1, 1 \rangle \geq \langle P1_B, 1 \rangle^2 = \mu(B)^2\]

as desired. \(\square\)

We obtain as a corollary the following result from [MW14], which uses the following terminology. Let again \(G\) be an abelian group. An ultrafilter \(p\) on \(G\) is **essential** if it is idempotent and \(d(A) > 0\) for all \(A \in p\). A **D set** in \(G\) is any subset of \(G\) that belongs to an essential ultrafilter on \(G\), and a subset of \(G\) is \(D^*\) if its intersection with any \(D\) is non-empty.

**Corollary 5.7.11** ([MW14, Corollary 5]). Let \(F\) be a countable field of finite characteristic and let \(p : F \to F^n\) be a polynomial mapping. For any action \(T\) of \(F^n\) on a probability space \((X, \mathcal{B}, \mu)\), any \(B\) in \(\mathcal{B}\) and any \(\varepsilon > 0\) the set

\[\{ x \in F : \mu(B \cap T^{p(x)}B) \geq \mu(B)^2 - \varepsilon \}\]

is \(D^*\).
Proof. It follows from the proof of Theorem 5.7.2 that (5.7.12) is of the form $A \setminus B$ where $A$ is $IP^*_r$ for some $r \in \mathbb{N}$ and $B$ has zero upper Banach density. Any $IP^*_r$ subset of $G$ is $IP^*$ and therefore belongs to every idempotent ultrafilter on $G$, so $A$ certainly belongs to every essential ultrafilter on $G$. By the filter property, removing from $A$ a set of zero upper Banach density does not change this fact, because every set in an essential idempotent has positive upper Banach density.

It has recently been shown [MZ14] that there are $D^*$ subsets of $\mathbb{Z}$ that are not $AIP^*$. This is also the case in countable fields of finite characteristic [McC14]. Thus our result constitutes a genuine strengthening of Corollary 5.7.11.
CHAPTER 6

SATEDNESS

We have given in Chapter 4 a description of characteristic factors for certain correlations arising from commuting actions of amenable groups. However, the factors described there do not seem to provide enough structure to deduce a proof of multiple recurrence. In this chapter we describe an alternative approach, introduced by Austin [Aus10], [Aus13], to obtaining characteristic factors. This technique will be employed in subsequent chapters to deduce more combinatorial results.

6.1 Relative independence

In this section we reproduce some basic facts about relatively independent \( \sigma \)-algebras, which we will need to discuss Austin’s notion of satedness. See [Tao07, Appendix A] for more on relatively independence.

Let \((X, \mathcal{B}, \mu)\) be a probability space. Two sub-\( \sigma \)-algebras \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) are relatively independent over a sub-\( \sigma \)-algebra \( \mathcal{D} \) if

\[
\mathbb{E}(f_1 \cdot f_2 | \mathcal{D}) = \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 | \mathcal{D})
\]

(6.1.1)

whenever \( f_1 \) is in \( L^\infty(X, \mathcal{C}_1, \mu) \) and \( f_2 \) is in \( L^\infty(X, \mathcal{C}_2, \mu) \). The following characterizations of relative independence will be useful later.

**Theorem 6.1.2.** Let \((X, \mathcal{B}, \mu)\) be a probability space and let \( \mathcal{C}_1, \mathcal{C}_2, \mathcal{D} \) be sub-\( \sigma \)-algebras of \( \mathcal{B} \). The following statements are equivalent:

1. \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) are relatively independent over \( \mathcal{D} \);

2. \( \mathbb{E}(f_1 | \mathcal{D} \vee \mathcal{C}_2) = \mathbb{E}(f_1 | \mathcal{D}) \) whenever \( f_1 \) is in \( L^\infty(X, \mathcal{C}_1, \mu) \);

3. \( \|\mathbb{E}(f_1 | \mathcal{D} \vee \mathcal{C}_2)\| = \|\mathbb{E}(f_1 | \mathcal{D})\| \) whenever \( f_1 \) is in \( L^\infty(X, \mathcal{C}_1, \mu) \).
Proof. The implication (2) ⇒ (3) is immediate. For (3) ⇒ (2) note that

\[ L^2(X, \mathcal{D} \lor \mathcal{C}_2, \mu) = L^2(X, \mathcal{D}, \mu) \oplus L^2(X, \mathcal{D}, \mu) \perp \]

so the result follows from the Pythagoras theorem.

We prove (2) ⇒ (1). Fix \( f_1 \) in \( L^\infty(X, \mathcal{C}_1, \mu) \), \( f_2 \) in \( L^\infty(X, \mathcal{C}_2, \mu) \) and \( h \) in \( L^\infty(X, \mathcal{D}, \mu) \).

Then

\[
\int \mathbb{E}(f_1 \cdot f_2 | \mathcal{D}) \cdot h \, d\mu = \int f_1 \cdot f_2 \cdot h \, d\mu = \int \mathbb{E}(f_1 | \mathcal{D} \lor \mathcal{C}_2) \cdot f_2 \cdot h \, d\mu
\]

\[
= \int \mathbb{E}(f_1 | \mathcal{D}) \cdot f_2 \cdot h \, d\mu = \int \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 | \mathcal{D}) \cdot h \, d\mu
\]

as desired. Lastly, we prove (1) ⇒ (2). Fix \( f_1 \) in \( L^\infty(X, \mathcal{C}_1, \mu) \). For any \( f_2 \) in \( L^\infty(X, \mathcal{C}_2, \mu) \) and any \( h \) in \( L^\infty(X, \mathcal{D}, \mu) \) we have

\[
\int \mathbb{E}(f_1 | \mathcal{D} \lor \mathcal{C}_2) \cdot f_2 \cdot h \, d\mu = \int f_1 \cdot f_2 \cdot h \, d\mu = \int \mathbb{E}(f_1 \cdot f_2 | \mathcal{D}) \cdot h \, d\mu
\]

\[
= \int \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 | \mathcal{D}) \cdot h \, d\mu = \int \mathbb{E}(f_1 | \mathcal{D}) \cdot f_2 \cdot h \, d\mu
\]

concluding the proof.

We will be concerned with relative independence in the case \( \mathcal{D} \subset \mathcal{C}_2 \). The following corollary describes how the equivalent definitions in Theorem 6.1.2 simplify under this additional assumption.

**Corollary 6.1.3.** Let \( (X, \mathcal{B}, \mu) \) be a probability space and let \( \mathcal{C}_1, \mathcal{C}_2, \mathcal{D} \) be sub-\( \sigma \)-algebras of \( \mathcal{B} \) with \( \mathcal{D} \subset \mathcal{C}_2 \). Then the following are equivalent:

1. \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) are relatively independent over \( \mathcal{D} \);
2. \( \mathbb{E}(f_1 | \mathcal{C}_2) = \mathbb{E}(f_1 | \mathcal{D}) \) for all \( f_1 \) in \( L^2(X, \mathcal{C}_1, \mu) \);
3. we have

\[
\int f_1 \cdot f_2 \, d\mu = \int \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 | \mathcal{D}) \, d\mu
\]

whenever \( f_1 \) is in \( L^\infty(X, \mathcal{C}_1, \mu) \) and \( f_2 \) is in \( L^\infty(X, \mathcal{C}_2, \mu) \).

**Proof.** Since \( \mathcal{D} \lor \mathcal{C}_2 = \mathcal{C}_2 \) in this case, the equivalence (1) ⇔ (2) follows from Theorem 6.1.2. That (1) ⇒ (3) follows by integrating (6.1.1). For (3) ⇒ (1) note that, for any \( h \) in \( L^\infty(X, \mathcal{D}, \mu) \) we have

\[
\int \mathbb{E}(f_1 \cdot f_2 | \mathcal{D}) \cdot h \, d\mu = \int f_1 \cdot f_2 \cdot h \, d\mu
\]

\[
= \int \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 \cdot h | \mathcal{D}) \, d\mu = \int \mathbb{E}(f_1 | \mathcal{D}) \cdot \mathbb{E}(f_2 | \mathcal{D}) \cdot h \, d\mu
\]

as desired. \( \square \)
6.2  Sated systems and sated extensions

In this section we recall the definition of Austin’s notion of satedness, which was introduced in [Aus10], and give the proof that sated extensions always exist.

**Definition 6.2.1.** Fix a group $G$. By a functorial sub-$\sigma$-algebra we mean a map $F$ that assigns to every $G$ system $X$ a sub-$\sigma$-algebra $FX$ of $\mathcal{B}$ such that $\pi^{-1}FY \subset FZ$ whenever $\pi : Z \to Y$ is a factor map. (See Figure 6.1 for a schematic of the $\sigma$-algebras involved.)

![Figure 6.1: Functorial sub-$\sigma$-algebra.](image)

We begin by giving some examples of functorial sub-$\sigma$-algebras.

**Example 6.2.2.**

1. Fix a group $G$. The assignment $X \mapsto \mathcal{I}X$, where $\mathcal{I}X$ is the sub-$\sigma$-algebra of sets $B \in \mathcal{B}$ such that $\mu(B \triangle (T^g)^{-1}B) = 0$ for all $g \in G$, is a functorial sub-$\sigma$-algebra.

2. Given a subgroup $H < G$, the assignment $X \mapsto \mathcal{I}_H X$, where $\mathcal{I}X$ is the sub-$\sigma$-algebra of sets $B \in \mathcal{B}$ such that $\mu(B \triangle (T^h)^{-1}B) = 0$ for all $h \in H$, is a functorial sub-$\sigma$-algebra.

3. Given functorial sub-$\sigma$-algebras $F$ and $H$, the assignment $X \to FX \lor HX$ is a functorial sub-$\sigma$-algebra.

4. Given a $G^k$ system $X$ and $e \subset \{1, \ldots, k\}$, write $\mathcal{I}_e X$ for $\mathcal{I}_{H_e} X$, where $H_e$ is the subgroup of elements that are equal to the identity in all coordinates not belonging to $e$. Then $\mathcal{I}_e$ is a functorial sub-$\sigma$-algebra. Given $d, e \subset \{1, \ldots, k\}$ write $\mathcal{I}_{d,e} X$ for $\mathcal{I}_d X \lor \mathcal{I}_e X$. Then $\mathcal{I}_{d,e}$ is also a functorial sub-$\sigma$-algebra.

5. The sub-$\sigma$-algebras $\mathcal{G}_{i,j}$ described in Section 4.4 are all functorial.
6. The sub-$\sigma$-algebra $\mathcal{K}X$ corresponding to the Kronecker factor (see Definition 4.2.3) of a system $X$ is a functorial sub-$\sigma$-algebra.

Functorial sub-$\sigma$-algebras provide the setting for satedness.

**Definition 6.2.3.** Let $F$ be a functorial sub-$\sigma$-algebra. A system $X$ is **sated** for $F$ if, whenever $\pi : W \to X$ is an extension, the sub-$\sigma$-algebras $\pi^{-1}X$ and $FW$ are relatively independent over $\pi^{-1}FX$.

If $X$ is sated for $F$ and $\pi : W \to X$ is an extension then the sub-$\sigma$-algebras $\pi^{-1}X$ and $FW$ are relatively independent over $\pi^{-1}FX$. Since we also have $\pi^{-1}FX \subset FW$, relative independence implies that

\[
E(f \circ \pi|FW) = E(f \circ \pi|\pi^{-1}FX) = E(f|FX) \circ \pi \quad (6.2.4)
\]

for any $f$ in $L^2(X)$.

**Definition 6.2.5.** An extension $p : X \to Y$ is **sated** for $F$ if, whenever $\pi : W \to X$ is an extension, the sub-$\sigma$-algebras $\pi^{-1}p^{-1}Y$ and $FW$ are relatively independent over $\pi^{-1}FX$. (See Figure 6.2 for a schematic of the $\sigma$-algebras involved.)

If $p : X \to Y$ is an extension that is sated for $F$ and $\pi : W \to X$ is an extension then the sub-$\sigma$-algebras $\pi^{-1}p^{-1}Y$ and $FW$ are relatively independent over $\pi^{-1}FX$. Since we also have $\pi^{-1}FX \subset FW$, Corollary 6.1.3 implies that

\[
E(f \circ p \circ \pi|FW) = E(f \circ p \circ \pi|\pi^{-1}FX) = E(f \circ p|FX) \circ \pi \quad (6.2.6)
\]

for any $f$ in $L^2(Y)$.

When a system $Y$ is compact metric one can find, for any fixed functorial sub-$\sigma$-algebra $F$, an extension $X \to Y$ that is sated for $F$. This result is the first part of Theorem 3.5 in [Aus13]. We provide the proof for completeness.

---

**Figure 6.2:** The extension $X \to Y$ is sated with respect to the functorial sub-$\sigma$-algebra $F$ if $FW$ and $\pi^{-1}p^{-1}Y$ are relatively independent over $\pi^{-1}FX$. 

86
Theorem 6.2.7. Let \( F \) be a functorial sub-\( \sigma \)-algebra. Every compact, metric system \( X \) has a compact, metric extension \( W \) such that the extension \( W \to X \) is sated for \( F \).

Proof. Let \( \{f_n : n \in \mathbb{N}\} \) be a countable, dense subset of the unit ball in \( L^2(X) \) and fix \( r : \mathbb{N} \to \mathbb{N} \) with the property that every fiber is infinite. We will construct an inverse family \( \{X_i : i \in \mathbb{N}\} \) of compact, metric probability spaces, the inverse limit of which (guaranteed to exist by [Cho58, Theorem 2.2]) will be the desired extension.

We proceed inductively, putting \( X_1 = X \). Suppose by induction that for some \( m \in \mathbb{N} \) we have defined an inverse family \( \{X_i : 1 \leq i \leq m\} \) with factor maps \( \pi_j^i : X_i \to X_j \) for all \( m \geq i \geq j \). Consider the family \( \Omega_m \) of all extensions \( \pi : Z \to X_m \) of compact, metric probability spaces. To each \( Z \) in \( \Omega_m \) we assign the value

\[
\omega_m(Z) = \|E(f_{r(m)} \circ \pi_1^m \circ \pi|FZ)\| - \|E(f_{r(m)} \circ \pi_1^m |FX_m) \circ \pi\|
\]

and choose \( X_{m+1} \) from \( \Omega_m \) so that

\[
\omega_m(X_{m+1}) \geq \frac{1}{2} \sup \{\omega_m(Z) : Z \in \Omega_M\}
\]

holds. This is possible because \( \omega_m(Z) \leq 2|f_{r(m)}| \) for any \( Z \) in \( \Omega_m \).

Let \( W \) be the inverse limit of the family \( X_m \) and let \( \pi_m : W \to X_m \) be the associated factor maps. We prove that the extension \( W \to X \) is sated for \( F \). Let \( \pi : Z \to W \) be an extension of compact, metric probability spaces. It suffices to prove that

\[
\|E(f_n \circ \pi_1 \circ \pi|FZ)\| = \|E(f_n \circ \pi_1 |FW) \circ \pi\|
\]

in \( L^2(Z) \) for each \( n \) because \( \{f_n : n \in \mathbb{N}\} \) is dense in the unit ball of \( L^2(X) \). We have

\[
\|E(f_n \circ \pi_1 \circ \pi|FZ)\| \geq \|E(f_n \circ \pi_1 |FW) \circ \pi\|
\]

because \( F \) is functorial. Suppose the inequality is strict. Consider the sequence \( m \mapsto |E(f_n \circ \pi_1^m |FX_m)| \), which is non-decreasing and bounded by \( |f_n| \). Thus for \( m \) sufficiently large with \( r(m) = n \) we obtain

\[
\omega_m(X_{m+1}) = \|E(f_n \circ \pi_1^{m+1} |FX_{m+1})\| - \|E(f_n \circ \pi_1^m |FX_m)\|
\]

\[
\leq \frac{1}{4} \left( \|E(f_n \circ \pi_1 \circ \pi|FZ)\| - \|E(f_n \circ \pi_1 |FW) \circ \pi\| \right)
\]

\[
\leq \frac{1}{4} \left( \|E(f_n \circ \pi_1 \circ \pi|FZ)\| - \|E(f_n \circ \pi_1 |FX_m) \circ \pi\| \right) = \frac{1}{4} \omega_m(Z)
\]

contradicting (6.2.8) because \( Z \to X_m \) is an extension of probability spaces. \( \square \)

One can use the above result to prove that every system has an extension that is sated for a given functorial sub-\( \sigma \)-algebra. In order to do this one makes use of the following lemmas, which are part of the proof of Theorem 3.5 in [Aus13].
Lemma 6.2.9. Let $F$ be a functorial sub-$\sigma$-algebra and let $p : X \to Y$ be sated for $F$. If $q : \tilde{X} \to X$ is an extension then $p \circ q : \tilde{X} \to Y$ is sated for $F$.

Proof. Let $\pi : W \to \tilde{X}$ be an extension. (See Figure 6.3 for a schematic of the systems involved.) Fix $f$ in $L^\infty(Y)$. We have

$$\mathbb{E}(f \circ p \circ q \circ \pi | FW) = \mathbb{E}(f \circ p \circ q \circ \pi | \pi^{-1} F X)$$

and

$$\mathbb{E}(f \circ p \circ q | F \tilde{X}) = \mathbb{E}(f \circ p \circ q | q^{-1} F X)$$

using the extensions $\pi \circ q : W \to X$ and $q : \tilde{X} \to X$ respectively. Lifting the second equality to $W$ gives

$$\mathbb{E}(f \circ p \circ q | F \tilde{X}) \circ \pi = \mathbb{E}(f \circ p \circ q | q^{-1} F X) \circ \pi = \mathbb{E}(f \circ p \circ q \circ \pi | \pi^{-1} q^{-1} F X)$$

so

$$\mathbb{E}(f \circ p \circ q \circ \pi | FW) = \mathbb{E}(f \circ p \circ q | F \tilde{X}) \circ \pi = \mathbb{E}(f \circ p \circ q \circ \pi | \pi^{-1} F \tilde{X})$$

as desired. \hfill \Box

Lemma 6.2.10. Let $F$ be a functorial sub-$\sigma$-algebra and let $X_i$ be an inverse sequence such that each $\pi_{i-1}^i : X_i \to X_{i-1}$ is sated for $F$. Then the inverse limit $X_\infty$ of the sequence is sated for $F$.

Proof. Let $\pi : W \to X_\infty$ be an extension. We need to show that $\mathbb{E}(f \circ \pi | FW) = \mathbb{E}(f \circ \pi | \pi^{-1} F X_\infty)$ for every $f$ in $L^\infty(X_\infty)$. It suffices to do this on an $L^2$ dense subspace of $L^2(X_\infty)$. Thus fix $i$ in $\mathbb{N}$ and some $f_i$ in $L^2(X_i)$. Since $\pi_{i+1}^i : X_{i+1} \to X_i$ is sated for $F$, the previous lemma implies that $\pi_i : X_\infty \to X_i$ is also sated for $F$. Thus

$$\mathbb{E}((f \circ \pi_i) \circ \pi | FW) = \mathbb{E}((f \circ \pi_i) \circ \pi | \pi^{-1} F X_\infty)$$

as desired. \hfill \Box
We can now prove that every system has an extension that is sated.

**Theorem 6.2.11.** Let $F$ be a functorial sub-$\sigma$-algebra and let $Y$ be a compact metric system. There is an extension $X \rightarrow Y$ such that $X$ is a compact metric system that is sated for $F$.

**Proof.** Put $X_0 = Y$ and inductively define an extension $X_i$ of $X_{i-1}$ by applying Theorem 6.2.7. Let $X$ be the inverse limit of the resulting inverse sequence. Since each of the extensions $X_{i+1} \rightarrow X_i$ is sated for $F$, the previous lemma implies that $X$ is sated for $F$. □

The above results do not rely heavily on the properties of systems for their proofs. One could restrict one’s attention to any subcategory of compact, metric probability spaces that is closed under inverse limits and construct, for any functorial sub-$\sigma$-algebra $F$, an extension of a given object that is sated for $F$. Indeed, sated extensions at this level of generality were employed by Austin in his proof [Aus11] of Furstenberg and Katznelson’s density Hales-Jewett theorem [FK91].

### 6.3 Characteristic factors via satedness

Fix an amenable group $G$. Let $X$ be a $G^2$ system and let $T_1, T_2$ be the $G$ actions $T_1^g = T^{(g,1)}$ and $T_2^g = T^{(1,g)}$ respectively. Write $T_{12}^g$ for the composition $T_{1}^g T_{2}^g$. In this section we will outline, following in [Aus13], how satedness can be used to describe characteristic factors for the correlation

$$\int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu$$

for fixed $f_0, f_1, f_2 \in L^\infty(X)$ under the assumptions that $X$ is sated for the functorial sub-$\sigma$-algebra $\mathcal{A}_{1,2} = \mathcal{A}_1 \lor \mathcal{A}_2$. As in Chapter 4, the key tool in this endeavor is the van der Corput trick (Theorem 2.3.7), which implies that

$$\lim_{g \to \Phi} \left\| \int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu \right\|$$

$$\leq \lim_{h,l \to \Phi} \lim_{g \to \Phi} \int \left( T_1^h f_1 \cdot T_1^l f_1 \right) \cdot T_{12}^g \left( T_{12}^h f_2 \cdot T_{12}^l f_2 \right) \, d\mu$$

$$= \int (T_1^h f_1 \cdot T_1^l f_1) \otimes (T_{12}^h f_2 \cdot T_{12}^l f_2) \, d\nu_2$$

$$= \int (T_1 \times T_{12})^h (f_1 \otimes f_2) \cdot (T_1 \times T_{12})^l (f_1 \otimes f_2) \, d\nu_2$$

$$= \int f_1 \otimes f_2 \cdot \mathbb{E}(f_1 \otimes f_2 | \mathcal{F}) \, d\nu_2$$

where $\nu_2$ is the relatively independent self-joining of $\mu$ over $\mathcal{A}_2 X$ and $\mathcal{F}$ is the $\sigma$-algebra of $T_1 \times T_{12}$ invariant sets.
In order to appeal to satedness, we need to equip \((X \times X, \mathcal{B} \times \mathcal{B}, \nu_2)\) with the structure of a \(G^2\) system that extends \(X\) and choose an appropriate functorial sub-\(\sigma\)-algebra. Form a \(G^2\) system \(Z_2 = (X \times X, \mathcal{B} \otimes \mathcal{B}, \nu_2, S)\) by defining
\[
S_1 = T_1 \times T_{12} \quad S_2 = T_2 \times I
\]
and note that
\[
\lim_{g \to \Phi} \left\| \int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu \right\| \leq \int f_1 \otimes f_2 \cdot E(f_1 \otimes f_2 | \mathcal{F}_1 Z_2) \, d\nu_2
\]
for all \(f_1, f_2 \in L^\infty(X)\). Moreover, the projection \(\pi_1 : Z_2 \to X\) on the first coordinate is a factor map. Note that, with \(S\) defined as above, the projection \(\pi_2\) on the second coordinate will not be a factor map in general. Writing
\[
f_1 \otimes f_2 = (f_1 \circ \pi_1) \cdot (f_2 \circ \pi_2)
\]
and using the fact that \(f_2 \circ \pi_2\) is \(S_2\) invariant, we have
\[
\int f_1 \otimes f_2 \cdot E(f_1 \otimes f_2 | \mathcal{F}_1 Z_2) \, d\nu_2 = \int f_1 \circ \pi_1 \cdot E(f_2 \circ \pi_2 | \mathcal{F}_2 Z_2) \cdot E(f_1 \circ f_2 | \mathcal{F}_1 Z_2) \, d\nu_2
\]
so we can replace \(f_1 \circ \pi_1\) with its conditional expectation on \(\mathcal{F}_{1,2} Z_2\). At this stage we can appeal to satedness: if \(X\) is sated for \(\mathcal{F}_{1,2}\) then
\[
E(f_1 \circ \pi_1 | \mathcal{F}_{1,2} Z_2) = E(f_1 | \mathcal{F}_{1,2} X) \circ \pi_1
\]
by (6.2.4), and we may conclude that
\[
\lim_{g \to \Phi} \int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu = \lim_{g \to \Phi} \int f_0 \cdot T_1^g E(f_1 | \mathcal{F}_{1,2} X) \cdot T_{12}^g f_2 \, d\mu \quad (6.3.1)
\]
via the van der Corput trick.

It is possible to deduce characteristic factors in the other coordinates from (6.3.1).

**Theorem 6.3.2.** Let \(G\) be an amenable group and let \(X\) be a \(G^2\) system. Then
\[
\lim_{g \to \Phi} \int f_0 \cdot T_1^g E(f_1 | \mathcal{F}_{1,2} X) \cdot T_{12}^g f_2 \, d\mu = \lim_{g \to \Phi} \int E(f_0 | \mathcal{F}_{1,12} X) \cdot T_1^g E(f_1 | \mathcal{F}_{1,2} X) \cdot T_{12}^g E(f_2 | \mathcal{F}_{12,2} X) \, d\mu
\]
for every \(f_0, f_1, f_2 \) in \(L^\infty(X)\).

**Proof.** By approximating, we may assume without loss of generality that \(E(f_1 | \mathcal{F}_{1,2} X)\) is of the form \(h_1 f_2\) for some functions \(h_i \in L^\infty(\mathcal{F}_i X)\). Then
\[
\lim_{g \to \Phi} \int f_0 \cdot T_1^g E(f_1 | \mathcal{F}_{1,2} X) \cdot T_{12}^g f_2 \, d\mu = \lim_{g \to \Phi} \int f_0 h_1 \cdot T_{12}^g (h_2 f_2) \, d\mu = \int E(f_0 h_1 | \mathcal{F}_{12} X) \cdot h_2 f_2 \, d\mu
\]
by the mean ergodic theorem, and we can replace $f_2$ with its conditional expectation on $\mathcal{I}_{1,2}X$. Assuming now that $f_2 = k_{12}k_2$ for some $k_i \in L^\infty(\mathcal{I}_iX)$, we obtain

$$
\lim_{g \to \Phi} \int f_0 \cdot T^g f \mathbb{E}(f_1|\mathcal{I}_{1,2}X) \cdot T^g_{12} f_2 \, d\mu
$$

$$
= \lim_{g \to \Phi} \int f_0 h_{1,12} \cdot T^g_{1}(h_2 k_2) \, d\mu = \int f_0 h_{1,12} \cdot \mathbb{E}(h_2 k_2|\mathcal{I}_1X) \, d\mu
$$

and we can replace $f_0$ with its conditional expectation on $\mathcal{I}_{1,2}X$ as desired. \hfill \Box

One can also deduce (6.3.1) from (4.4.1) using what are sometimes referred to as cube spaces (see [CZK14]). Indeed, one has the following result.

**Theorem 6.3.3.** Let $G$ be a second countable, amenable group and let $X$ be a compact metric $G^2$ system that is sated for $\mathcal{I}_{1,2}$. Then $\mathcal{C}_{1,2} = \mathcal{I}_{1,2}X$.

**Proof.** Recall that $\mathcal{C}_{1,2}$ is the sub-$\sigma$-algebra corresponding to the functions that are almost periodic for $T_2$ over $\mathcal{I}_1X$. It is immediate that $\mathcal{I}_{1,2}X \subset \mathcal{C}_{1,2}$. For the reverse inclusion we show any $f \in L^2(X)$ that is orthogonal to $\mathcal{I}_{1,2}X$ is also orthogonal to $\mathcal{C}_{1,2}$. By Theorem 4.3.8 it suffices to show that any such $f$ is weakly mixing for $T_2$ over $\mathcal{I}_1X$. For every $\phi \in L^\infty(X)$ we have

$$
\lim_{g \to \Phi} \|\mathbb{E}(\phi \cdot T^g_2 f|\mathcal{I}_1X)\|^2
$$

$$
= \lim_{g \to \Phi} \int (\phi \otimes \phi) \cdot (T_2 \otimes T_2)^g (f \otimes f) \, d\nu_1
$$

$$
= \int \mathbb{E}(\phi \otimes \phi|\mathcal{I}_2Z_1) \cdot (f \otimes f) \, d\nu_1
$$

$$
= \int \mathbb{E}(\phi \otimes \phi|\mathcal{I}_2Z_1) \cdot \mathbb{E}(f \circ \pi_1|\mathcal{I}_{1,2}Z_1) \cdot (f \circ \pi_2) \, d\nu_1
$$

where

$$
Z_1 = (X \times X, \mathcal{B} \otimes \mathcal{B}, \nu_1, R)
$$

is the relatively independent self-joining of $X$ over $\mathcal{I}_1X$ and $R_1 = T_2 \times T_{12}$, $R_2 = T_1 \times I$. By satedness

$$
\mathbb{E}(f \circ \pi_1|\mathcal{I}_{1,2}Z_1) = \mathbb{E}(f|\mathcal{I}_{1,2}X) \circ \pi_1 = 0
$$

as desired. \hfill \Box

### 6.4 Approximate satedness

In this section we present a modification of the definition of satedness that does not require the use of inverse limits in order to prove that one can always pass to a sated extension.
Definition 6.4.1. Fix a topological group \( G \) and a functorial sub-\( \sigma \)-algebra \( F \). Let \( X \) be a \( G \) system. Given \( f \in L^2(X) \) and \( \varepsilon > 0 \) we say that \( X \) is \((\varepsilon, f)\) sated for \( F \) if

\[
\| \mathbb{E}(f \circ \pi | F_Y) - \mathbb{E}(f | F_X) \circ \pi \| \leq \varepsilon
\]  
(6.4.2)

in \( L^2(Y) \) for every extension \( \pi : Y \to X \).

Our goal is to show that every system \( X \) has an extension that is \((\varepsilon, f)\) sated for a functorial sub-\( \sigma \)-algebra. We will be able to do this without appealing to inverse limits because the energy increment (6.4.2) is only required to be small rather than to vanish.

Theorem 6.4.3. Fix a topological group \( G \) and a functorial sub-\( \sigma \)-algebra \( F \). For every \( G \) system \( X \), every \( f \in L^2(X) \) and every \( \varepsilon > 0 \) there is an extension \( \psi : Y \to X \) such that \( Y \) is \((\varepsilon, f \circ \psi)\) sated for \( F \).

Proof. Fix \( \varepsilon > 0 \) and \( f \in L^2(X) \). Assume \( f \neq 0 \) as otherwise the conclusion is immediate. We have

\[
\sup \{ \| \mathbb{E}(f \circ \psi | F_Y) \| : Y \to X \text{ an extension} \} \leq \| f \| < \infty
\]

so there is an extension \( \psi : Y \to X \) such that \( \| \mathbb{E}(f \circ \psi | F_Y) \| \) is within \( \delta = \varepsilon / (4|f|) \) of the supremum above.

Fix an extension \( \pi : Z \to Y \). We have

\[
\mathbb{E}(f \circ \psi | F_Y) \circ \pi = \mathbb{E}(f \circ \psi \circ \pi | \pi^{-1}(F_Y)) = \mathbb{E}(\mathbb{E}(f \circ \psi \circ \pi | F_Z) | \pi^{-1}(F_Y))
\]

because \( F \) is a functorial sub-\( \sigma \)-algebra, and

\[
\| \mathbb{E}(f \circ \psi | F_Y) \| > \| \mathbb{E}(f \circ \psi \circ \pi | F_Z) \| - 2\delta
\]

by choice of \( Y \). Combining these results gives

\[
\| \mathbb{E}(f \circ \psi | F_Y) \circ \pi - \mathbb{E}(f \circ \psi \circ \pi | F_Z) \| \leq 4\delta \| \mathbb{E}(f \circ \psi \circ \pi | F_Z) \| \leq 4\delta \| f \| \leq \varepsilon
\]

because conditional expectation is an orthogonal projection.

6.5 Ergodicity and satedness

In the previous sections we have seen how passing to an extension of a given system can simplify the description of characteristic factors. If the original system has additional properties we would like for this property to be preserved in the extension. In this section we describe how this can be done for ergodicity. We begin with the following lemma.
Lemma 6.5.1. Let $X = (X, \mathcal{B}, \mu, T)$ be a compact, Hausdorff $G$ system that is ergodic, and let $Y = (Y, \mathcal{D}, \nu, S)$ be a compact, Hausdorff $G$ system that is an extension of $X$ via a continuous factor map $\pi$. The collection

$$M_X(Y) = \{ \lambda \in M_S(Y) : \pi \lambda = \mu \}$$

of $S$ invariant Baire measure on $Y$ that extend $\mu$ is a compact, convex subset of $M(Y)$ and the extreme points of $M_X(Y)$ are ergodic for $S$.

Proof. It is immediate that $M_X(Y)$ is closed and convex. We show that any extreme point of $M_X(Y)$ that can be written non-trivially in the form $t\lambda_1 + (1-t)\lambda_2$ for some $\lambda_1, \lambda_2$ from $M(Y)$. Projecting onto $M(X)$ using $\pi$ shows that $\pi \lambda_1 = \pi \lambda_2 = \mu$ because $\mu$ is an extreme point of $M(X)$ by ergodicity and Lemma 2.1.1. This implies that $\lambda_1$ and $\lambda_2$ belong to $M_X(Y)$, contradicting extremality of $\lambda$ there. \qed

We begin by remarking that one can restrict ones attention to ergodic systems in Theorem 6.2.11 by working entirely within the category of ergodic $G$ systems.

Theorem 6.5.2. Let $G$ be an amenable group and let $F$ be a functorial sub-$\sigma$-algebra. For any compact, metric $G$ system $Y$ there is an extension $X \to Y$ such that $X$ is a compact metric system that is ergodic and has the property that (6.2.4) holds whenever $W$ is an ergodic, compact, metric system that extends $X$.

The main result of this section is that ergodic $G^2$ systems that are sated for $\mathcal{I}_{1,2}$ have the expected characteristic factors. The crux of the issue is the fact that relatively independent self-joinings will usually fail to be ergodic.

Theorem 6.5.3. Fix a $\sigma$-compact amenable group $G$. Let $X$ be a $G^2$ system $X$ that is sated for $\mathcal{I}_{1,2}$ in the category of ergodic systems. Then

$$\lim_{g \to \Phi} \int f_0 \cdot T^g_1 f_1 \cdot T^g_1 f_2 d\mu = \lim_{g \to \Phi} \int f_0 \cdot T^g_1 \mathbb{E}(f_1 | \mathcal{I}_{1,2} Y) \cdot T^g_1 f_2 d\mu$$

for each left Følner sequence $\Phi$ in $G$ and every $f_0, f_1, f_2$ in $C(Y)$.

Proof. Fix a Baire measurable representative $\psi$ of $f_1 - \mathbb{E}(f_1 | \mathcal{I}_{1,2} X)$. Let $Z_2$ be the relatively independent self-joining of $X$ over $\mathcal{I}_{2}X$. The projection $\pi_2$ on the second coordinate is a factor map $Z_2 \to X$. By the van der Corput trick we have

$$\left| \lim_{g \to \Phi} \int f_0 \cdot T^g_1 \psi \cdot T^g_1 f_2 d\mu \right| \leq \lim_{g \to \Phi} \int (\psi \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot (T_1 \times T_1) \mathbb{E}(\psi \otimes f_2) d\nu$$

where $\nu$ is the relatively independent self-joining of $X$ over $\mathcal{I}_2X$. Define the system

$$Z_\lambda = (X \times X, \mathcal{B} \otimes \mathcal{B}, \lambda, S)$$

93
for each $\lambda \in \mathcal{M}_X(X \times X)$. Let $\eta$ be a measure on $\mathcal{M}_X(X \times X)$ that represents $\nu$ in the sense of the Choquet-Bishop-de Leeuw theorem. Since $\psi$ is Baire measurable, we have

$$
\lim_{g \to \Phi} \int (\psi \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot (T_1 \times T_{12})^g (\psi \otimes f_2) \, d\nu
$$

$$
= \lim_{g \to \Phi} \iint (\psi \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot (T_1 \times T_{12})^g (\psi \otimes f_2) \, d\lambda 
\, d\eta(\lambda)
$$

$$
= \iint (\psi \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot \mathbb{E}(\psi \otimes f_2 | \mathcal{F}_1 \mathcal{Z}_\lambda) \, d\lambda 
\, d\eta(\lambda)
$$

$$
= \iint \mathbb{E}(\psi \circ \pi_1 | \mathcal{F}_{1,2} \mathcal{Z}_\lambda) \cdot (f_2 \circ \pi_2) \cdot \mathbb{E}(\psi \otimes f_2 | \mathcal{F}_1 \mathcal{Z}_\lambda) \, d\lambda 
\, d\eta(\lambda)
$$

for any left Følner sequence $\Phi$ by the dominated convergence theorem. For any extreme point of $\mathcal{M}_X(X \times X)$ the inner integral above is zero by satedness in the category of ergodic systems. Since the set

$$
\Lambda = \left\{ \lambda \in \mathcal{M}_X(X \times X) : \lim_{g \to \Phi} \int (\psi \circ \pi_1) \cdot (f_2 \circ \pi_2) \cdot (T_1 \times T_{12})^g (\psi \otimes f_2) \, d\lambda \neq 0 \right\}
$$

is a Baire measurable subset of $\mathcal{M}_X(Y)$ that is disjoint from the extreme points we must have $\eta(\Lambda) = 0$. 

\hfill \Box
CHAPTER 7
MIXING GROUPS

One can obtain stronger combinatorial results in groups with mixing properties because the description of characteristic factors is strengthened by the mixing assumption. In this chapter we describe a strengthening of this sort that was recently obtained in [BCRZK14].

7.1 Weakly mixing groups

Definition 7.1.1. A group $G$ is weakly mixing if every ergodic $G$ system is weakly mixing.

Whether a group is weakly mixing is related to whether it has non-trivial almost-periodic function. Write $C_b(G)$ for the Banach space of all bounded continuous functions $f : G \to \mathbb{C}$ equipped with the supremum norm.

Definition 7.1.2. A function $f \in C_b(G)$ is called almost periodic if either of the following equivalent conditions holds.

1. The subset $\{L_g f : g \in G\}$ of $C_b(G)$ is relatively compact.

2. The subset $\{R_g f : g \in G\}$ of $C_b(G)$ is relatively compact.

3. $f$ is the pullback of a continuous function under the Bohr compactification of $G$, which is the maximal topological group compactification $\iota : G \to \text{Bohr}(G)$.

For the equivalence of (1) with (2) see [BJM78, Theorem 9.2] and for the equivalence of (2) with (3) see [BJM78, Remark 9.8]. Denote by $\mathcal{A}(G)$ the space of all almost periodic functions on $G$. It follows from the characterization (3) that almost periodic functions are uniformly continuous.

The matrix coefficients of a finite-dimensional continuous representation give rise to almost-periodic functions on $G$. Specifically, given a continuous representation $\phi$ of $G$
on a finite-dimensional, complex Hilbert space $V$ and vectors $x, y$ in $V$ we can form the almost-periodic function $f(g) = \langle \phi(g)x, y \rangle$.

**Theorem 7.1.3** ([Neu34, Theorems 30 and 31]). *Matrix coefficients span a dense subspace of $A(G)$.*

The constant functions are always almost-periodic. There are groups having no other almost-periodic functions (see [NW40]). Such groups are said to be **minimally almost-periodic**. Theorem 7.1.3 implies that a group is minimally almost-periodic if and only if it is a WM group. In view of the Peter–Weyl theorem, non-trivial compact groups are never WM groups.

**Theorem 7.1.4.** A group is weakly mixing if and only if it is minimally almost periodic.

**Example 7.1.5.** Let $G$ be a countably infinite group that is the union of an increasing sequence of non-cyclic finite simple subgroups. Then $G$ is WM. This follows from [Gow08, Theorem 4.7], which states that the minimal dimension of a non-trivial representation of a non-cyclic finite simple group tends to infinity as the size of the group goes to infinity. Moreover, $G$ is amenable since it is locally finite.

This applies, for instance, to the finite alternating group of the integers $A(\mathbb{N})$, which is the subgroup of the finite symmetric group of the integers

$$S(\mathbb{N}) = \{\sigma : \mathbb{N} \to \mathbb{N} : \sigma \text{ is a bijection and } \{n \in \mathbb{N} : \sigma(n) \neq n\} \text{ is finite}\}$$

consisting of the even permutations. Indeed, $A(\mathbb{N})$ is the union $\bigcup \{A_n : n \in \mathbb{N}\}$ where $A_n$ is the alternating group on $n$ points. Now $[S(\mathbb{N}) : A(\mathbb{N})] = 2$, so $S(\mathbb{N})$ is a virtually WM group. On the other hand, $\phi(\sigma) = (-1)^{\text{sgn}(\sigma)}$ defines a non-trivial representation of $S(\mathbb{N})$, so $S(\mathbb{N})$ is not a WM group.

### 7.2 Characteristic factors

Throughout this section we assume that $G$ is $\sigma$-compact. The assumption that $G$ is weakly mixing simplifies the description of characteristic factors for averages of multiple correlations. In particular we have the following results.

**Lemma 7.2.1.** Let $G$ be a weakly mixing amenable group. Then $\mathcal{H}X = \mathcal{I}X$ and

$$\mathcal{I}(X \times X) = \mathcal{I}X \times \mathcal{I}X$$

for any $G$ system $X$.  

96
Proof. The inclusion \( \mathcal{X} \subseteq \mathcal{X} \) is immediate. For the reverse inclusion, let \( H \) be a finite-dimensional, \( T \) invariant subspace of \( L^2(X) \). It induces a finite-dimensional representation of \( G \), which must be trivial because \( G \) is weakly mixing. Thus \( H \subseteq L^2(\mathcal{X}) \). The second assertion follows from the first and the fact that \( \mathcal{X}(X \times X) = \mathcal{X}(X) \otimes \mathcal{X}(X) \), which is a special case of Theorem 4.3.6.

**Lemma 7.2.2.** Fix a second countable amenable group \( G \) and let \( X_1 \) and \( X_2 \) be \( G \) systems. Let \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \) be \( T_1 \) and \( T_2 \) invariant sub-\( \sigma \)-algebras of \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \) respectively. Then \( \mathcal{A}(X_1|\mathcal{B}_1) \ominus \mathcal{A}(X_2|\mathcal{B}_2) = \mathcal{A}(X_1 \times X_2|\mathcal{B}_1 \otimes \mathcal{B}_2) \).

**Proof.** The inclusion \( \subseteq \) follows from the definition. For the reverse inclusion, note that if \( f_1 \) is orthogonal to \( \mathcal{A}(X_1|\mathcal{B}_1) \) or \( f_2 \) is orthogonal to \( \mathcal{A}(X_2|\mathcal{B}_2) \) then \( f_1 \otimes f_2 \) is orthogonal to \( \mathcal{A}(X_1 \times X_2|\mathcal{B}_1 \otimes \mathcal{B}_2) \) by the characterization of the orthogonal complements of these spaces in Theorem 4.3.8.

**Lemma 7.2.3.** Let \( G \) be a weakly mixing group and let \( X \) be an ergodic \( G^2 \) system. The function

\[
g \mapsto \int \mathbb{E}(f_0|\mathcal{J}_{1,12}X) \cdot T_1^0 \mathbb{E}(f_1|\mathcal{J}_{1,2}X) \cdot T_1^{2,0} \mathbb{E}(f_2|\mathcal{J}_{12,2}X) \, d\mu
\]

is constant for any \( f_0, f_1, f_2 \in L^\infty(X) \).

**Proof.** By density and linearity it suffices to prove this with \( f_0 = h_{0,1}h_{0,12}, f_1 = h_{1,1}h_{1,2}, f_2 = h_{2,12}h_{2,2} \) where \( h_{i,j} \) is \( \mathcal{J}_jX \) measurable. The above integral becomes

\[
\int h_{0,1}h_{0,12} \cdot T_1^0(h_{1,1}h_{1,2}) \cdot T_1^{2,0}(h_{2,12}h_{2,2}) \, d\mu
\]

\[
= \int h_{0,1}h_{1,1} \cdot T_1^0(h_{1,2}h_{2,2}) \cdot h_{0,12}h_{2,12} \, d\mu
\]

\[
= \int h_{0,1}h_{1,1} \cdot T_1^0(h_{1,2}h_{2,2}) \cdot \mathbb{E}(h_{0,12}h_{2,12}|\mathcal{J}_{1,2}X) \, d\mu
\]

under this assumption. Write \( h \) for \( \mathbb{E}(h_{0,12}h_{2,12}|\mathcal{J}_{1,2}X) \). We claim that \( h \) is constant. Since \( \mathcal{J}_{1,2}X \) is \( T_{12} \) invariant so is \( h \). By ergodicity of \( X \) we have

\[
L^2(X, \mathcal{J}_{1,2}X, \mu) \cong L^2(X, \mathcal{J}_1X, \mu) \otimes L^2(X, \mathcal{J}_2X, \mu)
\]

and since \( h \) is \( T_{12} \) invariant, its image in the right-hand side is \( T_2 \times T_1 \) invariant. But any \( T_2 \times T_1 \) invariant function in the right-hand side is a linear combination of functions of the form \( \xi_1 \otimes \xi_2 \) where \( \xi_1 \) has pre-compact \( T_2 \) orbit in \( L^2(X, \mathcal{J}_1X, \mu) \) and \( \xi_2 \) has pre-compact \( T_1 \) orbit in \( L^2(X, \mathcal{J}_2X, \mu) \). Since \( G \) is weakly mixing, it must be the case that \( \xi_1 \) is \( T_2 \) invariant and \( \xi_2 \) is \( T_1 \) invariant. By ergodicity \( \xi_1 \) and \( \xi_2 \) are constant, so \( h \) is too.
Finally, this implies
\[
\int h_{0,1}h_{0,12} \cdot T_1^g(h_{1,1}h_{1,2}) \cdot T_{12}^g(h_{2,12}h_{2,2}) \, d\mu \\
= \int h_{0,1}h_{1,1} \cdot T_1^g(h_{1,2}h_{2,2}) \, d\mu \int h_{0,12}h_{12} \, d\mu \\
= \int h_{0,1}h_{1,1} \, d\mu \int T_1^g(h_{1,2}h_{2,2}) \, d\mu \int h_{0,12}h_{12} \, d\mu
\]
by another application of ergodicity, and the last line does not depend on \( g \) because \( \mu \) is \( T_1 \) invariant.

**Lemma 7.2.4.** Fix a weakly mixing, \( \sigma \)-compact, amenable group \( G \) and let \( X \) be an ergodic \( G^2 \) system. Then
\[
\lim_{g \to \Phi} \int f \cdot T_1^g f \cdot T_{12}^g f \, d\mu \geq \left( \int f \, d\mu \right)^4
\]
for any \( f \geq 0 \) in \( C(X) \) and any left Følner sequence \( \Phi \) in \( G \).

**Proof.** Fix \( f \geq 0 \) in \( C(X) \). By Theorem 6.5.3 we may assume by passing to an extension that
\[
\lim_{g \to \Phi} \int f \cdot T_1^g f \cdot T_{12}^g f \, d\mu = \lim_{g \to \Phi} \int f : T_1^g \mathbb{E}(f | \mathcal{F}_{1,2}X) \cdot T_{12}^g f \, d\mu
\]
holds. By Theorem 6.3.2 and Lemma 7.2.3 we have
\[
\lim_{g \to \Phi} \int f \cdot T_1^g \mathbb{E}(f | \mathcal{F}_{1,2}X) \cdot T_{12}^g f \, d\mu = \int \mathbb{E}(f | \mathcal{F}_{1,12}X) \cdot \mathbb{E}(f | \mathcal{F}_{1,2}X) \cdot \mathbb{E}(f | \mathcal{F}_{12,2}X) \, d\mu
\]
which is at least \( (f \, d\mu)^4 \) by [Chu11, Lemma 1.6].

Our next goal is to strengthen Lemma 7.2.4 by replacing convergence on average with convergence in density.

**Lemma 7.2.6.** Let \( G \) be a weakly mixing, second countable amenable group. If \( X \) is a \( G \) system that is sated for \( \mathcal{F}_{1,2} \) then
\[
\lim_{g \to \Phi} \int (f_0 \otimes f_0) \cdot (T_1 \times T_1)^g(f_1 \otimes f_1) \cdot (T_{12} \times T_{12})^g(f_2 \otimes f_2) \, d(\mu \otimes \mu)
\]
\[
= \lim_{g \to \Phi} \int (f_0 \otimes f_0) \cdot (T_1 \times T_1)^g \mathbb{E}(f_1 \otimes f_1 | \mathcal{F}_{1,2}(X \times X)) \cdot (T_{12} \times T_{12})^g(f_2 \otimes f_2) \, d(\mu \otimes \mu)
\]
for every \( f_0, f_1, f_2 \) in \( L^\infty(X \times X) \).

**Proof.** We know from (4.4.1) that
\[
\lim_{g \to \Phi} \int (f_0 \otimes f_0) \cdot (T_1 \times T_1)^g(f_1 \otimes f_1) \cdot (T_{12} \times T_{12})^g(f_2 \otimes f_2) \, d(\mu \otimes \mu)
\]
\[
= \lim_{g \to \Phi} \int (f_0 \otimes f_0) \cdot (T_1 \times T_1)^g \mathbb{E}(f_1 \otimes f_1 | \mathcal{G}_{2,2}(X \times X)) \cdot (T_{12} \times T_{12})^g(f_2 \otimes f_2) \, d(\mu \otimes \mu)
\]
and the result follows by combining Lemma 7.2.1 and Lemma 7.2.2 with Theorem 6.3.3.
Theorem 7.2.7. Let $G$ be a second countable, weakly mixing, amenable group and let $X$ be an ergodic $G^2$ system that is $\mathcal{J}_{1,2}$ sated. Then

$$\text{dlim}_{g \to \Phi} \int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu = \int \mathbb{E}(f_0 | \mathcal{J}_{1,12}X) \cdot \mathbb{E}(f_1 | \mathcal{J}_{1,2}X) \cdot \mathbb{E}(f_2 | \mathcal{J}_{12,2}X) \, d\mu$$

for any $f_0, f_1, f_2$ in $L^\infty(X, \mathcal{B}, \mu)$ and every left Følner sequence $\Phi$ in $G$.

Proof. By Lemma 7.2.3 the function

$$g \mapsto \int \mathbb{E}(f_0 | \mathcal{J}_{1,12}X) \cdot T_1^g \mathbb{E}(f_1 | \mathcal{J}_{1,2}X) \cdot T_{12}^g \mathbb{E}(f_2 | \mathcal{J}_{12,2}X) \, d\mu$$

is constant. Thus it suffices to prove that

$$\text{dlim}_{g \to \Phi} \int f_0 \cdot T_1^g f_1 \cdot T_{12}^g f_2 \, d\mu = 0 \quad (7.2.8)$$

for every left Følner sequence in $G$ and any $f_0, f_1, f_2$ in $L^\infty(X, \mathcal{B}, \mu)$ satisfying one of the following conditions:

1. $f_0 \perp \mathcal{J}_{1,12}X \iff f_0 \otimes f_0 \perp \mathcal{J}_1(X \times X) \vee \mathcal{J}_{12}(X \times X)$
2. $f_1 \perp \mathcal{J}_{1,2}X \iff f_1 \otimes f_1 \perp \mathcal{J}_1(X \times X) \vee \mathcal{J}_2(X \times X)$
3. $f_2 \perp \mathcal{J}_{12,2}X \iff f_2 \otimes f_2 \perp \mathcal{J}_2(X \times X) \vee \mathcal{J}_{12}(X \times X)$

where the equivalences all follow from Lemma 7.2.1. But (7.2.8) is equivalent to

$$\lim_{g \to \Phi} \int (f_0 \otimes f_0) \cdot (T_1 \times T_1)^g (f_1 \otimes f_1) \cdot (T_{12} \times T_{12})^g (f_2 \otimes f_2) \, d(\mu \otimes \mu) = 0$$

for every left Følner sequence $\Phi$, which is true under any of the above conditions by Lemma 7.2.6 and Theorem 6.3.2. \qed

Corollary 7.2.9. Let $G$ be a weakly mixing, second countable, amenable group and let $X = (X, \mathcal{B}, \mu, T)$ be an ergodic $G^2$ system. Then

$$\text{dlim}_{g} \int f \cdot T_1^g f \cdot T_{12}^g f \, d\mu \geq \left( \int f \, d\mu \right)^4$$

for any $1 \geq f \geq 0$ in $L^\infty(X, \mathcal{B}, \mu)$ and any left Følner sequence $\Phi$ in $G$.

Proof. As in the proof of Lemma 7.2.5, the system $X$ admits an ergodic extension that is sated for $\mathcal{J}_{1,2}$, and we may therefore assume that $X$ is sated for $\mathcal{J}_{1,2}$. Theorem 7.2.7 and [Chu11, Lemma 1.6] combined yield

$$\text{dlim}_{g} \int f \cdot T_1^g f \cdot T_{12}^g f \, d\mu$$

$$= \int \mathbb{E}(f | \mathcal{J}_{1,12}X) \cdot \mathbb{E}(f | \mathcal{J}_{1,2}X) \cdot \mathbb{E}(f | \mathcal{J}_{12,2}X) \, d\mu$$

$$\geq \int f \cdot \mathbb{E}(f | \mathcal{J}_{1,12}X) \cdot \mathbb{E}(f | \mathcal{J}_{1,2}X) \cdot \mathbb{E}(f | \mathcal{J}_{12,2}X) \, d\mu \geq \left( \int f \, d\mu \right)^4$$

as desired. \qed
7.3 Finite products sets

In this section we use the results from the previous section to prove that every substantial subset of a $\sigma$-compact, weakly mixing, amenable group contains a two-sided finite products set, and give some examples to show that the assumption of two-sided substantial is necessary. We will need the following connection between invariant means and left Følner sequences.

**Proposition 7.3.1.** Let $G$ be a weakly mixing, $\sigma$-compact, amenable group. For every two-sided invariant mean on $L^\infty(G, m)$ and every $\phi \in L^\infty(G, m)$ there is a left Følner sequence $\Phi$ on $G$ such that

$$M(\phi) = \lim_{\iota \to \infty} \frac{1}{m(\Phi_\iota)} \int_{\Phi_\iota} \phi \, dm$$

holds.

*Proof.* Fix a two-sided invariant mean $M$ on $UC(G)$. By [Cho70, Lemma 2.2] we can extend $M$ to a topologically two-sided invariant mean on $L^\infty(G, m)$. By [Pat88, Theorem 4.17] there is a left Følner sequence $\Phi$ in $G$ that agrees with $M$ on $\phi$. \qed

Here is the main combinatorial result of this chapter.

**Theorem 7.3.2.** Let $G$ be a weakly mixing, second countable, amenable group. Every substantial subset of $G$ contains a two-sided finite products set.

*Proof.* Fix $1 \geq f \geq 0$ in $UC(G)$ and fix a two-sided invariant mean $M$ on $UC(G)$ with $M(f) > 0$. It suffices to show that $S = f^{-1}((0,1])$ contains a finite products set by Proposition 7.3.1.

Put $g_0 = 1$, $f_0 = f$ and $M_0 = M$. We proceed inductively, constructing

- a sequence $n \mapsto g_n$ in $G$;
- a sequence $n \mapsto f_n$ in $UC(G)$;
- a sequence $n \mapsto M_n$ of two-sided invariant means on $UC(G)$;

such that $0 \leq f_n \leq 1$, $f_n(g_{n+1}) > 0$, $M_n(f_n) > 0$ and

$$f_{n+1} = f_n \cdot L_{g_{n+1}} f_n \cdot R_{g_{n+1}} f_n$$

for all $n$.

Assume by induction that for some $n \geq 0$ we have $g_i$, $f_i$ and $M_i$ defined and with the desired properties for all $0 \leq i \leq n$. Theorem 3.1.6 gives us an ergodic $G^2$ system

100
\[ X_n = (X, \mathcal{B}, \nu_n, T), \] a continuous function \( \xi_n : X \to [0, \infty) \), and a two-sided invariant mean \( M_{n+1} \) on \( \text{UC}(G) \) such that
\[
M_{n+1}(f_n \cdot L_g f_n \cdot R_g f_n) = \int \xi_n \cdot T_1^g \xi_n \cdot T_2^{g^{-1}} \xi_n \, d\nu_n = \int \xi_n \cdot T_2^g \xi_n \cdot T_1^g \xi_n \, d\nu_n
\]
for all \( g \in G \). By Proposition 7.3.1 we can find a left Følner sequence \( \Phi \) on \( G \) that agrees with \( M_n \) on \( f_n \). Corollary 7.2.9 implies that
\[
\text{dlim}_{g \to \Phi} M_{n+1}(f_n \cdot L_g f_n \cdot R_g f_n) = M_n(f_n)^4
\]
so the set
\[
\{ g \in G : M_{n+1}(f_n \cdot L_g f_n \cdot R_g f_n) > 0 \text{ and } f_n(g) > 0 \}
\] (7.3.4)
has positive upper density with respect to \( \Phi \). Fix \( g_{n+1} \) therein and define \( f_{n+1} \) by (7.3.3).
We have \( 0 \leq f_{n+1} \leq f_n \leq 1, \ f_n(g_{n+1}) > 0 \text{ and } M_{n+1}(f_{n+1}) > 0 \) so the inductive construction is complete.

Write \( g \) for the sequence \( n \mapsto g_n \). It remains to prove that \( \text{FP}(g) \) is contained in \( S \). Put \( S_n = f_n^{-1}((0, 1]) \). First note that \( g_n \) belongs to \( S \) for each \( n \) and that (7.3.3) implies
\[
S_n = \cap \{ I_\alpha(g)^{-1} S D_\beta(g)^{-1} : \alpha, \beta \subset \{1, \ldots, n\} \text{ and } \alpha \cap \beta = \emptyset \}
\]
for each \( n \) in \( \mathbb{N} \) by induction on \( n \). Thus \( I_\alpha(g) g_{n+1} D_\beta(g) \) belongs to \( S \) for any disjoint subsets \( \alpha, \beta \) of \( \{1, \ldots, n\} \) as desired.

We remark that since (7.3.4) has positive upper density with respect to \( \Phi \), one can at each step in the induction choose \( g_n \) to lie outside any fixed compact subset \( K_n \) of \( G \). Theorem 7.3.2 combined with Example 3.1.3 immediately implies the following partition result.

**Corollary 7.3.5.** For any measurable partition \( C_1 \cup \cdots \cup C_r \) of a weakly mixing, amenable, second countable group one can find \( 1 \leq i \leq r \) such that, for any open neighborhood \( U \) of the identity, the set \( UC_i \setminus U \) contains a two-sided finite products set.

The following example shows that Theorem 7.3.2 does not extend to sets that are substantial with respect to a left invariant mean on \( \text{UC}(G) \), even in the discrete case.

**Example 7.3.6.** Let \( A(\mathbb{N}) \) be the finite alternating group on \( \mathbb{N} \). As in Example 7.1.5 we can view \( A(\mathbb{N}) \) as the increasing union of the alternating groups \( A_n = A(\{1, \ldots, n\}) \). For \( n \geq 4 \) let \( h_n := (1, n)(2, 3) \in A_n \) and \( \Phi_n := A_{n-1} h_n \). This is a left Følner sequence, and we see from
\[
\Phi_n \subset \{ \sigma \in A_n : \sigma(1) = n \}
\]
that the sets \( \Phi_n \) are pairwise disjoint. Let \( E := \cup_{n \geq 5} \Phi_n \), so that \( d_{\Phi}(E) = 1 \). Suppose that there exists a sequence \( (g_k) \) with \( \text{FPD}(g_k) \subset E \). We will show the following statement by induction on \( n \).
Claim. Suppose that for some $\alpha \in F$ and $n \geq 4$ we have $D_{\alpha}(g_k) \in \Phi_n$. Then there exists $\beta \in F$ such that $D_{\beta}(g_k) \in \Phi_4$.

Since the assumption is clearly satisfied for some $n \geq 5$ and some $\alpha \in F$, we obtain a contradiction.

Proof of the claim. For $n = 4$ the conclusion holds with $\beta = \alpha$. Suppose now that the claim is known to hold up to some $n$ and assume $h := D_{\alpha}(g_k) \in \Phi_{n+1}$. Let $i > \max \alpha$, let $j$ be such that $g_i \in \Phi_j$, and consider $D_{\alpha \cup \{i\}}(g_k) = g_i h$. By the assumption this is an element of $E$. Consider now the following cases.

$j > n + 1$ We have $g_i h(1) = g_i(n + 1) < j$ and $g_i h(j) = g_i(j) < j$, so that $g_i h \notin E$, contradicting the assumption.

$j \leq n$ The conclusion follows from the inductive assumption.

$j = n + 1$ In this case we have $g_i h(1) = g_i(n + 1) < n + 1$, so that $g_i h \notin \Phi_{n+1}$. Since $g_i h \in E \cap A_{n+1}$, it follows that $g_i h \in \Phi_m$ for some $m \leq n$, and the conclusion again follows from the inductive hypothesis.

In view of [BF09, Theorem 2.4] this implies that the set $E$ in this example has density zero with respect to any right Følner sequence.

7.4 Limits along minimal idempotent ultrafilters

In this section we prove an analogue of (6.3.1) for limits along minimal left idempotent ultrafilters under the assumption that the group is weakly mixing. We begin with a suitable version of the mean ergodic theorem.

Theorem 7.4.1. Let $G$ be a countable, discrete, weakly mixing group, let $X$ be a $G$ system, and let $p$ be a left minimal left idempotent ultrafilter on $G$. For any $f \in L^2(X)$ we have

$$\lim_{g \to p} T^g f = \mathbb{E}(f | \mathcal{F} X)$$

in the weak topology of $L^2(X)$.

Proof. Fix $f \in L^2(X)$. Equipped with the weak topology, the unit ball of $L^2(X)$ is compact and Hausdorff, so the limit in (7.4.2) makes sense via Proposition 2.6.24. Let $\phi$ be the limit of the sequence $T^g f$ along $p$. We first show that $\phi$ belongs to $L^2(\mathcal{F} X)$.

We claim that the orbit $\{T^g \phi : g \in G\}$ is relatively compact in the norm topology. Fix $\varepsilon > 0$. We have

$$\lim_{h \to p} T^{h^{-1}} \phi = \lim_{h \to p} T^{h^{-1}} \lim_{g \to p} T^{g^{-1}} f = \lim_{g \to p \ast p} T^{g^{-1}} f = \phi$$
by Proposition 2.6.29 because \( p \) is idempotent. Combined with

\[
|T^g \phi - \phi|^2 = \langle T^g \phi, T^g \phi \rangle - \langle T^g \phi, \phi \rangle - \langle \phi, T^g \phi \rangle + \langle \phi, \phi \rangle
\]

we see that \( A := \{ g \in G : |T^g \phi - \phi| < \varepsilon/2 \} = A^{-1} \) belongs to \( p \). Thus \( AA^{-1} \) is left syndetic by Lemma 2.6.34. Let \( F \subset G \) be finite with \( F^{-1} AA^{-1} = G \). Fix \( g \in G \) and write \( g^{-1} = k^{-1} ab^{-1} \) accordingly. We have

\[
|T^g \phi - T^k \phi| = |T^{ab^{-1}} \phi - \phi| = |T^b \phi - T^a \phi| \leq \varepsilon
\]

so the orbit \( \{ T^g \phi : g \in G \} \) is covered by the balls of radius \( \varepsilon \) centered at \( T^k \phi \) as \( k \) runs through \( F \).

It follows that for any \( \xi \in L^2(X) \) the function \( g \mapsto \langle T^g \phi, \xi \rangle \) is almost periodic. It is therefore constant because \( G \) is weakly mixing. Thus \( T^g \phi = \phi \) in \( L^2(X) \) for every \( g \in G \).

Let \( \varphi \) be a representative of \( \phi \). We have \( T^g \varphi \sim \varphi \) for every \( g \in G \), where \( \sim \) denotes equality almost everywhere. Since \( \mathcal{I}X \) contains the measure zero sets, it follows that \( \varphi \) is \( \mathcal{I}X \) measurable and that \( \phi \in L^2(\mathcal{I}X) \).

Lastly, for any \( \psi \in L^2(\mathcal{I}X) \) we have

\[
\int \phi \cdot \psi \, d\mu = \lim_{g \to p} \int T^g f \cdot \psi \, d\mu = \lim_{g \to p} \int f \cdot (T^g)^{-1} \psi \, d\mu = \int f \cdot \psi \, d\mu
\]

so \( \phi \) is the orthogonal projection of \( f \) on \( L^2(\mathcal{I}X) \).

We will need the following construction of the relatively independent self-joining of a \( G^2 \) system \( X \) over \( \mathcal{I}X \). As in [Fur81a, Chapter 5], one could attempt to construct such a joining using a disintegration of \( \mu \) over \( \mathcal{I}X \), but the existence of such a disintegration is not clear when \( X \) is non-metrizable. In our setting, the need for such a disintegration can be circumvented by using limits along idempotent ultrafilters to give an explicit description of the ergodic projection.

**Lemma 7.4.3.** Let \( G \) be a countable, discrete, weakly mixing group and let \( X \) a \( G \) system. Then there exists a unique Baire measure \( \nu \) on \( X \times X \) such that

\[
\int f_1 \otimes f_2 \, d\nu = \int \mathcal{E}(f_1 | \mathcal{I}X) \cdot \mathcal{E}(f_2 | \mathcal{I}X) \, d\mu \tag{7.4.4}
\]

for any \( f_1, f_2 \in C(X) \).

**Proof.** Uniqueness follows immediately by density of \( C(X) \otimes C(X) \) in \( C(X \times X) \), so it remains to show the existence. To this end fix a minimal left idempotent ultrafilter \( p \) on \( G \). Since \( G \) is weakly mixing Theorem 7.4.1 implies that

\[
\lim_{g \to p} T^g f = \mathcal{E}(f | \mathcal{I}X)
\]
in the weak topology of $L^2(X)$ for every $f \in L^2(X)$.

Let $\delta : X \to X^2$ be the diagonal embedding and let $\lambda$ be the push-forward $\delta \mu$. Define an action $R$ of $G$ on $X^2$ by $R^g(x_1, x_2) = (x_1, T^g x_2)$. For any $f_1, f_2 \in C(X)$ we have

$$
\lim_{g \to p} \int f_1 \otimes f_2 \, d(R^g \lambda) = \lim_{g \to p} \int f_1 \cdot T^g f_2 \, d\mu = \int \mathbb{E}(f_1|\mathcal{F}X) \cdot \mathbb{E}(f_2|\mathcal{F}X) \, d\mu
$$

by the above. Since the space of Baire probability measures on $X^2$ is a compact, Hausdorff space the sequence $g \mapsto R^g \lambda$ has a limit along $p$. Let $\nu$ be this limit. The above calculation implies that

$$
\int f_1 \otimes f_2 \, d\nu = \int \mathbb{E}(f_1|\mathcal{F}X) \cdot \mathbb{E}(f_2|\mathcal{F}X) \, d\mu
$$

for all $f_1, f_2 \in C(X)$ as desired. \hfill \square

Recall from Section 2.5 that, given a $G$ system $X$, the measure $\nu$ obtained by applying Lemma 7.4.3 to the system is called the **relatively independent self-joining** of $\mu$ over $\mathcal{F}_2 X$. It follows immediately from (7.4.4) and the properties of conditional expectation that $\nu$ is invariant under the commuting $G$ actions $R_1 = T_1 \times T_{12}$ and $R_2 = T_2 \times I$. Thus we have produced a $G^2$ system. Lastly, writing $\pi_1$ and $\pi_2$ for the coordinate projections $X^2 \to X$, note that (7.4.4) implies that $\pi_1 \nu = \pi_2 \nu = \mu$ because all three measures agree on $C(X)$.

We now turn to the main result of this section.

**Theorem 7.4.5.** Let $G$ be a discrete, countable group that is weakly mixing group. Let $X$ be a $G^2$ system that is sated for $\mathcal{F}_{1,2}$. Then

$$
\lim_{g \to p} \int f_0 \cdot T_{1}^{g} f_1 \cdot T_{12}^{g} f_2 \, d\mu = \lim_{g \to p} \int f_0 \cdot T_{1}^{g} \mathbb{E}(f_1|\mathcal{F}_{1,2} X) \cdot T_{12}^{g} f_2 \, d\mu
$$

for any $f_0, f_1, f_2 \in L^\infty(X)$ and every minimal left idempotent ultrafilter $p$ on $G$.

**Proof.** Define $u : G \to L^2(X)$ by $u(g) = T_1^g \phi \cdot T_{12}^g f_2$ where $\phi = f_1 - \mathbb{E}(f_1|\mathcal{F}_{1,2} X)$. We have

$$
\lim_{g \to p} \langle u(hg), u(g) \rangle = \lim_{g \to p} \int (\phi \cdot T_1^h \phi) \cdot T_2^g (f_2 \cdot T_{12}^h f_2) \, d\mu
$$

$$
= \int \mathbb{E}(\phi \cdot T_1^h \phi|\mathcal{F}_2 X) \cdot \mathbb{E}(f_2 \cdot T_{12}^h f_2|\mathcal{F}_2 X) \, d\mu
$$

for every $h \in G$ by Theorem 7.4.1.

Let $\nu$ be the relatively independent self-joining of $\mu$ over $\mathcal{F}_2 X$ from Lemma 7.4.3. Write $\pi_1$ and $\pi_2$ for the coordinate projections $X \times X \to X$. Define

$$
Z_2 = (X \times X, \mathcal{B} \otimes \mathcal{B}, \nu, S)
$$

104
where $S_1 = T_1 \times T_{12}$ and $S_2 = T_2 \times I$. We know from the above that $Z_2$ is a $G^2$ system and that $\pi_1 : Z_2 \to X$ is a factor map. Next, notice that

$$\lim_{h \to p} \lim_{g \to p} \langle u(hg), u(g) \rangle$$

$$= \lim_{h \to p} \int (\phi \otimes f_2) \cdot S_1^h (\phi \otimes f_2) \, d\nu$$

$$= \int \phi \circ \pi_1 \cdot f_2 \circ \pi_2 \cdot E(\phi \otimes f_2 | \mathcal{I}_1 Z_2) \, d\nu$$

by another application of Theorem 7.4.1. We have $f_2 \circ \pi_2 \in \mathcal{I}_2 Z_2$, so we may replace $\phi \circ \pi_1$ with its conditional expectation on $\mathcal{I}_{1,2} Z_2$, which is zero by our satedness assumption. □
BIBLIOGRAPHY


