PARTITION REGULAR POLYNOMIAL PATTERNS IN COMMUTATIVE SEMIGROUPS

DISSERTATON

Presented in partial fulfillment of the requirements for the degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Joel Moreira, B.S.

Graduate Program in Mathematics

The Ohio State University

2016

Dissertation Committee:

Vitaly Bergelson, Advisor

Alexander Leibman

Nimish Shah

ABSTRACT

In 1933 Rado characterized all systems of linear equations with rational coefficients which have a monochromatic solution whenever one finitely colors the natural numbers. A natural follow-up problem concerns the extension of Rado's theory to systems of polynomial equations. While this problem is still wide open, significant advances were made in the last two decades. We present some new results in this direction, and study related questions for general commutative semigroups.

Among other things, we obtain extensions of a classical theorem of Deuber to the polynomial setting and prove that any finite coloring of the natural numbers contains a monochromatic triple of the form $\{x, x+y, xy\}$, settling an open problem.

We employ methods from ergodic theory, topological dynamics and topological algebra.

ACKNOWLEDGEMENTS

My thanks go first and foremost to my advisor Vitaly Bergelson, for his constant encouragement, contagious enthusiasm, and commitment to his students. I have learned a great deal from him: from what the word recalcitrant means to how to solve recalcitrant problems.

I want to thank Alexander Leibman and Nimish Shah for serving on my dissertation committee. I also thank the graduate school at the Ohio State University for awarding me the Presidential Fellowship, which gave me ample time to complete my dissertation. I also want to thank (in roughly the order we met) Afonso Bandeira, Donald Robertson, Daniel Glasscock, Florian Richter, Marc Carnovale, and Andreas Koutsogiannis for numerous interesting and fruitful mathematical discussions. Finally I thank my family for their constant support.

VITA

June 2006	High School Escola Secundária José Saramago
June 2009	B.S. in Mathematics University of Coimbra
September 2010 to Present	Graduate Teaching Associate Department of Mathematics Ohio State University

PUBLICATIONS

- V. Bergelson, J. Moreira. "Van der Corput's difference theorem: some modern developments". In: *Indag. Math. (N.S.)* 27 (2016), no. 2, 437–479.
- J. Moreira, F. Richter. "Large subsets of discrete hypersurfaces in $\mathbb{Z}^d$ contain arbitrarily many collinear points". In: *European J. Combin.* 54 (2016), 163–176..
- A. S. Bandeira, M. Fickus, D. G. Mixon, J. Moreira. "Derandomizing restricted isometries via the Legendre symbol". In: *Constr. Approx.* 43 (2016), no. 3, 409–424.
- J. Moreira. "Usando probabilidades para aproximar funções por polinómios" (in Portuguese), chapter in the book *Números, Cirurgias e Nós de Gravata: 10 Anos de Seminário Diagonal no IST*, IST press, 2012.

FIELDS OF STUDY

Major Field: Mathematics

Specialization: Ergodic Theory, Combinatorics

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In this thesis we investigate some questions originating in Ramsey theory by using tools and methods from ergodic theory, topological dynamics and topological algebra in the Stone-Čech compactifications of commutative semigroups.

1.1 Combinatorial problems

We are mainly concerned with results of the form: suppose one colors the natural numbers $\mathbb{N} = \{1, 2, \dots\}$ with a finite set of colors (so that each number has exactly one color) in an arbitrary fashion. Then one can find certain monochromatic (i.e. with all elements of the same color) "patterns". One of the oldest results in this direction is Schur's theorem [Sch16]:

Theorem 1.1 (Schur)**.** *For any finite partition of the natural numbers* $\mathbb{N} = C_1 \cup \dots \cup C_r$ *there exist* $x, y \in \mathbb{N}$ *and* $C \in \{C_1, \dots, C_r\}$ *such that* $\{x, y, x + y\} \subset C$*.*

Another famous result in arithmetic Ramsey theory is van der Waerden's theorem [Wae27] on arithmetic progressions:

Theorem 1.2 (van der Waerden)**.** *For any finite partition of the natural numbers* $\mathbb{N} = C_1 \cup \dots \cup C_r$ *and any* $k \in \mathbb{N}$ *there exist* $x, y \in \mathbb{N}$ *and* $C \in \{C_1, \dots, C_r\}$ *such that* $\{x, x + y, x + 2y, \dots, x + ky\} \subset C$*.*

Observe that both Schur's and van der Waerden's theorems deal with *linear* patterns, in the sense that these patterns are defined by linear relations (this point will be made more precise in Section 1.3). In 1933, Rado obtained a remarkable result, classifying all linear configurations which can be found in a single cell of any finite partition of $\mathbb{N}$, extending simultaneously Theorems 1.1 and 1.2 (cf. Theorem 4.31 below).

The next natural step in this line of investigation is to consider polynomial configurations. A far reaching extension of van der Waerden's theorem to the polynomial setting was obtained by Bergelson and Leibman in [BL96]:

Theorem 1.3 (Polynomial van der Waerden theorem, cf. [BL96, Corollary 1.11])**.** *Let $f_1, \dots, f_k \in \mathbb{Z}[x]$ be polynomials such that $f_i(0) = 0$ for all $i = 1, \dots, k$. Then for any finite coloring of $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exist a color $C \in \{C_1, \dots, C_r\}$ and $x, y \in \mathbb{N}$ such that*

$$\{x, x + f_1(y), x + f_2(y), \dots, x + f_k(y)\} \subset C$$

Note that Theorem 1.2 is a rather special case of Theorem 1.3, corresponding to the choice of polynomials $f_i(y) = iy$. The strength and generality of the polynomial van der Waerden theorem notwithstanding, several questions remain unanswered. For instance, it is an open problem whether for every finite coloring of $\mathbb{N}$ there exists a color C and $x, y \in \mathbb{N}$ such that $\{x^2, y^2, x^2 + y^2\} \subset C$.

This thesis sheds some new light on the class of polynomial patterns that can be found monochromatically within a single cell of any finite partition of $\mathbb{N}$ and, indeed, of more general countable commutative semigroups. We obtain in Chapter 4 a common generalization of Theorems 1.1 and 1.3 to countable commutative semigroups (the precise statement, Theorem 4.2, is postponed until Chapter 4 for it uses some terminology from that chapter), providing new families of polynomial patterns which can be found in a single cell of any finite partition. However, this extension does not include every polynomial pattern. While it unites the polynomial version of van der Waerden's theorem with Schur's theorem, some pieces are still missing. In particular, the following conjecture is still unsolved:

Conjecture 1.4. *For any finite partition of the natural numbers $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exist $x, y \in \mathbb{N}$ and $C \in \{C_1, \dots, C_r\}$ such that $\{x, y, x + y, xy\} \subset C$.*

An affirmative answer to the analogue of Conjecture 1.4 in finite fields was recently obtained by Green and Sanders [GS16], generalizing previous work by Shkredov [Shk10] and Cilleruelo [Cil12] (see also [Vin14] and [Han13] for related results).

A major part of this thesis arises from studying a weaker form of Conjecture 1.4 which, until very recently, was unsolved:

Theorem 1.5 ([Mor])**.** *For any finite partition of the natural numbers* $\mathbb{N} = C_1 \cup \cdots \cup C_r$ *there exist* $x, y \in \mathbb{N}$ *and* $C \in \{C_1, \ldots, C_r\}$ *such that* $\{x, x+y, xy\} \subset C$*.*

Theorem 1.5 is proved in Chapter 6, as a special case of a significantly more general statement (see Theorem 6.2 below). Theorem 1.5 is the culmination of ideas and techniques developed in [BM16a] and [BM16b], where the analogous problem in $\mathbb{Q}$ (and, in fact, in any countable field) was studied. Besides the invaluable lessons learned, in this earlier work we obtained results which in certain respect are stronger than Theorem 1.5. These developments are presented in Chapter 5. In particular we showed that any "large" subset of $\mathbb{Q}$ contains patterns of the form $\{x+y, xy\}$. By way of contrast, observe that the set of odd numbers (which is a "large" subset of $\mathbb{N}$ in several senses) can not contain such a pattern.

1.2 Connections with dynamics and topological algebra

In [Sze75] Szemerédi established a density version of van der Waerden's theorem (cf. Theorem 2.41 below), hereby proving a famous conjecture of Erdős and Turán [ET36]. Shortly afterwards, Furstenberg gave a new proof of Szemerédi's theorem using ergodic theory [Fur77]. This was the beginning of a long and fruitful interaction between dynamics and Ramsey theory. Indeed, Furstenberg's method was successfully applied to many other problems in Ramsey theory, including density and polynomial versions of the Hales-Jewett theorem.

While ergodic theory has proven useful in establishing density results, topological dynamics can be used to obtain directly partition results. This approach was first employed by

Furstenberg and Weiss to obtain a dynamical proof of van der Waerden's theorem [FW78]. A similar method was later used by Bergelson and Leibman to obtain the polynomial van der Waerden theorem (a result which was previously unknown). Topological dynamics can also be used to show partition regularity for configurations which are not present in every set with positive density, such as IP-sets (cf. Section 2.2 below) or solutions to Rado's systems of equations.

Another effective technique in modern Ramsey theory is provided by the topological algebra of the Stone-Čech compactification, realized as the space of ultrafilters. On the one hand, ultrafilter methods are helpful in obtaining various partition results. On the other hand, utilizing convergence along ultrafilters allows one to refine and extend results classically obtained via Cesàro averages.

1.3 Ramsey families

To put our results into perspective, we will now briefly review some of the relevant classical results in Ramsey theory. The notion of *Ramsey families* will be convenient to state both classical and new results in a more general framework.

Definition 1.6. Let G be a countable commutative semigroup, let $k, m \in \mathbb{N}$, and let $f_1, \dots, f_k : G^m \to G$. We say that $\{f_1, \dots, f_k\}$ is a *Ramsey family in G* if for any finite coloring $G = C_1 \cup \dots \cup C_r$, there exist $\mathbf{x} \in G^m$ and a color $C \in \{C_1, \dots, C_r\}$ such that $\{f_1(\mathbf{x}), \dots, f_k(\mathbf{x})\} \subset C$.

In this language, Schur's theorem (Theorem 1.1) states that the family $\{x, y, x+y\}$ [1] is Ramsey in $\mathbb{N}$, and van der Waerden's theorem (Theorem 1.2) states that, for any $k \in \mathbb{N}$, the family $\{x, x+y, \dots, x+(k-1)y\}$ is Ramsey in $\mathbb{N}$. On the other hand, the families $\{x, x+1\}$ and $\{x, y, 3x-y\}$ are *not* Ramsey (in $\mathbb{N}$): if one colors each $n \in \mathbb{N}$ depending on its parity, then x and $x+1$ must have different colors; and if we color each $n \in \mathbb{N}$ in one of

[1] By a slight abuse of notation, we represent by $\{x, y, x+y\}$ the family comprised of the three functions $(x, y) \mapsto x$, $(x, y) \mapsto y$ and $(x, y) \mapsto x+y$.

four colors, according to the last non-zero digit in its 5-adic expansion, one can check that if x and y are of the same color, then $3x - y$ must have a different color.

A common extension of Schur's and van der Waerden's theorems was obtained by Brauer [Bra28]:

Theorem 1.7 (Brauer's theorem)**.** *For every $p \in \mathbb{N}$, the family $\{x, y, x+y, x+2y, \ldots, x+py\}$ is Ramsey in $\mathbb{N}$.*

In a different direction, one can extend Schur's theorem by adding to the family $\{x, y, x+y\}$ a new variable z and the sums $\{z+y, z+y, z+x+y\}$. In general we have

Theorem 1.8 (Folkman's theorem)**.** *For every $m \in \mathbb{N}$, the family*

$$\left\{\begin{array}{lllll} x_0 & & & & \\ x_1, & x_1 + x_0 & & & \\ x_2, & x_2 + x_1, & x_2 + x_0, & x_2 + x_1 + x_0 & \\ \vdots & \vdots & \vdots & \ddots & \\ x_m, & x_m + x_{m-1}, & x_m + x_{m-2}, & \ldots & x_m + x_{m-1} + \cdots + x_0 \end{array}\right\}$$

is Ramsey in $\mathbb{N}$.

In other words, Folkman's theorem states that for any finite coloring of $\mathbb{N}$ and every $m \in \mathbb{N}$ there exists a color C and a set $A \subset \mathbb{N}$ with $|A| = m$ such that $FS(A) \subset C$ (cf. (2.4) below).

Remark 1.9. *The attribution of Theorem 1.8 to J. Folkman is made in [GRS90] (without a reference). However we remark that it follows as a corollary from the earlier work of Rado (cf. Theorem 4.31 below) and it was also independently discovered by Sanders in his thesis [San68] and by Arnautov [Arn70].*

In the same way that van der Waerden's and Schur's theorems were simultaneous generalized by Brauer's theorem, it is possible to unite van der Waerden's theorem with Folkmans's.

Theorem 1.10 (Deuber's theorem, [Deu73]). *For any $m, p, c \in \mathbb{N}$, the family*

$$\left\{ \begin{array}{lr} cx_0, & \\ ix_0 + cx_1, & i \in \{-p, \dots, p\} \\ ix_0 + jx_1 + cx_2, & i, j \in \{-p, \dots, p\} \\ \quad\vdots & \vdots \\ i_0x_0 + \cdots + i_{m-1}x_{m-1} + cx_m, & i_{m-1}, \dots, i_0 \in \{-p, \dots, p\} \end{array} \right\}$$

is Ramsey in $\mathbb{N}$.

Observe that Theorem 1.10 contains Schur's, van der Waerden's, Brauer's and Folkman's theorems as special cases. In fact, Deuber's theorem applies to every finite linear Ramsey family in $\mathbb{N}$:

Theorem 1.11. *Let $m \in \mathbb{N}$ and let $\{f_1, \dots, f_k\}$ be a finite family of linear functions (i.e. semigroup homomorphisms) $f_i : \mathbb{N}^{m+1} \to \mathbb{N}$. Then the family is Ramsey if and only if there exist $p, c \in \mathbb{N}$ such that the family $\{f_1, \dots, f_k\}$ is contained in the family described in Theorem 1.10.*

The original version of Theorem 1.11 is due to Rado (see Theorem 4.31 below) but is formulated in the (quite different) language of partition regularity of systems of linear equations. As stated, Theorem 1.11 is due to Deuber; although it is not hard to derive it from Rado's theorem. Actually, Deuber proved a stronger result (see the discussion surrounding Definition 4.10 below).

There are several ways one can hope to extend Theorem 1.11. One option is to extend the notion of Ramsey families to infinite families of functions. There are several positive results in this direction, including Hindman's theorem (Theorem 2.27) which can be formulated in language of Ramsey families. One could also try to extend the scope of Theorem 1.11 from linear Ramsey families in $\mathbb{N}$ to linear Ramsey families in other commutative semigroups. Yet another, and perhaps more natural, possibility is to relax the condition that the functions must be linear and ask instead for a similar result when the functions are polynomials. This leads naturally to the following, by now classical, problem.

Problem 1.12. *Describe necessary and sufficient conditions on the polynomials $f_1, \ldots, f_k \in \mathbb{Z}[x_1, \ldots, x_s]$ that guarantee that the family $\{f_1, \ldots, f_k\}$ is Ramsey in $\mathbb{N}$.*

It follows from Schur's theorem that the family $\{x, y, xy\}$ is Ramsey (simply compose any given coloring $\chi : \mathbb{N} \to \{1, \ldots, r\}$ with the map $n \mapsto 2^n$ to create a new coloring and apply Schur's theorem). Using the same idea, van der Waerden's theorem implies that for each $k \in \mathbb{N}$ the family $\{x, xy, \ldots, xy^k\}$ is Ramsey, and Deuber's theorem implies that many more families of the form $\{f_1, \ldots, f_k\}$, where each f_i is a monomial, are Ramsey.

Configurations which combine both addition and multiplication, however, tend to be significantly harder to deal with: only in 1977 did Furstenberg and Sárközy prove, independently, that the family $\{x, x+y^2\}$ is monochromatic (cf. [Fur77, Theorem 1.2] and [Sár78]), obtaining the first example of a non-linear Ramsey family which does not consist solely of monomials. Bergelson improved this result by showing that in fact the family $\{x, y, x+y^2\}$ is Ramsey [Ber87].

The next major advance towards Problem 1.12 was Bergelson and Leibman's polynomial extension of van der Waerden's theorem [BL96] (Theorem 1.3). In the language of Ramsey families, they showed in particular that for any polynomials $p_1, \ldots, p_k \in \mathbb{Z}[x_1, \ldots, x_m]$ with $p_i(\mathbf{0}) = 0$, the family $\{x_0, x_0 + p_1(x_1, \ldots, x_m), \ldots, x_0 + p_k(x_1, \ldots, x_m)\}$ is Ramsey. The polynomial van der Waerden theorem has now been extended in several directions (see, for instance, [BFM96; BJM; BLL08]), each revealing new examples of polynomial Ramsey families.

In the last decade, many interesting polynomial Ramsey families were found [BBHS06; BBHS08; Ber05; FH; McC10], however a complete solution to Problem 1.12 is still very far from reach.

Outline of the following chapters

In the next chapter we set up some common terminology and definitions. We start by reviewing some background facts on ultrafilters. Then we survey classical definitions and results from (arithmetic) Ramsey theory and discuss connections with ergodic theory and topological dynamics.

In Chapter 3 we deal with affine semigroups – semigroups of affine transformations of a ring – and explore some of their properties. There is a strong analogy between the way a semigroup acts on itself by translations, and the way the affine semigroup over a ring R acts on R via affine transformations. This analogy allows us to transfer several classical results to the affine setting. The theory developed in this chapter is used later in Chapters 5 and 6 to obtain new Ramsey theoretic results.

In Chapter 4 we study extensions of Deuber's theorem (Theorem 1.10). We obtain a common extension of Deuber's theorem and Bergelson-Leibman's polynomial van der Waerden theorem. We further generalize this result to arbitrary countable abelian groups. We also present a generalization of Deuber's theorem which holds in every countable commutative semigroup.

In Chapter 5 we investigate the presence of configurations $\{x+y, xy\}$ in large subsets of (countably infinite) fields and answer the question of how abundant such configurations are. To this end we employ methods from ergodic theory and end up obtaining results on the long term behaviour of measure preserving actions of affine semigroups. In particular we establish affine analogues of the mean ergodic theorem, Khintchine's recurrence theorem, as well as versions thereof involving limits along ultrafilters.

Chapter 6 is dedicated to the study of certain polynomial families over a general class of rings. Among other things, we show that the family $\{x+y, xy\}$ is Ramsey in any ring of that class and in $\mathbb{N}$. We provide two proofs, one using topological dynamics and another, purely elementary proof. In particular we prove Theorem 1.5 and, as a corollary, obtain partition regularity of certain polynomial equations.

CHAPTER 2

PRELIMINARIES

In this section we will review some well known combinatorial results which we both are inspired by and use throughout the thesis. We will also introduce some definitions and facts about ultrafilters, ergodic theory and topological dynamics for later use.

2.1 Ultrafilters on commutative semigroups

The theory of ultrafilters, and specially the relation between the algebra and the topology of the set βR of all ultrafilters over a countable set R has became a major component of Ramsey theory in the past decades. Since we will make extensive use of the theory of ultrafilters in all of the later chapters, in this section we present a fairly detailed introduction to this useful subject. The reader will find missing details in [Ber10] or [HS98].

Definition 2.1 (Ultrafilter)**.** Let R be a countable set. An *ultrafilter* p is a non-empty collection of subsets of R satisfying:

- $\varnothing \notin p$
- If $A \subset B$ and $A \in p$ then $B \in p$
- If $A \in p$ and $B \in p$ then also $A \cap B \in p$
- If $A \cup B \in p$ then either $A \in p$ or $B \in p$

In particular, the second condition implies that $R \in p$ (because p is non-empty), so from the last property it follows that for each $A \subset R$, either $A \in p$ or $R \setminus A \in p$. In fact, iterating the last property, we deduce that for any finite partition of R, exactly one of the cells belongs to p. Indeed we have:

Proposition 2.2. *An ultrafilter on R is a family p of subsets of R such that for any finite partition of R, exactly one of the cells of the partition belongs to p.*

It is not hard to present a concrete example of an ultrafilter: pick $a \in R$ and consider the family $p_a := \{A \subset R : a \in A\}$; one can promptly check that this family satisfies all the conditions from Definition 2.1. Ultrafilters of the form p_a with $a \in R$ as just described are called *principal*, an ultrafilter which is not principal is called *non-principal*. In order to prove existence of non-principal ultrafilters one requires (at least some weak form of) the axiom of choice. This means that one can not give an explicit construction of an non-principal ultrafilter; nevertheless, one is often able to establish the existence of ultrafilters with some nice properties.

A family of subsets of R satisfying the first three properties of the Definition 2.1 is called a *filter*. The family of all co-finite sets is an example of a filter which is not contained in any principal ultrafilter. On the other hand, ultrafilters turn out to be precisely maximal filters (for the inclusion relation).

Proposition 2.3. *Every filter is contained in an ultrafilter. In particular, there exist non-principal ultrafilters.*

Proof. We first show that any maximal filter (for the partial order of inclusion) is an ultrafilter. Let p be a maximal filter and let $A, B \subset R$ be non-empty sets such that $A \cup B \in p$. We need to show that either A or B belong to p. Assume $B \notin p$ and consider the family $p' := p \cup \{C \subset R : A \subset C\} \cup \{A \cap C : C \in p\}$. We claim that p' is also a filter, by maximality this will imply that $p' = p$ and so $A \in p$ as desired.

Assume, for the sake of a contradiction, that $\varnothing \in p'$. Since p is a filter and A is non-empty we have that $A \cap C = \varnothing$ for some $C \in p$. But since $A \cup B \in p$, we would have $(A \cup B) \cap C$ is simultaneously a member of p and also a subset of B, contradicting the assumption that $B \notin p$. The second and third conditions for a filter trivially hold for p' and this proves the claim.

Now, it is also easy to check that the union of any totally ordered subset of filters is again a filter, so we can apply Zorn's lemma to conclude that any filter is contained in some ultrafilter. $\square$

If one considers R to be endowed with the discrete topology (as we always do in this thesis), the set of all ultrafilters on R can be identified with the Stone-Čech compactification βR of R (see, for example, Theorem 3.27 in [HS98]). We will now briefly explain this identification. Let $\Omega := \{0,1\}^R$ be the set of all functions $f : R \to \{0,1\}$. By identifying a subset $A \subset R$ with its indicator function, we can also think of Ω as the family of all subsets of R. We endow $\{0,1\}$ with the discrete topology and give Ω the product topology, making it into a compact space (by Tychonoff's theorem). Next we consider the space $X = \{0,1\}^\Omega$ of all functions from Ω in $\{0,1\}$. Invoking Tychonoff's theorem one more time we deduce that X is also compact with respect to the product topology. One can think of a point in X as a subset of Ω or, equivalently, as a collection of subsets of R.

Now, for each $n \in R$, let $p_n \in X$ be the function $p_n : \Omega \to \{0,1\}$ given by $p_n(f) = f(n)$ for each $f : R \to \{0,1\}$. This gives an embedding of R into X. The Stone-Čech compactification of R is the closure of R in X and is denoted by βR. Observe that βR is compact because it is a closed subset of the compact Hausdorff space X.

A point $p \in \beta R \subset X$ is a map from Ω to $\{0,1\}$. By identifying points in Ω with subsets of R, we can associate p with the family of subsets $A \subset R$ for which $p(A) = 1$. By an abuse of language we will denote this collection of subsets of R also by p and we will use interchangeably the notations $A \in p$ and $p(A) = 1$. Note that the element $p_n \in X$ is the principal ultrafilters at n.

Proposition 2.4. *Let $p \in X$ be a collection of subsets of R. Then $p \in \beta R$ if and only if p is an ultrafilter on R.*

Proof. Using the definition of product topology we see that $p \in \beta R$ if and only if for any finite collection $A_1, ..., A_k$ of subsets of R there exists some $n \in R$ such that $p_n(A_i) = p(A_i)$ for each $i = 1, ..., k$.

First assume that $p \in \beta R$. Since $p_n(\varnothing) = 0$ for all $n \in R$, also $p(\varnothing) = 0$ and hence $\varnothing \notin p$, proving the first property of Definition 2.1. Next let $A \in p$ and suppose $B \supset A$. Let $n \in R$ be such that $p_n(A) = p(A)$ and $p_n(B) = p(B)$. Then $n \in A$ so $n \in B$ and thus $p(B) = 1$, proving the second property. If both A and B are in p, let $n \in R$ be such that p_n and p agree at A, B and $A \cap B$. Then we conclude that $A \cap B \in p$ proving the third property. Finally suppose that $A \cup B \in p$ and let $n \in R$ be such that p_n agree with p at A, B and $A \cup B$. Thus either A or B will be in p proving the fourth property. This implies that p is indeed an ultrafilter.

Next, suppose that p is an ultrafilter. Given any subsets $A_1, ..., A_k$ of R, assume that $A_1, ..., A_r$ are in p and $A_{r+1}, ..., A_k$ are not in p. Then the intersection

$$\left(\bigcap_{i=1}^{r} A_i\right) \cap \left[\bigcap_{i=r+1}^{k} (R \setminus A_i)\right]$$

is in p and in particular it is non-empty. Let n be in that intersection. Then p_n agrees with p at the sets $A_1, ..., A_k$. Since the sets $A_1, ..., A_k$ were arbitrary, we conclude that there exists some p_n at each neighborhood of p, and hence $p \in \beta R$. $\square$

For a set $A \subset R$ denote by

$$\bar{A} := \{p \in \beta R : A \in p\} \subset \beta R \tag{2.1}$$

It turns out that, identifying A with the subset of βR consisting of those principal ultrafilters p_n with $n \in A$, the set $\bar{A}$ is precisely the closure of A in the topology of $\beta\mathbb{N}$. In fact, we have:

Lemma 2.5. *The sets $\bar{A}$ are clopen and form a basis for the topology on βR.*

Proof. Let $p \in X$. By definition $p \in \bar{A} \iff p(A) = 1$. Since $\{1\}$ is a clopen subset of $\{0, 1\}$ and by the definition of product topology on X we conclude that $\bar{A}$ is a clopen set of X, hence intersecting with βR we get a clopen subset of βR.

To prove that the sets $\{\bar{A}, A \subset R\}$ form a basis for the topology on βR, let $k \in \mathbb{N}$, let $1 \leq r \leq k$, let $A_1, ..., A_k$ be subsets of R and let $C := \{p \in X : A_1, ..., A_r \in p, A_{r+1}, ..., A_k \notin p\}$. Sets of the form C are a basis for the topology of X, so sets of the form $C \cap \beta R$ form a basis for the topology of βR. Consider the intersection

$$B = \left(\bigcap_{i=1}^{r} A_i\right) \cap \left[\bigcap_{i=r+1}^{k} (R \setminus A_i)\right]$$

Then, by the ultrafilter property, $p \in C \cap \beta R \iff B \in p$, and hence $C \cap \beta R = \bar{B}$. $\square$

Given a compact Hausdorff space K and a function $f : R \to K$, the universal property of the Stone-Čech compactification implies that there exists a unique continuous extension $\beta f : \beta R \to K$. The continuity implies that $\beta f(p) = \lim_{n \to p} f(n)$. We will denote this by the more suggestive notation $p\text{-}\lim_n f(n) := \beta f(p)$.

Lemma 2.6. *Let K be a compact Hausdorff space, let R be a countable discrete set, let $f : R \to K$, $p \in \beta R$ and $x \in K$. Then*

$$p\text{-}\lim_{n} f(n) = x \qquad \iff \qquad \forall\, U \ni x \text{ open}, \quad \{n \in R : f(n) \in U\} \in p. \tag{2.2}$$

Proof. Assume $p\text{-}\lim_n f(n) = x$. The set $\{n \in R : f(n) \in U\}$ is precisely the intersection of $V := (\beta f)^{-1}(U)$ and R (as a subset of βR). Since βf is continuous, V is open, and because $p \in V$ it follows that there exists a basic open set $\bar{A} \subset V$ such that $p \in \bar{A}$, where $A \subset R$. In other words $A \in p$ and $A \subset V$, which implies that $V \cap R$ contains A and hence is a member of p.

Next we prove the converse: assume $p\text{-}\lim f(n) = y \neq x$. Since K is Hausdorff, take disjoint neighborhoods U_x and U_y of x and y respectively. It follows from the first part that $\{n : f(n) \in U_y\} \in p$. Since $\{n : f(n) \in U_y\} \cap \{n : f(n) \in U_x\} = \varnothing$ we conclude that $\{n : f(n) \in U_x\} \notin p$, finishing the proof. $\square$

One could simply take (2.2) as the definition of $p\text{-}\lim$ but the definition above makes it more clear that for *any* map $f : R \to K$ into a compact Hausdorff space, $p\text{-}\lim_n f(n)$ exists and is unique.

Ultrafilters on semigroups

So far we didn't use any property of the set R (except that it is infinite). Now we will see that a binary operation on R induces a binary operation in βR. To motivate the definition it helps to think of ultrafilters in a different way, namely as finitely additive $\{0,1\}$-valued measures on R. In that way, the operation on βR is just the usual convolution of measures.

Let $\bullet$ be an associative binary operation on R (we will only use $\bullet = +$ and $\bullet = \times$), let $A \subset R$ and $n \in R$. We use the notation $A \bullet n^{-1}$ to denote the set $\{x \in R : x \bullet n \in A\}$.

Definition 2.7. Let p and q be ultrafilters on a semigroup $(R, \bullet)$. We define the operation

$$p \bullet q := \{A \subset R : \{n \in R : A \bullet n^{-1} \in p\} \in q\}$$

We first need to check that $p \bullet q$ is indeed an ultrafilter on R:

Proposition 2.8. *If $p, q \in \beta R$ then also $p \bullet q \in \beta R$.*

Proof. It is clear that $\varnothing \notin p \bullet q$. Also, if $A \subset B$, then for each n we have $(A \bullet n^{-1}) \subset (B \bullet n^{-1})$. It is now easy to check that if $A \in p \bullet q$ then also $B \in p \bullet q$. Next assume that both $A, B \in p \bullet q$. Since $(A \bullet n^{-1}) \cap (B \bullet n^{-1}) = (A \cap B) \bullet n^{-1}$ for each $n \in R$, we have that $\{n : (A \cap B) \bullet n^{-1} \in p\} = \{n : A \bullet n^{-1} \in p\} \cap \{n : B \bullet n^{-1} \in p\} \in q$ and hence $A \cap B \in p \bullet q$.

Finally, if $A \cup B \in p \bullet q$ then using the fact that p is an ultrafilter (and the fact that $(A \cup B) \bullet n^{-1} = A \bullet n^{-1} \cup B \bullet n^{-1}$) we have that for each n in the set $C := \{n : (A \cup B) \bullet n^{-1} \in p\}$ either $A \bullet n^{-1} \in p$ or $B \bullet n^{-1} \in p$. Since q is also an ultrafilter and $C \in q$, either $\{n : A \bullet n^{-1} \in p\} \in q$ or $\{n : B \bullet n^{-1} \in p\} \in q$ which is equivalent, respectively to $A \in p \bullet q$ or $B \in p \bullet q$. $\square$

Moreover, this binary operation turns out to be associative (cf. Theorems 4.1, 4.4 and 4.12 in [HS98]) :

Proposition 2.9. *The binary operation on βR just defined is associative.*

Proof. We first note that for $A \subset R$ and $n, m \in R$ we have

$$x \in (A \bullet n^{-1}) \bullet m^{-1} \iff x \bullet m \bullet n \in A$$

and so $(A \bullet n^{-1}) \bullet m^{-1} = A \bullet (m \bullet n)^{-1}$.

Let $p, q, r \in \beta R$. Then $A \in (p \bullet q) \bullet r$ if and only if

$$\begin{aligned}
\{n : A \bullet n^{-1} \in p \bullet q\} \in r &\iff \{n : \{m : (A \bullet n^{-1}) \bullet m^{-1} \in p\} \in q\} \in r \\
&\iff \{n : \{m : A \circ (m \bullet n)^{-1} \in p\} \in q\} \in r \\
&\iff \{n : \{m : m \bullet n \in \{x : A \bullet x^{-1} \in p\}\} \in q\} \in r \\
&\iff \{n : \{x : A \circ x^{-1} \in p\} \bullet n^{-1} \in q\} \in r \\
&\iff \{x : A \bullet x^{-1} \in p\} \in q \bullet r \\
&\iff A \in p \bullet (q \bullet r)
\end{aligned}$$

□

Thus $(\beta R, \bullet)$ is a semigroup. It should be remarked that this operation extends the operation on R. More precisely, if $n, m \in R$ then $p_n \bullet p_m := p_{n \bullet m}$. However, the operation in βR may not be commutative, even if $\bullet$ is.

Another important property of the operation in βR is that it is left continuous (and yet, not right continuous):

Proposition 2.10. *For each $p \in \beta R$, the map $\lambda_p : \beta R \to \beta R$ defined by $\lambda_p : q \mapsto p \bullet q$ is continuous.*

Proof. We will use the Lemma 2.5. Fix $p, q \in \beta R$ and let $\bar{A}$ be a clopen neighborhood of $\lambda_p(q)$. We need to show that $\{r : \lambda_p(r) \in \bar{A}\}$ contains $\bar{B}$ for some $B \in q$. We observe that $\lambda_p(r) \in \bar{A} \iff A \in p \bullet r \iff \{n : A \bullet n^{-1} \in p\} \in r$. Thus making $B = \{n : A \bullet n^{-1} \in p\}$ we get that indeed $\{r : \lambda_p(r) \in \bar{A}\}$ contains $\bar{B}$. □

Quite special elements of the semigroup $(\beta R, \bullet)$ are idempotent elements, i.e, ultrafilters p such that $p \bullet p = p$. While the existence of such ultrafilters is not obvious, it can be establish using a theorem by Ellis [Ell58]:

Theorem 2.11 (Ellis Theorem)**.** *If $(S, \bullet)$ is a compact Hausdorff left topological semigroup, then S contains an idempotent.*

Proof. Consider the family

$$V := \{\varnothing \neq W \subset S : W \text{ is compact }, W \bullet W := \{w_1 \bullet w_2 : w_1, w_2 \in W\} \subset W\}$$

V is non-empty because $S \in V$. Also the intersection of any nested subfamily of V is still in V (because a nested intersection of compacts in a Hausdorff space is non-empty). Applying Zorn's lemma we find a minimal (for the partial order of inclusion) element $W \in V$.

For each $x \in W$ we have $(x \bullet W) \bullet (x \bullet W) \subset x \bullet W \bullet W \bullet W \subset x \bullet W$, and by left continuity $x \bullet W$ is compact, hence $x \circ W \in V$. Also $x \bullet W \subset W \bullet W \subset W$, so by minimality $x \bullet W = W$.

In particular $x = x \bullet y$ for some $y \in W$. Thus the set $Z := \{z \in W : x \bullet z = x\}$ is non-empty. By continuity we have that Z is closed (hence compact) and if $y, z \in Z$ then $x \bullet (y \bullet z) = x \bullet z = x$, hence $Z \bullet Z \subset Z$. This implies that $Z \in V$, again by minimality we have that $Z = W$ and in particular $x \in Z$. We conclude that $x \bullet x = x$, and this is our idempotent. $\square$

This implies that there are idempotent ultrafilters on βR.

2.2 Notions of largeness for subsets of semigroups

Throughout this section we let G be a countable commutative left cancelative semigroup. This means that there is an associative binary operation on G such that for every $a, b \in G$ we have $ab = ba$ and the map $x \mapsto ax$ is injective. Most of the definitions and results hold in bigger generality (perhaps after properly changing the order of the operations), but even though non-commutative semigroups will make an appearance in this thesis, we will not make use of these notions of largeness in the non-commutative setting.

Given a set $A \subset G$ and an element $x \in G$ we denote by $xA := \{xa : a \in A\}$ and $x^{-1}A := \{y \in G : xy \in A\}$. Given two subsets $F, A \subset G$ we employ the notation $F^{-1}A := \bigcup_{x \in F} x^{-1}F = \{y \in G : (\exists x \in F) xy \in A\}$.

Upper density

In this subsection we write the operation in G multiplicatively, to draw the analogy with the similar theory for actions of affine semigroups which we will explore later in Chapter 3.

A *Følner sequence* in a countable commutative semigroup G is a sequence $(F_N)_{N\in\mathbb{N}}$ of finite subsets of G which is asymptotically invariant under the semigroup action, in the sense that

$$\forall g \in G \qquad \lim_{N\to\infty} \frac{\big|(gF_N) \cap F_N\big|}{|F_N|} = 1$$

It is a well known fact that in any countable commutative semigroup there exist Følner sequences [Pat88, Sec (4.22)].

Remark 2.12. *For cancelative semigroups, the size $|F_N|$ of the sets in a Følner sequence must grow to infinity as $N \to \infty$. However, certain non-cancelative semigroups admit only rather trivial Følner sequences. For instance, taking G to be the set of all finite subsets of $\mathbb{N}$ with the operation being the intersection, the (essentially unique) Følner sequence is obtained by letting each $F_N = \{\varnothing\}$.*

Example 2.13.

- For the semigroup $G = (\mathbb{N}, +)$, any sequence of intervals with increasing length will be a Følner sequence. The most common example is $F_N = \{1, 2, \ldots, N\}$.

- More generally, in the semigroup $G = (\mathbb{N}^d, +)$, for some $d \in \mathbb{N}$, one can take a Følner sequece to be any sequence of cubes, whose side length goes to infinity. In particular, the sequence $F_N := \{1, \ldots, N\}^d$ is a Følner sequence.

- In the multiplicative semigroup $(\mathbb{N}, \cdot)$, one can use the prime numbers to describe a Følner sequence as follows. Let $p_1, p_2, \ldots$ be an arbitrary enumeration of (all) the primes and let $F_N = \{p_1^{e_1} \cdots p_N^{e_N} : 0 \leq e_1, \ldots, e_N \leq N\}$. One can easily check from the definition that $(F_N)_{N\in\mathbb{N}}$ is indeed a Følner sequence.

- For a positive integer $N \in \mathbb{N}$ let R_N be the product of the first N prime numbers and, using the symbol $x|y$ to represent the statement that x divides y, define:

$$F_N := \left\{ \frac{a}{b} : \ a, b \in \mathbb{N}, \ R_N^{N+1}|b, \ b|R_N^{2N}, \ 1 \leq a \leq R_N^{4N}, \ \gcd(a,b) = 1 \right\} \subset \mathbb{Q}$$

  The sequence $(F_N)_{N\in\mathbb{N}}$ is a Følner sequence in the groups $(\mathbb{Q}, +)$ and $(\mathbb{Q}^{>0}, \cdot)$ simultaneously.

Given a Følner sequence $(F_N)_{N\in\mathbb{N}}$ in a semigroup, one can define the *upper density with respect to* $(F_N)_{N\in\mathbb{N}}$ via the formula

$$\bar{d}_{(F_N)}(E) := \limsup_{N\to\infty} \frac{|E \cap F_N|}{|F_N|} \tag{2.3}$$

When the Følner sequence is tacit, we denote $\bar{d}_{(F_N)}(E)$ simply by $\bar{d}(E)$. In particular, in $\mathbb{N}$ or $\mathbb{Z}$, we will denote by $\bar{d}$ the upper density with respect to the Følner sequence $(\{1, \ldots, N\})_{N\in\mathbb{N}}$.

The upper density satisfies the following properties:

Proposition 2.14. *Let G be a countable commutative cancelative semigroup, let $(F_N)_{N\in\mathbb{N}}$ be a Følner sequence in G and let $\bar{d}$ be the upper density with respect to $(F_N)_{N\in\mathbb{N}}$. Then for any $A, B \subset G$ and $g \in G$ we have*

1. $\bar{d}(G) = 1$,
2. $\bar{d}(gA) = \bar{d}(g^{-1}A) = \bar{d}(A)$,
3. $\bar{d}(A \cup B) \leq \bar{d}(A) + \bar{d}(B)$.

Syndetic, thick and piecewise syndetic sets

In this subsection we continue to denote the (commutative) operation on the semigroup G multiplicatively.

Definition 2.15. A subset S of a commutative semigroup (written multiplicatively) is *syndetic* if there exists a finite set $F \subset G$ such that $F^{-1}A = G$.

Example 2.16. The following are syndetic sets in the semigroup $(\mathbb{N}, +)$:

- Any co-finite subset $S \subset \mathbb{N}$,
- Any infinite progression of the form $a\mathbb{N} + b$,
- The set $\{n \in \mathbb{N} : d(n\alpha, \mathbb{N}) < \varepsilon\}$ for any $\alpha, \varepsilon > 0$, where $d(x, \mathbb{N})$ denotes the distance between the positive real number x and the lattice of natural numbers.
- The set $\{n \in \mathbb{N} : d(f(n), \mathbb{N}) < \varepsilon\}$ for any $\varepsilon > 0$ and any polynomial $f \in \mathbb{R}[x]$ with an irrational coefficient (other than the constant term).

One can show that a set S is syndetic if and only if for every Følner sequence $(F_N)_{N\in\mathbb{N}}$ in G, the upper density $\bar{d}_{(F_N)}(S)$ is positive.

Definition 2.17. A subset S of a commutative semigroup (written multiplicatively) is *thick* if for every finite set $F \subset G$ there exists $x \in G$ such that $Fx \subset T$.

Example 2.18. The following are thick sets in $(\mathbb{N}, +)$:

- Any co-finite subset $T \subset \mathbb{N}$,
- The set $\{n \in \mathbb{N} : |\hat{\mu}(n)| < \varepsilon\}$ where $\varepsilon > 0$ and $\hat{\mu}$ is the Fourier transform of a non-atomic measure μ on $[0, 1]$.

Observe that given any syndetic set S and any thick set T, the intersection $S \cap T$ is non-empty. In fact, the notions of thick and syndetic sets are dual, in the sense that a set is thick if and only if it has non-empty intersection with every syndetic sets, and conversely, a set is syndetic if and only if it has non-empty intersection with every thick set. One can show that a set $T \subset G$ is thick if and only if there exists a Følner sequence $(F_N)_{N\in\mathbb{N}}$ in G for which the upper density $\bar{d}_{(F_N)}(T)$ equals 1.

Definition 2.19. A set $A \subset G$ is a *piecewise syndetic set* if it is the intersection of a syndetic set and a thick set.

It follows from the observations above that given any piecewise syndetic set A there exists a Følner sequence $(F_N)_{N\in\mathbb{N}}$ in G such that $\bar{d}_{(F_N)}(A) > 0$. However, the converse is not true, as the following example shows:

Example 2.20. The set of squarefree numbers (i.e., numbers which are the product of distinct primes) has positive upper density with respect to the Følner sequence $F_N = \{1, \ldots, N\}$ but is not piecewise syndetic in $\mathbb{N}$.

For completeness, we provide a proof of this well known fact.

Proof. Denote by Q the set of squarefree numbers. Let $p_1, p_2, \ldots$ be an enumeration of the prime numbers and observe that

$$Q = \bigcap_{n=1}^{\infty} \left(\mathbb{N} \setminus p_n^2\mathbb{N}\right)$$

For each $N \in \mathbb{N}$, the intersection $Q_N := \bigcap_{n=1}^{N} (\mathbb{N} \setminus p_n^2\mathbb{N})$ is a periodic set (with period $p_1^2 \cdots p_N^2$) and hence its density can be easily computed (invoking the Chinese remainder theorem) to be $\bar{d}(Q_N) = \prod_{n=1}^{N}(1 - p_n^{-2})$. On the other hand, $Q_N \setminus Q \subset \bigcup_{n=N+1}^{\infty} p_n^2\mathbb{N}$, so for each large M we have

$$\begin{aligned}
\frac{|Q \cap \{1, \ldots, M\}|}{M} &\geq \prod_{n=1}^{N}\left(1 - \frac{1}{p_n^2}\right) - \sum_{n=N+1}^{\infty} \frac{|p_n^2\mathbb{N} \cap \{1, \ldots, M\}|}{M} \\
&= \prod_{n=1}^{N}\left(1 - \frac{1}{p_n^2}\right) - \sum_{n=N+1}^{\infty} \frac{1}{p_n^2}
\end{aligned}$$

so taking $N \to \infty$ we conclude that $\bar{d}(Q) \geq 1/\zeta(2) = 6/\pi^2$ (where ζ denotes the Riemann zeta function). In fact we have $\frac{|Q\cap\{1,\ldots,M\}|}{M} \to 6/\pi^2$ as $M \to \infty$. This shows that Q has positive upper density.

Next we show that Q is not piecewise syndetic. Assume, for the sake of a contradiction, that $Q = S \cap T$, where S is syndetic and T is thick. Let $n \in \mathbb{N}$ be such that any interval of length n has non-empty intersection with S, let $p_1, \ldots, p_n$ be distinct primes and let $N = (p_1 \cdots p_n)^2 + n$. Since T is thick we can find $a \in \mathbb{N}$ such that $\{a, a+1, \ldots, a+N\} \subset T$. Invoking the Chinese remainder theorem there exists $M \in \{a+1, \ldots, a+N\}$ such that $M \equiv -i \bmod p_i^2$ for every $i = 1, \ldots, n$. Finally, let $x \in \{M+1, \ldots, M+n\}$ be an element of S. Observe that also $x \in \{a, a+1, \dots, a+N\} \subset T$ and hence $x \in Q$. However, $i := x - M \in \{1, \dots, n\}$ and hence $x \equiv M + i \equiv -i + i = 0 \bmod p_i^2$, which implies that x is a multiple of a perfect square. This yields the desired contradiction. $\square$

The following natural result states that piecewise syndetic sets are precisely the *broken syndetic* sets.

Lemma 2.21. *Let G be a countable commutative semigroup. A set $A \subset G$ is piecewise syndetic if and only if there exists a syndetic set $S \subset G$ such that for any finite subset $F \subset S$ there exists a shift $m = m(F) \in G$ such that $m(F) \cdot F \subset A$.*

A fundamental fact about piecewise syndeticity is that this property can not be destroyed by taking finite partitions (see also Proposition 2.30 below for a stronger statement):

Lemma 2.22 (Brown [Bro68])**.** *Let G be a countable commutative semigroup and let $A \subset G$ be a piecewise syndetic set. Then for any finite partition $A = C_1 \cup \dots \cup C_r$ there exists $C \in \{C_1, \dots, C_r\}$ which is piecewise syndetic.*

Lemma 2.22 implies that any configuration which is present in every piecewise syndetic set can be found monochromatically in any finite coloring of G. The following result can be seen as a converse of this fact for shift invariant configurations.

Theorem 2.23. *Let G be a countable commutative semigroup and let $\mathcal{P}$ be a collection of finite subsets of G. Then the following are equivalent:*

(1) For every finite coloring $G = C_1 \cup \dots \cup C_r$ there exists $C \in \{C_1, \dots, C_r\}$, $P \in \mathcal{P}$ and $g \in G$ such that $gP \subset C$.

(2) For every $r \in \mathbb{N}$ there exists a finite subset $F \subset G$ such that for any coloring of F with r colors there exists $P \in \mathcal{P}$ and $g \in G$ such that gP is contained in F and is monochromatic.

(3) For every piecewise syndetic set A there exists $P \in \mathcal{P}$ and $g \in G$ such that $gP \subset A$.

(4) For every piecewise syndetic set $A \subset \mathbb{N}$ there exists $P \in \mathcal{P}$ such that the set

$$\{g \in G : gP \subset A\}$$

is piecewise syndetic.

Proof. Clearly (4)⇒(3). In view of Lemma 2.22, (3)⇒(1). We will prove (1)⇒(2)⇒(4) and this will finish the proof.

Assume $\mathcal{P}$ satisfies (1) and let $r \in \mathbb{N}$. Suppose, for the sake of a contradiction, that for each finite set $F \subset G$ there exists a coloring $\chi_F : F \to \{1, \ldots, r\}$ without a monochromatic configuration of the form gP with $P \in \mathcal{P}$ and $g \in G$. Extend each the χ_F to a coloring of the whole semigroup G by assigning $\chi_F(x) = 0$ for all $x \notin F$.

Let $(F_N)_{N\in\mathbb{N}}$ be sequence of finite subsets of G such that $F_N \subset F_{N+1}$ and $\bigcup F_N = G$. Then take a convergent subsequence of $(\chi_{F_N})_{N\in\mathbb{N}}$ in the compact metric space $\{0, 1, \ldots, r\}^{\mathbb{N}}$ and call the limit χ. Since $F_N \subset F_{N+1}$ and $\bigcup F_N = G$, we have that $\chi : \mathbb{N} \to \{1, \ldots, r\}$ (i.e. it does not map any element $x \in G$ to the "color" 0). Using (1) we can now find some $P \in \mathcal{P}$ and $g \in G$ such that gP is monochromatic with respect to χ. Let $N \in \mathbb{N}$ be such that χ_{F_N} and χ agree on the set gP. We conclude that gP is monochromatic for the coloring χ_{F_N}, which is a contradiction, and hence (2) holds.

Next we prove the implication (2)⇒(4). Assume that (2) holds and let A be an arbitrary piecewise syndetic set. Let S and T be a syndetic and a thick set, respectively, such that $A = S \cap T$. Let $H \subset G$ be a finite set such that $H^{-1}S = G$ and let $\chi : G \to H$ be a coloring such that $\chi(x)x \in S$ for every $x \in G$. Let $F \subset G$ be provided by (2), let $\tilde{T} := \{x \in \mathbb{N} : xFH \subset T\}$ and observe that $\tilde{T}$ is a thick set, and hence a piecewise syndetic set. Let $\tilde{\chi} : G \to H^F$ be the coloring defined by $\tilde{\chi}(x) = \big(\chi(xy)\big)_{y\in F}$. In view of Lemma 2.22 there exists a piecewise syndetic set $B \subset \tilde{T}$ such that $\tilde{\chi}|_B$ is constant.

Now, let $\chi' : F \to H$ be the coloring defined by $\chi'(x) = \chi(xb)$ for some (hence all) $b \in B$. From the construction of F (i.e. using (2)) we conclude that there exists $g \in G$ and some configuration $P \in \mathcal{P}$ such that $\chi'|_{gP}$ is constant. Let $h \in H$ be the value of $\chi'(gP)$. It follows that $\chi(gPB) = h$ and hence $hgPB \subset S$. On the other hand, $B \subset \tilde{T}$, $gP \subset F$ and $h \in H$, so $hgPB = BgPh \subset T$. We conclude that $(Bhg)P \subset A$, finishing the proof. □

IP-sets

In this subsection it will be convenient to denote the operation on the commutative semigroup G additively.

Given a subset A of G, we denote by $FS(A)$ the set of finite sums of A defined as

$$FS(A) := \left\{ \sum_{i \in I} i : \varnothing \neq I \subset A \right\} \tag{2.4}$$

A set which contains the finite sums of a set of cardinality m is called an *IP$_m$-set*. A set which is IP$_m$ for every $m \in \mathbb{N}$ is called an *IP$_0$-set*:

Definition 2.24. A set $A \subset G$ is an *IP$_0$-set* if

$$\forall m \in \mathbb{N} \quad \exists F \subset G \quad |F| = m \qquad FS(F) = \left\{ \sum_{i \in I} i : \varnothing \neq I \subset F \right\} \subset A$$

There is an infinite version of IP$_0$-sets, called IP-sets. Given an infinite set X, we denote by $\mathcal{F}(X)$ the family of all finite non-empty subsets of X, i.e., $\mathcal{F}(X) := \{\alpha \subset X : 0 < |\alpha| < \infty\}$. We denote simply by $\mathcal{F}$ the family $\mathcal{F}(\mathbb{N})$ of all non-empty finite subsets of $\mathbb{N}$. Let G be a countable commutative semigroup and let $(x_n)_{n\in\mathbb{N}}$ be an injective sequence in G. For each $\alpha \in \mathcal{F}$ define $x_\alpha = \sum_{n\in\alpha} x_n$. The *IP-set* generated by $(x_n)_{n\in\mathbb{N}}$ is the set $FS(x_n) = \{x_\alpha : \alpha \in \mathcal{F}\}$. Clearly $x_{\alpha\cup\beta} = x_\alpha + x_\beta$ for any disjoint $\alpha, \beta \in \mathcal{F}$. Moreover, if $(y_\alpha)_{\alpha\in\mathcal{F}}$ is any 'sequence' indexed by $\mathcal{F}$ such that $x_{\alpha\cup\beta} = x_\alpha + x_\beta$ for any disjoint $\alpha, \beta \in \mathcal{F}$, then the set $\{y_\alpha : \alpha \in \mathcal{F}\}$ is an IP-set (generated by $(y_{\{n\}})_{n\in\mathbb{N}})$). For this reason we will denote IP-sets by $(y_\alpha)_{\alpha\in\mathcal{F}}$, with the understanding that they are generated by the singletons y_n, $n \in \mathbb{N}$. Therefore it is convenient to think of IP-sets as both the map $\mathcal{F} \to G$ and the image of that map. Observe that (the image of) any IP-set $\{x_\alpha : \alpha \in \mathcal{F}\}$ is an IP$_0$-set; but not every IP$_0$-set contains (the image of) an IP-set.

An example of an IP-set is the set A of all numbers which, when written in base 10, only use the digits 0 and 1. Indeed $A = FS(\{10^n : n \geq 0\})$.

Proposition 2.25. *Every thick set contains (the image of) an IP-set.*

Proof. Let G be a countable commutative semigroup and let $T \subset G$ be a thick set. We will construct a sequence $(x_n)_{n\in\mathbb{N}}$ recursively as follows. Let $x_1 \in T$ be arbitrary. For each $n \geq 1$ assume we already constructed $x_1, \ldots, x_n$ and let $x_{n+1} \in T$ be such that $x_{n+1} + FS(\{x_1, \ldots, x_n\}) \subset T$ (the existence of such x_{n+1} follows directly from the Definition 2.17).

It is clear that $FS\big((x_i)_{i=1}^n\big) \subset T$ for every $n \in \mathbb{N}$, and hence $x_\alpha \in T$ for every $\alpha \in \mathcal{F}$. $\square$

Definition 2.26. Let $(x_\alpha)_{\alpha\in\mathcal{F}}, (y_\alpha)_{\alpha\in\mathcal{F}}$ be IP-sets in a countable commutative semigroup G.

1. For $\alpha, \beta \in \mathcal{F}$ we write $\alpha < \beta$ as a shortcut to $\max_{i\in\alpha} i < \min_{j\in\beta} j$.

2. We say that $(x_\alpha)_{\alpha\in\mathcal{F}}$ is a *sub-IP-set* of $(y_\alpha)_{\alpha\in\mathcal{F}}$ if there exist $\alpha_1 < \alpha_2 < \cdots$ in $\mathcal{F}$ such that $x_n = y_{\alpha_n}$ for all $n \in \mathbb{N}$.

A remarkable property of IP-sets is that they are *partition regular* in the following sense:

Theorem 2.27 (Hindman's theorem, [Hin74])**.** *Let $(x_\alpha)_{\alpha\in\mathcal{F}}$ be an IP-set in a countable commutative semigroup G and let $G = C_1 \cup \cdots \cup C_r$ be an arbitrary finite coloring. Then there exists a color $C \in \{C_1, \ldots, C_r\}$ and a sub-IP-set $(y_\alpha)_{\alpha\in\mathcal{F}}$ of $(x_\alpha)_{\alpha\in\mathcal{F}}$ such that $y_\alpha \in C$ for every $\alpha \in \mathcal{F}$.*

One way to prove Hindman's theorem is by showing that given any countable commutative semigroup $(G, +)$ and any idempotent ultrafilter $p = p+p \in \beta G$, each member $A \in p$ contains an IP-set. Conversely, given any IP-set $(x_\alpha)_{\alpha\in\mathcal{F}}$ in G, there exists an idempotent ultrafilter $p \in \beta G$ such that $\{x_\alpha : \alpha \in \mathcal{F}\} \in p$. In fact we have

Proposition 2.28 (cf. [Ber10, Theorem 2.6])**.** *Let G be a countable commutative semigroup and let $p \in \beta G$ be an ultrafilter. Then p belongs to the closure (in βG) of the set of idempotent ultrafilters if and only if each $A \in p$ contains an IP-set.*

A ring R has two semigroup structures (addition and multiplication) on the same underlying set. The following result establishes a relation between notions of largeness with respect to both structures.

Theorem 2.29 (cf. [BH94, Theorem 3.5])**.** *Let $(R, +, \cdot)$ be a (commutative) integral domain and let $A \subset (R \setminus \{0\}, \cdot)$ be a (multiplicative) piecewise syndetic set. Then A is an (additive) IP_0-set in $(R, +)$.*

Proof. For every $m \in \mathbb{N}$ consider the following collection of finite subsets of R:

$$\mathcal{P}_m := \{FS(F) : F \subset G,\ |F| = m\}.$$

Observe that whenever $P \in \mathcal{P}_m$ and $r \in R \backslash \{0\}$, also the set $rP \in \mathcal{P}_m$. In view of Hindman's theorem, the collection $\mathcal{P}_m$ satisfies the first property of Theorem 2.23, and hence the third. In particular A contains a configuration P from $\mathcal{P}_m$ for every $m \in \mathbb{N}$; which is equivalent to say that A is an (additive) IP_0-set. □

We remark that not every multiplicatively piecewise syndetic set is an additive IP-set (cf. [BH94, Theorem 3.6]); in this sense Theorem 2.29 is the best possible.

Central sets

Central sets were introduced by Furstenberg in $(\mathbb{N}, +)$ in [Fur81], using the language of topological dynamics (cf. Section 2.5 below). A characterization in terms of ultrafilters was discovered later by Bergelson and Hindman (with the help of Weiss) in [BH90], and this spurred the study of central sets.

A *right ideal* in βG is a closed subset $I \subset \beta G$ satisfying $I + \beta G \subset I$ (we maintain the additive notation for the operations in G and in βG). By Zorn's Lemma, there exist minimal (with respect to the inclusion relation) right ideals in βG. A *minimal* ultrafilter is an ultrafilter $p \in \beta G$ which belongs to some minimal right ideal. One can show that if $p \in \beta G$ is a minimal ultrafilter and $A \in p$, then A is piecewise syndetic. In fact we have

Proposition 2.30 (see [HS98, Corollary 4.41])**.** *Let G be a countable commutative semigroup and let $p \in \beta G$ be an ultrafilter. Then p belongs to the closure (in βG) of the set of minimal ultrafilters if and only if each $A \in p$ is piecewise syndetic.*

This fact, together with Proposition 2.2, provides a proof of Lemma 2.22.

Of special importance among minimal ultrafilters are the *minimal idempotent ultrafilters* i.e. ultrafilters which are simultaneously minimal and idempotent. In any countable commutative semigroup G there exist minimal idempotent ultrafilters $p \in \beta G$. More generally we have

Proposition 2.31 (cf. [BH94, Lemma 3.3])**.** *Let G be a countable commutative semigroup and let $T \subset G$ be a thick set. Then there exists a minimal idempotent ultrafilter $p \in \beta G$ such that $T \in p$.*

Definition 2.32. Let G be a countable commutative semigroup and let $A \subset G$. We say that A is a *central set* if there exists a minimal idempotent ultrafilter $p \in \beta G$ such that $A \in p$.

Since every countable commutative semigroup has a minimal idempotent, it follows from Remark 2.2 that for every finite partition of a countable commutative semigroup, one of the cells is a central set. Central sets are important in combinatorics because they are both IP-sets and piecewise syndetic sets; the combinatorial richness possessed by central sets is best illustrated by the central sets theorem (see Theorem 2.40 below).

Corollary 2.33. *Every thick set is central. Every central set is piecewise syndetic.*

Proof. The first assertion follows from Definition 2.32 and Proposition 2.31; and the second from Definition 2.32 and Proposition 2.30. □

In the spirit of Proposition 2.28 and Proposition 2.30 we have the following if and only if characterization of minimal idempotent ultrafilters, which follows directly from Definition 2.32:

Proposition 2.34. *Let G be a countable commutative semigroup and let $p \in \beta G$ be an ultrafilter. Then p belongs to the closure (in βG) of the set of minimal idempotent ultrafilters if and only if each $A \in p$ is central.*

2.3 Variations on the theme of van der Waerden's theorem

Van der Waerden's theorem has many different extensions in several directions. In this section we list some which will be useful later on.

Linear configurations

We start by recalling van der Waerden's theorem (Theorem 1.2 from the introduction):

Theorem 2.35 (van der Waerden's theorem)**.** *For any finite coloring* $\mathbb{N} = C_1 \cup \cdots \cup C_r$ *and any* $k \in \mathbb{N}$ *there exists* $C \in \{C_1, \ldots, C_r\}$ *and* $a, b \in \mathbb{N}$ *such that* $\{a, a+b, a+2b, \ldots, a+kb\} \subset C$*.*

One can interpret van der Waerden's theorem in a geometric way as stating that, given a finite coloring of $\mathbb{N}$, any finite subset of $\mathbb{N}$ has an homothetic[1] copy which is monochromatic (cf. [Ber96, Theorem 1.3]. There is a multidimensional analogue of van der Waerden's theorem, due to T. Grünwald/Galai, which can be formulated in geometric terms as saying that for any $d \in \mathbb{N}$, given a finite coloring of $\mathbb{N}^d$, any finite subset of $\mathbb{N}^d$ has an homothetic copy which is monochromatic. More precisely:

Theorem 2.36 (Multidimensional van der Waerden's theorem, cf. [GRS90, Theorem 2.8])**.** *For any* $d, k \in \mathbb{N}$ *and any finite coloring* $\mathbb{N}^d = C_1 \cup \cdots \cup C_r$ *there exists a color* $C \in \{C_1, \ldots, C_r\}$ *and* $a \in \mathbb{N}^d$*,* $b \in \mathbb{N}$ *such that* $\Big\{a + (i_1 b, \cdots, i_d b) : 0 \leq i_1, \ldots, i_d \leq k\Big\} \subset C$*.*

The following theorem can be thought of as a set theoretic version of van der Waerden's theorem.

Definition 2.37 (Combinatorial line)**.** Let A be a finite alphabet, let $* \notin A$ be a *wild card* element and let $n \in \mathbb{N}$. A *variable word* in A^n is an element of the set $(A \cup \{*\})^n \setminus A^n$. Given a variable word w and $a \in A$, let $w(a) \in A^n$ be the word obtained by replacing each instance of $*$ in w with a. The *combinatorial line* generated by a variable word w is the set $\{w(a) : a \in A\} \subset A^n$.

Theorem 2.38 (Hales-Jewett theorem [HJ63])**.** *For each* $k, r \in \mathbb{N}$ *there exists* $HJ(k, r) \in \mathbb{N}$ *such that for all* $n \geq HJ(k, r)$ *and any* r*-coloring of* $\{1, \ldots, k\}^n$*, there exists a monochromatic combinatorial line.*

[1] Recall that an homothety is the composition of a dilation with a translation.

To see how Theorem 2.38 implies Theorem 2.35, identify each $n \in \mathbb{N}$ with its expansion in base $k+1$ and consider only numbers with exactly $HJ(k,r)$ digits. However, Theorem 2.38 is quite more powerful and it implies the following IP-version of van der Waerden's theorem in an arbitrary commutative semigroup (which contains Theorem 2.36 as a special case).

Proposition 2.39 (IP van der Waerden theorem in commutative semigroups)**.** *Let $(G,+)$ be a countable commutative semigroup, let $j \in \mathbb{N}$, let $(y_\alpha)_{\alpha\in\mathcal{F}}$ be an IP-set in G^j, let $A \subset G$ be piecewise syndetic and let F be a finite set of semigroup homomorphisms[2] from G^j to G. Then there exists $\alpha \in \mathcal{F}$ and $x \in G$ such that $x + f(y_\alpha) \in A$ for each $f \in F$.*

Proof. In view of Theorem 2.23 it suffices to show that for any finite coloring of G there exist $\alpha \in \mathcal{F}$ and $x \in G$ such that $\{x + f(y_\alpha) : f \in F\}$ is monochromatic.

To show this, let r be the number of colors, let $n = HJ(|F|, r)$ be the number given by Theorem 2.38 and color each $(f_1, \dots, f_n) \in F^n$ with the color of $f_1(y_1) + \cdots + f_n(y_n) \in G$ (where $y_1, \dots, y_n$ are the first generators of the given IP-set $(y_\alpha)_{\alpha\in\mathcal{F}}$). Apply Theorem 2.38 to find a variable word $w \in (F \cup \{*\})^n$ whose corresponding combinatorial line is monochromatic. Let $B = \{i \in \{1, \dots, n\} : w_i \in F\}$, let $\alpha = \{1, \dots, n\} \setminus B$ be the positions of the wild card $*$ in w and let

$$x = \sum_{i\in B} w_i(y_i).$$

For any $f \in F$ we have that $w(f) \in F^n$ has the same color as

$$\sum_{i\in B} w_i(y_i) + \sum_{i\in\alpha} f(y_i) = x + f\left(\sum_{i\in\alpha} y_i\right) = x + f(y_\alpha)$$

and hence the set $\{x+f(x_\alpha) : f \in F\}$ is indeed monochromatic, which finishes the proof. □

Proposition 2.39 has an infinitary extension for central sets.

Theorem 2.40 (Central sets theorem)**.** *Let G be a countable commutative semigroup, let $j \in \mathbb{N}$, let $A \subset G$ be a central set and let $(y_\alpha)_{\alpha\in\mathcal{F}}$ be an IP-set in G^j. Then there exists an*

[2] A semigroup homomorphism is a function $f : H \to G$ such that $f(x+y) = f(x) + f(y)$ for every $x, y \in H$.

IP-set $(x_\beta)_{\beta\in\mathcal{F}}$ in G and a sub-IP-set $(z_\beta)_{\beta\in\mathcal{F}}$ of $(y_\alpha)_{\alpha\in\mathcal{F}}$ such that

$$\forall i \in \{1,\dots,j\} \qquad \forall \beta \in \mathcal{F} \qquad\qquad x_\beta + \pi_i(z_\beta) \in A \tag{2.5}$$

where $\pi_i : G^j \to G$ is the projection onto the i-th coordinate.

This theorem was obtained by Furstenberg for the case $G = \mathbb{N}$ in [Fur81]. In [BH90], Theorem 2.40 was proved for certain classes of countable commutative semigroups, and an alternative, dynamical characterization of central sets for arbitrary countable commutative semigroups was establish, which hinted at the full generality of Theorem 2.40. Theorem 2.40 was obtained as stated in [HMS96].

We remark that Theorem 2.40 has an important distinction from the other theorems in this subsection; namely that it does not have a *density version.* In other words, one can not replace "central set" with "set with positive upper density". Indeed, for $j = 1$, (2.5) says that A contains the IP-set $(x_\beta + z_\beta)_{\beta\in\mathcal{F}}$, and the set of odd numbers $2\mathbb{Z} - 1 \subset (\mathbb{Z}, +)$ which has positive density (and is in fact syndetic) can not contain an IP-set (or even a Schur triple $\{x, y, x + y\}$).

For contrast, we now formulate a density version of van der Waerden's theorem, established by Szemerédi in [Sze75].

Theorem 2.41 (Szemerédi's theorem)**.** *Let $A \subset \mathbb{N}$ be such that $\bar{d}(A) > 0$. Then A contains arbitrarily long arithmetic progressions.*

Polynomial theorems

We recall from the introduction the polynomial extension of van der Waerden's theorem due to Bergelson and Leibman [BL96].

Theorem 1.3. *Let $F \subset \mathbb{Z}[x]$ be a finite set of polynomials such that $p(0) = 0$ for all $p \in F$. For any finite coloring $\mathbb{N} = C_1 \cup \cdots \cup C_r$ and any $k \in \mathbb{N}$ there exists a color $C \in \{C_1, \dots, C_r\}$ and $a, b \in \mathbb{N}$ such that $\{a + f(b) : f \in F\} \subset C$.*

Bergelson and Leibman later obtained a common extension of Theorems 2.38 and 1.3, namely a polynomial extension of the Hales-Jewett theorem. We do not state this theorem here because it would require some setup and we do not directly use it. However, we will need a consequence of the polynomial Hales-Jewett theorem in the spirit of Proposition 2.39.

Definition 2.42. Given a map $f : H \to G$ between countable commutative groups we say that f is a *polynomial map of degree 0* if it is constant. We say that f is a *polynomial map of degree d*, $d \in \mathbb{N}$, if it is not a polynomial map of degree $d-1$ and for every $h \in H$, the map $x \mapsto f(x+h) - f(x)$ is a polynomial of degree $\leq d-1$. Finally we denote by $\mathbb{P}(H,G)$ the set of all polynomial maps $f : H \to G$ with $f(0_H) = 0_G$.

Theorem 2.43 (IP polynomial van der Waerden theorem for abelian groups, cf. [BL99, Corolary 8.8])**.** *Let G, H be countable abelian groups and let $F \subset \mathbb{P}(H,G)$ be a finite subset. Then for every finite partition $G = C_1 \cup \cdots \cup C_r$ and every IP set $(y_\alpha)_{\alpha \in \mathcal{F}}$ in H there exists $C \in \{C_1, \ldots, C_r\}$, $a \in C$ and $\alpha \in \mathcal{F}$ such that $a + f(y_\alpha) \in C$ for every $f \in F$.*

In particular, we record the simpler polynomial van der Waerden theorem for abelian groups, which does not require IP-sets to state.

Corollary 2.44. *Let G, H be countable abelian groups and let $F \subset \mathbb{P}(H,G)$ be a finite subset. Then for every finite partition $G = C_1 \cup \cdots \cup C_r$ there exists $C \in \{C_1, \ldots, C_r\}$, $a \in C$ and $b \in H \setminus \{0\}$ such that $a + f(b) \in C$ for every $f \in F$.*

Combining Theorem 2.43 with Theorem 2.23 we obtain:

Corollary 2.45. *Let $j \in \mathbb{N}$, let G be a countable abelian group and let F be a finite family of polynomial maps from G^j to G such that $f(\mathbf{0}) = 0$ for each $f \in F$. Then for every piecewise syndetic (in particular, central) set $A \subset G$ and every IP set $(y_\alpha)_{\alpha \in \mathcal{F}}$ in G^j there exists $a \in A$ and $\alpha \in \mathcal{F}$ such that $a + f(y_\alpha) \in A$ for every $f \in F$.*

2.4 Ergodic theory

In 1977, Furstenberg gave a second proof of Szemerédi's theorem (Theorem 2.41), using ergodic theory [Fur77; FKO79]. The ergodic theoretic proof has the advantage of being highly versatile and has in fact been used to establish several extensions of Szemerédi's theorem (including suitable density versions of Theorem 2.36, Theorem 2.43 and Theorem 2.38).

The general idea is to observe that $\{a, a+d, \cdots, a+kd\} \subset E$ if and only if $a \in E \cap (E-d) \cap \cdots \cap (E-kd)$, and then show that if E has positive density then in fact

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \bar{d}(E \cap (E-n) \cap \cdots \cap (E-kn)) > 0. \tag{2.6}$$

One can then choose some n for which the upper density of the intersection in (2.6) is positive, which implies that the intersection is non-empty and hence implies that E contains (several) arithmetic progressions of length $k+1$.

In order to establish (2.6), Furstenberg devised a *correspondence principle* which allows one to derive (2.6) from a statement in ergodic theory – the multiple recurrence theorem. Before we state both the correspondence principle and the multiple recurrence theorem we will need some notation and terminology.

Let $(X, \mathcal{B}, \mu)$ be a probability space (this means that $\mathcal{B}$ is a σ-algebra on X and $\mu : \mathcal{B} \to [0,1]$ is a probability (countably additive) measure). A measurable map $T : X \to X$ is called a *measure preserving transformation* if for every $B \in \mathcal{B}$ one has $\mu(T^{-1}B) = \mu(B)$, where as usual $T^{-1}B := \{x \in X : Tx \in B\}$. A quadruple $(X, \mathcal{B}, \mu, T)$, where $(X, \mathcal{B}, \mu)$ is a probability space and $T : X \to X$ is a measure preserving transformation, is called a measure preserving system. More generally, given a semigroup G and an action $(T_g)_{g\in G}$ of G on a probability space $(X, \mathcal{B}, \mu)$ (this means that for each $g \in G$ there is a map $T_g : X \to X$ and for any $g, h \in G$ we have $T_{gh} = T_g T_h$) by measure preserving transformations, the quadruple $\big(X, \mathcal{B}, \mu, (T_g)_{g\in G}\big)$ is also called a measure preserving system, or a measure preserving G-system.

Theorem 2.46 (Furstenberg's correspondence principle)**.** *For any set $E \subset \mathbb{N}$ there exists a measure preserving system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E)$ such that for any $n_1, \ldots, n_k \in \mathbb{N}$*

$$\bar{d}(E \cap (E-n_1) \cap \cdots \cap (E-n_k)) \geq \mu(A \cap T^{-n_1}A \cap \cdots \cap T^{-n_k}A)$$

Theorem 2.46 was implicitly obtained in [Fur77] and [FKO79] and can be found explicitly in [Ber96, Theorem 1.8]. Several versions of the correspondence principle have since been obtained and applied to different situations (in fact we will need a new version, presented in Section 3.5). In view of Theorem 2.46 and the observations in the beginning of this section, Szemerédi's theorem follows from the following multiple recurrence theorem.

Theorem 2.47 (Furstenberg's multiple recurrence theorem, [Fur77, Theorem 11.13])**.** *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving system and let $A\in\mathcal{B}$ be such that $\mu(A)>0$. Then for every $k\in\mathbb{N}$*

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

Theorem 2.47 is a far reaching extension of a classical recurrence theorem due to Poincaré:

Theorem 2.48 (Poincaré's recurrence theorem)**.** *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving system and let $A\in\mathcal{B}$ be such that $\mu(A)>0$. Then there exists $n\in\mathbb{N}$ such that $\mu(A\cap T^{-n}A)>0$.*

Theorem 2.48 was in turn considerably strengthened by Khintchine:

Theorem 2.49 (Khintchine's recurrence theorem, [Khi34])**.** *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving system and let $A\in\mathcal{B}$. For every $\varepsilon>0$, the set*

$$\{n\in\mathbb{N}:\mu(A\cap T^{-n}A)>\mu^2(A)-\varepsilon\}\tag{2.7}$$

is syndetic.

Observe that the quantity $\mu(A)^2$ is optimal, as seen by taking $(X,\mathcal{B},\mu)$ to be the unit interval $[0,1)$ with the Borel σ-algebra and the Lebesgue measure, $T:[0,1]\to[0,1)$ to be the measure preserving map $T:x\mapsto 2x \bmod 1$ and $A=[0,1/2)$. Khintchine's theorem implies, via Furstenberg's correspondence principle, that for any $E\subset\mathbb{N}$ with positive upper density and any $\varepsilon>0$, the set

$$\Big\{n\in\mathbb{N}:\bar{d}(E\cap(E-n))>\big(\bar{d}(E)\big)^2-\varepsilon\Big\}$$

is syndetic. One can easily derive Theorem 2.49 from von Neumann's mean ergodic theorem, which we now state in a general setting; see for instance Theorem 5.5 in [Ber06] for a proof of this version.

Theorem 2.50. *Let G be a countable abelian group and let (F_N) be a Følner sequence in G. Let H be a Hilbert space and let $(U_g)_{g\in G}$ be a unitary representation of G on H. Let P be the orthogonal projection onto the subspace of vectors fixed under G. Then*

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} U_g f = Pf \qquad\qquad \forall f \in H \tag{2.8}$$

in the strong topology of H.

We now briefly explain how Theorem 2.50 implies Theorem 2.49 (see also the proof of Theorem 5.11 below from a mean ergodic theorem, Theorem 5.4). Observe that, given a measure preserving system $(X, \mathcal{B}, \mu, T)$ the *Koopman operator* $U : L^2(X) \to L^2(X)$, obtained by composing a given function with T, is unitary (because T is measure preserving). Then it follows from (2.8) that for any $A \in \mathcal{B}$ and any Følner sequence $(F_N)_{N\in\mathbb{N}}$ we have

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{n\in F_N} \mu(A \cap T^{-n}A) = \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{n\in F_N} \langle 1_A, U^n 1_A\rangle = \langle 1_A, P1_A\rangle$$

where 1_A is the indicator function of A and $P : L^2(X) \to L^2(X)$ is an orthogonal projection. Finally, since $P1 = 1$ (where, by a slight abuse of notation, we are denoting by 1 the constant function equal to 1), an application of the Cauchy-Schwartz inequality implies that $\langle 1_A, P1_A\rangle = \langle P1_A, P1_A\rangle \geq \langle P1_A, 1\rangle^2 = \langle 1_A, 1\rangle^2 = \mu(A)^2$ and hence it follows that the set (2.7) has positive upper density with respect to any Følner sequence, which implies that it is syndetic (cf. Section 2.2).

The correspondence principle and the recurrence theorems above hold in much more general settings.

Definition 2.51. Let G be a semigroup. A set $R \subset G$ is a *set of recurrence* if for all probability preserving actions $(\Omega, \mu, (T_g)_{g\in G})$ and every measurable set $B \subset \Omega$ with positive measure, there exists some non-identity $g \in R$ such that $\mu(B \cap T_g^{-1}B) > 0$.

The significance of sets of recurrence is revealed by the correspondence principle: they are precisely the sets which have non-empty intersection with every set of the form $A - A$, where $\bar{d}(A) > 0$. For instance one can show that the set $\{n^2 : n \in \mathbb{N}\}$ of perfect squares is a set of recurrence in $\mathbb{N}$ (this fact was first established by Furstenberg in [Fur77, Proposition 1.3]). As a consequence one derives a result of Sárközy [Sár78], stating that given any set $A \subset \mathbb{N}$ with positive upper density $\bar{d}(A) > 0$, there exists $a \in A$ and $n \in \mathbb{N}$ such that $a + n^2 \in A$.

The following lemma is well known; we include the proof for the convenience of the reader.

Lemma 2.52. *Let G be a semigroup and let $R \subset G$ be a set of recurrence. Then for every finite partition $R = R_1 \cup \cdots \cup R_r$, one of the sets R_i is also a set of recurrence.*

Proof. The proof goes by contradiction. Assume that none of the sets $R_1, \ldots, R_r$ is a set of recurrence. Then for each $i = 1, \ldots, r$ there is some probability preserving action $(\Omega_i, \mu_i, (T_g)^{(i)}_{g\in G})$ and a set $B_i \subset \Omega_i$ with $\mu_i(B_i) > 0$ and such that $\mu_i(B_i \cap (T_g^{(i)})^{-1}B_i) = 0$ for all $g \in R_i$.

Let $\Omega = \Omega_1 \times \cdots \times \Omega_r$, let $\mu = \mu_1 \otimes \cdots \otimes \mu_r$, let $B = B_1 \times \cdots \times B_r$ and, for each $g \in G$, let $T_g(\omega_1, \ldots, \omega_r) = (T_g^{(1)}\omega_1, \ldots, T_g^{(r)}\omega_r)$. Then $(T_g)_{g\in G}$ is a probability preserving action of G on Ω and $\mu(B) = \mu_1(B_1) \cdots \mu_r(B_r) > 0$.

Since R is a set of recurrence, there exists some $g \in R$ such that $\mu(B \cap T_g^{-1}B) > 0$. Since $\mu(B \cap T_g^{-1}B) = \prod_{i=1}^r \mu_i(B_i \cap (T_g^{(i)})^{-1}B_i)$ we conclude that $\mu_i(B_i \cap (T_g^{(i)})^{-1}B_i) > 0$ for all $i = 1, \ldots, r$. But this implies that $g \notin R_i$ for all $i = 1, \ldots, r$, which contradicts the fact that $g \in R = R_1 \cup \cdots \cup R_r$. $\square$

One of the main tools in establishing (multiple) recurrence results in ergodic theory is a version of van der Corput's trick on uniform distribution. We record here the version which we will need later on. For a proof see, for instance, Lemma 2.9 in [BLM05]. For a detailed discussion of many forms of the van der Corput trick, see the recent survey [BM16c].

Proposition 2.53. *Let H be an Hilbert space, let G be a commutative semigroup (written multiplicatively), let $(F_N)_{N\in\mathbb{N}}$ be a Følner sequence in G and let $(a_u)_{u\in G}$ be a bounded sequence in H indexed by G. If for all b in a co-finite subset of G we have*

$$\limsup_{N\to\infty}\left|\frac{1}{|F_N|}\sum_{u\in F_N}\langle a_{bu},a_u\rangle\right|=0$$

then also

$$\lim_{N\to\infty}\frac{1}{|F_N|}\sum_{u\in F_N}a_u=0$$

2.5 Topological dynamics

Let G be a semigroup. A *G-topological dynamical system* is a pair $\big(X,(T_g)_{g\in G}\big)$ where X is a compact Hausdorff space and $(T_g)_{g\in G}$ is an action of G on X by continuous functions. A system $(X,(T_g)_{g\in G})$ is *minimal* if X contains no proper non-empty closed invariant subsets. A point $x\in X$ is a *minimal point* if its orbit closure $Y:=\overline{\{T_gx:g\in G\}}$ is a minimal subsystem of X (i.e., if $\big(Y,(T_g|_Y)_{g\in G}\big)$ is a minimal system).

We record for future reference in the following proposition the fact that any topological dynamical system contains a minimal subsystem. The proof, which we omit, consists of standard application of Zorn's lemma (see for instance [Aus88, Proposition 1.3] or [Gla03, Exercise 1.1.3]).

Proposition 2.54. *Let X be a compact Hausdorff space, let G be a countable semigroup and let $(T_g)_{g\in G}$ be an action of G on X such that each map $T_g:X\to X$ is continuous. Then there exists a non-empty closed subset $Y\subset X$ such that $T_g(Y)\subset Y$ for every $g\in G$ and such that the system $\big(Y,(T_g)_{g\in G}\big)$ is minimal.*

An important example of dynamical systems are systems of *isometries*, i.e., when X is a metric space and each map $T_g:X\to X$ preserves the distance. In this case we have the following:

Lemma 2.55. *Let $(G,+)$ be a countable commutative semigroup and let (X,d) be a metric space and let $\big(X,(T_g)_{g\in G}\big)$ be a topological dynamical system where $d(T_gx,T_gy)=d(x,y)$ for every $g\in G$ and $x,y\in X$.*

Then for any $x \in X$, $\varepsilon > 0$ and any IP_0-set $A \subset G$, there exists $g \in A$ such that $d(T_g x, x) < \varepsilon$.

Proof. Let $Y = \overline{\{T_g x : g \in G\}}$ and, for each $g \in G$, let $B_g := \{y \in Y : d(T^g x, y) < \varepsilon/2\}$. Clearly $Y = \bigcup_{g\in G} B_g$. Hence, by compactness, there exists some finite set $F \subset G$ such that the union $\bigcup_{g\in F} B_g$ contains all of Y.

Since A is an IP_0-set, there exists a set $Z \subset G$ with cardinality $|Z| = |F| + 1$ and such that $FS(Z) \subset A$. List the elements $Z = \{z_1, ..., z_r\}$, let $z_i' = z_1 + ... + z_i$ for each $i = 1, ..., r$ and note that $z_i' - z_j' \in A$ for each $i > j$. By the pigeonhole principle, there are $1 \leq i < j \leq r$ such that $T_{z_i'} x$ and $T_{z_j'} x$ are in the same ball B_g for some $g \in F$. Thus $d(T_{z_i'} x - T_{z_j'} x\| < \varepsilon$ and since the action $(T_g)_{g\in G}$ preserves the metric we conclude that $d(T_{z_i' - z_j'} x, x) < \varepsilon$. $\square$

CHAPTER 3

AFFINE SEMIGROUPS

In this chapter we study affine semigroups of certain rings, which we call LID, and their natural actions on the corresponding rings. The definitions and results presented in this chapter were originally obtained in [BM16a] and [BM16b] and will be used in Chapters 5 and 6.

3.1 Large Ideal Domains and the affine semigroup

Definition 3.1. A ring R is called a *large ideal domain* (LID) if it is a countably infinite integral domain and for each $x \in R \setminus \{0\}$, the ideal xR is a finite index additive subgroup of R.

Every field is trivially an LID. The following proposition gives some non-trivial examples of LID rings.

Proposition 3.2. *The following rings are LID:*

1. *Any integral domain R whose underlying additive group is finitely generated. In particular, the ring of integers $\mathcal{O}_K$ of a number field K satisfies this property.*

2. *The ring of polynomials $\mathbb{F}[x]$ over a finite field $\mathbb{F}$.*

Proof.

1. Since $(R, +)$ is an infinite finitely generated abelian group, it contains torsion-free elements and therefore the identity 1_R of R has infinite order in $(R, +)$. If some

element $x \in R$ had torsion, say $nx = 0$ for some $n \in \mathbb{N}$, then $(n1_R)x = 0$, contradicting the absence of 0 divisors. Using the classification of finitely generated abelian groups we can now represent $(R, +)$ as $\mathbb{Z}^d$ for some $d \in \mathbb{N}$.

For any non-zero $x \in R$, the map $\phi : y \mapsto xy$ is an injective endomorphism of $(R, +)$ (injectivity follows from the absence of divisors of 0) whose image $\phi(R)$ is the ideal xR. We claim that the image of any injective homomorphism $\phi : \mathbb{Z}^d \to \mathbb{Z}^d$ has a finite index in $\mathbb{Z}^d$, which will finish the proof.

Indeed, representing ϕ as a matrix, injectivity implies that the determinant of ϕ is non-zero. Therefore it has an inverse ϕ^{-1} with entries in $\mathbb{Q}$. Multiplying ϕ^{-1} by the least common multiple n of its entries we obtain a matrix $n\phi^{-1}$ with coefficients in $\mathbb{Z}$. Therefore $n\mathbb{Z}^d = (n\phi^{-1})\phi(\mathbb{Z}^d) \subset \phi(\mathbb{Z}^d)$, so $[\mathbb{Z}^d : \phi(\mathbb{Z}^d)] \leq [\mathbb{Z}^d : n\mathbb{Z}^d] = n^d < \infty$, proving the claim.

2. Let $f \in \mathbb{F}[x]$ have degree d. For any $g \in \mathbb{F}[x]$ one can divide g by f and obtain $g = fq + r$ where $\deg r < d$. Therefore $g - r$ belongs to the ideal $f\mathbb{F}[x]$. It follows that the set of polynomials r with degree smaller than d form a complete set of coset representatives for $f\mathbb{F}[x]$. Since $\mathbb{F}$ is finite, there are only finitely many such representatives and hence the index of $f\mathbb{F}[x]$ is finite as desired. $\square$

Remark 3.3. *There are number fields whose ring of integers is not a principal ideal domain (PID). Hence, part (1) of Proposition 3.2 includes some LID which are not PID. We also observe that not every PID is a LID. Indeed, the ring $\mathbb{Q}[x]$ of all polynomials with rational coefficients is a PID, but the ideal $x\mathbb{Q}[x]$ has infinite index as an additive subgroup of $\mathbb{Q}[x]$, so $\mathbb{Q}[x]$ is not a LID.*

Lemma 3.4. *Let R be a LID and let $B \subset (R, +)$ be piecewise syndetic. Then for any $a \in R \setminus \{0\}$, the dilation aB is also piecewise syndetic.*

Proof. Let S and T be such that $B = S \cap T$ and S is syndetic and T is thick. Let $T' = aT \cup (R \setminus aR)$ and let $S' = aS$. Then clearly $aB = T' \cap S'$. We now claim that T' is thick and S' is syndetic, which will finish the proof.

Let $F \subset R$ be a finite set such that $S - F = R$. Then $S' - aF = aR$. Since R is a LID, the ideal aR has finite index in R. Let $\tilde{F}$ be a (finite) set of co-set representatives. Then $aR - \tilde{F} = R$ and hence $S' - (aF + \tilde{F}) = R$. Taking $F' := aF + \tilde{F}$ we deduce that $S' - F' = R$ and S' is syndetic, as desired.

Next we show that T' is thick. Let $F \subset R$ be an arbitrary finite set; we will find $x \in R$ such that $x + F \subset T'$. Split $F = F_1 \cup F_2$ where $F_1 = F \cap aR$ and $F_2 = F \setminus F_1$. If F is disjoint from aR then it is already contained in T'. Let $F' = F_1/a$ and let $x' \in R$ be such that $x' + F' \subset T$. Then, taking $x = ax'$ we have $x + F = a(x' + F') \cup ax' + F_2$. Since $x' + F' \subset T$, the first term $a(x' + F')$ is inside $aT \subset T'$. Since F_2 is disjoint from aR, also $ax' + F_2$ is disjoint from aR, and hence contained in T'. Therefore $x + F \subset T'$, as desired. $\square$

Observe that Lemma 3.4 does not hold in general rings, not even in every principal ideal domain. Borrowing the example from Remark 3.3, the ring $\mathbb{Q}[x]$ of polynomials with rational coefficients is a piecewise syndetic set within itself but the ideal $x\mathbb{Q}[x]$ has infinite index as an additive subgroup and hence can not be piecewise syndetic.

Given a ring R, we denote by R^* the set of its non-zero elements. An *affine transformation* of R is a map $f : R \to R$ of the form $f(x) = ux + v$ with $u \in R^*, v \in R$. The *affine semigroup* of R is the semigroup of all affine transformations of R (the semigroup operation being composition of functions) and will be denoted by $\mathcal{A}_R$. Observe that $\mathcal{A}_R$ is a group if and only if R is a field.

For each $v \in R$, the map $x \mapsto x + v$ will be denoted by A_v (add v) and, for each $u \in R^*$, the map $x \mapsto ux$ will be denoted by M_u (multiply by u). Note that the distributive law in R can be expressed as:

$$M_u A_v = A_{uv} M_u \tag{3.1}$$

The affine transformations A_v with $v \in R$ form the *additive subgroup* of $\mathcal{A}_R$, denoted by S_A. The affine transformations M_u with $u \in R^*$ form the *multiplicative sub-semigroup* of $\mathcal{A}_R$, denoted by S_M. Observe that S_A is isomorphic to the additive group $(R, +)$ and S_M is isomorphic to the multiplicative semigroup $(R^*, \cdot)$.

Note that the map $x \mapsto ux + v$ is the composition $A_v M_u$. Thus the sub-semigroups S_M and S_A generate the semigroup $\mathcal{A}_R$. When K is a field, $\mathcal{A}_K$ is the semidirect product of the (abelian) groups S_A and S_M and hence is amenable. However, as it was pointed out in Remark 6.2 in [BM16a], the semigroup $\mathcal{A}_{\mathbb{Z}}$ is not amenable. In fact we have:

Proposition 3.5. *Let R be a countable integral domain. The affine semigroup $\mathcal{A}_R$ is amenable if and only if R is a field.*

Proof. As was explained above, if R is a field then $\mathcal{A}_R$ is the semidirect product of two abelian groups, and hence is solvable. It is a well known fact that solvable groups are amenable (cf. [Pat88, (0.15) and (0.16)]) so this proves one direction.

Assume now that $\mathcal{A}_R$ is amenable. The semigroup $\mathcal{A}_R$ acts naturally on R by affine transformations, therefore the amenability of $\mathcal{A}_R$ implies the existence of a finitely additive mean $\lambda : \mathcal{P}(R) \to [0, 1]$ defined on all the subsets of R which is invariant under all affine transformations (this means that $\lambda(\{x \in R : g(x) \in E\}) = \lambda(E)$ for any $E \subset R$ and $g \in \mathcal{A}_R$). Given $x \in R^*$, we have $1 = \lambda(R) = \lambda(xR)$ (because the map $y \mapsto xy$ belongs to $\mathcal{A}_R$).

Assume, for the sake of a contradiction, that R is not a field and let $x \in R^*$ be a non-invertible element. The ideal xR is not the whole ring and hence there is a shift $xR+a$ which is disjoint from xR. The invariance of λ implies that $\lambda(xR) = \lambda(xR + a)$, but disjointness implies that $\lambda\big(xR \cup (xR + a)\big) = \lambda(xR) + \lambda(xR + a) = 2\lambda(xR)$. We now conclude that

$$1 = \lambda(xR) = \frac{1}{2}\lambda\big(xR \cup (xR + a)\big) \leq \frac{1}{2}\lambda(R) = \frac{1}{2}$$

which gives the desired contradiction. $\square$

3.2 Double Følner sequences

As mentioned in Proposition 3.5, when K is a (countable, discrete) field, its affine group $\mathcal{A}_K$ is a (discrete) countable amenable group. This suggests the existence of a sequence of finite sets (F_N) in K asymptotically invariant under the action of $\mathcal{A}_K$. Indeed we have the following:

Proposition 3.6. *Let K be a countable field. There exists a sequence of non-empty finite sets (F_N) in K which forms a Følner sequence for the actions of both the additive group $(K,+)$ and the multiplicative group $(K^*,\times)$. In other words, for each $u \in K^*$ we have:*

$$\lim_{N\to\infty} \frac{|F_N \cap (F_N + u)|}{|F_N|} = \lim_{N\to\infty} \frac{|F_N \cap (uF_N)|}{|F_N|} = 1$$

We call such a sequence (F_N) a double Følner sequence.

Proof. Let $(G_N)_{N\in\mathbb{N}}$ be a (left) Følner sequence in $\mathcal{A}_K$. This means that G_N is a non-empty finite subset of $\mathcal{A}_K$ for each $N \in \mathbb{N}$, and that for each $g \in \mathcal{A}_K$ we have

$$\lim_{N\to\infty} \frac{|G_N \cap (gG_N)|}{|G_N|} = 1$$

Note that for $g_1, g_2 \in \mathcal{A}_K$, if $g_1 \neq g_2$ then there is at most one solution $x \in K$ to the equation $g_1 x = g_2 x$. Thus, for each $N \in \mathbb{N}$, we can find a point x_N in the (infinite) field K such that $g_i x_N \neq g_j x_N$ for all pairs $g_i, g_j \in G_N$ with $g_i \neq g_j$. It follows that $F_N := \{g x_N : g \in G_N\}$ has $|G_N|$ elements.

Since $F_N \cap gF_N \supset \{h x_N : h \in G_N \cap gG_N\}$ we have $|F_N \cap gF_N| \geq |\{h x_N : h \in G_N \cap gG_N\}| = |G_N \cap gG_N|$. Therefore:

$$1 \geq \limsup_{N\to\infty} \frac{|F_N \cap gF_N|}{|F_N|} \geq \liminf_{N\to\infty} \frac{|F_N \cap gF_N|}{|F_N|} \geq \lim_{N\to\infty} \frac{|G_N \cap gG_N|}{|G_N|} = 1$$

Finally, putting $g = M_u$ and $g = A_u$ in the previous equation we get that (F_N) is a Følner sequence for $(K^*,\times)$ and for $(K,+)$. $\square$

Observe that, according to (the proof of) Proposition 3.5, an LID ring R which is not a field can *not* possess a double Følner sequence.

Given a double Følner sequence $(F_N)_{N\in\mathbb{N}}$ in K and a set $E \subset K$, the lower density of E with respect to (F_N) is defined by the formula:

$$\underline{d}_{(F_N)}(E) := \liminf_{N\to\infty} \frac{|F_N \cap E|}{|F_N|}$$

and the upper density of E with respect to (F_N) is defined by the formula:

$$\bar{d}_{(F_N)}(E) := \limsup_{N\to\infty} \frac{|F_N \cap E|}{|F_N|}$$

Several basic properties of the upper and lower densities with respect to a Følner sequence in a group remain true for densities with respect to double Følner sequences, and the proofs carry over to this setting. We list some of these facts in the next lemma.

When $g \in \mathcal{A}_R$ is an affine transformation of a ring R and $E \subset R$ is any subset, we define

$$\theta_g E = \{g(x) : x \in E\} \qquad \text{and} \qquad \theta_g^{-1} E = \{x \in R : g(x) \in E\} \tag{3.2}$$

Lemma 3.7. *Let K be a field, let (F_N) be a double Følner sequence in K, let $E_1, E_2 \subset K$ and let $g \in \mathcal{A}_K$.*

1. $\bar{d}_{(F_N)}(\theta_g E) = \bar{d}_{(F_N)}(E)$ *and* $\underline{d}_{(F_N)}(\theta_g E) = \underline{d}_{(F_N)}(E)$.
2. $\bar{d}_{(F_N)}(E_1 \cup E_2) \leq \bar{d}_{(F_N)}(E_1) + \bar{d}_{(F_N)}(E_2)$
3. $\underline{d}_{(F_N)}(E_1 \cup E_2) \geq \underline{d}_{(F_N)}(E_1) + \underline{d}_{(F_N)}(E_2)$.
4. *If* $E_2 = K \setminus E_1$, *then* $\bar{d}_{(F_N)}(E_1) + \underline{d}_{(F_N)}(E_2) = 1$.

Note that part 1 of this lemma implies in particular that for every $u \in K^*$ and $E \subset K$ we have $\underline{d}_{(F_N)}(E/u) = \underline{d}_{(F_N)}(E-u) = \underline{d}_{(F_N)}(E)$ and $\bar{d}_{(F_N)}(E/u) = \bar{d}_{(F_N)}(E-u) = \bar{d}_{(F_N)}(E)$.

We will need the following lemma, which, roughly speaking, asserts that certain transformations of Følner sequences are still Følner sequences.

Lemma 3.8. *Let (F_N) be a double Følner sequence in a field K and let $b \in K^*$. Then the sequence (bF_N) is also a double Følner sequence. Also, if (F_N) is a Følner sequence for the multiplicative group $(K^*, \times)$, then the sequence $\left(F_N^{-1}\right)$, where $F_N^{-1} = \{g^{-1} : g \in F_N\}$, is still a Følner sequence for that group.*

Proof. The sequence (bF_N) is trivially a Følner sequence for the multiplicative group. To prove that it is also a Følner sequence for the additive group, let $x \in F$, we have

$$\lim_{N \to \infty} \frac{|bF_N \cap (x + bF_N)|}{|bF_N|} = \lim_{N \to \infty} \frac{\left|b\big(F_N \cap (x/b + F_N)\big)\right|}{|F_N|} = 1$$

To prove that $\left(F_N^{-1}\right)$ is a Følner sequence for the multiplicative group note that for any finite sets $A, B \subset K$ we have $\left|A^{-1}\right| = |A|$, $(A \cap B)^{-1} = A^{-1} \cap B^{-1}$ and if $x \in K^*$ then $(xA)^{-1} = x^{-1}A^{-1}$. Putting all together we conclude that

$$\begin{aligned}
\lim_{N\to\infty} \frac{\left|F_N^{-1} \cap (xF_N^{-1})\right|}{\left|F_N^{-1}\right|} &= \lim_{N\to\infty} \frac{\left|\left(F_N \cap (x^{-1}F_N)\right)^{-1}\right|}{|F_N|} \\
&= \lim_{N\to\infty} \frac{\left|F_N \cap (x^{-1}F_N)\right|}{|F_N|} = 1
\end{aligned}$$

□

3.3 Ultrafilters with nice affine properties

Let R be a LID. Since R is endowed with two operations (addition and multiplication), also its Stone-Čech compactification βR has two semi-continuous operations. Recall from Section 2.1 that these operations are not commutative (even though R is a commutative ring). Moreover, the operations in βR also fail to satisfy the distributive law in general. Nevertheless, we have

Proposition 3.9. *Let $u \in R$ and $p, q \in \beta R$. Then*

- *$u + p = p + u$ and $up = pu$.*
- *$(p + q)u = pu + qu$.*

One can easily check that for each $p, q \in \beta R$ we have (cf. Remark 4.2 in [HS98]):

$$p + q = p\text{-}\lim_u (u + q) \qquad\qquad pq = p\text{-}\lim_u (uq) \tag{3.3}$$

An ultrafilter $p \in \beta R$ is an *additive idempotent* if it is an idempotent ultrafilter with respect to the semigroup structure $(R, +)$ (equivalently if $p+p = p$), and it is a *multiplicative idempotent* if it is an idempotent ultrafilter with respect to the semigroup structure $(R^*, \times)$ (equivalently if $pp = p$). Observe that the principal ultrafilter $1_R \in \beta R$ is a multiplicative idempotent and $0_R \in \beta R$ is both an additive idempotent and a multiplicative idempotent.

Since R is an integral domain and $\beta(R^*) = (\beta R) \setminus \{0\}$ is closed in βR, it follows from (3.3) that $\beta(R^*)$ is closed under multiplication. In view of Proposition 3.9 and (3.3) we have that, for each $u \in R$, both maps $A_u : p \mapsto p + u$ and $M_u : p \mapsto pu$ are continuous. Therefore we can define topological dynamical systems $(\beta R, S_A)$ and $(\beta(R^*), S_M)$, where S_A and S_M are the additive and multiplicative sub-semigroups of $\mathcal{A}_R$, respectively. Invoking again (3.3) one can check that any closed S_A-invariant subset of βR is a semigroup for addition, and any closed S_M-invariant subset of βR^* is a semigroup for multiplication.

In view of Proposition 2.54 there exist minimal non-empty compact S_A-invariant subsets of βR and minimal non-empty compact S_M-invariant subsets of $\beta(R^*)$. An *additive minimal idempotent* is a non-principal ultrafilter $p \in \beta R$ which belongs to a minimal compact S_A-invariant set and such that $p+p = p$. A *multiplicative minimal idempotent* is a non-principal ultrafilter $p \in \beta(R^*)$ which belongs to a minimal compact S_M-invariant set and such that $pp = p$.

Definition 3.10. Let R be a ring. We denote by $\mathcal{AMI}$ the set of all additive minimal idempotents in βR and we denote by $\mathcal{MMI}$ the set of all multiplicative minimal idempotents in $\beta(R^*)$.

A set $C \subset R$ is called *additively central* if there exists $p \in \mathcal{AMI}$ such that $C \in p$. Similarly, any member of an ultrafilter $p \in \mathcal{MMI}$ is called *multiplicatively central*. We will be interested in sets $C \subset R$ which are simultaneously additively and multiplicatively central.

Unfortunately, the sets $\mathcal{AMI}$ and $\mathcal{MMI}$ are in general disjoint (cf [HS98, Corollary 13.15]). However, at least when R is an LID, the closure $\overline{\mathcal{AMI}}$ has non-trivial intersection with $\mathcal{MMI}$ (see Proposition 3.22 below).

Definition 3.11.

- Let $\mathcal{G} = \overline{\mathcal{AMI}} \cap \mathcal{MMI}$.
- A set $C \subset R$ is called *DC* (doubly central) if there exists an ultrafilter $p \in \mathcal{G}$ such that $C \in p$.

- A set $C \subset R$ is called DC^* if it has non-empty intersection with every DC set[1].

Observe that a set $C \subset R$ is DC^* if and only if it is contained in every ultrafilter $p \in \mathcal{G}$ (this follows directly from Definition 3.11 and the definition of ultrafilters).

Remark 3.12. *Observe that a DC set $A \subset R$ in a LID is central with respect to both the additive and the multiplicative structure, hence the name doubly central. On the other hand, it follows from Theorem 5.35 that not every set A which is central both additively and multiplicatively is a DC set.*

We will need the following technical lemma

Lemma 3.13. *Let G be a group and let $H \subset G$ be a normal subgroup with finite index. Then for any ultrafilter $p \in \beta G$ in the closure of the idempotents we have $H \in p$.*

Proof. The set of ultrafilters containing H is a closed set, hence we can assume that p is itself an idempotent. Since H has only finitely many cosets, exactly one of them, say aH is in p. Therefore, given $g \in G$ we have $g^{-1}aH \in p$ if and only if $g^{-1}a \in aH$. This is equivalent to $g \in aHa^{-1} = H$ (because H is normal). Since $aH \in p = p + p$ we conclude

$$\{g \in G : g^{-1}aH \in p\} \in p \iff H \in p$$

□

Corollary 3.14. *Let G be a countable commutative group, let $(x_\alpha)_{\alpha\in\mathcal{F}}$ be an IP-set in G and let $H \subset G$ be a subgroup with finite index. Then there exists a sub-IP-set $(y_\beta)_{\beta\in\mathcal{F}}$ of $(x_\alpha)_{\alpha\in\mathcal{F}}$ taking values in H.*

Proof. In view of Proposition 2.28, $(x_\alpha)_{\alpha\in\mathcal{F}}$ is a member of an ultrafilter p in the closure of the idempotents. By Lemma 3.13, H is in p and hence so is $\{x_\alpha : \alpha \in \mathcal{F}\} \cap H$. Invoking again Proposition 2.28 it follows that the intersection contains a sub-IP-set $(y_\beta)_{\beta\in\mathcal{F}}$ of $(x_\alpha)_{\alpha\in\mathcal{F}}$. □

---

[1] We call the reader's attention to the fact that there is no relation between the * in DC^* and the * in R^*.

A particular case of Lemma 3.13 is when R is a LID, H is a non-trivial ideal and $p \in \mathcal{G}$. If $p \in \beta(R^*)$ contains an ideal bR for some $b \in R^*$, then one can define an ultrafilter $b^{-1}p$ as the family q of those sets $E \subset R$ such that $bE \subset p$. Observe that in this case $bq = p$.

The following lemma is the analogue of Theorem 5.4 in [BH90] (where it is stated and proved for $\mathbb{N}$).

Lemma 3.15. *Let R be a LID, let $p \in \overline{\mathcal{AMI}}$ and let $u \in R^*$. Then both up and $u^{-1}p$ belong to $\overline{\mathcal{AMI}}$.*

Proof. Since $M_u : p \mapsto up$ and $M_u^{-1} : p \mapsto u^{-1}p$ are continuous (on their respective domains), it suffices to show that if $p \in \mathcal{AMI}$ then also both up and $u^{-1}p$ are in $\mathcal{AMI}$. It follows directly from Proposition 3.9 that $up + up = u(p+p) = up$, so up is an additive idempotent. Checking the definitions easily yields that also $u^{-1}p$ is an additive idempotent.

All that remains to show is that up and $u^{-1}p$ are both (additively) minimal.

(1) $up \in \mathcal{AMI}$

In view of Proposition 2.30, it suffices to show that every $E \in up$ is additively piecewise syndetic. From the definition of multiplication, $E \in up \iff E/u \in p$. Since R is a LID, Lemma 3.13 implies that $uR \in p$, hence $u \cdot (E/u) = E \cap uR \in p$ and therefore is piecewise syndetic (because of Proposition 2.30). It now follows from Lemma 3.4 that E/u and hence E itself must be (additively) piecewise syndetic.

(2) $u^{-1}p \in \mathcal{AMI}$

We use a similar argument: a set $E \in u^{-1}p \iff uE \in p$. Therefore Proposition 2.30 and Lemma 3.4 imply that every member of $u^{-1}p$ is piecewise syndetic, which in view of Proposition 2.30 is equivalent to the statement that $u^{-1}p$ belong to the closure of (additively) minimal ultrafilters.

□

Lemma 3.16. *Let X be a compact space and let $(x_u)_{u \in R}$ be a sequence in X indexed by a countable ring R. Then for each $k \in R^*$ and $p \in \beta R$ we have $p\text{-}\lim\limits_{u} x_{ku} = kp\text{-}\lim\limits_{u} x_u$.*

Proof. Let $x = p\text{-}\lim_u x_{ku}$ and let $U \subset X$ be a neighborhood of x. By definition, the set $E = \{u \in R : x_{ku} \in U\} \in p$. Note that $E = \{u \in R : x_u \in U\}/k$, and hence $\{u \in R : x_u \in U\} \in kp$. Since U is an arbitrary neighborhood of x we conclude that $kp\text{-}\lim_u x_u = x$. □

3.4 Affine syndeticity and thickness

In this section we will develop the notions of affinely syndetic and affinely thick subsets of R. The definitions and proofs are parallel to the usual notions of syndetic and thick. Recall that, for a discrete semigroup G, a set $S \subset G$ is *syndetic* if finitely many translates of S cover G (see Definition 2.15).

Recall from equation (3.2) the notation $\theta_g E = \{g(x) : x \in E\}$ for a set $E \subset R$ and $g \in \mathcal{A}_R$. When $F \subset \mathcal{A}_R$, $S \subset R$ and $x \in R$ we write

$$\theta_F^{-1} S := \bigcup_{g \in F} \theta_g^{-1} S \qquad \text{and} \qquad \theta_F x := \bigcup_{g \in F} g(x)$$

Definition 3.17. Let R be a ring. A set $S \subset R$ is *affinely syndetic* if there exists a finite set $F \subset \mathcal{A}_R$ such that $\theta_F^{-1} S = R$.

Observe that if a set $S \subset R^*$ is syndetic in either the group $(R, +)$ or the semigroup $(R^*, \cdot)$, then S is affinely syndetic. Indeed, assume, for instance, that S is syndetic in $(R, +)$ and let $F \subset R$ be a finite set such that $S - F = R$. Then considering the subset $\{A_u : u \in F\} \subset \mathcal{A}_R$ we deduce that $\theta_F^{-1} S = R$ and hence S is affinely syndetic. On the other hand, S can be affinely syndetic and not be syndetic for neither the group $(R, +)$ nor the semigroup $(R^*, \cdot)$ (this follows from Example 3.19 and Proposition 3.20 below).

Recall that, for a discrete semigroup G, a set $T \subset G$ is *thick* if it contains a shift of an arbitrary finite set (see Definition 2.17).

Definition 3.18. A set $T \subset R$ is *affinely thick* if for every finite set $F \subset \mathcal{A}_R$ there exists $x \in R$ such that $\theta_F x \subset T$.

Observe that if $T \subset R$ is affinely thick, then it is thick in both the group $(R, +)$ and the semigroup $(R^*, \cdot)$. The following example shows that there exist sets T which are not affinely thick (even when R is a field) but thick in both $(R, +)$ and $(R^*, \cdot)$:

Example 3.19. We take the ring $R = \mathbb{Q}$ of rational numbers. Let (G_N) be an increasing sequence of finite subsets of $\mathbb{Q}$ whose union is $\mathbb{Q}$. For any sequence $(a_N) \subset \mathbb{Q}^*$, the set

$$E = \left(\bigcup_{N=1}^{\infty} (a_{2N-1} + G_{2N-1}) \right) \cup \left(\bigcup_{N=1}^{\infty} (a_{2N} G_{2N}) \right) = \bigcup_{N=1}^{\infty} E_N$$

is additively thick and multiplicatively thick, where $E_N = a_N + G_N$ when N is odd and $E_N = a_N G_N$ when N is even. However, if (a_N) is growing sufficiently fast, then E is not affinely thick. Indeed, for every point $x \in \mathbb{Q}$ we may have

$$\theta_{\{Id, A_1 M_2\}} x = \{x, x+1, 2x\} \not\subset E$$

To see this, let $a_0 = 1$ and $E_0 := \{0\}$. Let ΔG_N denote the set defined by $\Delta G_N = \{x_2 - x_1, x_3 - x_2, \ldots, x_k - x_{k-1}\}$ where $x_1 < x_2 < \cdots < x_k$ is an ordering of the elements of G_N. Let $M_N = \min \{|x| : x \in G_N \setminus \{0\}\}$. Define recursively

$$a_N = \begin{cases} 2\max(E_{N-1}) + \max(G_N) - 2\min(G_N) & \text{if} \quad N \text{ is odd} \\ \frac{1}{\min(\Delta G_N)} + \frac{2\max(E_{N-1})}{M_N} & \text{if} \quad N \text{ is even} \end{cases}$$

Note that if N is even and $x \in E_N$, then $x + 1 \notin E_N$. If N is odd and $x \in E_N$, then $x \geq \min(G_N) + a_N$ which implies that $2x > \max(G_N) + a_N$ and hence $2x \notin E_N$. Thus for any $N \in \mathbb{N}$ and $x \in \mathbb{Q}$, the set $\{x, x+1, 2x\}$ is not a subset of E_N.

Since $\min\{|x| : x \in E_{N+1} \setminus \{0\}\} > 2\max\{|x| : x \in E_N\}$, if $x \in E_N$, then $2x \notin E_{N+1}$ (and in fact $2x \notin E_L$ for any $L > N$) and hence $\{x, x+1, 2x\}$ is not a subset of E for any $x \in \mathbb{Q}$.

The following proposition is an immediate consequence of the definitions.

Proposition 3.20. *A set $S \subset R$ is affinely syndetic if and only if it has non-empty intersection with every affinely thick set. A set $T \subset R$ is affinely thick if and only if it has non-empty intersection with every affinely syndetic set.*

Now we connect affine syndeticity and thickness in countable fields with upper and lower density with respect to double Følner sequences.

Theorem 3.21. *Let K be a countable field. A set $S \subset K$ is affinely syndetic if and only if for every double Følner sequence (F_N) in K, we have $\bar{d}_{(F_N)}(S) > 0$. A set $S \subset K$ is affinely thick if and only if there exists a double Følner sequence (F_N) in K such that $\underline{d}_{(F_N)}(T) = 1$.*

Proof. Assume $S \subset K$ is affinely syndetic and let $F \subset \mathcal{A}_K$ be a finite set such that $\theta_F^{-1} S = K$. Then for any double Følner sequence (F_N), using parts (1) and (2) of Lemma 3.7 we have

$$1 = \bar{d}_{(F_N)}(K) = \bar{d}_{(F_N)}\left(\bigcup_{g \in F} \theta_{g^{-1}} S\right) \leq \sum_{g \in F} \bar{d}_{(F_N)}(\theta_{g^{-1}} S) = |F| \bar{d}_{(F_N)}(S)$$

and hence $\bar{d}_{(F_N)}(S) \geq 1/|F| > 0$.

Now assume that $T \subset K$ is affinely thick and let (G_N) be an arbitrary (left) Følner sequence in $\mathcal{A}_K$. For each $N \in \mathbb{N}$ let $x_N \in K$ be such that $F_N := \theta_{G_N} x_N \subset T$ and $|F_N| = |G_N|$. To see why this is possible, note that for any affine transformations $g_1, g_2 \in \mathcal{A}_K$ with $g_1 \neq g_2$, there is at most one solution $x \in K$ to the equation $g_1(x) = g_2(x)$. Thus there are only finitely many $x \in K$ such that $g_1 x = g_2 x$ for some pair $g_1 \neq g_2 \in G_N$. On the other hand, since T is affinely thick, there are infinitely many $x \in K$ such that $\theta_{G_N} x \subset T$ (and indeed an affinely thick set of such x).

We now show that (F_N) is a double Følner sequence in K. For any fixed $g \in \mathcal{A}_K$ we have

$$F_N \cap \theta_g F_N = \theta_{G_N} x_N \cap \theta_g(\theta_{G_N} x_N) \supset \theta_{G_N \cap g G_N} x_N$$

and hence

$$1 \geq \limsup_{N \to \infty} \frac{|F_N \cap g F_N|}{|F_N|} \geq \liminf_{N \to \infty} \frac{|F_N \cap g F_N|}{|F_N|} \geq \lim_{N \to \infty} \frac{|G_N \cap g G_N|}{|G_N|} = 1$$

because (G_N) is a left Følner sequence in $\mathcal{A}_K$. This implies that (F_N) is a double Følner sequence in K. Since for each $N \in \mathbb{N}$ we have $F_N \subset T$ we conclude that $\underline{d}_{(F_N)}(T) = 1$.

Now if S is not syndetic then it follows from Proposition 3.20 that $K \setminus S$ is thick. Therefore there exits a double Følner sequence (F_N) such that $\underline{d}_{(F_N)}(K \setminus S) = 1$. From part (4) of Lemma 3.7 if follows that $\bar{d}_{(F_N)}(S) = 0$.

Finally, if T is not thick, then $K \setminus T$ is syndetic and hence for every double Følner sequence (F_N) we have $\bar{d}_{(F_N)}(K \setminus T) > 0$. By part (4) of Lemma 3.7 we have $\underline{d}_{(F_N)}(T) < 1$ for every double Følner sequence in K.

□

In every countable semigroup, any thick set is central (cf. Proposition 2.31). The same phenomenon occurs in our affine setting:

Proposition 3.22. *Let R be a LID. Then every affinely thick set in R is DC (see Definition 3.11).*

Proof. Let $T \subset R$ be an affinely thick set. For $g \in \mathcal{A}_R$ define $\overline{\theta_{g^{-1}}T} \subset \beta R$ by equations (3.2) and (2.1). Note that, for any finite set $F \subset \mathcal{A}_R$:

$$\bigcap_{g \in F} \overline{\theta_{g^{-1}}T} = \overline{\bigcap_{g \in F} \theta_{g^{-1}}T} = \overline{\{x \in R : \theta_F x \subset T\}}$$

Since T is affinely thick, the family of compact sets $\left\{\overline{\theta_{g^{-1}}T} : g \in \mathcal{A}_R\right\}$ has the finite intersection property, and hence the intersection $\mathcal{T} := \bigcap_{g \in \mathcal{A}_R} \overline{\theta_{g^{-1}}T}$ is a non-empty compact subset of βR. We have the following description of $\mathcal{T}$:

$$p \in \mathcal{T} \iff (\forall g \in \mathcal{A}_R) p \in \overline{\theta_{g^{-1}}T} \iff (\forall g \in \mathcal{A}_R) \theta_{g^{-1}}T \in p$$

If $p, q \in \mathcal{T}$, we claim that both $p + q \in \mathcal{T}$ and $pq \in \mathcal{T}$. Indeed, for all $g \in \mathcal{A}_R$ and $u \in R$ we have $A_u^{-1}\theta_{g^{-1}}T = (\theta_g A_u)^{-1}T$. Therefore we have:

$$\theta_{g^{-1}}T \in p + q \iff \{u \in R : A_u^{-1}\theta_{g^{-1}}T \in q\} \in p \iff \{u \in R : (\theta_g A_u)^{-1}T \in q\} \in p$$

Since $q \in \mathcal{T}$ the set $\{u \in R : (\theta_g A_u)^{-1}T \in q\} = R \in p$, so we conclude that $p + q \in \mathcal{T}$. The same argument with obvious modifications implies that $pq \in \mathcal{T}$ proving the claim.

We now have that $(\mathcal{T}, S_A)$ is a topological dynamical system. Hence by Proposition 2.54 there exists a minimal subsystem. It follows from (3.3) that each minimal subsystem is actually an (additive) left ideal in βR, and hence, in view of Theorem 2.11, there exist (additive) minimal idempotents in $\mathcal{T}$. Therefore the intersection $\mathcal{T}_1 := \overline{\mathcal{AMI}} \cap \mathcal{T}$ is a non-empty compact subset of $\mathcal{T}$.

If $u \in R^*$ and $p \in \mathcal{T}_1$, it follows from Lemma 3.15 that $up \in \overline{\mathcal{AMI}}$, and thus $up \in \mathcal{T}_1$. This means that $(\mathcal{T}_1, S_M)$ is a topological dynamical system and hence by Proposition 2.54 it has minimal subsystems. By Ellis lemma (Theorem 2.11) each minimal system contains some multiplicative idempotent. Let p be a multiplicative minimal idempotent in $\mathcal{T}_1$. Since $\mathcal{T}_1 \subset \mathcal{T}$ we conclude that $T \in p$. Since $\mathcal{T}_1 \subset \overline{\mathcal{AMI}}$ we conclude that $p \in \overline{\mathcal{AMI}}$, and hence $p \in \mathcal{G}$. □

Remark 3.23. *An immediate consequence of Propositions 3.22 and 3.20 is that every DC* set is affinely syndetic.*

3.5 An affine version of Furstenberg's correspondence principle

We will need an extension of Furstenberg's Correspondence Principle for an action of a group on a set (the classical versions deal with the case when the group acts on itself by translations, cf. [Fur77]).

Theorem 3.24. *Let X be a set, let G be a countable group and let $(\tau_g)_{g\in G}$ be an action of G on X. Assume that there exists a sequence (G_N) of finite subsets of X such that for each $g \in G$ we have the property:*

$$\frac{|G_N \cap (\tau_g G_N)|}{|G_N|} \to 1 \text{ as } N \to \infty \tag{3.4}$$

Let $E \subset X$ and assume that $\bar{d}_{(G_N)}(E) := \limsup_{N\to\infty} \frac{|G_N \cap E|}{|G_N|} > 0$.

Then there exists a compact metric space Ω, a probability measure μ on the Borel sets of Ω, a μ-preserving G-action $(T_g)_{g\in G}$ on Ω, a Borel set $B \subset \Omega$ such that $\mu(B) = \bar{d}_{(G_N)}(E)$, and for any $k \in \mathbb{N}$ and $g_1, \ldots, g_k \in G$ we have

$$\bar{d}_{(G_N)}\left(\tau_{g_1} E \cap \cdots \cap \tau_{g_k} E\right) \geq \mu\left(T_{g_1} B \cap \ldots \cap T_{g_k} B\right)$$

Proof. Define the family of sets

$$\mathcal{S} := \left\{ \bigcap_{j=1}^{k} \tau_{g_j} E : k \in \mathbb{N}, g_j \in G \ \forall j = 1, \dots, k \right\} \cup \{X\}$$

Note that $\mathcal{S}$ is countable, so using a diagonal procedure we can find a subsequence $(\tilde{G}_N)$ of the sequence (G_N) such that $\bar{d}_{(G_N)}(E) = \lim_{N\to\infty} |E \cap \tilde{G}_N|/|\tilde{G}_N|$ and, for each $S \in \mathcal{S}$, the following limit exists

$$\lim_{N\to\infty} \frac{|S \cap \tilde{G}_N|}{|\tilde{G}_N|}$$

Note that (3.4) holds for any subsequence of (G_N), and in particular for $(\tilde{G}_N)$. Let $B(X)$ be the space of all bounded complex-valued functions on X. The space $B(X)$ is a Banach space with respect to the norm $\|f\| = \sup_{x\in X} |f(x)|$. Let $\rho \in \ell^\infty(\mathbb{N})^*$ be a Banach limit[2].

Define the linear functional $\lambda : B(X) \to \mathbb{C}$ by

$$\lambda(f) = \rho\left(\left(\frac{1}{|\tilde{G}_N|} \sum_{x\in\tilde{G}_N} f(x)\right)_{N\in\mathbb{N}}\right)$$

The functional λ is positive (i.e. if $f \geq 0$ then $\lambda(f) \geq 0$) and $\lambda(1) = 1$. For any $f \in B(X)$, $g \in G$ and $x \in X$, the equation $f_g(x) = f(\tau_g x)$ defines a new function $f_g \in B(X)$. By (3.4) we have that $\lambda(f_g) = \lambda(f)$ for all $g \in G$, so λ is an invariant mean for the action $(\tau_g)_{g\in G}$. Moreover, $\bar{d}_{(G_N)}(E) = \lambda(1_E)$ and, for any $S \in \mathcal{S}$, we have $\bar{d}_{(G_N)}(S) \geq \lambda(1_S)$.

Note that the Banach space $B(X)$ is a commutative C^*-algebra (with the involution being pointwise conjugation). Now let $Y \subset B(X)$ be the (closed) subalgebra generated by the indicator functions of sets in $\mathcal{S}$. Then Y is itself a C^*-algebra. It has an identity (the constant function equal to 1) because $X \in \mathcal{S}$. If $f \in Y$ then $f_g \in Y$ for all $g \in G$. Moreover, since $\mathcal{S}$ is countable, Y is separable. Thus, by the Gelfand representation theorem (see, for instance, [Arv76], Theorem 1.1.1), there exists a compact metric space Ω and a map $\Phi : Y \to C(\Omega)$ which is simultaneously an algebra isomorphism and a homeomorphism.

The linear functional λ induces a positive linear functional L on $C(\Omega)$ by $L\big(\Phi(f)\big) = \lambda(f)$. Applying the Riesz Representation Theorem we have a measure μ on the Borel sets

[2]This means that $\rho : \ell^\infty(\mathbb{N}) \to \mathbb{C}$ is a shift invariant positive linear functional such that for any convergent sequence $\mathbf{x} = (x_n) \in \ell^\infty(\mathbb{N})$ we have $\rho(\mathbf{x}) = \lim x_n$.

of Ω such that

$$\lambda(f) = L(\Phi(f)) = \int_\Omega \Phi(f) d\mu \qquad \forall f \in B(X)$$

The action $(\tau_g)_{g\in G}$ induces an anti-action (or right action) $(U_g)_{g\in G}$ of G on $C(\Omega)$ by $U_g\Phi(f) = \Phi(f_g)$, where $f_g(x) = f(\tau_g x)$ for all $g \in G$, $f \in Y$ and $x \in X$. It is not hard to see that, for each $g \in G$, U_g is a positive invertible isometry of $C(\Omega)$. By the Banach-Stone theorem ([Sto37]), for each $g \in G$, there is a homeomorphism $T_g : \Omega \to \Omega$ such that $U_g\phi = \phi \circ T_g$ for all $\phi \in C(\Omega)$. Moreover for all $g, h \in G$ we have $\phi \circ T_{gh} = U_{gh}\phi = U_h U_g \phi = U_h(\phi \circ T_g) = \phi \circ (T_g \circ T_h)$. This means that $(T_g)_{g\in G}$ is an action of G on Ω. For every $f \in Y$ we have $\lambda(f_g) = \lambda(f)$ and hence

$$\begin{aligned}
\int_\Omega \Phi(f) \circ T_g d\mu &= \int_\Omega U_g \Phi(f) d\mu = \int_\Omega \Phi(f_g) d\mu \\
&= \lambda(f_g) = \lambda(f) = \int_\Omega \Phi(f) d\mu
\end{aligned}$$

Therefore the action (T_g) preserves measure μ.

Note that the only idempotents of the algebra $C(\Omega)$ are indicator functions of sets. Therefore, given any set $S \in \mathcal{S}$, the Gelfand transform $\Phi(1_S)$ of the characteristic function 1_S of S is the characteristic function of some Borel subset (which we denote by $\Phi(S)$) in Ω. In other words, $\Phi(S)$ is such that $\Phi(1_S) = 1_{\Phi(S)}$. Let $B = \Phi(E)$. We have

$$\bar{d}_{(G_N)}(E) = \lambda(1_E) = \int_\Omega \Phi(1_E) d\mu = \int_\Omega 1_B d\mu = \mu(B)$$

Since the indicator function of the intersection of two sets is the product of the indicator functions, we conclude that for any $k \in \mathbb{N}$ and any $g_1, ..., g_k \in G$ we have

$$\begin{aligned}
\bar{d}_{(G_N)}\left(\bigcap_{i=1}^k \tau_{g_i} E\right) &\geq \lambda\left(\prod_{i=1}^k 1_{\tau_{g_i} E}\right) = \int_\Omega \Phi\left(\prod_{i=1}^k 1_{\tau_{g_i} E}\right) d\mu \\
&= \int_\Omega \prod_{i=1}^k \Phi\left(1_{\tau_{g_i} E}\right) d\mu = \int_\Omega \prod_{i=1}^k U_{g_i^{-1}} \Phi\left(1_E\right) d\mu \\
&= \int_\Omega \prod_{i=1}^k 1_B \circ T_{g_i^{-1}} d\mu = \int_\Omega \prod_{i=1}^k 1_{T_{g_i} B} d\mu = \mu\left(\bigcap_{i=1}^k T_{g_i} B\right)
\end{aligned}$$

$\square$

3.6 An affine topological correspondence principle

The elegant idea of using topological dynamics to study partition regular configurations on $\mathbb{N}$ was developed by Furstenberg and Weiss in [FW78]. They considered each coloring $\chi : \mathbb{N} \to \{1, \ldots, r\}$ as a point in the symbolic system $(\{1, \ldots, r\}^{\mathbb{N}}, T)$ (where T is the left shift), and observed that it is possible to reformulate van der Waerden's theorem (and several other statements) as a multiple recurrence result on minimal subsystems of $(\{1, \ldots, r\}^{\mathbb{N}}, T)$. By proving the resulting multiple recurrence theorem ([FW78, Theorem 1.5]), they obtained a new proof of van der Waerden's theorem (and indeed of it's multidimensional version, Theorem 2.36). This correspondence is now a standard technique; for instance it was used by Bergelson and Leibman in their proof of the polynomial van der Waerden's theorem [BL96, Corollary 1.11] (see Corollary 6.7).

Unfortunately, the procedure described in the previous paragraph does not allow one to obtain dynamical formulations regarding the partition regularity of certain polynomial configurations. This is essentially because configurations such as $\{x+y, xy\}$ are not invariant under shifts (additive or multiplicative): if P is a set of the form $\{xy, x+y\}$ and $c \in \mathbb{N}$, then in general neither $P + c$ nor Pc is of the same form. By contrast, observe that arithmetic progressions are invariant under *both* addition and multiplication, in the sense that for any arithmetic progression P and any $c \in \mathbb{N}$, both $P + c$ and Pc are arithmetic progressions of the same length.

Nevertheless we have the following correspondence principle. Observe that any $\mathcal{A}_R$-topological system $(X, (T_g)_{g\in\mathcal{A}_R})$ naturally induces an additive $(R, +)$-topological system $(X, (S_u)_{u\in R})$, by letting $S_u := T_{A_u}$. A point $x \in X$ is called *additively minimal* if it is a minimal point for the system $(X, (S_u)_{u\in R})$ (cf. Section 2.5).

Theorem 3.25. *Let R be a LID and let $\mathcal{A}_R$ denote the semigroup of all affine transformations of R. There exists an $\mathcal{A}_R$-topological system $(X, (T_g)_{g\in\mathcal{A}_R})$ with a dense set of additively minimal points, such that each map $T_g : X \to X$ is open and injective, and with the property that for any finite coloring $R = C_1 \cup \cdots \cup C_r$ there exists an open cover*

$X = U_1 \cup \cdots \cup U_r$ *such that for any* $g_1, \ldots, g_k \in \mathcal{A}_R$ *and* $t \in \{1, \ldots, r\}$,

$$\bigcap_{\ell=1}^{k} T_{g_\ell}(U_t) \neq \varnothing \qquad \Longrightarrow \qquad \bigcap_{\ell=1}^{k} g_\ell(C_t) \neq \varnothing. \tag{3.5}$$

Remark 3.26. *It follows from the proof of Theorem 3.25 that the system* $(X, (T_g)_{g\in\mathcal{A}_R})$ *also has the property that for any piecewise syndetic set* $C_t \subset R$ *there exists a non-empty open set* $U_t \subset X$ *such that* (3.5) *holds for any* $g_1, \ldots, g_k \in \mathcal{A}_R$.

Remark 3.27. *It follows from the proof of Theorem 3.25 that the intersection* $\bigcap_{j=1}^{k} g_j(C_t)$ *(both in the theorem and in Remark 3.26) is not only non-empty but is in fact piecewise syndetic.*

The remainder of this section is dedicated to the proof of Theorem 3.25. The construction of X is quite explicit as a subset of the Stone-Čech compactification βR of R. In this setting, the action of $\mathcal{A}_R$ on X is natural. The idea of using the Stone-Čech compactification to prove the correspondence principle was inspired by its implicit use in [Bei11] (in the setting of measurable dynamics).

There is a natural action $(T_g)_{g\in\mathcal{A}_R}$ of $\mathcal{A}_R$ on the set $\beta R \setminus R$ of non-principle ultrafilters, described as follows. For $g \in \mathcal{A}_R$, the map $T_g : \beta R \setminus R \to \beta R \setminus R$ takes $p \in \beta R \setminus R$ to

$$T_g(p) := \{E \subset R : \theta_g^{-1}(E) \in p\} = \{E \subset \mathbb{N} : \{x \in R : g(x) \in E\} \in p\} \tag{3.6}$$

Remark 3.28. *An equivalent way to define* T_g *is to start with a map* $T_g : \beta R \to \beta R$, *defined on principal ultrafilters via the formula* $T_g(p_x) = p_{g(x)}$ *and then extend it to* βR *using the universal property of the Stone-Čech compactification. One can then check that for a non-principle ultrafilter* $p \in \beta R \setminus R$, *the image* $T_g(p)$ *is in fact in* $\beta R \setminus R$ *and corresponds to the ultrafilter described in* (3.6). *We will not make use of this fact.*

Lemma 3.29. *For each* $g \in \mathcal{A}_R$, *the map* $T_g : \beta R \setminus R \to \beta R \setminus R$ *is continuous, open and injective. Moreover, for* $g, h \in \mathcal{A}_R$ *one has* $T_g \circ T_h = T_{gh}$.

Proof. One can easily check (using only the definitions) that $T_g(p)$ is indeed a non-principle ultrafilter and that $T_g \circ T_h = T_{gh}$. To show that T_g is continuous, take a basic open set

$\overline{E} \subset \beta R$ for $E \subset R$ infinite; we need to show that $T_g^{-1}(\overline{E})$ is open. We have

$$p \in T_g^{-1}(\overline{E}) \iff E \in T_g(p) \iff g^{-1}(E) \in p$$

therefore $T_g^{-1}(\overline{E}) = \overline{g^{-1}(E)}$ is open and T_g is continuous.

To show that T_g is injective, let $p \neq q$ be in $\beta R \setminus R$ and let $E \in p \setminus q$. Since $g : R \to R$ is injective we have that $g^{-1}(g(E)) = E$; since $E \notin q$, it follows that $g(E) \notin T_g(q)$. Moreover, $g^{-1}(g(E)) = E \in p$, therefore $g(E) \in T_g(p)$ and hence $T_g(p) \neq T_g(q)$, proving injectivity.

Finally we show that T_g is open. Let $E \subset R$ be infinite; we will show that $T_g(\overline{E} \setminus R) = \overline{g(E)} \setminus R$, which will imply that $T_g : \beta R \setminus R \to \beta R \setminus R$ is indeed open. As in the proof of injectivity, if $p \in \overline{E}$ is non-principal, then $g(E) \in T_g(p)$, proving one of the inclusions. Conversely, if $p \in \beta R \setminus R$ is such that $g(E) \in T_g(p)$, then $E = g^{-1}(g(E)) \in p$, establishing the other inclusion and finishing the proof. □

Lemma 3.29 implies that $(T_g)_{g \in \mathcal{A}_R}$ is an action on $\beta R \setminus R$ and hence $\big(\beta R \setminus R, (T_g)_{g \in \mathcal{A}_R}\big)$ is an $\mathcal{A}_R$-topological dynamical system. We are now ready to prove Theorem 3.25.

Proof of Theorem 3.25. Let $Y \subset \beta R \setminus R$ be the set of all additively minimal points in $\big(\beta R \setminus R, (T_g)_{g \in \mathcal{A}_R}\big)$ and let $X := \overline{Y}$ be its closure. It is usual to denote $Y = K(\beta R, +)$. We will show that for each $g \in \mathcal{A}_R$, T_g maps X into X.

According to Proposition 2.30, an ultrafilter $p \in \beta R$ is in $X = \overline{K(\beta R, +)}$ if and only if every member $E \in p$ is piecewise syndetic. Take $p \in X$ and $g \in \mathcal{A}_R$; we claim that $T_g(p) \in X$. Using the definition, it suffices to show that if $g^{-1}(E)$ is piecewise syndetic, then so is E. It follows from Lemma 3.4 that if $g^{-1}(E)$ is piecewise syndetic, then so is $g(g^{-1}(E)) = E$. This shows that each $g \in \mathcal{A}_R$ induces a natural continuous map $T_g : X \to X$. Moreover, a similar argument shows that if $p \in \beta R \setminus R$ and $g \in \mathcal{A}_R$ are such that $T_g(p) \in X$, then $p \in X$; therefore $T_p : X \to X$ is also open.

So far we constructed a compact Hausdorff space X together with an action $(T_g)_{g \in \mathcal{A}_R}$ of $\mathcal{A}_R$ on X by continuous injective open maps with a dense set of additively minimal points. To finish the proof, consider a coloring $\mathbb{N} = C_1 \cup \cdots \cup C_r$ and let $U_t := \{p \in X : C_t \in p\} = \overline{C_t} \cap X$ for each $t \in \{1, \ldots, r\}$. Then each U_t is a (possibly empty) open subset of

X and each $p \in X$ belongs to some U_t. Now let $g_1, \ldots, g_k \in \mathcal{A}_R$ and $t \in \{1, \ldots, r\}$ be such that $\bigcap_{\ell=1}^{k} T_{g_\ell}(U_t) \neq \varnothing$. Then, since the maps $T_{g_\ell} : X \to X$ are open, it follows that $\bigcap_{\ell=1}^{k} T_{g_\ell}(U_t)$ is a non-empty open subset of X. Take any p in this intersection; we claim that $g_\ell(C_t) \cap \mathbb{N} \in p$ for any $\ell \in \{1, \ldots, k\}$.

Indeed, for each $\ell \in \{1, \ldots, k\}$, there exists $p_\ell \in U_t \subset \overline{C_t}$ such that $p = T_{g_\ell}(p_\ell)$. Since $g_\ell^{-1}(g_\ell(C_t)) = C_t \in p_\ell$, it follows that indeed $g_\ell(C_t) \in p$, as desired. Finally, it follows that the finite intersection $\bigcap_{\ell=1}^{k} g_\ell(C_t)$ is also in p and hence is non-empty. $\square$

CHAPTER 4

POLYNOMIAL EXTENSION OF DEUBER'S THEOREM

4.1 Introduction

In this chapter we present recent work from [BJM] extending the scope of Deuber's theorem (Theorem 1.10) to the setting of polynomial configurations in abelian groups (Theorem 4.7) and to (linear configurations on) arbitrary commutative semigroups (Theorem 4.6). Our proofs are based on a multidimensional polynomial extension of the central sets theorem, Theorem 4.8.

In the spirit of Deuber's original result in [Deu73], we also obtain a partition regularity result concerning the (linear) extension of Deuber's theorem to commutative semigroups. Roughly speaking, we show that for any finite coloring of a "large enough (but finite) configuration" there exists a monochromatic "smaller configuration" (see Theorem 4.15 below for a precise formulation). However, it is not clear how to obtain a similar partition regularity result for polynomial Deuber sets (our Theorem 4.7 requires one to finitely color the whole (infinite) group in order to obtain the desired monochromatic configuration).

In order to formulate our results we will need the following definitions.

Definition 4.1. Let G be a countable commutative semigroup.

1. A *shape* in G is a triple $(m,\vec{F},c)$ where $m\in\mathbb{N}$, $c:G\to G$ is a homomorphism and $\vec{F}$ is an m-tuple $\vec{F}=(F_1,\dots,F_m)$ where each F_j is a finite set of functions from G^j to G.

2. Given a shape $(m,\vec{F},c)$ and $\mathbf{s}=(s_0,\dots,s_m)\in(G\setminus\{0\})^{m+1}$, the $(m,\vec{F},c)$-*set gener-*

ated by $\mathbf{s}$ is the set

$$D(m,\vec{F},c;\mathbf{s}) := \left\{\begin{array}{ll} c(s_0) & \\ f(s_0)+c(s_1), & f\in F_1 \\ f(s_0,s_1)+c(s_2), & f\in F_2 \\ \vdots & \vdots \\ f(s_0,\ldots,s_{m-1})+c(s_m), & f\in F_m \end{array}\right\}$$

Observe that Schur triples $\{x,y,x+y\}$ are precisely the $(m,\vec{F},c)$ sets when $m=1$, c is the identity map and F_1 consists only of the identity map. More generally, the family which appears in Deuber's theorem (Theorem 1.10) for a given $m,p,c\in\mathbb{N}$ corresponds to the shape $(m,\vec{F},\tilde{c})$, where $\tilde{c}$ is the map $\tilde{c}:x\mapsto cx$ and F_j is the set of all maps $f:\mathbb{N}^j\to\mathbb{Z}$ of the form $f:\mathbf{x}\mapsto\langle\mathbf{x},\xi\rangle$ with $\xi\in\{-p,\ldots,p\}^j$.

Polynomial and multidimensional versions of Deuber's theorem

The following theorem is a joint extension of Folkman's theorem and the polynomial van der Waerden theorem, in the same way Theorem 1.10 is a joint extension of Folkman's theorem and (classical) van der Waerden's theorem.

Theorem 4.2 (Polynomial Deuber theorem)**.** *Let $m\in\mathbb{N}$ and, for each $i=1,2,\ldots,m$, let F_i be a finite set of polynomials $f:\mathbb{Z}^i\to\mathbb{Z}$ such that $f(\mathbf{0})=0$. Let $\vec{F}=(F_1,\ldots,F_m)$ and let $c:\mathbb{Z}\to\mathbb{Z}$ be a multiplication by some constant. Then the family associated with the shape $(m,\vec{F},c)$ is Ramsey. In other words, for any finite coloring $\mathbb{N}=C_1\cup\cdots\cup C_r$ there exists a color $C\in\{C_1,\ldots,C_r\}$ and a vector $\mathbf{s}\in\mathbb{N}^{m+1}$ such that $D(m,\vec{F},c;\mathbf{s})\subset C$.*

Theorem 4.2 follows from its multidimensional version, Theorem 4.4 below. We remark that Theorem 4.2 contains Theorems 1.1, 1.2, 1.3, 1.7, 1.8 and 1.10 as special cases. One could wonder whether Theorem 4.2 applies to every polynomial Ramsey family (hence solving Problem 1.12). Unfortunately, this is not the case, as the following result, obtained independently by Bergelson [Ber10, Theorem 6.1] and Hindman [Hin11], shows:

Theorem 4.3. *For any finite coloring of* $\mathbb{N}$ *there exist* x, y, z, t *of the same color satisfying* $xy = z + t$*. In other words, the family* $\{x, y, z, xy - z\}$ *is Ramsey.*

It is not hard to see that the family $\{x, y, z, xy - z\}$ is not contained in any shape, and hence Theorem 4.3 can not be derived from Theorem 4.2.

Next we combine Theorem 4.2 with the multidimensional van der Waerden's theorem (Theorem 2.36).

Theorem 4.4 (Multidimensional polynomial Deuber theorem)**.** *Let* $d, m \in \mathbb{N}$ *and, for each* $i = 1, 2, \ldots, m$ *let* F_i *be a finite set of polynomials* $f : (\mathbb{Z}^d)^i \to \mathbb{Z}^d$ *such that* $f(\mathbf{0}) = \mathbf{0}$*. Let* $\vec{F} = (F_1, \ldots, F_m)$ *and let* $c : \mathbb{Z}^d \to \mathbb{Z}^d$ *be a scalar homomorphism (i.e., multiplication by a constant). For any finite coloring of* $\mathbb{Z}^d$*, there exists* $\mathbf{s} \in (\mathbb{Z} \setminus \{0\})^{m+1}$ *such that the set* $D(m, \vec{F}, c; \mathbf{s})$ *is monochromatic.*

As mentioned in Chapter 2, for every finite partition of a countable commutative semigroup G, one of the colors is a central set (see Definition 2.32). We prove Theorem 4.4 by showing that in fact every central set contains a set of the form $D(m, \vec{F}, c; \mathbf{s})$.

Theorem 4.5 (Multidimensional polynomial Deuber theorem in central sets)**.** *Let* $d, m \in \mathbb{N}$ *and, for each* $i = 1, 2, \ldots, m$ *let* F_i *be a finite set of polynomials* $f : (\mathbb{Z}^d)^i \to \mathbb{Z}^d$ *such that* $f(\mathbf{0}) = \mathbf{0}$*. Let* $\vec{F} = (F_1, \ldots, F_m)$ *and let* $c : \mathbb{Z}^d \to \mathbb{Z}^d$ *be a scalar homomorphism. For any central set* $A \subset \mathbb{Z}^d$*, there exists* $\mathbf{s} \in (\mathbb{Z} \setminus \{0\})^{m+1}$ *such that* $D(m, \vec{F}, c; \mathbf{s}) \subset A$*.*

Theorem 4.5 is in turn a corollary to Theorem 4.7 below. The version of Theorem 4.5 for $d = 1$ and linear polynomials (i.e. multiplication by a constant) was first proved by Furstenberg in [Fur81, Theorem 8.22].

Deuber's theorem in commutative semigroups

We state now a version of Deuber's theorem which holds in any countable commutative semigroup. For semigroups G, H, denote by $\mathrm{Hom}(H, G)$ the set of all semigroup homomorphisms from H to G. Let also $\mathrm{End}(G)$ denote $\mathrm{Hom}(G, G)$ (elements of $\mathrm{End}(G)$ are often referred to as endomorphisms).

Theorem 4.6. *Let G be a countable commutative semigroup, let $A \subset G$ be a central set and let $(m, \vec{F}, c)$ be a shape in G. Assume that the map $c : G \to G$ is the identity map and, for each $j = 1, \ldots, m$, we have $F_j \subset \mathrm{Hom}(G^j, G)$. Then there exists $\mathbf{s} \in (G \setminus \{0\})^{m+1}$ such that $D(m, \vec{F}, c; \mathbf{s}) \subset A$.*

Next we move to polynomial maps. Recall from Definition 2.42 the notation $\mathbb{P}(G, H)$ to denote the set of all polynomial maps $f : G \to H$ with $f(0) = 0$.

Theorem 4.7. *Let G be a countable abelian group, let $A \subset G$ be a central set and let $(m, \vec{F}, c)$ be a shape in G. Assume that the image of c has finite index in G and, for each $j = 1, \ldots, m$, $F_j \subset \mathbb{P}(G^j, G)$. Then there exists $\mathbf{s} \in (G \setminus \{0\})^{m+1}$ such that $D(m, \vec{F}, c; \mathbf{s}) \subset A$.*

Theorems 4.6 and 4.7 are proved at the end of Section 4.2.

Polynomial extensions of the central sets theorem in abelian groups.

The main tool employed by Furstenberg in his proof of the special case of Theorem 4.5 mentioned above was his central sets theorem (cf. Theorem 2.40). Our proof of Theorem 4.7 is based on the following polynomial version of the central sets theorem, which we believe is of independent interest.

Theorem 4.8 (Multidimensional polynomial central sets theorem)**.** *Let G be a countable abelian group, let $j \in \mathbb{N}$ and let $(y_\alpha)_{\alpha \in \mathcal{F}}$ be an IP-set in G^j. Let $F \subset \mathbb{P}(G^j, G)$ be a finite set and let $A \subset G$ be a central set. Then there exist an IP-set $(x_\beta)_{\beta \in \mathcal{F}}$ in G and a sub-IP-set $(z_\beta)_{\beta \in \mathcal{F}}$ of $(y_\alpha)_{\alpha \in \mathcal{F}}$ such that*

$$\forall f \in F \quad \forall \beta \in \mathcal{F} \qquad x_\beta + f(z_\beta) \in A$$

When $G = \mathbb{Z}$ and the polynomial maps in F are homomorphisms, this reduces to the classical central sets theorem. Theorem 4.8 will be derived as a corollary of the more general Theorem 4.24 below.

Since for any finite coloring of a countable commutative group G one of the colors is a central set, as a corollary of Theorem 4.8 we deduce that

Theorem 4.9. *Let G be a countable abelian group, let $j \in \mathbb{N}$ and let $(y_\alpha)_{\alpha\in\mathcal{F}}$ be an IP-set in G^j. Let $F \subset \mathbb{P}(G^j, G)$ be a finite set and let $G = C_1 \cup \cdots \cup C_r$ be a finite coloring of G. Then there exist a color $C \in \{C_1, \ldots, C_r\}$, an IP-set $(x_\beta)_{\beta\in\mathcal{F}}$ in G and a sub-IP-set $(z_\beta)_{\beta\in\mathcal{F}}$ of $(y_\alpha)_{\alpha\in\mathcal{F}}$ such that*

$$\forall f \in F \quad \forall \beta \in \mathcal{F} \qquad x_\beta + f(z_\beta) \in C$$

Theorem 4.9 (and in fact a stronger version of it) was obtained by McCutcheon in [McC99, Theorem A]. The method used by McCutcheon also gives Theorem 4.8, although it does not appear to have been explicitly stated.

Partition regularity

Call a set $A \subset \mathbb{N}$ *rich* if for every linear Ramsey family $\{f_1, \ldots, f_k\}$ there exists $\mathbf{x}$ such that $\{f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \subset A$. In view of Theorem 1.11, a set is rich if and only if for every $m, p, c \in \mathbb{N}$ there exists $\mathbf{s} \in \mathbb{N}^{m+1}$ such that $D(m, p, c; \mathbf{s}) \subset A$. In view of Furstenberg's theorem [Fur81, Theorem 8.22] mentioned above, every central set in $\mathbb{N}$ is rich.

One of the main motivations for Deuber to introduce (m, p, c)-sets was to solve a conjecture of Rado stating that for a finite partition of a rich set, one of the cells is still rich. We obtain an analogous result for certain $(m, \vec{F}, c)$-sets. Before we state the main result in this direction (Theorem 4.12 below) we need a few definitions.

Definition 4.10. Let G be a countable commutative semigroup. A *clique* in G is an infinite (not necessarily countable) set of shapes. Given a clique Λ in G, we say that a set $A \subset G$ is Λ-*rich* if for every shape $(m, \vec{F}, c) \in \Lambda$ there exists an $(m, \vec{F}, c)$-set contained in A.

For example, let Λ be the clique in $\mathbb{N}$ consisting of the shapes that arise from all possible triples $(m, p, c) \in \mathbb{N}^3$. Then a set $A \subset \mathbb{N}$ is Λ-rich if and only if it is rich in the sense defined above. Here are more examples.

Example 4.11.

1. Let $k \in \mathbb{N}$, let $c : \mathbb{N} \to \mathbb{N}$ be the identity map, let $F_{1,k} = \{x \mapsto ix : i = 0, \ldots, k - 1\} \subset \mathrm{End}(\mathbb{N})$ and make $\vec{F_k} = (F_{1,k})$. Then any $(1, \vec{F_k}, c)$-set contains a "Brauer configuration" of length k (i.e. an arithmetic progression of length k together with its common difference, cf. Theorem 1.7).

2. Let c and $\vec{F_k}$ be as in part (1) above. Consider the clique $\Lambda = \{(1, \vec{F_k}, c) : k \in \mathbb{N}\}$. A set $A \subset \mathbb{N}$ is Λ-rich if and only it contains Brauer configurations of arbitrary length.

3. Let again $m \in \mathbb{N}$ and let $c : \mathbb{N} \to \mathbb{N}$ be the identity map. For each $j = 1, \ldots, m$, let $F_{j,m}$ be the set of all maps $f : \mathbb{N}^j \to \mathbb{N}$ of the form $f : \mathbf{x} \mapsto \langle \mathbf{x}, \xi \rangle$ where $\xi \in \{0, 1\}^j$. Let $\vec{F_m} = (F_{1,m}, \ldots, F_{m,m})$. Then any $(m, \vec{F_m}, c)$-set is a set of the form $FS(A)$ for some set $A \subset \mathbb{N}$ with cardinality $m + 1$.

4. Let c and $\vec{F_m}$ be as in part (3) of this example. Define the shape $\Lambda = \{(m, \vec{F_m}, c) : m \in \mathbb{N}\}$. A set $A \subset \mathbb{N}$ is Λ-rich if and only if it is an IP_0 set (cf. Definition 2.24).

One can reinterpret both Theorem 4.6 and Theorem 4.7 as providing an example each of a clique Λ such that every central set is Λ-rich. Our next theorem provides a natural example of a clique Λ with the stronger property that, for any finite partition of a Λ-rich set, one of the cells is still Λ-rich.

Theorem 4.12. *Let G be a countable commutative semigroup and let Λ_t be the clique consisting of all shapes $(m, \vec{F}, c)$ with $m \in \mathbb{N}$, c in the center of $\mathrm{End}(G)$ and $\vec{F} = (F_1, \ldots, F_m)$ where each $F_j \subset \mathrm{Hom}(G^j, G)$. In other words*

$$\Lambda_t = \left\{ (m, \vec{F}, c) : \begin{array}{l} m \in \mathbb{N},\ c \textit{ is in the center of } \mathrm{End}(G), \\ \vec{F} = (F_1, \ldots, F_m),\ F_j \subset \mathrm{Hom}(G^j, G)\ \forall j \end{array} \right\}$$

For any finite partition of a Λ_t-rich set, one of the cells is still Λ_t-rich.

If we take $G = \mathbb{N}$ then $\mathrm{End}(\mathbb{N})$ is isomorphic to the multiplicative semigroup $(\mathbb{N}, \times)$ and hence it is commutative; this means that any shape $(m, \vec{F}, c)$ arising from a triple $(m, p, c) \in \mathbb{N}^3$ as explained after Definition 4.1, is in Λ_t. Therefore Theorem 4.12 recovers

Deuber's solution of Rado's conjecture (that for any finite partition of a rich set, one of the cells must be rich).

Theorem 4.12 will be derived in Section 4.3 from its finitistic version, Theorem 4.15 below.

Definition 4.13. Let G be a countable commutative semigroup. Let $m \in \mathbb{N}$ and, for each $i = 1, \dots, m$, let $F_i \subset \mathrm{Hom}(G^i, G)$ be finite. Also, let $\vec{F} = (F_1, \cdots, F_m)$ and let $c \in \mathrm{End}(G)$. We say that c is *concordant* with $\vec{F}$ if there exists a non-zero homomorphism $b \in \mathrm{End}(G)$ and, for each $i \in \{1, \dots, m\}$ and $f \in F_i$, there is a homomorphism $a_f \in \mathrm{Hom}(G^i, G)$ such that $c \circ a_f = f \circ \mathbf{b}$, where $\mathbf{b} : G^i \to G^i$ is the homomorphism $\mathbf{b}(g_1, \dots, g_i) = \big(b(g_1), \dots, b(g_i)\big)$.

Observe that the identity homomorphism $c : x \mapsto x$ is concordant with any $\vec{F}$. More generally, if c is in the center of the semigroup $\mathrm{End}(G)$, then c is concordant with any $\vec{F}$ (by taking $b = c$ and $a_f = f$).

When c is an automorphism, it is concordant with any $\vec{F}$. Indeed, one can take b to be the identity map and $a_f = c^{-1} \circ f$. In the following example, c is neither in the center of $\mathrm{End}(G)$ nor is it an automorphism.

Example 4.14. Let $G = \mathbb{Z}^2$, let $m = 1$, let $c \in \mathrm{End}(\mathbb{Z}^2)$ be the projection onto the first coordinate and let $\vec{F} = (F_1)$ where F_1 consists of finitely many endomorphisms of $\mathbb{Z}^2$ whose image is contained in $c(\mathbb{Z}^2)$. Then c is concordant with $\vec{F}$.

Indeed, take $f \in F_1$. We let $b \in \mathrm{End}(\mathbb{Z}^2)$ be the identity map and $a_f = f$. Since the restriction of c to its image is the identity map, we have $c \circ a_f = f \circ b$.

Theorem 4.15. *Let G be a countable commutative semigroup and let Λ_c be the clique of all shapes $(m, \vec{F}, c)$ where $m \in \mathbb{N}$, $F_i \subset \mathrm{Hom}(G^i, G)$ for all $i = 1, \dots, m$ and c is concordant with $\vec{F}$.*

For any $r \in \mathbb{N}$ and any shape $(m, \vec{F}, c) \in \Lambda_c$ there exists another shape $(M, \vec{H}, C) \in \Lambda_c$ such that any partition of an $(M, \vec{H}, C)$-set into r-cells, one of the cells contains an $(m, \vec{F}, c)$-set.

Moreover, if c is the identity, we can take C to be the identity as well, and if c is in the center of $\operatorname{End}(G)$ *we can take C to be in the center of* $\operatorname{End}(G)$*.*

The proof of Theorem 4.15 occupies most of Section 4.3.

4.2 Idempotent ultrafilters and $(m, \vec{F}, c)$-sets

Theorems 4.7 and 4.6 have similar proofs. To avoid repetition, we unify both results into a single abstract result; this is Theorem 4.23 below. Before formulating it, we need to introduce some definitions.

Definition 4.16 (R-family)**.** Let G, H be countable commutative semigroups and let $p \in \beta G$ be an ultrafilter. Let Γ be a set of functions from $H \to G$. We say that Γ is an *R-family*[1] *with respect to p* if for every finite set $F \subset \Gamma$, every $A \in p$ and every IP-set $(y_\alpha)_{\alpha \in \mathcal{F}}$ in H, there exist $x \in G$ and $\alpha \in \mathcal{F}$ such that

$$x + f(y_\alpha) \in A \qquad\qquad \forall f \in F$$

Example 4.17. Let G be a countable abelian group, let $j \in \mathbb{N}$ and let $H = G^j$. Then the family $\Gamma = \mathbb{P}(G^j, G)$ is an R-family with respect to any minimal idempotent ultrafilter. Indeed, let $p \in \beta G$ be a minimal idempotent ultrafilter and let $A \in p$. In view of Proposition 2.30, A is a piecewise syndetic set. Fix a finite set $F \subset \Gamma$ and an IP-set $(y_\alpha)_{\alpha \in \mathcal{F}}$ in G^j. It follows from Corollary 2.45 that there exists $a \in A$ and $\alpha \in \mathcal{F}$ such that $a + f(y_\alpha) \in A$ for all $f \in F$, and hence Γ is an R-family.

Example 4.18. Let G be a countable commutative semigroup, let $j \in \mathbb{N}$ and let $H = G^j$. Then the family $\Gamma = \operatorname{Hom}(G^j, G)$ is an R-family with respect to any minimal idempotent ultrafilter. Indeed, let $p \in \beta G$ be a minimal idempotent ultrafilter and let $A \in p$. By definition, A is a central set, hence a piecewise syndetic set. Fix a finite set $F \subset \Gamma$ and an IP-set $(y_\alpha)_{\alpha \in \mathcal{F}}$ in G^j. It follows from Proposition 2.39 that there exists $a \in A$ and $\alpha \in \mathcal{F}$ such that $a + f(y_\alpha) \in A$ for all $f \in F$, and hence Γ is an R-family.

[1] R stands for *returns.*

Definition 4.19. Let G, H be countable commutative semigroups and let Γ be a set of functions from H to G. We say that Γ is *licit* if for any $f \in \Gamma$ and any $z \in H$, there exists a function $\phi_z \in \Gamma$ such that $f(y+z) = \phi_z(y) + f(z)$.

Example 4.20. Let G, H be countable commutative semigroups and let $\Gamma \subset \mathrm{Hom}(H, G)$. It is not hard to see that Γ is licit. Indeed, note that for every $f \in \Gamma$ and any $z \in H$ one can take $\phi_z = f$ in the definition.

Example 4.21. If G, H are countable abelian groups, the set $\Gamma = \mathbb{P}(H, G)$ is licit. Indeed, for each $f \in \Gamma$ and $z \in H$ one can define $\phi_z(y) := f(y+z) - f(z)$. Clearly $\phi(0) = 0$. For any $h \in H$, we have

$$\begin{aligned}\phi_z(y+h) - \phi_z(y) &= f(y+z+h) - f(z) - f(y+z) + f(z)\\ &= f\big((y+z)+h\big) - f(y+z)\end{aligned}$$

If $f \in \mathbb{P}(H, G)$ has degree d, then $f\big((y+z)+h\big) - f(y+z)$ is a polynomial map of degree at most $d-1$ in the variable y (now both h and z are constants), and hence ϕ_z is also a polynomial map of degree at most d.

Definition 4.22. Let G be a countable commutative semigroup. An endomorphism $c \in \mathrm{End}(G)$ is called *IP-regular* if for every IP-set $(x_\alpha)_{\alpha\in\mathcal{F}}$ in G there exists an IP-set $(y_\alpha)_{\alpha\in\mathcal{F}}$ such that $\big(c(y_\alpha)\big)_{\alpha\in\mathcal{F}}$ is a sub-IP-set of $(x_\alpha)_{\alpha\in\mathcal{F}}$ (and in particular $\big(c(y_\alpha)\big)_{\alpha\in\mathcal{F}}$ is itself an IP-set).

When $G = \mathbb{Z}$, any nontrivial endomorphism $c \in \mathrm{End}(\mathbb{Z})$ is IP-regular. It's not hard to see that when G is an arbitrary countable abelian group, any endomorphism whose image has finite index is IP-regular. We can now formulate our abstract theorem (which has theorems 4.6 and 4.7 as corollaries):

Theorem 4.23. *Let G be a countable commutative semigroup, let $p \in \beta G$ be an idempotent ultrafilter and let $\Gamma_1, \Gamma_2, \dots$ be R-families with respect to p which are licit, where Γ_j consists of maps from G^j to G. Let $c : G \to G$ be IP-regular, let $m \in \mathbb{N}$ and, for each $j = 1, \dots, m$,* *let $F_j \subset \Gamma_j$ be finite. Finally, put $\vec{F} = (F_1, \ldots, F_m)$. Then for any $A \in p$ there exists an IP-set $(\mathbf{s}_\alpha)_{\alpha\in\mathcal{F}}$ in G^{m+1} such that $D(m, \vec{F}, c; \mathbf{s}_\alpha) \subset A$ for every $\alpha \in \mathcal{F}$.*

In order to prove Theorem 4.23 we first need to establish an abstract version of the central sets theorem.

Theorem 4.24. *Let G, H be countable commutative semigroups, let $p \in \beta G$ be an idempotent ultrafilter, let Γ be an R-family with respect to p which is licit. Then for any finite set $F \subset \Gamma$, any $A \in p$ and any IP set $(y_\alpha)_{\alpha\in\mathcal{F}}$ in H, there exists a sub-IP-set $(z_\beta)_{\beta\in\mathcal{F}}$ of $(y_\alpha)_{\alpha\in\mathcal{F}}$ and an IP-set $(x_\beta)_{\beta\in\mathcal{F}}$ in G such that*

$$\forall f \in F \qquad \forall \beta \in \mathcal{F} \qquad x_\beta + f(y_\beta) \in A$$

Proof. Let $B = \{n \in A : A - n \in p\}$. Because p is an idempotent ultrafilter, $B \in p$. Moreover, by Lemma 4.14 in [HS98], for any $n \in B$, we have $B - n \in p$. We will construct sequences $x_1, x_2, \ldots$ in G and $\alpha_1 < \alpha_2 < \cdots$ in $\mathcal{F}$ inductively, so that for each n we have

$$\forall f \in F \qquad \forall \beta \subset [n],\ \beta \neq \varnothing \qquad x_\beta + f(z_\beta) \in B \tag{4.1}$$

where $z_\beta = \sum_{i\in\beta} y_{\alpha_i}$.

Since Γ is an R-family with respect to p, we can find $\alpha_1 \in \mathcal{F}$ and $x_1 \in G$ such that $x_1 + f(y_{\alpha_1}) \in B$ for all $f \in F$; in other words we get (4.1) for $n = 1$.

Now assume we have found $x_1, \ldots, x_n$ in G and $\alpha_1 < \cdots < \alpha_n$ in $\mathcal{F}$ such that (4.1) is true. Let

$$C = B \cap \left(\bigcap_{\substack{\varnothing\neq\beta\subset[n] \\ f\in F}} B - x_\beta - f(z_\beta) \right)$$

Each of the sets of the intersection is in p, and because p is closed under finite intersections, also $C \in p$. We now take advantage of the fact that Γ is licit to find, for each $f \in F$ and each nonempty $\beta \subset [n]$, a map $\phi_\beta^f \in \Gamma$ such that $f(z_\beta + y) = \phi_\beta^f(y) + f(z_\beta)$. Let $\Phi = F \cup \{\phi_\beta^f : \varnothing \neq \beta \subset [n]; f \in F\}$. We can now use again the fact that Γ is an R-family with respect to p and find $x_{n+1} \in G$ and $\alpha_{n+1} > \alpha_n$ in $\mathcal{F}$ such that $x_{n+1} + f(z_{n+1}) \in C$

for all $f \in \Phi$, where $z_{n+1} := y_{\alpha_{n+1}}$. We claim that (4.1) holds for $n+1$ with these choices, which will complete the induction and finish the proof.

Indeed, let $f \in F$ and let $\beta \subset [n+1]$ be non-empty. If $\beta \subset [n]$, then $x_\beta + f(z_\beta) \in B$ by the induction hypothesis. If $\beta = \{n+1\}$, then $x_{n+1} + f(z_{n+1}) \in C \subset B$ because $F \subset \Phi$. Otherwise the set defined by $\gamma := \beta \setminus \{n+1\} \subset [n]$ is nonempty. Recalling that $x_\beta = x_\gamma + x_{n+1}$ and $z_\beta = z_\gamma + z_{n+1}$, we have

$$x_{n+1} + \phi^f_\gamma(z_{n+1}) \in C \subset B - x_\gamma - f(z_\gamma),$$

so

$$x_\gamma + x_{n+1} + f(z_\gamma) + \phi^f_\gamma(z_{n+1}) \in B$$

which is equivalent to

$$x_\beta + f(z_\beta) \in B.$$

□

A concrete corollary of this general result is Theorem 4.8, which can be interpreted as a polynomial version of the central sets theorem. It follows from Theorem 4.24 by taking G to be a group and letting $H = G^j$, $\Gamma = \mathbb{P}(G^j, G)$, and p to be a minimal idempotent. According to Example 4.17, Γ is an R-family so Theorem 4.8 follows.

We are now in position to prove Theorem 4.23.

Proof of Theorem 4.23. What we need to show is that there exists some IP-set $(\mathbf{s}_\alpha)_{\alpha \in \mathcal{F}}$ in G^{m+1} such that, for all $\alpha \in \mathcal{F}$,

$$\begin{array}{llll} & c(s_{\alpha,0}) & \in & A \\ \forall f \in F_1 & f(s_{\alpha,0}) + c(s_{\alpha,1}) & \in & A \\ \forall f \in F_2 & f(s_{\alpha,0}, s_{\alpha,1}) + c(s_{\alpha,2}) & \in & A \\ \vdots & \vdots & \vdots & \vdots \\ \forall f \in F_m & f(s_{\alpha,0}, \ldots, s_{\alpha,m-1}) + c(s_{\alpha,m}) & \in & A \end{array} \tag{4.2}$$

The proof goes by induction on m; assume first that $m = 0$. Since A belongs to an idempotent ultrafilter, it contains an IP-set, say $(\tilde{x}_\alpha)_{\alpha \in \mathcal{F}}$. Since c is IP-regular, we can find an IP-set $(x_\alpha)_{\alpha\in\mathcal{F}}$ such that $\big(c(x_\alpha)\big)$ is a sub-IP-set of $(\tilde{x}_\alpha)_{\alpha\in\mathcal{F}}$ and hence $c(x_\alpha)\in A$ for each $\alpha\in\mathcal{F}$. Let $\mathbf{s}_\alpha^{(0)} := x_\alpha$ for each $\alpha\in\mathcal{F}$.

Now suppose that $m\geq 1$ and we have an IP-set in G^m

$$(\mathbf{s}_\alpha^{(m-1)})_{\alpha\in\mathcal{F}} = \Big(\big(s_{\alpha,0}^{(m-1)}, s_{\alpha,1}^{(m-1)}, \ldots, s_{\alpha,m-1}^{(m-1)}\big)\Big)_{\alpha\in\mathcal{F}}$$

such that for any $\alpha\in\mathcal{F}$ we have $D(m-1,\vec{F},c;\mathbf{s}_\alpha^{(m-1)})\subset A$; in other words, if we take $s_i = s_{\alpha,i}^{(m-1)}$ for each $i=0,\ldots,m-1$ we get the first m lines of (4.2), for any $\alpha\in\mathcal{F}$.

Now apply Theorem 4.24 with $H=G^m$, $\Gamma=\Gamma_m$, $F=F_m$ and $(y_\alpha)_{\alpha\in\mathcal{F}} = (\mathbf{s}_\alpha^{(m-1)})_{\alpha\in\mathcal{F}}$. We obtain a sub-IP-set $(\mathbf{t}_\alpha)$ of $(\mathbf{s}_\alpha^{(m-1)})$ in G^m and some IP set $(x_\alpha)_{\alpha\in\mathcal{F}}$ in G such that

$$\forall\alpha\in\mathcal{F}\quad\forall f\in F_m\qquad\qquad x_\alpha + f(\mathbf{t}_\alpha)\in A. \tag{4.3}$$

Since c is IP-regular we can find an IP-set $(y_\beta)_{\beta\in\mathcal{F}}$ in G such that $\big(c(y_\beta)\big)_{\beta\in\mathcal{F}}$ is a sub-IP-set of $(x_\alpha)_{\alpha\in\mathcal{F}}$; in other words, there exist $\alpha_1<\alpha_2<\cdots$ such that $c(y_\beta)=\sum_{i\in\beta}x_{\alpha_i}$ for all $\beta\in\mathcal{F}$. To ease the notation, let α_β denote the set $\alpha_\beta := \bigcup_{i\in\beta}\alpha_i\in\mathcal{F}$. Then

$$\forall\beta\in\mathcal{F}\qquad\qquad c(y_\beta) = x_{\alpha_\beta} \tag{4.4}$$

Now define $(\mathbf{s}_\beta^{(m)})_{\beta\in\mathcal{F}}$ by taking the corresponding sub-IP-set of $(\mathbf{t}_\alpha)_{\alpha\in\mathcal{F}}$ for the first m coordinates and letting $(y_\beta)_{\beta\in\mathcal{F}}$ be the last coordinate. More precisely we have:

$$\mathbf{s}_\beta^{(m)} = (\mathbf{t}_{\alpha_\beta}, y_\beta)\in G^{m+1}$$

Now fix $\beta\in\mathcal{F}$; we need to show that $D(m,\vec{F},c;\mathbf{s}_\beta^{(m)})\subset A$. If $j\in\{0,1,\ldots,m-1\}$ and $f\in F_j$ then

$$f\Big(s_{\beta,0}^{(m)},\ldots,s_{\beta,j-1}^{(m)}\Big) + c(s_{\beta,j}^{(m)}) = f\Big(s_{\alpha_\beta,0}^{(m-1)},\ldots,s_{\alpha_\beta,j-1}^{(m-1)}\Big) + c(s_{\alpha_\beta,j}^{(m-1)}) \tag{4.5}$$

and the expression in (4.5) is in A by induction. If $j=m$ then

$$f\Big(s_{\beta,0}^{(m)},\ldots,s_{\beta,j-1}^{(m)}\Big) + c(s_{\beta,j}^{(m)}) = f(\mathbf{t}_{\alpha_\beta}) + c(y_\beta) = c(y_\beta) + f(\mathbf{t}_{\alpha_\beta}) \tag{4.6}$$

By (4.4), the expression in (4.6) is equal to $x_{\alpha_\beta} + f(\mathbf{t}_{\alpha_\beta})$ and hence, by (4.3), it is also in A. We conclude that $D(m,\vec{F},c;\mathbf{s}_\beta^{(m)})\subset A$. This finishes the induction process and the proof. $\square$

We notice that Theorem 4.23 allows for repeated terms in $D(m, \vec{F}, c; \mathbf{s})$, in other words, one could have $1 \leq i \leq j \leq m$ and $f \in F_i, g \in F_j$ such that

$$f(s_0, \dots, s_{i-1}) + c(s_i) = g(s_0, \dots, s_{j-1}) + c(s_j)$$

In fact, under the same conditions as Theorem 4.23, one may not be able to find $\mathbf{s}$ for which $D(m, \vec{F}, c; \mathbf{s})$ has no repeated terms. However, if one makes the additional assumption that for every $j \in \{1, \dots, m\}$ and every $f, g \in F_j$ the set $\{\mathbf{x} \in G^j : f(\mathbf{x}) = g(\mathbf{x})\}$ is finite, then one can modify the above proof to guarantee the additional property that $D(m, \vec{F}, c; \mathbf{s})$ has no repeated terms.

Indeed, observe that this condition implies that, for every $j \in \{1, \dots, m\}$, the set

$$\{\mathbf{x} \in G^j : (\exists f, g \in F_j) : f(\mathbf{x}) = g(\mathbf{x})\}$$

is finite. Thus, given any IP-set $(\mathbf{x}_\alpha)_{\alpha \in \mathcal{F}}$ in G^j there exists a sub-IP-set $(\mathbf{y}_\beta)_{\beta \in \mathcal{F}}$ such that for all $\beta \in \mathcal{F}$ and $f, g \in F_j$ one has $f(y_\beta) \neq g(y_\beta)$. Only one modification of the proof of Theorem 4.23 is needed to obtain this condition: after choosing the sub-IP-set $(\mathbf{t}_\alpha)$ of $(\mathbf{s}_\alpha^{(m-1)})$ with the property (4.3), pass to a further sub-IP-set $(\mathbf{y}_\beta)$ of $(\mathbf{t}_\alpha)$ with the property that for all $\beta \in \mathcal{F}$ and all $f, g \in F_j$ one has $f(y_\beta) \neq g(y_\beta)$.

The following theorem summarizes the above discussion.

Theorem 4.25. *Let $G, p, c, m, F_1, \dots, F_m, \vec{F}$ be as in Theorem 4.23. Assume that for every $j \in \{1, \dots, m\}$ and every $f, g \in F_j$ the set $\{\mathbf{x} \in G^j : f(\mathbf{x}) = g(\mathbf{x})\}$ is finite. Then for any $A \in p$ there exists an IP-set $(\mathbf{s}_\alpha)_{\alpha \in \mathcal{F}}$ in G^{m+1} such that $D(m, \vec{F}, c; \mathbf{s}_\alpha)$ is contained in A and has no repeated terms. More precisely, for every $\alpha \in \mathcal{F}$ and for all i, j with $1 \leq i \leq j \leq m$ and $f \in F_i, g \in F_j$ we have*

$$f(s_{\alpha,0}, \dots, s_{\alpha,i-1}) + c(s_{\alpha,i}) \neq g(s_{\alpha,0}, \dots, s_{\alpha,j-1}) + c(s_{\alpha,j})$$

We will now deduce Theorems 4.6 and 4.7 from our abstract Theorem 4.23.

Proof of Theorem 4.6. Let G be a countable commutative semigroup and let $A \subset G$ be a central set. Thus, there exists a minimal idempotent $p \in \beta G$ with $A \in p$.

Assume $(m, \vec{F}, c)$ is a shape in G where c is the identity map and that $F_j \subset \mathrm{Hom}(G^j, G)$ for each $j = 1, \ldots, m$. The endomorphism c is trivially IP-regular. For each $j \in \mathbb{N}$ let $\Gamma_j = \mathrm{Hom}(G^j, G)$; it follows from Example 4.18 that each Γ_j is an R-family with respect to p. Finally, by Example 4.20 each Γ_j is licit. We can now apply Theorem 4.23 to find $\mathbf{s} \in G^{m+1}$ with $D(m, \vec{F}, c; \mathbf{s}) \subset A$ as desired. □

Proof of Theorem 4.7. Let G be a group and $(m, \vec{F}, c)$ is a shape in G where c is an endomorphism whose image has finite index in G and, for each $j = 1, \ldots, m$, $F_j \subset \mathbb{P}(G^j, G)$. To see that c is IP-regular, observe that in view of Corollary 3.14 any IP-set has a sub-IP-set contained in the image of c, and that IP-sets carry through homomorphisms. For each $j \in \mathbb{N}$ let $\Gamma_j = \mathbb{P}(G^j, G)$; it follows from Example 4.17 that each Γ_j is an R-family with respect to p. By Example 4.21 each Γ_j is licit. We can now apply Theorem 4.23 to find $\mathbf{s} \in G^{m+1}$ with $D(m, \vec{F}, c; \mathbf{s}) \subset A$ as desired. □

4.3 Proof of partition regularity of $(m, \vec{F}, c)$-sets

In this section we prove Theorems 4.12 and 4.15. Our proof of Theorem 4.15 is inspired by a proof of Deuber's original result presented in [Gun02]. Before we start with the proofs we need a definition.

Definition 4.26. Let G be a countable commutative semigroup, let $(m, \vec{F}, c)$ be a shape in G, let $\mathbf{s} \in G^{m+1}$ and let $k \in \{0, 1, \ldots, m\}$. The *$k$-th line* of the $(m, \vec{F}, c)$-set $D(m, \vec{F}, c; \mathbf{s})$ is the set

$$\{f(s_0, \ldots, s_{k-1}) + c(s_k) : f \in F_k\}$$

Observe that $D(m, \vec{F}, c; \mathbf{s})$ is the union of its $m + 1$ lines.

The proof of Theorem 4.15 goes by induction. Due to its complicated nature it is convenient to isolate the induction step as a separate lemma.

Lemma 4.27. *Let G be a countable commutative semigroup with identity 0, let Λ_c be the clique defined in Theorem 4.15, let $(m, \vec{F}, c) \in \Lambda_c$ and let $r \in \mathbb{N}$. Then there exists a shape $(M, \vec{H}, C) \in \Lambda_c$ such that for any r-coloring of an $(M, \vec{H}, C)$-set such that the last k lines* *are each monochromatic (but different lines can have different colors) there exists a subset which is an* $(m, \vec{F}, c)$*-set whose last* $k+1$ *lines are each monochromatic.*

Moreover, if c *is the identity map, we can take* C *to be the identity map as well, and if* c *is in the center of* $\mathrm{End}(G)$ *we can take* C *to be in the center of* $\mathrm{End}(G)$*.*

Proof. Since any subset of a monochromatic set is monochromatic, we can work with conveniently chosen supersets of the F_i's. Hence we may and will assume that each F_i contains the projection homomorphisms $\pi_j : G^i \to G$ (in each coordinate) and the zero homomorphism. We will also add to each F_i all the homomorphisms of the form

$$\phi(x_0, \ldots, x_{i-1}) = f(x_0, \ldots, x_{j-1}) \qquad \text{with } f \in F_j \text{ and } j < i$$

The main technical tool of our proof is Hales-Jewett's theorem (Theorem 2.38). Let $n = HJ(|F_{m-k}|, r)$ be such that any r-coloring of F_{m-k}^n contains a monochromatic combinatorial line. Since c is concordant with $\vec{F}$, there exists an endomorphism $b : G \to G$ and, for each $f \in F_{m-k}$, there exists $a_f \in \mathrm{Hom}(G^{m-k}, G)$ such that $c \circ a_f = f \circ \mathbf{b}$ (where $\mathbf{b} \in \mathrm{End}(G^{m-k})$ is defined by $\mathbf{b}(x_1, \ldots, x_{m-k}) = b(x_1) + \cdots + b(x_{m-k})$).

For convenience we denote by N the product $N = n(m-k)$ and let $M = N + k$. For each $j = 1, \ldots, M$, let H_j be a finite set of homomorphisms from $G^j \to G$ that will be determined later. Let H_N be the set of all homomorphisms $\phi : G^N \to G$ of the form

$$\phi(t_0, \ldots, t_{N-1}) = \sum_{i=0}^{n-1} f_i \circ \mathbf{b}(t_{i(m-k)}, t_{i(m-k)+1}, \ldots, t_{i(m-k)+m-k-1})$$

with $f_0, \ldots, f_{n-1} \in F_{m-k}$. Finally, make $\vec{H} = (H_1, \ldots, H_M)$ and $C = c \circ b$. Observe that if c is in the center of $\mathrm{End}(G)$, then $b = c$, and hence C is also in the center of $\mathrm{End}(G)$. Moreover, if c is the identity map, then b is also the identity map, and so is C.

Let $t_0, \ldots, t_M \in G$ be arbitrary and let S_H be the $(M, \vec{H}, C)$-set they induce. It will simplify considerably the notation to let

$$T_i := (t_{i(m-k)}, t_{i(m-k)+1}, \ldots, t_{i(m-k)+m-k-1}) \in G^{m-k}$$

for each $i = 0, \ldots, n-1$. Thus, in particular, we can write

$$H_N = \left\{ \phi : (t_0, \ldots, t_{N-1}) \mapsto \sum_{i=0}^{n-1} f_i \circ \mathbf{b}(T_i) : f_0, \ldots, f_{n-1} \in F_{m-k} \right\}$$

Assume that we are given a coloring of S_H into r colors such that each of the last k lines are monochromatic (but not necessarily of the same color).

Color $w = (f_0, \ldots, f_{n-1}) \in F_{m-k}^n$ with the color of

$$\sum_{i=0}^{n-1} f_i \circ \mathbf{b}(T_i) + C(t_N) \tag{4.7}$$

Observe that the elements in (4.7) are in the Nth line of S_H. It follows from the Hales-Jewett theorem that one can find a variable word $w \in (F_{m-k} \cup \{*\})^n$ which induces a monochromatic combinatorial line. We let $\langle n \rangle = \{0, \ldots, n-1\}$, let $A = \{i \in \langle n \rangle : w_i = *\}$ and let $B = \langle n \rangle \setminus A$. Now define

$$u_j = \begin{cases} b(t_{M-m+j}) & \text{if} \quad m-k < j \leq m \\ \sum_{i \in B} a_{w_i}(T_i) + b(t_N) & \text{if} \quad j = m-k \\ \sum_{i \in A} b(t_{i(m-k)+j}) & \text{if} \quad 0 \leq j < m-k \end{cases} \tag{4.8}$$

Note that, for each $\ell = 0, \ldots, m$, the point $u_{m-\ell}$ depend only on $t_0, \ldots, t_{M-\ell}$.

We claim that, with the right choice of $\vec{H}$, the $(m, \vec{F}, c)$-set S_F generated by $u_0, \ldots, u_m$ is a subset of S_H and that each of the last $k+1$ lines of S_F are monochromatic. Indeed, for $m-k < j \leq m$, the j-th line of S_F is the set

$$\{f(u_0, \ldots, u_{j-1}) + c(u_j) : f \in F_j\} = \{f(u_0, \ldots, u_{j-1}) + C(t_{M-m+j}) : f \in F_j\}$$

This will be a subset of the line $M-m+j$ of S_H if we make H_{M-m+j} contain all the homomorphisms ϕ of the form

$$\phi(t_0, \ldots, t_{M-m+j-1}) = f(u_0, \ldots, u_{j-1})$$

for any $f \in F_j$, any possible choice of $A, B \subset \langle n \rangle$ and any $w_i \in F_{m-k}$ (with the u_j's being determined by (4.8)). Hence the j-th line of S_F is monochromatic.

The $(m-k)$-th line of S_F is the set

$$\begin{aligned} & \{f(u_0, \ldots, u_{m-k-1}) + c(u_{m-k}) : f \in F_{m-k}\} \\ = \quad & \left\{ f\left(\sum_{i \in A} \mathbf{b}(T_i)\right) + \sum_{i \in B} c \circ a_{w_i}(T_i) + C(t_N) : f \in F_{m-k} \right\} \\ = \quad & \left\{ \sum_{i \in A} f \circ \mathbf{b}(T_i) + \sum_{i \in B} w_i \circ \mathbf{b}(T_i) + C(t_N) : f \in F_{m-k} \right\} \end{aligned}$$

which is precisely the monochromatic combinatorial line found by applying the Hales-Jewett's theorem. Hence the $(m-k)$-th line of S_F is inside S_H and it is monochromatic.

For $j < m-k$, the j-th line of S_F is the set

$$\begin{aligned}&\{f(u_0,\dots,u_{j-1})+c(u_j) : f\in F_j\}\\ &=\left\{f(u_0,\dots,u_{j-1})+c\left(\sum_{i\in A} b(t_{i(m-k)+j})\right) : f\in F_j\right\}\end{aligned}$$

Let $a=\max A$. Then the j-th line of S_F can be written as

$$\left\{f(u_0,\dots,u_{j-1})+\sum_{i\in A\setminus\{a\}} C(t_{i(m-k)+j})+C(t_{a(m-k)+j}) : f\in F_j\right\}$$

which will be contained in the $a(m-k)+j$-th line of S_H if we make $H_{a(m-k)+j}$ contain all the homomorphisms ϕ of the form

$$\phi(t_0,\dots,t_{a(m-k)+j})=\sum_{i\in A} f(u_0,\dots,u_{j-1})+C(t_{i(m-k)+j})$$

for any $f\in F_j$ and any possible choice of $A\subset\langle a\rangle$, where the dependence of u_i on t_i is given by (4.8).

It is routine to verify that C is concordant with H. This finishes the proof. $\square$

We move now to proving Theorem 4.15.

Proof of Theorem 4.15. If $r=1$ there is nothing to prove so we assume $r>1$. Let $(m,\vec{F},c)\in\Lambda_c$, let $n=m(r-1)$ and, for each $j=1,\dots,n$, let $H_j^{(0)}$ be a finite set of homomorphisms $\phi: G^j\to G$ of the following form. Take $\ell\in\{1,\dots,j\}$ and let $0\le i_1<\dots<i_\ell<j$ be arbitrary. Let $f\in F_\ell$ and define

$$\phi_{f,i_1,\dots,i_\ell}(x_0,\dots,x_{j-1})=f(x_{i_1},x_{i_2},\dots,x_{i_\ell})$$

We let $H_j^{(0)}=\left\{\phi_{f,i_1,\dots,i_\ell}\middle|\ell\in\{1,\dots,j\}, f\in F_\ell, 0\le i_1<\dots<i_\ell<j\right\}$. Now let $\vec{H}^{(0)}=(H_1^{(0)},\dots,H_{m_0}^{(0)})$. Finally, put $c_0=c$ and $m_0=n$.

Applying repeatedly Lemma 4.27, we construct inductively sequences $(m_i)_{i=0}^n$, $(\vec{H}^{(i)})_{i=0}^n$ and $(c_i)_{i=0}^n$, such that the shape $(m_i,\vec{H}^{(i)},c_i)$ satisfies the conclusion of Lemma 4.27 when we input the shape $(m_{i-1},\vec{H}^{(i-1)},c_{i-1})$ and set $k=n-i$.

Let $M = m_n$, $\vec{H} = \vec{H}^{(n)}$ and $C = c_n$. By construction, for any r-coloring of an $(M, \vec{H}, C)$-set S_H we can find a subset which is a $(m_{n-1}, \vec{H}^{(n-1)}, c_{n-1})$-set with the last line monochromatic. Iterating, we obtain for each $i = 0, \ldots, n$, a sub $(m_i, \vec{H}^{(i)}, c_i)$-set with the last $n - i$ lines monochromatic. In particular, setting $i = 0$ we obtain a $(n, \vec{H}^{(0)}, c)$-set with each line monochromatic (but different lines can have different colors).

Let $\mathbf{t} = (t_0, \ldots, t_n)$ be the generator of this $(n, \vec{H}^{(0)}, c)$-set. Applying the pigeonhole principle one can find, among the $n+1$ lines of $D(n, \vec{H}^{(0)}, c; \mathbf{t})$, $m+1$ lines of the same color, say the lines $\ell_0, \ell_2, \ldots, \ell_m$. For each $j = 0, \ldots, m$ let $s_j = t_{\ell_j}$ and let $\mathbf{s} = (s_0, \ldots, s_m)$. By the construction of $H_j^{(0)}$ we deduce that the j-th line of $D(m, \vec{F}, c; \mathbf{s})$ is contained in ℓ_j-th line of $D(n, \vec{H}^{(0)}, c; \mathbf{t})$. Therefore $D(m, \vec{F}, c; \mathbf{s})$ is monochromatic as desired. $\square$

To derive Theorem 4.12 from Theorem 4.15 we need first to establish a lemma.

Definition 4.28. Given two shapes $\lambda_1 = (m_1, \vec{F}^{(1)}, c_1)$ and $\lambda_2 = (m_2, \vec{F}^{(2)}, c_2)$ in a countable commutative semigroup G, we say that λ_1 *contains* λ_2 if for every $\mathbf{s}_1 \in G^{m_1+1}$ there exists $\mathbf{s}_2 \in G^{m_2+1}$ such that

$$D(m_1, \vec{F}^{(1)}, c_1; \mathbf{s}_1) \supset D(m_2, \vec{F}^{(2)}, c_2; \mathbf{s}_2).$$

Lemma 4.29. *Let G be a countable commutative semigroup and let Λ_t be the clique defined in Theorem 4.12. For any two shapes $\lambda_1, \lambda_2 \in \Lambda_t$ there exists some shape $\lambda \in \Lambda$ which contains both λ_1 and λ_2.*

Proof. Let $(m_i, F^{(i)}, c_i) = \lambda_i$ for $i = 1, 2$. Let $c = c_1 \circ c_2 = c_2 \circ c_1$ and let $m = \max(m_1, m_2)$. We can assume that $m_1 = m_2 = m$, putting $F_k^{(i)} = \varnothing$ for $k > m_i$ if necessary. For each $i = 1, 2$ and $n = 1, \ldots, m$, let $\mathbf{c}_i \in \mathrm{End}(G^n)$ be the map $\mathbf{c}_i : (g_0, \ldots, g_{n-1}) \mapsto (c_i(g_0), \ldots, c_i(g_{n-1}))$ and let

$$F_n = \left\{ f \circ \mathbf{c}_2 : f \in F_n^{(1)} \right\} \cup \left\{ f \circ \mathbf{c}_1 : f \in F_n^{(2)} \right\}$$

Let $F = (F_1, \ldots, F_m)$ and let $\lambda = (m, \vec{F}, c)$. Since both c_1 and c_2 are in the center of $\mathrm{End}(G)$, so is c and hence $\lambda \in \Lambda_t$.

Finally, given any $\mathbf{s} \in G^m$ we need to show that $D(m, \vec{F}, c; \mathbf{s})$ contains an $(m_i, F^{(i)}, c_i)$-set for each $i = 1, 2$. Let $\mathbf{s}^{(1)} = \mathbf{c}_2(\mathbf{s}) = (c_2(s_0), \ldots, c_2(s_m))$ and let $\mathbf{s}^{(2)} = \mathbf{c}_1(\mathbf{s}) = (c_1(s_0), c_1(s_1), \ldots, c_1(s_m))$. We claim that

$$D(m_i, F^{(i)}, c_i; \mathbf{s}^{(i)}) \subset D(m, \vec{F}, c; \mathbf{s})$$

Indeed, for any $i = 1, 2$, any $n = 0, 1, \ldots, m$ and any $f \in F_n^{(i)}$ we have

$$c_i(s_n^{(i)}) + f(s_{n-1}^{(i)}, \ldots, s_0^{(i)}) = c_i(c_{3-i}(s_n)) + f(\mathbf{c}_{3-i}(s_{n-1}, \ldots, s_0)) \tag{4.9}$$

Since $c_i \circ c_{3-i} = c$ and for $f \in F_n^{(i)}$ we have $f \circ \mathbf{c}_{3-i} \in F_n$, we deduce that the element (4.9) is in $D(m, \vec{F}, c; \mathbf{s})$ as desired. $\square$

Proof of Theorem 4.12. Let A be a Λ_t-rich set and consider an arbitrary finite partition $A = A_1 \cup \cdots \cup A_r$. Assume none of the A_i is Λ-large. Then for each $i \in \{1, \ldots, r\}$ there exists a shape $\lambda_i \in \Lambda_t$ such that A_i does not contain an $(m, \vec{F}, c)$-set of shape λ_i.

Applying Lemma 4.29 $r - 1$ times, one can find a shape $\lambda \in \Lambda_t$ that contains each of the shapes $\lambda_1, \ldots, \lambda_r$. Therefore, none of the A_i can contain an $(m, \vec{F}, c)$-set of shape λ.

It follows from Theorem 4.15 that there exists a shape $(M, \vec{H}, C) \in \Lambda_t$ such that any partition of an $(M, \vec{H}, C)$-set into r cells contains a $(m, \vec{F}, c)$-set in a single cell. On the one hand, because A was assumed to be Λ_t-large, it will contain an $(M, \vec{H}, C)$-set. On the other hand, this implies that some A_i contains an $(m, \vec{F}, c)$-set, contradicting the construction above. This contradiction implies that some A_i must be Λ_t-large. $\square$

Theorems 4.15 and 4.12 deal only with shapes $(m, \vec{F}, c)$ where each component F_i of $\vec{F}$ is a set of homomorphisms. It is not clear if the methods used to prove them can be adapted to more general cliques, such as those where the F_i are allowed to contain polynomial maps.

4.4 Applications to systems of equations in commutative semigroups

In this section we derive some corollaries of our results that pertain to partition regularity of homogeneous systems of equations. In particular we show that the sufficient condition in Rado's theorem, when appropriately formulated, applies to any countable commutative semigroup. Our departure point is Rado's theorem itself.

Definition 4.30. Let $d, k \in \mathbb{N}$, let A be a $k \times d$ matrix with integer coefficients and let $c_1, \ldots, c_d \in \mathbb{Z}^k$ be the columns of A. We say that A satisfies the *columns condition* if there exist $m \in \mathbb{N}$ and integers $0 = d_0 < d_1 < d_2 < \cdots < d_m < d_{m+1} = d$ such that for every $0 \leq j \leq m$, the sum

$$c_{d_j+1} + c_{d_j+2} + \cdots + c_{d_{j+1}}$$

is in the linear span (over $\mathbb{Q}$) of the set $\{c_i : i \leq d_j\}$ (with the understanding that the only vector in the linear span of the empty set is $\mathbf{0}$).

Theorem 4.31 (Rado [Rad33])**.** *Let $d, k \in \mathbb{N}$ and let A be a $k \times d$ matrix with integer entries. Then for any finite coloring of $\mathbb{N}$ there exists $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{N}^d$ with all coordinates in the same color and $A\mathbf{x} = \mathbf{0}$ if and only if A satisfies the columns condition (possibly after some permutation of the columns of A).*

The 'if' direction of Rado's theorem follows directly from Deuber's Theorem 1.11. The idea is that the columns condition implies the existence of a triple $(m, p, c) \in \mathbb{N}^3$ such that any (m, p, c)-set contains a solution to $A\mathbf{x} = \mathbf{0}$.

More precisely, using the columns condition one can find a $d \times (m + 1)$ matrix B such that $AB = 0$ and, for any $\mathbf{s} \in \mathbb{N}^{m+1}$, the entries of the vector $B\mathbf{s}$ are contained in the (m, p, c)-set $D(m, p, c; \mathbf{s})$ for some $c, p \in \mathbb{N}$ that only depend on A. Then, for any finite coloring of $\mathbb{N}$ one can find $\mathbf{s} \in \mathbb{N}^{m+1}$ such that $D(m, p, c; \mathbf{s})$ is monochromatic, and in particular, all coordinates of $B\mathbf{s}$ are monochromatic. Since $AB = 0$, also $A(B\mathbf{s}) = 0$. The details of this deduction can be found, for instance, in [GRS90].

We now turn to linear systems of equations in countable commutative semigroups and establish an analogue of the columns condition in this setting.

Definition 4.32. Let G be a countable commutative semigroup with identity 0, let $k, d \in \mathbb{N}$ and let $A : G^d \to G^k$ be a homomorphism. For each $i = 1, \ldots, d$ let $c_i : G \to G^k$ be the map defined by $c_i(x) = A(0, \ldots, 0, x, 0, \ldots, 0)$,where the x appears in the i-th position.

We say that A satisfies the *columns condition* if there exist $c \in \mathrm{End}(G)$, $m \in \mathbb{N}$ and $0 = d_0 < d_1 < \cdots < d_{m+1} = d$ such that

1. The composition $(c_1 + c_2 + \cdots + c_{d_1}) \circ c$ is the zero map;
2. For each $1 \leq t \leq m$ there are $f_1^{(t)}, \ldots, f_{d_t}^{(t)} \in \mathrm{End}(G)$ such that

$$(c_{d_t+1} + \cdots + c_{d_{t+1}}) \circ c + \left(c_1 \circ f_1^{(t)} + \cdots + c_{d_t} \circ f_{d_t}^{(t)}\right) = 0 \tag{4.10}$$

This definition can be seen as a direct extension of Definition 4.30. Indeed, when $G = \mathbb{Z}$, the only homomorphisms are multiplication by a fixed integer and equation (4.10) expresses the fact the sum $c_{d_t+1} + \cdots + c_{d_{t+1}}$ is a linear combination of $c_1, \ldots, c_{d_t}$.

The next proposition is an extension of the 'if' part of Rado's theorem to countable commutative semigroups.

Proposition 4.33. *Let G be a countable commutative semigroup with identity 0, let $k, d \in \mathbb{N}$ and let $A : G^d \to G^k$ be a homomorphism which satisfies the columns condition for some $c \in \mathrm{End}(G)$ that is either in the center of $\mathrm{End}(G)$ or is IP-regular. Then for any finite coloring of G there exists $\mathbf{x} = (x_1, \ldots, x_d)$ with all entries in the same color such that $A(\mathbf{x}) = \mathbf{0}$.*

Proof. Let $m \in \mathbb{N}$ and $c \in \mathrm{End}(G)$ be given by the columns condition. For each $j = 1, \ldots, m$ let

$$F_j = \left\{ f : (s_0, \ldots, s_{j-1}) \mapsto \sum_{\ell=0}^{j-1} f_i^{(m-\ell)}(s_\ell) : d_{m-j} < i \leq d_{m+1-j} \right\}$$

and let $\vec{F} = (F_1, \ldots, F_m)$. Assume we are given a finite coloring of G. Appealing to either Theorem 4.15 or Theorem 4.23 (according to whether c is in the center of $\mathrm{End}(G)$ or IP-regular) we can find $\mathbf{s} \in G^{m+1}$ such that the $(m, \vec{F}, c)$-set $D(m, \vec{F}, c; \mathbf{s})$ is monochromatic. For each $i = 1, \ldots, d$, let $j \in \{0, \ldots, m\}$ be such that $d_{m-j} < i \leq d_{m+1-j}$ (and observe that j is uniquely determined). Let

$$x_i = \sum_{\ell=0}^{j-1} f_i^{(m-\ell)}(s_\ell) + c(s_j)$$

Observe that $x_i \in D(m, \vec{F}, c; \mathbf{s})$ and hence all the entries of the vector $\mathbf{x} = (x_1, \ldots, x_d) \in G^d$ are of the same color. Finally we need to check that $A(\mathbf{x}) = \mathbf{0}$. Let $c_1, \ldots, c_d$ be as in Definition 4.32 and observe that each $c_i : G \to G^k$ is a homomorphism. We have

$$
\begin{aligned}
A(\mathbf{x}) &= \sum_{i=1}^{d} c_i(x_i) = \sum_{j=0}^{m} \sum_{i=d_{m-j}+1}^{d_{m-j+1}} c_i(x_i) \\
&= \sum_{j=0}^{m} \sum_{i=d_{m-j}+1}^{d_{m-j+1}} c_i \left(\sum_{\ell=0}^{j-1} f_i^{(m-\ell)}(s_\ell) + c(s_j) \right) \\
&= \sum_{j=0}^{m} \sum_{i=d_{m-j}+1}^{d_{m-j+1}} (c_i \circ c)(s_j) + \sum_{j=0}^{m} \sum_{i=d_{m-j}+1}^{d_{m-j+1}} \sum_{\ell=0}^{j-1} (c_i \circ f_i^{(m-\ell)})(s_\ell) \\
&= \sum_{\ell=0}^{m} \left[\sum_{i=d_{m-\ell}+1}^{d_{m-\ell+1}} (c_i \circ c) + \sum_{j=0}^{m-\ell-1} \sum_{i=d_{m-j}+1}^{d_{m-j+1}} (c_i \circ f_i^{(m-\ell)}) \right] (s_\ell) \\
&= \sum_{t=0}^{m} \left[\left(\sum_{i=d_t+1}^{d_{t+1}} c_i \right) \circ c + \sum_{i=1}^{d_t} c_i \circ f_i^{(t)} \right] (s_\ell) \\
&= 0
\end{aligned}
$$

where the last equality follows from the columns conditions. □

While Proposition 4.33 provides a quite satisfactory extension of the sufficient condition in Rado's theorem to a general seting, it is not even clear how to formulate the necessary condition.

Problem 4.34. *Let G be a countable commutative cancelative semigroup, let $k, d \in \mathbb{N}$ and let $A : G^d \to G^k$ be a homomorphism. Give necessary and sufficient conditions for A so that for any finite partition of G there exists a non-zero $\mathbf{x} = (x_1, \ldots, x_d)$ with all entries in the same cell of the partition and such that $A(\mathbf{x}) = \mathbf{0}$.*

We conclude by remarking that an analogue of the columns condition can be concocted for polynomial equations in such a way that an analogue of Proposition 4.33 holds, but the condition is cumbersome and so it appears to be of little practical value.

CHAPTER 5

PATTERNS $\{x+y, xy\}$ IN LARGE SETS OF COUNTABLE FIELDS

In this chapter we present work from [BM16a] and [BM16a] on the presence (and abundance) of $\{x+y, xy\}$ patterns in large subsets of countable fields.

5.1 Introduction

Theorem 4.2, stated in the previous section, is a polynomial extension of Deuber's theorem, which in turn characterizes the linear Ramsey families in $\mathbb{N}$. However, as mentioned above, not every polynomial Ramsey family is included in the $(m, \vec{F}, c)$-sets which appear in Theorem 4.2. One reason for this is that Theorem 4.2 was obtained as the joint extension of the polynomial van der Waerden theorem and the (linear) Folkman's theorem. Recalling that the classical Deuber's theorem is itself the outcome of combining the (linear) van der Waerden's theorem with Folkman's theorem, perhaps in order to obtain every polynomial Ramsey family one has to combine the polynomial van der Waerden's theorem with a polynomial version of Folkman's theorem.

Unfortunately, it is not clear what a polynomial version of Folkman's theorem would look like. One option was suggested in the book [GRS90] of Graham, Rothschild and Spencer in the late 1970's:

Conjecture 5.1 ([GRS90, end of Section 3.4])**.** *For every finite coloring of* $\mathbb{N}$ *and every* $m \in \mathbb{N}$ *there exists a color* C *and a set* $A \subset \mathbb{N}$ *with* $|A| = m$ *such that* $FS(A) \cup FP(A) \subset C$*, where we denote by* $FP(A)$ *the* finite product set*, i.e., the analogue of* $FS(A)$ *for the multiplicative semigroup structure* $(\mathbb{N}, \times)$*.*

We can state Conjecture 5.1 in the language of Ramsey families as saying that for every $m \in \mathbb{N}$ the family $\{S_A : \varnothing \neq A \subset \{1, \ldots, m\}\} \cup \{P_A : \varnothing \neq A \subset \{1, \ldots, m\}\}$ is Ramsey in $\mathbb{N}$, where for any non-empty $A \subset \{1, \ldots, m\}$ the functions $S_A, P_A : \mathbb{N}^m \to \mathbb{N}$ are defined as

$$S_A : (x_1, \ldots, x_m) \mapsto \sum_{i \in A} x_i \qquad \text{and} \qquad P_A : (x_1, \ldots, x_m) \mapsto \prod_{i \in A} x_i$$

The case $m = 2$ of Conjecture 5.1 is precisely Conjecture 1.4 from the introduction and corresponds to a polynomial version of Schur's theorem. As even the case $m = 2$ is still open, it is possible that Conjecture 5.1 will be very hard to solve. On the other hand, Conjecture 5.1 is only a special case of Problem 1.12.

Very recently we have established a weaker version of Conjecture 1.4 (see Theorem 1.5). In this chapter we study certain density analogues of this kind of questions in countable fields and certain subrings. Besides its intrinsic interest, the ideas and techniques developed in this chapter also paved the way for the proof of Theorem 1.5 in [Mor] (presented in the next chapter).

$\{x + y, xy\}$ patterns in large subsets of countable fields

Not every set of positive (additive) upper density in $\mathbb{N}$ contains a configuration $\{x+y, xy\}$. Indeed no such pair can consist solely of odd numbers and in fact there are sets $A \subset \mathbb{N}$ which are simultaneously thick with respect to addition and to multiplication (hence having density one with respect to suitable Følner sequences), and yet contain no non-trivial configurations $\{x + y, xy\}$ (see Theorem 5.35 below). However, as we saw in Chapter 3, the affine structures of $\mathbb{N}$ and of $\mathbb{Q}$ are significantly different: the affine semigroup $\mathcal{A}_{\mathbb{N}}$ of $\mathbb{N}$ is non-amenable, while $\mathcal{A}_{\mathbb{Q}}$ is a solvable (hence amenable) group. In particular, $\mathbb{Q}$ possesses double Følner sequences (see Proposition 3.6) and hence there are finitely additive probability measures on all subsets of $\mathbb{Q}$ which are invariant under any affine transformation.

We show that any large set in $\mathbb{Q}$ with respect to an affinely invariant mean contains the sought-after configurations, which leads to a partition result involving three-element sets having the form $\{x, y + x, yx\}$. In fact, the ergodic method that we employ works equally well in the framework of arbitrary countable fields.

Theorem 5.2. *Let K be a countable field, let $(F_N)_{N\in\mathbb{N}}$ be a double Følner sequence in K and let $E\subset K$ be such that $\bar{d}_{(F_N)}(E)>0$. Then there are infinitely many pairs $x,y\in K^*$ with $x\neq y$ such that*

$$\{x+y,xy\}\subset E \tag{5.1}$$

A precise formulation of how large is the set of pairs (x,y) that satisfy equation (5.1) is given by Theorem 5.10 below. In view of the correspondence principle (Theorem 3.24 in Chapter 3), Theorem 5.2 can be easily derived from the following ergodic result, which can be seen as an affine version of Khintchine's recurrence theorem.

Corollary 5.3. *Let K be an infinite countable field, let $\big(X,\mathcal{B},\mu,(T_g)_{g\in\mathcal{A}_K}\big)$ be a probability measure preserving system and let $B\in\mathcal{B}$. Then, for any $\delta\in(0,1)$, the set*

$$\mathcal{R}(B,\delta):=\Big\{u\in K:\mu(T_{M_u}^{-1}B\cap T_{A_u}^{-1}B)>\delta\mu(B)^2\Big\} \tag{5.2}$$

is affinely syndetic.

It is not hard to see that the quantity $\mu(B)^2$ in (5.2) is the largest possible (consider for example the case when the action of $\mathcal{A}_K$ is strongly mixing).

Corollary 5.3 is in turn derived from the following analogue of von Neumann's mean ergodic theorem:

Theorem 5.4. *Let K be an infinite countable field, let $(U_g)_{g\in\mathcal{A}_K}$ be a unitary representation of $\mathcal{A}_K$ on a Hilbert space $\mathcal{H}$, let $I=\{f\in\mathcal{H}:(\forall g\in\mathcal{A}_K)\ U_gf=f\}$ be the invariant subspace and let $P:\mathcal{H}\to I$ be the orthogonal projection onto I. Then for any $f\in\mathcal{H}$ and any double Følner sequence $(F_N)_{N\in\mathbb{N}}$ in K we have*

$$\lim_{N\to\infty}\frac{1}{|F_N|}\sum_{u\in F_N}U_{M_uA_{-u}}f=Pf \tag{5.3}$$

Employing a method developed by Bergelson in [Ber86] we can then quickly deduce from Corollary 5.3 the following partition statement:

Theorem 5.5 (see Theorem 5.18 for a more precise formulation)**.** *Let K be a countable field. Given a finite coloring $K=\bigcup C_i$, there exists a color C_i and $x,y\in K$ such that $\{x,x+y,xy\}\subset C_i$.*

We remark that it follows from Theorem 5.18 below that x and y can in fact be chosen from outside any prescribed finite set.

When trying to obtain analogues of Corollary 5.3 or Theorem 5.4 for $\mathbb{Z}$ or other rings one runs into serious difficulties, the main problem being the lack of amenability of the affine semigroup. Therefore it is a priori not clear what kind of statement similar to Theorem 5.4 can be formulated (and proved) if one replaces fields by more general rings. In particular, one would like to know if the corresponding set $\mathcal{R}(B,\delta)$ from (5.2) is non-empty (or indeed affinely syndetic) for any measure preserving action of the affine semigroup $\mathcal{A}_{\mathbb{Z}}$ of $\mathbb{Z}$.

We developed an alternative approach to obtaining Corollary 5.3 which does not rely on the existence of a double Følner sequence. This approach, based on convergence along ultrafilters, not only allows one to a have reasonable analogue of Theorem 5.4 to actions of affine semigroups of general LID's, but also leads to a strong generalization of Corollary 5.3 which guarantees that the sets $\mathcal{R}(B,\delta)$ are not only affinely syndetic but actually possess the filter property.

Theorem 5.6. *Let R be an LID, let $t \in \mathbb{N}$ and, for each $i = 1, ..., t$, let (Ω_i, μ_i) be a probability space, let $(T_g^{(i)})_{g\in\mathcal{A}_R}$ be a measure preserving action of the affine semigroup $\mathcal{A}_R$ of K on (Ω_i, μ_i) and let $B_i \subset \Omega_i$ be a measurable set with positive measure. Let $\delta \in (0, 1)$ and let $\mathcal{R}(B_i, \delta)$ be defined as in equation (5.2) with respect to the action $(T_g^{(i)})_{g\in\mathcal{A}_R}$. Then the intersection*

$$\mathcal{R}(B_1,\delta) \cap ... \cap \mathcal{R}(B_t,\delta) \tag{5.4}$$

is affinely syndetic (and, in particular, nonempty).

Observe that, in general, affinely syndetic sets do not have the finite intersection property. For example, the subsets of rational numbers defined by

$$E_1 = \bigcup_{n\in\mathbb{Z}} [2n, 2n+1) \subset \mathbb{Q} \qquad\qquad E_2 = \bigcup_{n\in\mathbb{Z}} [2n-1, 2n) \subset \mathbb{Q}$$

are both additively (hence affinely) syndetic, but have empty intersection.

Remark 5.7. *To appreciate the power of the ultrafilter approach, one should note that the Cesàro convergence results established in [BM16a] imply only the affine syndeticity of the* *intersections*

$$\mathcal{R}(B_1, 0) \cap ... \cap \mathcal{R}(B_t, 0) \tag{5.5}$$

of return sets $\mathcal{R}(B_i, 0)$*, rather than the affine syndeticity of the intersection of the 'optimal' return sets* $\mathcal{R}(B_i, \delta)$*, as in* (5.4).

The filter property achieved in Theorem 5.6 is deduced from the fact that the sets of the form $\mathcal{R}(B, \delta)$ are DC* (see Definition 3.11):

Theorem 5.8. *Let* R *be an LID, let* (Ω, μ) *be a probability space, let* $(T_g)_{g \in \mathcal{A}_R}$ *be a measure preserving action of* $\mathcal{A}_R$ *on* Ω*, let* $B \subset \Omega$ *be a measurable set and let* $\varepsilon > 0$*. Then the set*

$$\left\{u \in R : \mu(A_u^{-1} B \cap M_u^{-1} B) > \mu(B)^2 - \varepsilon\right\}$$

is DC and, in particular, affinely syndetic.*

This in turn allows us to obtain, as a corollary, the following analogue of formula (5.3) for measure preserving actions $(T_g)_{g \in \mathcal{A}_R}$ of the affine semigroup of an LID and a ultrafilter $p \in \mathcal{G}$ (see Theorem 5.19 below for a more precise formulation):

$$p\text{-}\lim_u \mu(T_{A_u}^{-1} B \cap T_{M_u}^{-1} B) \geq \mu(B)^2$$

In view of Theorem 5.8 we can now extend Theorem 5.2 to a more general setting, in a sense halfway towards a proof that the family $\{x + y, xy\}$ is Ramsey in $\mathbb{N}$. The following theorem lists some special cases of a more general Theorem 5.32, to be found in Chapter 5:

Theorem 5.9.

1. *For any finite partition* $\mathbb{Q} = C_1 \cup \cdots \cup C_r$ *of the rational numbers, there exists a cell* $i \in \{1, \ldots, r\}$ *and many*[1] $x \in \mathbb{Q}$, $n \in \mathbb{N}$ *such that* $\{x + n, xn\} \subset C_i$.
2. *More generally, if* K *is a number field and* $\mathcal{O}_K$ *is its ring of integers, for any finite partition* $K = C_1 \cup \cdots \cup C_r$*, there exists a cell* $i \in \{1, \ldots, r\}$ *and many* $x \in K$, $n \in \mathcal{O}_K$ *such that* $\{x + n, xn\} \subset C_i$.

[1] In Theorem 5.32 we describe more precisely how large is the set of such x and n.

3. *Let $\vec{F}$ be a finite field, let K denote the field of rational functions (i.e. quotients of polynomials) over $\vec{F}$ and let $\vec{F}[x]$ denote the ring of polynomials. Then for any finite partition $K = C_1 \cup \cdots \cup C_r$, there exists a cell $i \in \{1, \ldots, r\}$ and many $f \in K$, $g \in \vec{F}[x]$ such that $\{f+g, fg\} \subset C_i$.*

5.2 An affine Khintchine theorem

The following is a more precise version of Theorem 5.2:

Theorem 5.10. *Let K be a countable field and let $E \subset K$ be such that $\bar{d}_{(L_N)}(E) > 0$ for some double Følner sequence $(L_N)_{N\in\mathbb{N}}$. Then for each $\varepsilon > 0$ there is a set $D \subset K^*$ such that for every double Følner sequence $(F_N)_{N\in\mathbb{N}}$*

$$\underline{d}_{(F_N)}(D) \geq \frac{\varepsilon}{\varepsilon + \bar{d}_{(L_N)}(E) - \bar{d}_{(L_N)}(E)^2}$$

and for all $u \in D$ we have

$$\bar{d}_{(L_N)}\big((E-u) \cap (E/u)\big) > \bar{d}_{(L_N)}(E)^2 - \varepsilon$$

In this section we derive Theorem 5.10 from ergodic results concerning actions of the affine groups $\mathcal{A}_K$ of a groups. The proof of the ergodic results will be given in Section 5.3. We begin with the following analogue of the mean ergodic theorem for affine actions.

Theorem 5.11. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and suppose that $\mathcal{A}_K$ acts on Ω by measure preserving transformations. Let (F_N) be a double Følner sequence on K. Then for each $B \in \mathcal{B}$ we have*[2]

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}B) \geq \mu(B)^2$$

and, in particular the limit exists.

In the case when the action of $\mathcal{A}_K$ is ergodic, we can replace one of the sets B with a potentially distinct set C. This is the content of the next theorem.

[2]By slight abuse of language we use the same symbol to denote the elements (such as $M_{1/u}$ and A_{-u}) of $\mathcal{A}_K$ and the measure preserving transformation they induce on Ω.

Theorem 5.12. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and suppose that $\mathcal{A}_K$ acts ergodically on Ω by measure preserving transformations. Let (F_N) be a double Følner sequence on K. Then for any $B, C \in \mathcal{B}$ we have*

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}C) = \mu(B)\mu(C)$$

and, in particular, the limit exists.

Remark 5.13. *We note that Theorem 5.12 fails without ergodicity. Indeed, take the normalized disjoint union of two copies of the same measure preserving system. Choosing B to be one of the copies and C the other we get*

$$A_{-u}B \cap M_{1/u}C = \varnothing \text{ for all } u \in K^*$$

We can extract some quantitative bounds from Theorems 5.11 and 5.12. This is summarized in the next corollary (which is an enhanced version of Corollary 5.3).

Corollary 5.14. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and suppose that $\mathcal{A}_K$ acts on Ω by measure preserving transformations. Let (F_N) be a double Følner sequence on K, let $B \in \mathcal{B}$ and let $\varepsilon > 0$. Then we have*

$$\underline{d}_{(F_N)}\left(\left\{u \in K^* : \mu(A_{-u}B \cap M_{1/u}B) > \mu(B)^2 - \varepsilon\right\}\right) \geq \frac{\varepsilon}{\varepsilon + \mu(B) - \mu(B)^2}$$

Moreover, if the action of $\mathcal{A}_K$ is ergodic and $B, C \in \mathcal{B}$, the set

$$D_\varepsilon := \left\{u \in K^* : \mu(A_{-u}B \cap M_{1/u}C) > \mu(B)\mu(C) - \varepsilon\right\}$$

satisfies

$$\underline{d}_{(F_N)}(D_\varepsilon) \geq \max\left(\frac{\varepsilon}{\varepsilon + \mu(B)(1-\mu(C))}, \frac{\varepsilon}{\varepsilon + \mu(C)(1-\mu(B))}\right) \tag{5.6}$$

Corollary 5.14 will be proved in Section 5.3. We will use it now, together with the correspondence principle, to deduce Theorem 5.10.

Proof of Theorem 5.10. Let $X = K$, let $G = \mathcal{A}_K$ and let $(G_N) = (L_N)$. Applying the correspondence principle (Theorem 3.24), we obtain, for each $E \subset K$, a measure preserving action $(T_g)_{g\in\mathcal{A}_K}$ of $\mathcal{A}_K$ on a probability space $(\Omega,\mathcal{B},\mu)$, a set $B\in\mathcal{B}$ such that $\mu(B)=\bar{d}_{(L_N)}(E)$, and for all $u\in K^*$ we have $\bar{d}_{(F_L)}(A_{-u}E\cap M_{1/u}E)\geq\mu(T_{A_{-u}}B\cap T_{M_{1/u}}B)$. To simplify notation we will denote the measure preserving transformations $T_{A_{-u}}$ and $T_{M_{1/u}}$ on Ω by just A_{-u} and $M_{1/u}$. Also, recalling that $A_{-u}E=E-u$ and $M_{1/u}E=E/u$ we can rewrite the previous equation as

$$\bar{d}_{(L_N)}(E-u\cap E/u)\geq\mu(A_{-u}B\cap M_{1/u}B)\qquad\qquad\forall u\in K^*$$

Now assume that $\bar{d}_{(L_N)}(E)>0$ and let $\varepsilon>0$. Let

$$D_\varepsilon:=\{u\in K^*:\bar{d}_{(L_N)}((E-u)\cap(E/u))>\bar{d}_{(L_N)}(E)^2-\varepsilon\}$$

By Corollary 5.14 we have, for any double Følner sequence $(F_N)_{N\in\mathbb{N}}$,

$$\begin{aligned}\underline{d}_{(F_N)}(D_\varepsilon) &\geq \underline{d}_{(F_N)}\left(\left\{u\in K^*:\mu(A_{-u}B\cap M_{1/u}B)>\mu(B)^2-\varepsilon\right\}\right)\\ &\geq \frac{\varepsilon}{\varepsilon+\mu(B)-\mu(B)^2}\\ &= \frac{\varepsilon}{\varepsilon+\bar{d}_{(L_N)}(E)-\bar{d}_{(L_N)}(E)^2}\end{aligned}$$

□

5.3 Proof of the affine ergodic theorem

In this section we will prove Theorems 5.11 and 5.12 and Corollary 5.14. Throughout this section let K be a countable field, let $(\Omega,\mathcal{B},\mu)$ be a probability space, let $(T_g)_{g\in\mathcal{A}_K}$ be a measure preserving action of $\mathcal{A}_K$ on Ω and let (F_N) be a double Følner sequence on K.

Let $H=L^2(\Omega,\mu)$ and let $(U_g)_{g\in\mathcal{A}_K}$ be the unitary Koopman representation of $\mathcal{A}_K$ (this means that $(U_gf)(x)=f(g^{-1}x)$). By a slight abuse of notation we will write A_uf instead of $U_{A_u}f$ and M_uf instead of $U_{M_u}f$.

Let P_A be the orthogonal projection from H onto the subspace of vectors which are fixed under the action of the additive subgroup S_A and let P_M be the orthogonal projection from H onto the subspace of vectors which are fixed under the action of the multiplicative subgroup S_M.

We will show that the orthogonal projections P_A and P_M commute, which can be surprising considering that the subgroups S_A and S_M do not. The reason for this is that for each $k \in K^*$, the map $M_k : K \to K$ is an isomorphism of the additive group.

Lemma 5.15. *For any $f \in H$ we have*

$$P_A P_M f = P_M P_A f$$

Proof. We first prove that for any $k \in K^*$, the projection P_A commutes with M_k. For this we will use Theorem 2.50, Lemma 3.8 and equation (3.1):

$$\begin{aligned} M_k P_A f &= M_k \left(\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} A_u f \right) = \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_k A_u f \\ &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} A_{ku} M_k f = \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in kF_N} A_u M_k f = P_A M_k f \end{aligned}$$

Now we can conclude the result:

$$P_M P_A f = \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u P_A f = P_A \left(\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u f \right) = P_A P_M f$$

□

Lemma 5.15 implies that $P_M P_A f$ is invariant under both S_A and S_M. Since those two subgroups generate $\mathcal{A}_K$, this means that $P_M P_A$ is the orthogonal projection onto the space of functions invariant under $\mathcal{A}_K$.

Let $P : H \to H$ be the orthogonal projection onto the space of functions invariant under the action of the group $\mathcal{A}_K$. We have $P = P_A P_M = P_M P_A$.

We can now prove the Hilbert space version of the ergodic theorem:

Theorem 5.4. *Let K be an infinite countable field, let $(U_g)_{g\in\mathcal{A}_K}$ be a unitary representation of $\mathcal{A}_K$ on a Hilbert space $\mathcal{H}$, let $I = \{f \in \mathcal{H} : (\forall g \in \mathcal{A}_K)\ U_g f = f\}$ be the invariant subspace and let $P : \mathcal{H} \to I$ be the orthogonal projection onto I. Then for any $f \in \mathcal{H}$ and any double Følner sequence $(F_N)_{N\in\mathbb{N}}$ in K we have*

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} U_{M_u A_{-u}} f = Pf$$

In particular, the limit exists.

Proof. We assume first that $P_A f = 0$. For $u \in K^*$, let $a_u = M_u A_{-u} f$. Then for each $b \in K^*$ we have

$$\begin{aligned}
\langle a_{ub}, a_u \rangle &= \langle M_{ub} A_{-ub} f, M_u A_{-u} f \rangle \\
&= \langle M_b A_{-ub} f, A_{-u} f \rangle \\
&= \langle A_{-ub} f, M_{1/b} A_{-u} f \rangle \\
&= \langle A_{-ub} f, A_{-u/b} M_{1/b} f \rangle \\
&= \langle A_{-u(b-1/b)} f, M_{1/b} f \rangle
\end{aligned}$$

where we used equation (3.1) and the fact that the operators are unitary. Now if $b \neq \pm 1$ then $b - \frac{1}{b} = \frac{b^2-1}{b} \neq 0$ and so the sequence of sets $\left(-\frac{b^2-1}{b} F_N\right)_N$ is again a double Følner sequence on K, by Lemma 3.8. Thus, applying Theorem 2.50 we get (keeping $b \neq \pm 1$ fixed)

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \langle a_{ub}, a_u \rangle &= \left\langle \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} A_{-u(b-1/b)} f, M_{1/b} f \right\rangle \\
&= \left\langle \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in -\frac{b^2-1}{b} F_N} A_u f, M_{1/b} f \right\rangle \\
&= \langle P_A f, M_{1/b} f \rangle = 0
\end{aligned}$$

Thus it follows from Proposition 2.53 that

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u A_{-u} f = 0$$

Now, for a general $f \in H$, we can write $f = f_1 + f_2$, where $f_1 = P_A f$ and $f_2 = f - P_A f$ satisfies $P_A f_2 = 0$. Note that f_1 is invariant under A_u. Therefore

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u A_{-u} f &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u A_{-u} f_1 \\
&= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} M_u f_1 \\
&= P_M f_1 = P_M P_A f = Pf
\end{aligned}$$

□

Remark 5.16. *Theorem 5.4 can be interpreted as an ergodic theorem along a sparse subset of $\mathcal{A}_K$ (namely the subset $\{M_u A_{-u} : u \in K^*\}$).*

Proof of Theorem 5.11. Let $B \in \mathcal{B}$. By Theorem 5.4 applied to the characteristic function 1_B of B we get that

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}B) &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \int_\Omega A_{-u}1_B M_{1/u}1_B d\mu \\
&= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \int_\Omega (M_u A_{-u}1_B)1_B d\mu \\
&= \int_\Omega (P1_B)1_B d\mu
\end{aligned}$$

We can use the Cauchy-Schwartz inequality with the functions $P1_B$ and the constant function 1, and the trivial observation that $P1 = 1$, to get

$$\int_\Omega (P1_B)1_B d\mu = \|P1_B\|^2 \geq \langle P1_B, 1\rangle^2 = \langle 1_B, P1\rangle^2 = \langle 1_B, 1\rangle^2 = \mu(B)^2$$

Putting everything together we obtain

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}B) \geq \mu(B)^2$$

□

Proof of Theorem 5.12. Let $B, C \in \mathcal{B}$. By Theorem 5.4 applied to the characteristic function 1_B of B we get that

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}C) &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \int_\Omega A_{-u}1_B M_{1/u}1_C d\mu \\
&= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \int_\Omega (M_u A_{-u}1_B)1_C d\mu \\
&= \int_\Omega (P1_B)1_C d\mu
\end{aligned}$$

Since the action of $\mathcal{A}_K$ is ergodic, $P1_B = \mu(B)$, and hence

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}C) = \mu(B) \int_\Omega 1_C d\mu = \mu(B)\mu(C)$$

□

Proof of Corollary 5.14. Let $B, C \subset \mathcal{B}$ be arbitrary and observe that

$$\mu(A_{-u}B \cap M_{1/u}B) \leq \mu(M_{1/u}B) = \mu(B).$$

For each $\varepsilon > 0$ let D_ε be the set $D_\varepsilon := \{u \in K : \mu(A_{-u}B \cap M_{1/u}B) > \mu(B)^2 - \varepsilon\}$.

Now let $(\tilde{F}_N)_{N\in\mathbb{N}}$ be a subsequence of $(F_N)_{N\in\mathbb{N}}$ such that

$$\underline{d}_{(F_N)}(D_\varepsilon) = \lim_{N\to\infty} \frac{|D_\varepsilon \cap \tilde{F}_N|}{|\tilde{F}_N|}$$

Thus $\underline{d}_{(F_N)}(D_\varepsilon) = \underline{d}_{(\tilde{F}_N)}(D_\varepsilon) = \bar{d}_{(\tilde{F}_N)}(D_\varepsilon)$. By Theorem 5.11, we now have

$$\begin{aligned}
\mu(B)^2 &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}B) \\
&= \lim_{N\to\infty} \frac{1}{|\tilde{F}_N|} \sum_{u\in \tilde{F}_N} \mu(A_{-u}B \cap M_{1/u}B) \\
&= \lim_{N\to\infty} \frac{1}{|\tilde{F}_N|} \left(\sum_{u\in \tilde{F}_N \cap D_\varepsilon} \mu(A_{-u}B \cap M_{1/u}B) + \sum_{u\in \tilde{F}_N \setminus D_\varepsilon} \mu(A_{-u}B \cap M_{1/u}B) \right) \\
&\leq \mu(B)\bar{d}_{(\tilde{F}_N)}(D_\varepsilon) + \big(\mu(B)^2 - \varepsilon\big)\big(1 - \underline{d}_{(\tilde{F}_N)}(D_\varepsilon)\big) \\
&= \mu(B)\underline{d}_{(F_N)}(D_\varepsilon) + \big(\mu(B)^2 - \varepsilon\big)\big(1 - \underline{d}_{(F_N)}(D_\varepsilon)\big)
\end{aligned}$$

From this we conclude that $\underline{d}_{(F_N)}(D_\varepsilon) \geq \varepsilon / \Big(\varepsilon + \mu(B)\big(1 - \mu(B)\big)\Big)$.

Now assume that the action of $\mathcal{A}_K$ is ergodic. Note that trivially $\mu(A_{-u}B \cap M_{1/u}C) \leq \mu(M_{1/u}C) = \mu(C)$. For each $\varepsilon > 0$ let D_ε be the set $D_\varepsilon := \{u \in K : \mu(A_{-u}B \cap M_{1/u}C) > \mu(B)\mu(C) - \varepsilon\}$.

Now let $(\tilde{F}_N)_{N\in\mathbb{N}}$ be a subsequence of $(F_N)_{N\in\mathbb{N}}$ such that

$$\underline{d}_{(F_N)}(D_\varepsilon) = \lim_{N\to\infty} \frac{|D_\varepsilon \cap \tilde{F}_N|}{|\tilde{F}_N|}$$

Thus $\underline{d}_{(F_N)}(D_\varepsilon) = \underline{d}_{(\tilde{F}_N)}(D_\varepsilon) = \bar{d}_{(\tilde{F}_N)}(D_\varepsilon)$. By Theorem 5.12 we now have

$$\begin{aligned}
\mu(B)\mu(C) &= \lim_{N\to\infty} \frac{1}{|F_N|} \sum_{u\in F_N} \mu(A_{-u}B \cap M_{1/u}C) \\
&= \lim_{N\to\infty} \frac{1}{|\tilde{F}_N|} \sum_{u\in \tilde{F}_N} \mu(A_{-u}B \cap M_{1/u}C) \\
&= \lim_{N\to\infty} \frac{1}{|\tilde{F}_N|} \left(\sum_{u\in \tilde{F}_N \cap D_\varepsilon} \mu(A_{-u}B \cap M_{1/u}C) + \sum_{u\in \tilde{F}_N \setminus D_\varepsilon} \mu(A_{-u}B \cap M_{1/u}C) \right) \\
&\leq \mu(C)\bar{d}_{(\tilde{F}_N)}(D_\varepsilon) + \big(\mu(B)\mu(C) - \varepsilon\big)\big(1 - \underline{d}_{(\tilde{F}_N)}(D_\varepsilon)\big) \\
&= \mu(C)\underline{d}_{(F_N)}(D_\varepsilon) + \big(\mu(B)\mu(C) - \varepsilon\big)\big(1 - \underline{d}_{(F_N)}(D_\varepsilon)\big)
\end{aligned}$$

From this we conclude that $\underline{d}_{(F_N)}(D_\varepsilon) \geq \varepsilon / \Big(\varepsilon + \mu(C)\big(1 - \mu(B)\big)\Big)$. Switching the roles of B and C we obtain Equation (5.6). □

Remark 5.17. *Note that the lower bound on $\underline{d}_{(F_N)}(D_\varepsilon)$ does not depend on the set B, only on the measure $\mu(B)$. Moreover, it does not depend on the double Følner sequence (F_N).*

5.4 An application of the coloring trick

In this section we give a proof of Theorem 5.5. We start by giving a more precise statement:

Theorem 5.18. *For any finite coloring $K = \bigcup C_i$ there exists a color C_i, a subset $D \subset K$ satisfying $\bar{d}_{(F_N)}(D) > 0$ and, for each $u \in D$, there is a set $D_u \subset K$ also satisfying $\bar{d}_{(F_N)}(D_u) > 0$ such that for any $v \in D_u$ we have $\{u, u+v, uv\} \subset C_i$.*

The proof of Theorem 5.18 uses the fact that sets of recurrence (cf. Definition 2.51) are partition regular (Lemma 2.52). For other similar applications of this phenomenon see for instance [Ber86], the discussion before Question 11 in [Ber96] and Theorem 0.4 in [BM96].

Proof of Theorem 5.18. Let $K = C_1 \cup C_2 \cup ... \cup C_{r'}$ be a finite partition of K. Assume without loss of generality that, for some $r \leq r'$, the upper density $\bar{d}_{(F_N)}(C_i)$ is positive for $i = 1, ..., r$ and $\bar{d}_{(F_N)}(C_i) = 0$ for $i = r+1, ..., r'$.

For a set $C \subset K$ and each $u \in C$ define the set $D_u(C) = (C - u) \cap (C/u)$. Let $D(C) = \Big\{u \in C : \bar{d}_{(F_N)}(D_u(C)) > 0\Big\}$. We want to show that for some $i = 1, ..., r$ we have $\bar{d}_{(F_N)}(D(C_i)) > 0$.

If for some $1 \leq i \leq r$ we have $\bar{d}_{(F_N)}(D(C_i)) = 0$ but $D(C_i) \neq \varnothing$, we can consider the more refined coloring obtained by distinguish $D(C_i)$ and $C_i \setminus D(C_i)$. Since

$$D(C_i \setminus D(C_i)) \subset (C_i \setminus D(C_i)) \cap D(C_i)$$

we conclude that $D(C_i \setminus D(C_i)) = \varnothing$. Thus, without loss of generality, we can assume that either $D(C_i) = \varnothing$ or $\bar{d}_{(F_N)}(D(C_i)) > 0$. Therefore it suffices to show that for some $i = 1, ..., r$ we have $D(C_i) \neq \varnothing$.

For each $i = 1, ..., r$ let $R_i = \{M_u A_{-u} : u \in C_i\} \subset \mathcal{A}_K$ and let $R = R_1 \cup ... \cup R_r$. We claim that R is a set of recurrence. Indeed, given any probability preserving action $(\Omega, \mu, (T_g)_{g \in \mathcal{A}_K})$ of $\mathcal{A}_K$ and any measurable set $B \subset \Omega$ with positive measure, by Theorem 5.11 we find that the set $\{u \in K^* : \mu(A_{-u}B \cap M_{1/u}B) > 0\}$ has positive upper density. In particular, for some $u \in C_1 \cup ... \cup C_r$ we have that $\mu(M_u A_{-u} B \cap B) = \mu(A_{-u}B \cap M_{1/u}B) > 0$. Since $M_u A_{-u} \in R$ we conclude that R is a set of recurrence.

By Lemma 2.52 we conclude that for some $i = 1, ..., r$ the set R_i is a set of recurrence. We claim that $D(C_i) \neq \varnothing$.

To see this, apply the correspondence principle (Theorem 3.24) with $X = K$, $G = \mathcal{A}_K$, $G_N = F_N$ and $E = C_i$ to find a probability preserving action $(T_g)_{g \in \mathcal{A}_K}$ of $\mathcal{A}_K$ on some probability space (Ω, μ) and a measurable set $B \subset \Omega$ satisfying $\mu(B) = \bar{d}_{(F_N)}(C_i)$ and

$$\bar{d}_{(F_N)}\left(A_{-u}C_i \cap M_{1/u}C_i\right) \geq \mu\left(T_{A_{-u}}B \cap T_{M_{1/u}}B\right)$$

for all $u \in K^*$. Since R_i is a set of recurrence, there is some $u \in C_i$ such that

$$\begin{aligned} 0 &< \mu(T_{M_u A_{-u}}B \cap B) = \mu\left(T_{A_{-u}}B \cap T_{M_{1/u}}B\right) \\ &\leq \bar{d}_{(F_N)}\left(A_{-u}C_i \cap M_{1/u}C_i\right) = \bar{d}_{(F_N)}(D_u(C_i)) \end{aligned}$$

We conclude that $u \in D(C_i)$, hence $\bar{d}_{(F_N)}(D^i) > 0$.

Let $D = D(C_i) \subset C_i$ and for each $u \in D$ let $D_u = D_u(C_i)$. Now let $v \in D_u$. Then we have $u + v \in C_i$ and $uv \in C_i$. We conclude that $\{u, u+v, uv\} \subset C_i$ as desired. $\square$

5.5 Finite intersection property of sets of return times

In this section we study isometric anti-representations[3] $(U_g)_{g\in\mathcal{A}_R}$ of the affine semigroup $\mathcal{A}_R$ of a ring R on a Hilbert space $\mathcal{H}$ (this means that $\langle U_g\phi, U_g\psi\rangle = \langle\phi,\psi\rangle$ and $U_g(U_h\phi) = U_{hg}\phi$ for any $g, h \in \mathcal{A}_R$ and $\phi, \psi \in \mathcal{H}$).

Recall that if G is a semigroup and $(U_g)_{g\in G}$ is an isometric (anti-)representation of G on a Hilbert space $\mathcal{H}$, then a vector $\phi \in \mathcal{H}$ is called *compact* if the orbit $\{U_g\phi : g \in G\} \subset \mathcal{H}$ is pre-compact in the norm topology. The set of compact vectors forms a closed subspace.

When G is the additive sub-semigroup S_A of the affine semigroup $\mathcal{A}_R$, we denote the orthogonal projection onto the space of compact vectors by V_A and when G is the multiplicative sub-semigroup S_M of the affine semigroup $\mathcal{A}_R$, we denote the orthogonal projection onto the space of compact vectors by V_M. Our main ergodic-theoretic result is the following analogue of Theorem 5.4, with Cesàro averages (which are unavailable in our current situation) replaced with limits along ultrafilters $p \in \mathcal{G} = \overline{\mathcal{AMI}} \cap \mathcal{MMI}$ (see Definition 3.11).

Theorem 5.19. *Let R be an LID (see Definition 3.1), let $\mathcal{H}$ be a Hilbert space and let $(U_g)_{g\in\mathcal{A}_R}$ be an isometric anti-representation of $\mathcal{A}_R$ on $\mathcal{H}$. Then, for any $\phi, \psi \in \mathcal{H}$ and $p \in \mathcal{G}$ (see Definition 3.11) we have*

$$p\text{-}\lim_u \langle A_u\phi, M_u\psi\rangle = \langle V_A\phi, V_M\psi\rangle.$$

In this section we will always work under the assumptions of Theorem 5.19.

Projection onto the space of compact vectors

We have the following result:

Lemma 5.20. *If $p \in \mathcal{G}$ (see Definition 3.11) and $\phi \in \mathcal{H}$ then*

$$V_M\phi = p\text{-}\lim_u M_u\phi \qquad \textit{in the topology of weak convergence}$$

[3] We deal here with anti-representations instead of (a priori more natural) representations because a measure preserving action $(T_g)_{g\in G}$ of a non-commutative semigroup G induces a natural anti-representation of G by isometries on the corresponding L^2 space. Of course, the results obtained in this section hold true for isometric representations as well.

If $p \in \overline{\mathcal{AMI}}$ and $k \in R^*$ then

$$V_A\phi = p\text{-}\lim_u A_{ku}\phi \qquad \textit{in the topology of weak convergence.}$$

Proof. Since $p \in \mathcal{MMI}$, the first equality follows[4] from Corollary 4.6 on [Ber03]. By the same corollary we have that $V_A\phi = q\text{-}\lim_u A_u\phi$ for every additive minimal idempotent q.

It follows from Lemma 3.16 that $p\text{-}\lim_u A_{ku}\phi = kp\text{-}\lim_u A_u\phi$. In view of Lemma 3.15 we have that $kp \in \overline{\mathcal{AMI}}$. Since the map $q \mapsto q\text{-}\lim_u A_u\phi$ is continuous we conclude that

$$p\text{-}\lim_u A_{ku}\phi = kp\text{-}\lim_u A_u\phi = V_A\phi$$

□

Lemma 5.21. *For every $\phi \in \mathcal{H}$ we have $V_AV_M\phi = V_MV_A\phi$.*

Proof. Let $p \in \mathcal{G}$. For each $k \in R^*$, it follows from Lemma 5.20 that

$$M_kV_A\phi = M_k\left(p\text{-}\lim_u A_u\phi\right) = p\text{-}\lim_u M_kA_u\phi = p\text{-}\lim_u A_{ku}M_k\phi = V_AM_k\phi$$

Therefore

$$V_MV_Af = p\text{-}\lim_k M_kV_A\phi = p\text{-}\lim_k V_AM_k\phi = V_A\left(p\text{-}\lim_k M_k\phi\right) = V_AV_M\phi$$

□

In view of Lemma 5.21, the operator $V := V_AV_M$ is an orthogonal projection. This gives the following simple corollary of Lemma 5.21 which will be needed in the proof of Theorem 5.8 below.

Corollary 5.22. *Let $\phi, \psi \in \mathcal{H}$ and assume that $U_g\psi = \psi$ for every $g \in \mathcal{A}_R$. Then*

$$|\langle \phi, \psi \rangle|^2 \leq \|\psi\|^2 \cdot \langle V_A\phi, V_M\phi \rangle$$

[4]In [Ber03] the results are stated and proved for groups only, but it is easy to check that the proofs work for discrete semigroups as well (as is observed in the first paragraph after the remark following Theorem 4.1 in [Ber03]).

Proof. We have

$$\begin{aligned}\|\psi\|^2 \cdot \langle V_A\phi, V_M\phi\rangle &= \|\psi\|^2 \cdot \langle V\phi, \phi\rangle = \|\psi\|^2 \cdot \|V\phi\|^2 \\ &\geq \left|\langle V\phi, \psi\rangle\right|^2 = \left|\langle \phi, V\psi\rangle\right|^2 = \left|\langle \phi, \psi\rangle\right|^2\end{aligned}$$

where the inequality follows from Cauchy-Schwarz inequality. □

Dealing with $V_A\phi$

The scheme of the proof of Theorem 5.19 is as follows: first we decompose $\phi = V_A\phi + \phi^\perp$ (where $\phi^\perp := \phi - V_A\phi$) into its 'additively compact' and 'additively weak mixing' components. Observe that, since V_A is an orthogonal projection, $V_A(V_A\phi) = V_A\phi$ and $V_A\phi^\perp = 0$. The two main steps are to show that $p\text{-}\lim_u\langle A_u V_A\phi, M_u\psi\rangle = \langle V_A\phi, V_M\psi\rangle$ and that $p\text{-}\lim_u\langle A_u V_A\phi^\perp, M_u\psi\rangle = 0$. In this subsection we deal with the first step.

Lemma 5.23. *Let $\phi \in \mathcal{H}$ be additively compact (i.e. such that $V_A\phi = \phi$). Then for any $p \in \mathcal{G}$*

$$p\text{-}\lim_u \|A_u\phi - \phi\| = 0$$

In other words, for all $\varepsilon > 0$ the set $S := \{u \in K : \|A_u\phi - \phi\| < \varepsilon\}$ is DC.*

Proof. From the definition of V_A, the orbit closure $X = \overline{\{A_u\phi : u \in R\}}$ of ϕ is compact. Hence it follows from Lemma 2.55 (applied to the system $(X, (A_u)_{u\in R})$) that S intersects non-trivially every IP_0 set in $(R, +)$. Invoking Theorem 2.29 it follows that S intersects every piecewise syndetic set in $(R^*, \times)$. Since every DC set is central (and hence piecewise syndetic, according to Corollary 2.33) in $(R^*, \times)$, it follows that S is DC*. □

Lemma 5.24. *For all $p \in \mathcal{G}$ and $\phi, \psi \in \mathcal{H}$ we have*

$$p\text{-}\lim_u\langle A_u(V_A\phi), M_u\psi\rangle = \langle V_A\phi, V_M\psi\rangle$$

Proof. We will assume, without loss of generality, that $\|\phi\|, \|\psi\| \leq 1$. In view of Lemma 5.20 we have

$$p\text{-}\lim_u\langle V_A\phi, M_u\psi\rangle = \left\langle V_A\phi, (p\text{-}\lim_u M_u\psi)\right\rangle = \langle V_A\phi, V_M\psi\rangle$$

Therefore, for every $\varepsilon > 0$, the set

$$S_1 = \left\{ u \in R : |\langle V_A\phi, M_u\psi\rangle - \langle V_A\phi, V_M\psi\rangle| < \frac{\varepsilon}{2} \right\}$$

belongs to p.

Applying Lemma 5.23 with $V_A\phi$ we get that the set $S_2 := \{u \in R : \|A_u V_A\phi - V_A\phi\| < \varepsilon/2\}$ is also in p. Using the Cauchy-Schwarz inequality we have that for any $u \in S_2$

$$|\langle V_A\phi, M_u\psi\rangle - \langle A_u V_A\phi, M_u\psi\rangle| < \frac{\varepsilon}{2}$$

Finally let $S := S_1 \cap S_2 \in p$ and let $u \in S$. We conclude that

$$|\langle A_u V_A\phi, M_u\psi\rangle - \langle V_A\phi, V_M\psi\rangle| < \varepsilon$$

which finishes the proof. $\square$

Dealing with $\phi^\perp$ when R is a field

We now turn our attention to the weak mixing component $\phi^\perp := \phi - V_A\phi$. Dealing with this component in the general case requires some technical steps which obscure the main ideas. In order to clarify these ideas we restrict our attention in this subsection to the case where R is a field; the general case is treated in the next subsection. (Of course the results of this subsection also follow logically from the results in the next one.)

We will use the following version of the van der Corput trick.

Proposition 5.25 (cf. [BM07, Theorem 2.3])**.** *Let $p \in \mathcal{G}$, let $\mathcal{H}$ be a Hilbert space, let $(a_u)_{u\in R^*}$ be a bounded sequence in $\mathcal{H}$ indexed by R^*. If $p\text{-}\lim_u \langle a_{bu}, a_u\rangle = 0$ for all b in a co-finite subset of R^* then $p\text{-}\lim_u a_u = 0$ in the weak topology of $\mathcal{H}$.*

Lemma 5.26. *Let K be a field, let $\mathcal{H}$ be a Hilbert space, let $(U_g)_{g\in\mathcal{A}_K}$ be a unitary anti-representation of $\mathcal{A}_K$ on $\mathcal{H}$ and let $\phi^\perp, \psi \in \mathcal{H}$, where we assume that $V_A\phi^\perp = 0$. Then, for all $p \in \mathcal{G}$ we have*

$$p\text{-}\lim_u \langle A_u\phi^\perp, M_u\psi\rangle = 0$$

Proof. Observe that, since we deal with an anti-representation, the distributive law (see (3.1)) takes the form

$$A_v M_u = M_u A_{vu} \tag{5.7}$$

for any $v \in K$ and $u \in K^*$. Let $a_u = M_{1/u} A_u \phi^\perp$. Then for all $b \in K \setminus \{-1, 0, 1\}$, using (5.7) and the fact that isometries preserve scalar products we have

$$\langle a_{ub}, a_u \rangle = \langle M_{1/ub} A_{ub} \phi^\perp, M_{1/u} A_u \phi^\perp \rangle = \langle A_{u(b-1/b)} \phi^\perp, M_b \phi^\perp \rangle$$

Therefore, it follows from Lemma 5.20 that for every $p \in \mathcal{G}$ we have

$$p\text{-}\lim_u \langle a_{ub}, a_u \rangle = \left\langle p\text{-}\lim_u A_{u(b-1/b)} \phi^\perp, M_b \phi^\perp \right\rangle = \langle V_A \phi^\perp, M_b \phi^\perp \rangle = 0$$

By Proposition 5.25 we conclude that $p\text{-}\lim_u M_{1/u} A_u \phi^\perp = p\text{-}\lim_u a_u = 0$. Hence we have

$$\begin{aligned} p\text{-}\lim_u \langle A_u \phi^\perp, M_u \psi \rangle &= p\text{-}\lim_u \langle M_{1/u} A_u \phi^\perp, \psi \rangle \\ &= \left\langle p\text{-}\lim_u M_{1/u} A_u \phi^\perp, \psi \right\rangle = 0 \end{aligned}$$

□

Dealing with $\phi^\perp$ when R is a general LID

In this subsection we extend the scope of Lemma 5.26 from the previous sub-section to the case when we have a general LID (not necessarily a field). Namely, we will prove:

Lemma 5.27. *Assume R is an LID, let $\mathcal{H}$ be a Hilbert space, let $(U_g)_{g \in \mathcal{A}_R}$ be an isometric anti-representation of $\mathcal{A}_R$ on $\mathcal{H}$ and let $\phi^\perp, \psi \in \mathcal{H}$. Assume that $V_A \phi^\perp = 0$. Then, for all $p \in \mathcal{G}$ we have*

$$p\text{-}\lim_u \langle A_u \phi^\perp, M_u \psi \rangle = 0$$

In the proof of this lemma we will need a few facts about isometric anti-representations of $\mathcal{A}_R$. First observe that, unlike the case when R is a field, M_u is not necessarily invertible. Thus its adjoint M_u^T (defined so that $\langle M_u \phi, \psi \rangle = \langle \phi, M_u^T \psi \rangle$ for all $\phi, \psi \in \mathcal{H}$) may not be in $\mathcal{A}_R$. However, since A_u is invertible (and hence unitary) we have the following distributivity relation:

Lemma 5.28. *Under the assumptions of Lemma 5.27 we have*

$$A_{uv}M_u^T = M_u^T A_v$$

Proof. We have, for any $\phi, \psi \in \mathcal{H}$

$$\langle A_{uv}M_u^T\phi, \psi\rangle = \langle \phi, M_u A_{-uv}\psi\rangle = \langle \phi, A_{-v}M_u\psi\rangle = \langle M_u^T A_v\phi, \psi\rangle.$$

This implies the identity in question. □

Another difficulty which is present in our current context is the fact that the composition $M_nM_n^T$ is not necessarily the identity map. The following lemma allows us to circumvent this difficulty when R is an LID.

Lemma 5.29. *Under the assumptions of Lemma 5.27, there exists an orthogonal projection $P : \mathcal{H} \to \mathcal{H}$ such that for every $\phi \in \mathcal{H}$ we have*

$$p\text{-}\lim_u \|M_uM_u^T\phi - P\phi\| = 0$$

Proof. Let $P_u = M_uM_u^T$. Since M_u is an isometry, P_u is the orthogonal projection onto the image of M_u. Observe that, in particular, the image of $P_{u_1u_2}$ is contained in the image of each P_{u_i}, $i = 1, 2$.

Let $\{r_1, r_2, \dots\}$ be an arbitrary enumeration of the elements of R^* and let $u_n = \prod_{i=1}^n r_i$. Let S_n be the image of M_{u_n}, so that P_{u_n} is the orthogonal projection onto S_n. Note that $S_{n+1} \subset S_n$. Let $S = \bigcap_{n\geq 1} S_n$ and let $P : \mathcal{H} \to S$ be the orthogonal projection. Let E_0 be an orthonormal basis for S and, for each $n \geq 1$ let E_n be an orthonormal basis for $S_n \cap (S_{n+1})^\perp$. Thus $E = \bigcup_{n\geq 0} E_n$ is an orthonormal basis for $\mathcal{H}$. Write ϕ in terms of the basis E as $\phi = \sum_{n\geq 0}\sum_{e\in E_n} c_e e$. For a fixed $\varepsilon > 0$ let $m \in \mathbb{N}$ be such that $\sum_{n\geq m}\sum_{e\in E_n} |c_e|^2 < \varepsilon^2$.

Next, let u be in the ideal u_mR. We have that the image of P_u is contained in the image of P_{u_m}, so $P_uh \in S_m$ and hence

$$P_u\phi = \sum_{e\in E_0} c_e e + \sum_{n=m}^{\infty}\sum_{e\in E_n} c_e e = P\phi + \sum_{n=m}^{\infty}\sum_{e\in E_n} c_e e$$

Therefore $\|P_u\phi - P\phi\| < \varepsilon$. Since the ideal u_mR has finite index as an additive group, it follows from Lemma 3.13 that it belongs to p. We conclude that $p\text{-}\lim M_nM_n^T\phi = p\text{-}\lim P_n\phi = P\phi$ in the strong topology, as desired. □

Finally, we need a strengthening of Lemma 5.20.

Definition 5.30. Let R be an integral domain, let $b \in R$ and let $p \in \beta R$. Assume that $bR \in p$. Given a sequence $(x_u)_{u \in R}$ in a compact space X we define $p\text{-}\lim_u x_{u/b}$ to be the point $x \in X$ such that for every neighborhood U of x, the set $\{u \in bR : x_{u/b} \in U\} \in p$.

Lemma 5.31. *Let R be an LID, let $p \in \mathcal{G}$ and let $k, b \in R^*$. For any unitary anti-representation $(U_g)_{g \in \mathcal{A}_R}$ of the semigroup $\mathcal{A}_R$ on a Hilbert space $\mathcal{H}$ and any $\phi \in \mathcal{H}$ we have*

$$p\text{-}\lim_u A_{ku/b}\phi = V_A\phi \qquad \textit{in the weak topology}$$

Proof. First observe that the p-lim is well defined since the ideal bR has finite index in R, p belongs to the closure $\overline{\mathcal{AMI}}$ of the additive minimal idempotents and hence, in view of Lemma 3.13, $bR \in p$.

It follows from Lemma 3.16 that $p\text{-}\lim_u A_{ku/b}\phi = kp\text{-}\lim_u A_{u/b}\phi$. Since, in view of Lemma 3.15, $kp \in \overline{\mathcal{AMI}}$, we can and will assume that $k = 1$. Next, let $q = b^{-1}p$ be the ultrafilter defined so that $E \in q \iff bE \in p$. It follows from Lemma 3.15 that $q \in \overline{\mathcal{AMI}}$. Therefore, it follows from Lemma 5.20 that for any $\psi \in \mathcal{H}$ and $\varepsilon > 0$ the set

$$E = \{u \in R : \big|\langle A_u\phi - V_A\phi, \psi\rangle\big| < \varepsilon\} \in q$$

We conclude that

$$bE = \{u \in bR : \big|\langle A_{u/b}\phi - V_A\phi, \psi\rangle\big| < \varepsilon\} \in p$$

□

We can now give a proof of Lemma 5.27:

Proof of Lemma 5.27. Denoting by M_u^T the adjoint of M_u, we can write $\langle A_u\phi^\perp, M_u\psi\rangle = \langle M_u^T A_u\phi^\perp, \psi\rangle$, so the lemma will follow if we show that $p\text{-}\lim M_u^T A_u\phi^\perp = 0$ (in the weak topology). To do this we will use the van der Corput trick (Proposition 5.25), and so it suffices to show that

$$p\text{-}\lim_u \langle M_{ub}^T A_{ub}\phi^\perp, M_u^T A_u\phi^\perp\rangle = 0 \qquad \forall b \in R \setminus \{-1, 0, 1\} \tag{5.8}$$

Since the operator A_u is unitary we can rewrite the inner product in (5.8) as

$$\langle M_{ub}^T A_{ub}\phi^\perp, M_u^T A_u\phi^\perp\rangle = \langle A_{-u}M_u M_{ub}^T A_{ub}\phi^\perp, \phi^\perp\rangle.$$

By (5.7) we have $A_{-u}M_u = M_u A_{-u^2}$ (recall this is an anti-representation). Also, assuming that $u \in bR$ and evoking Lemma 5.28 we conclude that

$$\langle M_{ub}^T A_{ub}\phi^\perp, M_u^T A_u\phi^\perp\rangle = \langle M_u M_{ub}^T A_{ub-u/b}\phi^\perp, \phi^\perp\rangle = \langle A_{ub-u/b}\phi^\perp, M_b M_n M_n^T\phi^\perp\rangle$$

By Lemma 5.31 we have that $p\text{-}\lim A_{ub-u/b}\phi^\perp = V_A\phi^\perp = 0$ in the weak topology. By Lemma 5.29 we have that $p\text{-}\lim_u M_b M_u M_u^T\phi^\perp$ exists in the strong topology. Thus we conclude that $p\text{-}\lim\langle A_{ub-u/b)}\phi^\perp, M_b M_n M_n^T\phi^\perp\rangle = 0$, which gives (5.8) and finishes the proof. □

Proofs of 5.19 and some corollaries

We have now gathered all the ingredients necessary to prove Theorem 5.19:

Proof of Theorem 5.19. Let $\phi^\perp = \phi - V_A\phi$, so that $V_A\phi^\perp = 0$. Using Lemmas 5.24 and 5.27 we deduce that

$$p\text{-}\lim\langle A_u\phi, M_u\psi\rangle = p\text{-}\lim\langle A_u V_A\phi, M_u\psi\rangle + \langle A_u\phi^\perp, M_u\psi\rangle = \langle V_A\phi, V_M\psi\rangle$$

□

As a corollary we now deduce Theorem 5.8.

Proof of Theorem 5.8. Let R be an LID, let (Ω,μ) be a probability space, let $(T_g)_{g\in\mathcal{A}_R}$ be a measure preserving action of $\mathcal{A}_R$ on Ω, let $B \subset \Omega$ be a measurable set and let $\varepsilon > 0$. We need to show that the set

$$R(B,\varepsilon) := \left\{u \in R : \mu(A_u^{-1}B \cap M_u^{-1}B) \geq \mu(B)^2 - \varepsilon\right\}$$

is DC*.

Let $\mathcal{H} = L^2(\Omega, \mu)$ and, for each $g \in \mathcal{A}_R$, define the operator $(U_g\phi)(x) = \phi(T_gx)$. Observe that $U_gU_h = U_{hg}$, so this induces an isometric anti-representation $(U_g)_{g\in\mathcal{A}_R}$ of $\mathcal{A}_R$ in $\mathcal{H}$. Let $B \subset \Omega$. Observe that

$$1_{T_g^{-1}B}(x) = 1 \iff T_gx \in B \iff 1_B(T_gx) = 1 \iff U_g1_B(x) = 1$$

Therefore $\mu(A_u^{-1}B \cap M_u^{-1}B) = \int_\Omega A_u1_B \cdot M_u1_B d\mu = \langle A_u1_B, M_u1_B\rangle$. It follows from Theorem 5.19 that for any $\varepsilon > 0$ the set

$$\{u \in R : \langle A_u1_B, M_u1_B\rangle \geq \langle V_A1_B, V_M1_B\rangle - \varepsilon\}$$

is DC^*. Finally, it follows from Corollary 5.22 (applied with $\phi = 1_B$ and $\psi \equiv 1$) that

$$\langle V_A1_B, V_M1_B\rangle \geq \mu(B)^2.$$

$\square$

Observe that Theorem 5.6 easily follows from Theorem 5.8. Indeed, given $p \in \mathcal{G}$ it follows from the definition of DC^* sets and Theorem 5.8 that $R(B_i, \delta) \in p$ for every i. Therefore also the intersection $R = R(B_1, \delta) \cap \cdots \cap R(B_t, \delta)$ belongs to p. Since $p \in \mathcal{G}$ was arbitrary, it follows that R is itself a DC^* set. Finally, Remark 3.23 implies that R must be affinely syndetic.

We now present the main combinatorial corollary of Theorem 5.8:

Theorem 5.32. *Let K be a countable field and let $R \subset K$ be a sub-ring which is a LID. Let $E \subset K$ with $\bar{d}_{(F_N)}(E) > 0$ for some double Følner sequence (F_N) and let $\varepsilon > 0$. Then the set*

$$\left\{u \in R : \bar{d}_{(F_N)}((E-u) \cap (E/u)) > \bar{d}_{(F_N)}(E)^2 - \varepsilon\right\} \tag{5.9}$$

is DC^* and, in particular, affinely syndetic in R.

Proof. Using the correspondence principle (Theorem 2.8 in [BM16a]) one can construct a measure preserving action $(T_g)_{g\in\mathcal{A}_K}$ of $\mathcal{A}_K$ on a probability space $(\Omega, \mathcal{B}, \mu)$ and a set $B \in \mathcal{B}$ such that $\mu(B) = \bar{d}_{(F_N)}(E)$ and, for each $u \in K^*$

$$\bar{d}_{(F_N)}((E-u) \cap (E/u)) \geq \mu(A_u^{-1}B \cap M_u^{-1}B)$$

The result now follows from Theorem 5.8. $\square$

One can deduce parts (2) and (3) of Theorem 5.9 from Theorem 5.32 using the fact that for any finite partition of a countable field, one of the cells of the partition has positive upper density with respect to a double Følner sequence. Then using that cell C_i of the partition as E, for any element n of the (non-empty) set defined in (5.9) and for any x in the (non-empty) intersection $(C_i - n) \cap (C_i/n)$ we have $\{x+n, xn\} \subset C_i$.

To deduce part (1) of Theorem 5.9, one needs an additional fact:

Proposition 5.33. *The subset $\mathbb{N}$ of the ring $\mathbb{Z}$ belongs to every non-principal multiplicative idempotent.*

Proof. Let $p \in \beta\mathbb{Z}$ be a non-principal multiplicative idempotent. Assume, for the sake of a contradiction, that $\mathbb{N} \notin p$. Then $-\mathbb{N} \in p = pp$, which by definition implies that $\{n \in \mathbb{Z}^* : -\mathbb{N}/n \in p\} \in p$. Observe that

$$-\mathbb{N}/n = \{a \in \mathbb{Z}^* : an \in -\mathbb{N}\} = \begin{cases} \mathbb{N} & \text{if } n \in -\mathbb{N} \\ -\mathbb{N} & \text{if } n \in \mathbb{N} \end{cases}$$

Therefore $\{n \in \mathbb{Z}^* : -\mathbb{N}/n \in p\} = \mathbb{N} \notin p$, which is the desired contradiction. $\square$

To deduce part (1) of Theorem 5.9 one applies Theorem 5.32 with $K = \mathbb{Q}$, $R = \mathbb{Z}$ and E being a cell of the partition with positive upper density with respect to a double Følner sequence. The set S defined by (5.9) is DC^* in $\mathbb{Z}$, which means that for any $p \in \mathcal{G}$ we have $S \in p$. Since any $p \in \mathcal{G}$ is a non-principal multiplicative idempotent, it follows from Proposition 5.33 that also $\mathbb{N} \in p$, and therefore $S \cap \mathbb{N} \in p$ and hence is non-empty. For any n in that intersection the set $(E - n) \cap (E/n)$ is non-empty and any x in this intersection yields $\{x+n, xn\} \subset E$.

5.6 Notions of largeness and configurations $\{xy, x+y\}$ in $\mathbb{N}$

In this section we discuss notions of largeness which guarantee the presence of configurations of the form $\{x+y, xy\}$. In the next chapter we will show that for any finite partition of the natural numbers, one of the cells of the partition must contain such a configuration (and indeed significantly more general configurations), however there is no analogue of Theorem 2.23 in the affine setting and hence it is not clear which notion(s) of largeness imply the existence of such a configuration (we are being purposefully vague about the term "notion of largeness" to accommodate any potential candidate). The understanding of which notions of largeness imply the presence of $\{x+y, xy\}$ patterns would likely lead to new results on partition regularity of other configurations, for instance by combining it with procedures similar to those employed in Section 5.4.

It is a trivial observation that the set of odd numbers in $\mathbb{N}$ or in $\mathbb{Z}$ does not contain pairs $\{x+y, xy\}$. Therefore, additively syndetic sets (i.e. sets which are syndetic with respect to the additive semigroup) do not contain, in general, configurations $\{x+y, xy\}$. It is thus somewhat surprising that *multiplicatively* syndetic subsets in any integral domain do contain such patterns:

Theorem 5.34. *Let R be an infinite countable integral domain and let $S \subset R^*$ be multiplicatively syndetic (i.e. syndetic as a subset of the semigroup $(R^*, \cdot)$). Then S contains (many) pairs of the form $\{x+y, xy\}$.*

Proof. Let $F \subset R^*$ be a finite set such that $R^* = \bigcup_{n \in F} S/n$ (the existence of such F is equivalent, by definition, to the statement that S is multiplicatively syndetic). Thus R^* is finitely partitioned into multiplicative shifts of S and hence there exist (many) $a, b \in R^*$ such that $a + bF \subset S$ [5]. Since $ab \in R^* = \bigcup_{n \in F} S/n$, there exist some $n \in F$ such that $abn \in S$. We conclude that $\{a+bn, a(bn)\} \subset S$ as desired. □

While it is not hard to see that there exist partitions of $\mathbb{N}$ or $\mathbb{Z}$ with none of the cells of the partition being multiplicatively syndetic, it is a classical fact that for any finite partition of a semigroup, one of the cells is piecewise syndetic. One could then hope that any multiplicatively piecewise syndetic subset of R^* contains a pattern $\{x+y, xy\}$. Unfortunately, the next example refutes this assertion.

[5] This can be easily proved with the help of the Hales-Jewett theorem. Alternatively, one may combine Proposition 2.39 with Lemma 2.22.

Theorem 5.35. *There exists a set $E \subset \mathbb{N}$ which is additively thick and multiplicatively thick (and so, in particular, E is a multiplicatively piecewise syndetic subset of $\mathbb{N}$) but does not contain a pair $\{x+y, xy\}$ with $x, y > 2$.*

Proof. Let (p_N) be a sequence of primes such that $p_1 = 5$ and, for each $N \in \mathbb{N}$, we have $p_{N+1} > 4(Np_N)^4$. For each $N \in \mathbb{N}$, let

$$E_{2N-1} = p_N[1, N] \qquad \text{and} \qquad E_{2N} = [(Np_N)^2 + 1, 2(Np_N)^2 - 3]$$

where we use the notation $[a, b]$ to denote the set $\{a, a+1, \ldots, b\}$. Let $E = \bigcup E_N$. It follows directly from the construction that E is additively thick as a subset of either $\mathbb{N}$ or $\mathbb{Z}$ and is multiplicatively thick as a subset of $\mathbb{N}$. Moreover, $E \cup (-E)$ is a multiplicatively thick subset of $\mathbb{Z}^*$. Since $\mathbb{N}$ is a multiplicatively syndetic subset of $\mathbb{Z}^*$, it follows that E is a multiplicative piecewise syndetic subset of $\mathbb{Z}^*$.

We first show that no set E_{2N} contains a pair $\{x+y, xy\}$: assume that $a = x+y \in E_{2N}$ and $x, y \geq 2$. Let $b = xy$. Then $b \geq 2(a-2) \geq 2[(Np_N)^2 + 1 - 2] = 2(Np_N)^2 - 2$, so b is too large to be in E_{2N}.

Next we show that no set E_{2N-1} contains such a pair. Assume $xy \in E_{2N-1}$, say $xy = np_N$, then without loss of generality we have $x = p_N d$ and $y = n/d$ for some divisor d of n. But then $x + y < p_N(d+1)$ because $n/d \leq N < p_N$. Hence $x + y \notin E_{2N-1}$.

For each $N \in \mathbb{N}$ we have $(\max E_{2N-1})^2 = (Np_N)^2 < (Np_N)^2 + 1 = \min E_{2N}$ and $(\max E_{2N})^2 = (2(Np_N)^2 - 3)^2 < 4(Np_N)^4 < p_{N+1} = \min E_{2N+1}$. Fix a pair $x, y \in \mathbb{N}$ with both $x, y \geq 2$, let $a = xy$ and $b = x + y$. We observe that $b \leq a \leq (b/2)^2$.

If $b \in E$, say $b \in E_n$, then $\min E_n \leq b \leq a \leq (b/2)^2 < [(\max E_n)/2]^2 < \min E_{n+1}$ so a can not be in E_m for any $m \neq n$. Since we already showed that $a \notin E_n$ (otherwise E_n would contain $\{b, a\} = \{x+y, xy\}$), we conclude that $a \notin E$ and this finishes the proof. $\square$

We observe that the complement $\tilde{E} = \mathbb{N} \setminus E$ of the set constructed in Theorem 5.35 is also rather large. In particular $\bar{d}(\tilde{E}) = 1$, where, as usual, for a subset $S \subset \mathbb{N}$, $\bar{d}(S)$ denotes its upper density (see (2.3)). The next result shows that sets having upper density 1 are large not only additively, but also multiplicatively.

Theorem 5.36. *Let $E \subset \mathbb{N}$ satisfy $\bar{d}(E) = 1$. Then E is affinely thick.*

Proof. Since $\bar{d}$ is the upper density with respect to an additive Følner sequence, it is not hard to see that $\bar{d}\big((E-n)\cap E\big) = 1$ for any $n \in \mathbb{N}$. We claim that also $\bar{d}\big((E/n)\cap E\big) = 1$ for any $n \in \mathbb{N}$.

Assuming the claim for now, let $F = \{g_1, \dots, g_k\} \subset \mathcal{A}_\mathbb{N}$ be an arbitrary finite set. We can write each g_i as the map $g_i : x \mapsto a_i x + b_i$. Let $E_0 = E$ and, for each $i = 1, \dots, k$, let $A_i = \big((E_{i-1} - b_i) \cap E_{i-1}\big)$ and $E_i = \big((A_i/a_i) \cap A_i\big)$. It follows by induction that each of the sets E_i, A_i satisfies $\bar{d}(E_i) = \bar{d}(A_i) = 1$. Take $x \in E_k$, we will show that $g_i(x) \in E$ for every i. Indeed, $x \in E_k \subset E_i = \big((A_i/a_i) \cap A_i\big)$, so $a_i x \in A_i = \big((E_{i-1} - b_i) \cap E_{i-1}\big)$ and hence $a_i x + b_i = g_i(x) \in E_{i-1} \subset E$ as desired.

Now we prove the claim. We will write $[1, x]$ to denote the set $\{1, 2, \dots, \lfloor x \rfloor\}$, where $\lfloor x \rfloor$ is the largest integer no bigger than x.

Let $n \in \mathbb{N}$ and take $\varepsilon > 0$ arbitrary. For some arbitrarily large $N \in \mathbb{N}$ we have

$$|E \cap [1, N]| > \left(1 - \frac{\varepsilon}{2n}\right) N = N - \frac{\varepsilon N}{2n}$$

This implies that

$$|nE \cap [1, N]| = \left| E \cap \left[1, \tfrac{N}{n}\right] \right| > \frac{N}{n} - \frac{\varepsilon N}{2n}$$

Using the general fact that $|X \cup Y| + |X \cap Y| = |X| + |Y|$ we deduce that $nE \cap E \cap [1, N] = (nE \cap [1, N]) \cap (E \cap [1, N])$ has cardinality

$$\begin{aligned}
\big|nE \cap E \cap [1, N]\big| &= \big|E \cap [1, N]\big| + \big|nE \cap [1, N]\big| - \Big|\big(nE \cap [1, N]\big) \cup \big(E \cap [1, N]\big)\Big| \\
&\geq \quad N - \frac{\varepsilon N}{2n} \quad + \quad \frac{N}{n} - \frac{\varepsilon N}{2n} \quad - \quad N \\
&= \frac{N}{n}(1 - \varepsilon)
\end{aligned}$$

Dividing by n (and observing that every number in the intersection $nE \cap E \cap [1, N]$ is divisible by n) we deduce that

$$|E \cap (E/n) \cap [1, N/n]| = |nE \cap E \cap [1, N]| \geq \frac{N}{n}(1 - \varepsilon)$$

As N can be taken arbitrarily large and ε arbitrarily small we conclude that $\bar{d}\big(E \cap (E/n)\big) = 1$, proving the claim. $\square$

It is clear that, for any $y \in \mathbb{N}$, any affinely thick set contains configurations of the form $\{x+y, xy\}$. This observation applies, in particular, to the complement $\tilde{E}$ of the set E constructed in Theorem 5.35.

Recall now the notion of DC set (see Definition 3.11) and observe that for any finite partition of $\mathbb{N}$ one of the cells is a DC set. It follows from Corollary 2.33 that any DC set is both additively piecewise syndetic and multiplicatively piecewise syndetic. For a partition of $\mathbb{N}$ into two cells, one has the following dichotomy: either one of the cells has upper density 1 (in which case Theorem 5.36 assures us that it contains configurations $\{x+y, xy\}$) or both cells have positive lower density. In view of this observation we make the following conjecture:

Conjecture 5.37. *Let $E \subset \mathbb{N}$ be additively and multiplicatively piecewise syndetic and have positive lower density. Then E contains many configurations of the form $\{x+y, xy\}$.*

While Conjecture 5.37 implies that for any partition of $\mathbb{N}$ into two cells, one of the cells contains many configurations $\{x+y, xy\}$, the property of having positive lower density is not stable under partitions. Indeed it is not hard to construct a partition of $\mathbb{N}$ into two sets, both with 0 lower density. However, for any finite partition of a DC set, one of the cells is still a DC set. Observe that the example E constructed in the proof of the Theorem 5.35 can be split into two sets $E = E_A \cup E_M$ such that E_A is additively thick, but has density 0 with respect to any multiplicative Følner sequence, and E_M is multiplicatively thick but has density 0 with respect to any additive Følner sequence. Therefore E is very far from being a DC set. This observation leads to the following conjecture:

Conjecture 5.38. *Every DC set in $\mathbb{N}$ contains a configuration $\{x+y, xy\}$.*

Observe that Conjecture 5.38 is a strengthening of the fact (implied by Theorem 1.5) that any finite coloring of $\mathbb{N}$ yields a monochromatic pair $\{x+y, xy\}$. Assuming Conjecture 5.38 is true, the additional knowledge it provides, potentially coupled with a coloring trick similar to the one used in Section 5.4, may allow one to solve Conjecture 1.4.

5.7 Some concluding remarks

Iterating Theorem 5.10 one can obtain more complex configurations. For instance, if $E \subset K^*$ is such that $\bar{d}_{(F_N)}(E) > 0$, then there exist $x, y \in K^*$ such that

$$\begin{aligned} \bar{d}_{(F_N)}\left(\Big(((E-x)\cap(E/x)) - y\Big) \cap \Big(((E-x)\cap(E/x))/y\Big)\right) &= \\ \bar{d}_{(F_N)}\left((E-x-y)\cap(E/x-y)\cap((E-x)/y)\cap(E/(xy))\right) &> 0 \end{aligned}$$

In particular there exist $x, y, z \in K^*$ such that $\{z+y+x, (z+y)x, zy+x, zyx\} \subset E$. Iterating once more we get $x, y, z, t \in K^*$ such that

$$\left\{ \begin{array}{llll} ((t+z)+y)+x & ((t+z)+y)\times x & ((t+z)\times y)+x & ((t+z)\times y)\times x \\ ((t\times z)+y)+x & ((t\times z)+y)\times x & ((t\times z)\times y)+x & ((t\times z)\times y)\times x \end{array} \right\} \subset E$$

More generally, for each $k \in \mathbb{N}$, applying k times Theorem 5.10 we find, for a given set $E \subset K^*$ with $\bar{d}_{(F_N)}(E) > 0$, a finite sequence $x_0, x_2, ..., x_k$ such that

$$(\ldots(((x_0 \circ_1 x_1) \circ_2 x_2) \circ_3 x_3) \ldots) \circ_k x_k \in E$$

for each of the 2^k possible choices of operations $\circ_i \in \{+, \times\}$. Note that the sequence $x_0, ..., x_k$ depends on k, so we do not necessarily have an infinite sequence $x_0, x_1, ...$ which works for every k (in the same way that we have arbitrarily long arithmetic progressions on a set of positive density but not an infinite arithmetic progression).

This pattern obtained by iteration should be compared with the (less general) patterns obtained below in Theorem 6.2 (if one restricts the functions involved to be projections on some coordinate). We remark that such an iterative procedure is not available in the setting of Theorem 6.2 because of the lack of a notion of largeness responsible for the presence of the monochromatic patterns.

Note that the stipulation about arbitrarily 'large' in Theorem 5.5 is essential since we want to avoid the case when the configuration $\{x+y, xy\}$ degenerates to a singleton. To better explain this point, let $x \in K$, $x \neq 1$ and let $y = \frac{x}{x-1}$. Then $xy = x + y$ and hence the configuration $\{x+y, xy\}$ is rather trivial. We just showed that for any finite coloring of K there are infinitely many (trivial) monochromatic configurations of the form $\{x+y, xy\}$. Note that our Theorem 5.18 is much stronger than this statement, not only because we have configurations with 3 terms $\{x, x+y, xy\}$, but also because for each of "many" x (indeed an affinely syndetic set of x's) there exist "many" y (indeed an affinely syndetic set of y's) such that $\{x, x+y, xy\}$ is monochromatic.

Our main ergodic result (Theorem 5.11) raises the question of whether, under the same assumptions, one has a triple intersection of positive measure $\mu(B \cap A_{-u}B \cap M_{1/u}B) > 0$ for some $u \in K^*$. This would imply that, given any set $E \subset K$ with $\bar{d}_{(F_N)}(E) > 0$, one can find $u, y \in K^*$ such that $\{y, y+u, yu\} \subset E$. Using the methods of Section 5.4, one could then show that for every finite coloring of K, one color contains a configuration of the form $\{u, y, y+u, yu\}$.

On the other hand, not every set $E \subset K$ with $\bar{d}_{(F_N)}(E) > 0$ contains a configuration $\{u, y, y+u, yu\}$. In fact, in every abelian group there exists a syndetic set (hence of positive density for any Følner sequence) not containing a configuration of the form $\{u, y, y+u\}$. Indeed, let G be an abelian group and let $\chi : G \to \mathbb{R}/\mathbb{Z}$ be a non-principal character (a non-zero homomorphism; it exists by Pontryagin duality). Then the set $E := \{g \in G : \chi(g) \in [1/3, 2/3)\}$ has no triple $\{u, y, y+u\}$. However it is syndetic because the intersection $[1/3, 2/3) \cap \chi(G)$ is syndetic in the group $\chi(G)$. (This is true and easy to check with $\chi(G)$ replaced by any subgroup of $\mathbb{R}/\mathbb{Z}$.)

CHAPTER 6

POLYNOMIAL RAMSEY FAMILIES IN LIDS

In this chapter we present recent work from [Mor] on polynomial Ramsey families in LID rings. Our main result is Theorem 6.2 below.

6.1 Introduction

Very recently we established in [Mor] that the family $\{x+y, xy\}$ is Ramsey in $\mathbb{N}$, settling an old open problem and establishing an important step towards solving Conjecture 1.4 or even Conjecture 5.1. More generally, we obtain the following.

Theorem 6.1. *Let $s \in \mathbb{N}$ and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $\mathbb{N}^i \to \mathbb{N}$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in \mathbb{N}$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Then the family*

$$\{x_0 \cdots x_s\} \cup \left\{x_0 \cdots x_j + f(x_{j+1}, \ldots, x_i) : 0 \leq j < i \leq s, f \in F_{i-j}\right\}$$

is Ramsey in $\mathbb{N}$.

In particular, taking $s = 1$ and $F_1 = \{x \mapsto 0, x \mapsto x\}$ consisting only of the zero function and the identity function, we obtain Theorem 1.5.

Our method from [Mor] works equally well in the scope of general LID rings.

Theorem 6.2. *Let R be a LID, let $s \in \mathbb{N}$ and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $R^i \to R$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in R$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Then the family*

$$\{x_0 \cdots x_s\} \cup \left\{x_0 \cdots x_j + f(x_{j+1}, \ldots, x_i) : 0 \leq j < i \leq s, f \in F_{i-j}\right\}$$

is Ramsey in R.

As an illustration of the applications of Theorem 6.2, setting $s = 5$ in Theorem 6.2 and letting each F_i consist only of the function $f_i : (x_1, \ldots, x_i) \mapsto x_1 \cdots x_i$, we obtain the following (aesthetically pleasing) Ramsey family.

Example 6.3. The following family is Ramsey:

$$\left\{\begin{array}{lllll} x & & & & \\ xy, & x+y & & & \\ xyz, & x+yz, & xy+z & & \\ xyzt, & x+yzt, & xy+zt, & xyz+t & \\ xyztw, & x+yztw, & xy+ztw, & xyz+tw & xyzt+w \end{array}\right\}$$

Theorem 6.1 can also be used to obtain new partition regular equations:

Corollary 6.4. *Let $k \in \mathbb{N}$ and $c_1, \ldots, c_k \in \mathbb{Z} \setminus \{0\}$ be such that $c_1 + \cdots + c_k = 0$. Then for any finite coloring of $\mathbb{N}$ there exist pairwise distinct $a_0, \ldots, a_k \in \mathbb{N}$, all of the same color, such that*

$$c_1 a_1^2 + \cdots + c_k a_k^2 = a_0.$$

In particular, setting $k = 2$ and $c_1 = 1$, $c_2 = -1$, we deduce:

Corollary 6.5. *For any finite coloring of $\mathbb{N}$ there exists a solution a, b, c of the equation $a^2 - b^2 = c$ with all a, b and c of the same color.*

Note that the similar equation $a^2 - b = c$ is *not* partition regular (cf. [CGS12, Theorem 3]).

Our proof of Theorem 6.2 proceeds by first transferring the problem to the language of topological dynamics using the correspondence principle (Theorem 3.25). We are then left to prove the following:

Theorem 6.6. *Let $(X, (T_g)_{g \in \mathcal{A}_R})$ be an $\mathcal{A}_R$-topological system with a dense set of additively minimal points, and assume that each map $T_g : X \to X$ is open and injective. Let $s \in \mathbb{N}$*

and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $R^i \to R$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in R$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Then for any open cover $\mathcal{U}$ of X there exists an open set $U \in \mathcal{U}$ in that cover and infinitely many s-tuples $x_1, \ldots, x_s \in R$ such that

$$U \cap \bigcap_{0 \leq j < i \leq s} \bigcap_{f \in F_{i-j}} M_{x_{j+1} \cdots x_s} A_{f(x_{j+1}, \ldots, x_i)} U \neq \varnothing$$

The proof of Theorem 6.6 uses ideas developed in [BM16a] and presented in Chapter 5 together with a "complexity reduction" method inspired by [BL96].

The proof of Theorem 6.2 can be made elementary; to illustrate this, we present in Section 6.4 a short and purely combinatorial proof of Theorem 6.1 which is independent from the rest of this chapter. While strictly speaking Theorem 6.2 does not directly apply to $\mathbb{N}$ (since $\mathbb{N}$ is not a ring and hence not a full-fledged LID) the same proofs apply simultaneously to Theorems 6.1 and 6.2. To avoid repetition, we chose to present a dynamical proof of Theorem 6.2 and a combinatorial proof of Theorem 6.1, but of course one can also obtain a dynamical proof of Theorem 6.1, as was done in [Mor], as well as a combinatorial proof of Theorem 6.2 (in [Mor] we only a combinatorial proof of the special case Theorem 1.5).

6.2 Reducing Theorem 6.2 to a dynamical statement

In this section we reduce Theorem 6.2 to its dynamical formulation, Theorem 6.6.

The proof of Theorem 6.6 is presented in Section 6.3. In order to derive Theorem 6.2 from its topological counterpart we will make use of the affine topological correspondence principle, Theorem 3.25.

Proof of Theorem 6.2. Let $s \in \mathbb{N}$ and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $R^i \to R$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in R$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Let $R = C_1 \cup \cdots \cup C_r$ be a finite coloring of R. We need to show that there exists a color $C \in \{C_1, \ldots, C_r\}$ and (infinitely many) $s+1$-tuples $x_0, \ldots, x_s \in R$ such that $x_0 \cdots x_s \in C$ and, for every $0 \leq j < i \leq s$ and $f \in F_{i-j}$, we have $x_1 \cdots x_j + f(x_{j+1}, \ldots, x_i) \in C$.

We append to F_s the zero function $f : R^s \to \{0_R\}$ if necessary. Invoking Theorem 3.25 and then Theorem 6.6, we find a color C and (infinitely many) s-tuples $x_1, \ldots, x_s \in \mathbb{N}$ such that the intersection

$$C \cap \bigcap_{0 \le j < i \le s} \bigcap_{f \in F_{i-j}} M_{x_{j+1} \cdots x_s} A_{-f(x_{j+1}, \ldots, x_i)} C \tag{6.1}$$

is non-empty. Take x in the intersection (6.1) and observe that $x \in x_1 \cdots x_s C$ (letting $j = 0$, $i = s$ and $f \equiv 0$). Therefore $x_0 := x/(x_1 \cdots x_s) \in C$ (and in particular is an integer).

Finally, for $0 \le j < i \le s$ and $f \in F_{i-j}$, we have $x \in x_{j+1} \cdots x_s (C - f(x_{j+1}, \ldots, x_i))$, so $x_0 \cdots x_j + f(x_{j+1}, \ldots, x_i) = x/(x_{j+1} \cdots x_s) + f(x_{j+1}, \ldots, x_i) \in C$. $\square$

6.3 Proof of Theorem 6.6

We will make use of a version of the polynomial van der Waerden theorem of Bergelson and Leibman in general abelian groups (Corollary 2.44).

We recall that the proof of the polynomial van der Waerden theorem in [BL96] is derived from a topological statement. While this topological statement (namely, [BL96, Theorem C]) is only proved for metrizable spaces, it is remarked in [BL96, Proposition 1.10] that the result holds in the non-metrizable setting, either by running a similar proof or by applying the combinatorial version of polynomial van der Waerden directly. We use the second approach to derive the following corollary, which is a dynamical version of Corollary 2.44 in the form we will use.

Corollary 6.7. *Let R be a LID, let $(X, (T_g)_{g \in \mathcal{A}_R})$ be an $\mathcal{A}_R$-topological dynamical system, and assume that X contains a dense set of additively minimal points. Let $F \subset R[x]$ be a finite set such that $p(0) = 0$ for all $p \in F$. Then for any nonempty open set $U \subset X$ there exists $n \in R$ such that*

$$\bigcap_{p \in F} A_{p(n)} U \neq \varnothing$$

Proof. Let $y \in U$ be an additively minimal point, and let $Y = \overline{\{A_n y : n \in R\}}$ be its additive orbit closure. Since $(Y, (A_n)_{n \in R})$ is a minimal topological system, the union $\bigcup_n A_n U$ covers

Y, and by compactness there exists $r \in \mathbb{N}$ and $n_1, \ldots, n_r \in R$ such that the finite union $\bigcup_{t=1}^r A_{n_t} U$ covers Y. We define a coloring $\chi : R \to \{1, \ldots, r\}$ of R by letting $\chi(n)$ be such that $A_n y \in A_{n_{\chi(n)}} U$.

We invoke Corollary 2.44 (with $-F$) to find some $t \in \{1, \ldots, r\}$ and $x, n \in R$ such that $\chi(x - p(n)) = t$ for every $p \in F$. In other words, $A_{x-p(n)} y \in A_{n_t} U$ for all $p \in F$ and hence $A_{x-n_t} y \in A_{p(n)U}$ for every $p \in F$. We conclude that

$$A_{x-n_t} y \quad \in \quad \bigcap_{p \in F} A_{p(n)} U,$$

proving the intersection to be non-empty. $\square$

Outline of the proof

There are two main ingredients in the proof of Theorem 6.6. One is a "complexity reduction" technique similar to the one used by Bergelson and Leibman in [BL96] to prove the polynomial van der Waerden theorem (and also used in [BM16c, Lemma 8.5]). The other main ingredient is a fact about the algebraic behaviour of the expression $g : n \mapsto M_n A_{f(n)} \in \mathcal{A}_R$ discovered (and explored) in [BM16a] and [BM16b], namely that the "multiplicative derivative" $n \mapsto g(nm) g(n)^{-1}$ becomes a purely additive expression whenever f is a polynomial. This fact is also the heart of Theorem 5.4 and Theorem 5.19 above.

Before we delve into the full details of the proof of Theorem 6.6 in the next subsection, we explain the main steps of the proof in the special case when $R = \mathbb{Z}$, $s = 1$ and F_1 is a singleton consisting only of the map $x \mapsto -x$. In other words, we will show that for any finite cover of a nice $\mathcal{A}_{\mathbb{Z}}$-topological system X, there is a set U in the cover and some $y \in \mathbb{Z}$ such that $U \cap M_y A_{-y} U \neq \varnothing$ (after applying the correspondence principle this special case corresponds essentially to Theorem 1.5).

The idea is to construct a sequence (B_n) of non-empty open sets of X, each contained inside some member U_n of the open cover, such that

$$\forall\, n < m, \quad \exists\, y = y(n, m) \in \mathbb{Z}^*, \qquad M_y A_{-y} B_n \supset B_m. \tag{6.2}$$

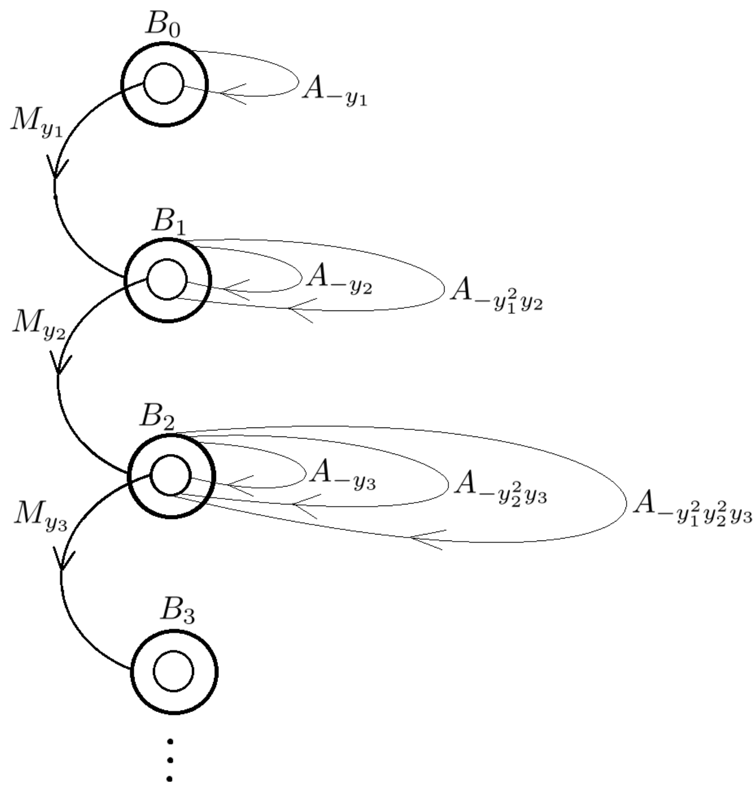

Figure 6.1: Construction of the sequence (B_n)

Assuming we construct such sequence, since the open cover is finite we can find $n < m$ for which both B_n and B_m are contained inside the same member U of the open cover; it then follows from (6.2) that $U \cap M_y A_{-y} U \neq \varnothing$, finishing the proof.

The construction of the sequence (B_n) is natural and is illustrated by Figure 6.1: starting with an arbitrary non-empty open set B_0, we find some y_1 such that $B_0 \cap A_{-y_1} B_0 \neq \varnothing$ (such y_1 exists since B_0 contains some additively minimal points), and then we "push" that intersection by M_{y_1} to create $B_1 := M_{y_1}(B_0 \cap A_{-y_1} B_0)$. In particular, (6.2) holds for $n = 0, m = 1$ with $y = y_1$. For the next step, we start similarly: assume $y_2 \in \mathbb{Z}$ is such that $B_1 \cap A_{-y_2} B_1 \neq \varnothing$. As long as we take $B_2 \subset M_{y_2}(B_1 \cap A_{-y_2} B_1)$, we will indeed have $B_2 \subset M_{y_2} A_{-y_2} B_1$ (and hence (6.2) holds for $n = 1$ and $m = 2$). Next we need to force B_2 to satisfy (6.2) for $n = 0$ and $m = 2$. Since we know how to control the "multiplicative derivative" of the expression $M_y A_{-y}$, we seek to obtain (6.2) with $y(0, 2) = y_1 y_2$; in other words, we want $B_2 \subset M_{y_1 y_2} A_{-y_1 y_2} B_0$. Putting both conditions together, we are left to find

$y_2 \in \mathbb{N}$ so that

$$M_{y_2}(B_1 \cap A_{-y_2}B_1) \cap M_{y_1y_2}A_{-y_1y_2}B_0 \neq \varnothing.$$

Applying $M_{y_2}^{-1}$ it suffices to make $B_1 \cap A_{-y_2}B_1 \cap M_{y_1}A_{-y_1y_2}B_0 \neq \varnothing$. Using the distributivity law (3.1), we have that $M_{y_1}A_{-y_1y_2} = A_{-y_1^2y_2}M_{y_1}$, and since $M_{y_1}B_0 \supset M_1$, we see that it is sufficient to find $y_2 \in \mathbb{N}$ such that

$$B_1 \cap A_{-y_2}B_1 \cap A_{-y_1^2y_2}B_1 \neq \varnothing.$$

The existence of such a $y_2 \in \mathbb{Z}$ is a consequence of Corollary 6.7, so setting $B_2 := M_{y_2}(B_1 \cap A_{-y_2}B_1 \cap A_{-y_1^2y_2}B_1)$ we have successfully constructed B_2 and y_2 satisfying (6.2) whenever $n \leq 2$.

Proceeding in this fashion we can construct the sequence B_n, each time invoking Corollary 6.7 to choose $y_n \in \mathbb{Z}$ so that

$$B_n := M_{y_n}(B_{n-1} \cap A_{-y_n}B_{n-1} \cap A_{-y_{n-1}^2y_n}B_{n-1} \cap \cdots \cap A_{-y_1^2\cdots y_{n-1}^2y_n}B_{n-1})$$

is non-empty. One can see, using the distributivity law (3.1), that (6.2) indeed holds with $y(n,m) = y_{n+1}\cdots y_m$. For instance, to see why $M_{y_2y_3y_4}A_{-y_2y_3y_4}B_1 \supset B_4$, observe that

$$M_{y_2y_3y_4}A_{-y_2y_3y_4}B_1 = M_{y_4}A_{-y_2^2y_3^2y_4}M_{y_3}M_{y_2}B_1 \subset M_{y_4}A_{-y_2^2y_3^2y_4}B_3 \subset B_4.$$

Proof of Theorem 6.6

Let $(X, (T_g)_{g\in\mathcal{A}_R})$ be an $\mathcal{A}_R$-topological system with a dense set of additively minimal points and assume that each map $T_g : X \to X$ is open and injective. Let $s \in \mathbb{N}$ and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $R^i \to R$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in R$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Let $\mathcal{U}$ be an open cover of X. We need to find $U \in \mathcal{U}$ and infinitely many s-tuples $x_1, \ldots, x_s \in R$ such that

$$U \cap \bigcap_{0\leq j<i\leq s} \bigcap_{f\in F_{i-j}} M_{x_{j+1}\cdots x_s}A_{f(x_{j+1},\ldots,x_i)}U \neq \varnothing. \tag{6.3}$$

Since X is compact, we can find a finite subcover $U_1, \ldots, U_r$ of $\mathcal{U}$ with each $U_t \neq \varnothing$.

We will construct, inductively, four sequences:

- $(t_n)_{n\geq 0}$ in $\{1,\ldots,r\}$,
- $(y_n)_{n\geq 1}$ in R injective,
- $(B_n)_{n\geq 0}$ of non-empty open subsets of X,
- $(D_n)_{n\geq 1}$ of non-empty open subsets of X,

such that $B_n \subset U_{t_n}$ (the set D_n corresponds to the smaller circle inside B_{n-1} in Figure 6.1). It will be convenient to denote by $y(m,n)\in R$ the product $y(m,n) := y_{m+1}y_{m+2}\cdots y_n$ for any $0\leq m\leq n$, with the convention that the (empty) product $y(n,n)$ equals 1.

Initiate $t_0 = 1$ and $B_0 = U_1$. Using Corollary 6.7 we find $y_1\in R$ such that

$$D_1 := B_0 \cap \bigcap_{f\in F_1} A_{f(y_1)}B_0 \neq \varnothing.$$

Since $U_1,\ldots,U_r$ forms an open cover of X and $M_n : X\to X$ is an open map, we can find $t_1\in\{1,\ldots,r\}$ such that $B_1 := M_{y_1}D_1\cap U_{t_1}$ is open and nonempty. Next we invoke Corollary 6.7 again to find $y_2\in R$ such that

$$D_2 := B_1 \cap \left(\bigcap_{f\in F_1} A_{f(y_2)}B_1 \cap A_{y_1 f(y_1y_2)}B_1\right) \cap \left(\bigcap_{f\in F_2} A_{y_1 f(y_1,y_2)}B_1\right) \neq \varnothing.$$

We then choose $t_2\in\{1,\ldots,r\}$ such that $B_2 := M_{y_2}D_2\cap U_{t_2}\neq\varnothing$. The third step of the iteration becomes a little more complicated. Using Corollary 6.7 one more time we find $y_3\in R$ such that

$$\begin{aligned} D_3 := B_2 \cap &\left(\bigcap_{f\in F_1} A_{f(y_3)}B_2 \cap A_{y_2 f(y_2y_3)}B_2 \cap A_{y_1y_2f(y_1y_2y_3)}\right) \\ &\cap \left(\bigcap_{f\in F_2} A_{y_2f(y_2,y_3)}B_2 \cap A_{y_1y_2f(y_1y_2,y_3)}B_2 \cap A_{y_1y_2f(y_1,y_2y_3)}B_2\right) \\ &\cap \left(\bigcap_{f\in F_3} A_{y_1y_2f(y_1,y_2,y_3)}B_2\right) \neq \varnothing. \end{aligned}$$

We then choose $t_3\in\{1,\ldots,r\}$ such that $B_3 := M_{y_3}D_3\cap U_{t_3}\neq\varnothing$.

In general, for $n\geq 2$, assume that $(t_m)_{m=0}^{n-1}$, $(y_m)_{m=1}^{n-1}$, $(B_m)_{m=0}^{n-1}$ and $(D_m)_{m=1}^{n-1}$ have been constructed. For each $i\in\{1,\ldots,s\}$ and each $f\in F_i$, we define the collection $G_n(f)$

of all functions $g : R \to R$ of the form

$$g : z \mapsto y(m_1, n-1) f\big(y(m_1, m_2),\ y(m_2, m_3),\ \ldots,\ y(m_i, n-1) \cdot z\big)$$

for any choice $0 \le m_1 < m_2 < \cdots < m_i < n$. If $i > n$ then we set $G_n(f)$ to be empty. Observe that each $g \in G_n(f)$ is a polynomial satisfying $g(0) = 0$.

Invoking Corollary 6.7, we can find $y_n \in R$ satisfying

$$D_n := B_{n-1} \cap \bigcap_{i=1}^{s} \bigcap_{f \in F_i} \bigcap_{g \in G_n(f)} A_{g(y_n)} B_{n-1} \neq \varnothing. \tag{6.4}$$

Let $t_n \in \{1, \ldots, r\}$ be such that the intersection $B_n := M_{y_n} D_n \cap U_{t_n} \neq \varnothing$ (observe that B_n is open because M_{y_n} is an open map). This finishes the construction of y_n, t_n, D_n, B_n. It is immediate from the construction that $B_n \subset U_{t_n}$ for every $n \ge 0$. Moreover, $B_n \subset M_{y_n} D_n \subset M_{y_n} B_{n-1}$. Iterating this observation we obtain

$$\forall m \le n, \qquad B_n \subset M_{y(m,n)} B_m. \tag{6.5}$$

Since the sequence $(t_n)_{n \ge 0}$ takes only finitely many values, there exists $t \in \{1, \ldots, r\}$ and infinitely many tuples of natural numbers $n_0 < \cdots < n_s$ such that $t_{n_i} = t$. For each $i \in \{1, \ldots, s\}$, let $x_i = y(n_{i-1}, n_i)$. We claim that (6.3) is satisfied with $U = U_t$ and with this choice of x_i. We will show that the intersection in (6.3) is non-empty by proving that it contains B_{n_s}. Since $B_{n_j} \subset U_t$ for every $j \in \{0, \ldots, s\}$, it suffices to show that

$$\forall\ 0 \le j < i \le s, \qquad \forall\ f \in F_{i-j}, \qquad B_{n_s} \subset M_{x_{j+1} \cdots x_s} A_{f(x_{j+1}, \ldots, x_i)} B_{n_j}. \tag{6.6}$$

Now fix $0 \le j < i \le s$ and $f \in F_{i-j}$. Observe that there exists some $g \in G_{n_i}(f)$ such that $y(n_j, n_i - 1) f(x_{j+1}, \ldots, x_i) = g(y_{n_i})$. Using (6.5), we conclude

$$\begin{aligned}
B_{n_s} &\subset M_{y(n_i, n_s)} B_{n_i} \subset M_{y(n_i, n_s)} M_{y_{n_i}} D_{n_i} \\
\text{using (6.4)} \quad &\subset M_{y(n_i - 1, n_s)} \Big(A_{g(y_{n_i})} B_{n_i - 1} \Big) \\
\text{using (6.5)} \quad &\subset M_{y(n_i - 1, n_s)} A_{g(y_{n_i})} M_{y(n_j, n_i - 1)} B_{n_j} \\
\text{using (3.1)} \quad &= M_{y(n_i - 1, n_s)} M_{y(n_j, n_i - 1)} A_{g(y_{n_i}) / y(n_j, n_i - 1)} B_{n_j} \\
&= M_{x_{j+1} \cdots x_s} A_{f(x_{j+1}, \ldots, x_i)} B_{n_j}.
\end{aligned}$$

This proves (6.6) and finishes the proof of Theorem 6.6.

6.4 An elementary proof

In this section we present an elementary rendering of the above proof of Theorem 6.2. To keep things slightly more concrete, we prove Theorem 6.1; the proof in this section can be easily adapted to obtain Theorem 6.2 instead. We remark that, while this proof is short and essentially self contained, it is, in essence, a combinatorial rephrasing of the dynamical proof.

We will use the following strengthening of the polynomial van der Waerden's theorem which can be obtained by combining Theorem 1.3 with Theorem 2.23.

Theorem 6.8 (cf. [BH01, Theorem 4.5])**.** *Let $E \subset \mathbb{N}$ be piecewise syndetic, and let $F \subset \mathbb{Z}[x]$ be finite. Then there exists $n \in \mathbb{N}$ such that the intersection*

$$E \cap \bigcap_{f \in F} (E - f(n))$$

is piecewise syndetic.

Proof of Theorem 6.1. Let $s \in \mathbb{N}$ and, for each $i = 1, \ldots, s$, let F_i be a finite set of functions $\mathbb{N}^i \to \mathbb{N}$ such that for all $f \in F_i$ and any $x_1, \ldots, x_{i-1} \in \mathbb{N}$, the function $x \mapsto f(x_1, \ldots, x_{i-1}, x)$ is a polynomial with 0 constant term. Let $r \in \mathbb{N}$ and let $\mathbb{N} = C_1 \cup \cdots \cup C_r$ be an arbitrary coloring of $\mathbb{N}$. We need to find $t \in \{1, \ldots, r\}$ and (infinitely many) $x_0, x_1, \ldots, x_s \in \mathbb{N}$ satisfying

$$\{x_0 \cdots x_s\} \cup \Big\{x_0 \cdots x_j + f(x_{j+1}, \ldots, x_i) : 0 \leq j < i \leq s, f \in F_{i-j}\Big\} \subset C_t. \tag{6.7}$$

As above, we will construct inductively four sequences:

- an increasing sequence $(y_n)_{n\geq 1}$ of natural numbers,
- two sequences $(B_n)_{n\geq 0}$ and $(D_n)_{n\geq 1}$ of piecewise syndetic subsets of $\mathbb{N}$,
- a sequence $(t_n)_{n\geq 0}$ of colors in $\{1, \ldots, r\}$,

such that $B_n \subset C_{t_n}$ for every $n \geq 0$.

It will be convenient to denote by $y(m,n)\in\mathbb{N}$ the product $y(m,n):=y_{m+1}y_{m+2}\cdots y_n$ for any $0\leq m\leq n$, with the convention that the (empty) product $y(n,n)$ equals 1. Initiate by choosing $t_0\in\{1,\ldots,r\}$ such that C_{t_0} is piecewise syndetic (using Lemma 2.22), and let $B_0:=C_{t_0}$.

Assume now that $n\geq 1$ and that we have already defined $(t_m)_{m=0}^{n-1}$, $(y_m)_{m=1}^{n-1}$, $(B_m)_{m=0}^{n-1}$ and $(D_m)_{m=1}^{n-1}$. For each $i\in\{1,\ldots,s\}$ and each $f\in F_i$, we define the collection $G_n(f)$ of all functions $g:\mathbb{N}\to\mathbb{N}$ of the form

$$g: z\mapsto y(m_1,n-1)f\big(y(m_1,m_2),\ y(m_2,m_3),\ \ldots,\ y(m_i,n-1)\cdot z\big)$$

for any choice $0\leq m_1<m_2<\cdots<m_i<n$. If $i>n$ then we set $G_n(f)$ to be empty. Observe that each $g\in G_n(f)$ is a polynomial with rational coefficients satisfying $g(0)=0$.

We apply Theorem 6.8 to find $y_n\in\mathbb{N}$ such that

$$D_n:=B_{n-1}\cap\bigcap_{i=1}^{s}\bigcap_{f\in F_i}\bigcap_{g\in G_n(f)}\Big(B_{n-1}-g(y_n)\Big) \tag{6.8}$$

is piecewise syndetic. Observe that y_nD_n is also piecewise syndetic, and therefore Lemma 2.22 provides some $t_n\in\{1,\ldots,r\}$ such that $B_n:=y_nD_n\cap C_{t_n}$ is piecewise syndetic. This finishes the construction of the sequences.

Note that $B_n\subset y_nD_n\subset y_nB_{n-1}$; iterating this fact we obtain

$$\forall\ 0\leq m<n,\qquad B_n\subset y(m,n)B_m. \tag{6.9}$$

Since the sequence $(t_n)_{n\geq 0}$ takes only finitely many values, there exists $t\in\{1,\ldots,r\}$ and infinitely many tuples of natural numbers $n_0<\cdots<n_s$ such that $t_{n_i}=t$. For each $i\in\{1,\ldots,s\}$, let $x_i=y(n_{i-1},n_i)$. Also, let $\tilde{x}\in B_{n_s}$ be arbitrary and let $x_0:=\tilde{x}/(x_1x_2\cdots x_s)$. Observe that, in view of (6.9), $x_0\in B_{n_0}$ and in particular is an integer. Moreover, (6.9) also implies that any initial product $x_0\cdots x_j\in B_{n_j}$.

We claim that (6.7) is satisfied with this choice of t and x_i, which will finish the proof. Since $x_0\cdots x_s=\tilde{x}$ was chosen to belong to $B_{n_s}\subset C_t$, all that remains to prove is that for every $0\leq j<i\leq s$ and every $f\in F_{i-f}$,

$$x_0\cdots x_j+f(x_{j+1},\ldots,x_i)\in C_t \tag{6.10}$$

Observe that there exists some $g \in G_{n_i}(f)$ such that $y(n_j, n_i - 1)f(x_{j+1}, \ldots, x_i) = g(y_{n_i})$. We have

$$\begin{aligned}
x_{j+1} \cdots x_i(x_0 \cdots x_j + f(x_{j+1}, \ldots, x_i)) &= x_0 \cdots x_i + y_{n_i} g(y_{n_i}) \\
&\in B_{n_i} + y_{n_i} g(y_{n_i}) \\
&\subset y_{n_i} D_{n_i} + y_{n_i} g(y_{n_i}) \\
&\subset y_{n_i} B_{n_i - 1} \\
&\subset y_{n_i} y(n_j, n_i - 1) B_j \\
&= x_{j+1} \cdots x_i B_j \\
&\subset x_{j+1} \cdots x_i C_t
\end{aligned}$$

Dividing by $x_{j+1} \cdots x_i$ we obtain precisely (6.10). □

Remark 6.9. *As an alternative approach, one could replace piecewise syndetic sets with sets having positive upper density and replace the polynomial van der Waerden's theorem with (a suitable form of) the polynomial Szemerédi's theorem in [BM96].*

6.5 Applications to Ramsey theory

In this section we derive some corollaries of Theorem 6.1, by specifying values of s and sets of functions F_i of interest.

By letting $s = 1$ in Theorem 6.1 we obtain the following result:

Corollary 6.10. *Let $k \in \mathbb{N}$ and let $f_1, \ldots, f_k \in \mathbb{Z}[x]$ satisfy $f_\ell(0) = 0$ for each ℓ. Then for any finite coloring of $\mathbb{N}$ there exist $x, y \in \mathbb{N}$ such that the set*

$$\{xy, x + f_1(y), \ldots, x + f_k(y)\}$$

is monochromatic.

Observe that by putting $f_1(y) = 0$, the monochromatic configuration in the previous corollary contains x.

In a different direction, letting s be arbitrary but requiring each F_i to consist of only the zero function and the function $f_i(x_1, \ldots, x_i) = x_1 \cdots x_i$ we deduce:

Corollary 6.11. *For any $s \in \mathbb{N}$ and any finite coloring of $\mathbb{N}$, there exist $x_0, \ldots, x_s \in \mathbb{N}$ such that the set*

$$\left\{ \prod_{\ell=0}^{j} x_\ell \ : \ 0 \leq j \leq s \right\} \cup \left\{ \prod_{\ell=0}^{j} x_\ell + \prod_{\ell=j+1}^{i} x_\ell \ : \ 0 \leq j < i \leq s \right\}$$

is monochromatic.

Observe that we do not require that each function $f \in F_i$ in Theorem 6.1 be a polynomial in all its variables (but only in the last variable). In particular, we obtain the following examples:

Example 6.12. The following are Ramsey families:

1. $\{x, x+y, xy, xyz, x+z, x+z^y\}$;
2. $\{x, xy, xyz, x+f(y)z\}$ for any function $f : \mathbb{N} \to \mathbb{Z}$;
3. $\{x, xy, xyz, xyzt, x+z^y, x+t^z, x+f(y)t^{g(z)}\}$ for any functions $f, g : \mathbb{N} \to \mathbb{N}$.

Finally, we prove Corollary 6.4 from the introduction.

Corollary 6.4. *Let $k \in \mathbb{N}$ and $c_1, \ldots, c_k \in \mathbb{Z} \setminus \{0\}$ be such that $c_1 + \cdots + c_k = 0$. Then for any finite coloring of $\mathbb{N}$ there exist pairwise distinct $a_0, \ldots, a_k \in \mathbb{N}$, all of the same color, such that*

$$c_1 a_1^2 + \cdots + c_k a_k^2 = a_0. \tag{6.11}$$

Proof. Consider the quadratic polynomials

$$p(t) = \sum_{\ell=1}^{k} c_\ell (1+\ell t)^2, \qquad q(t) = \sum_{\ell=1}^{k-1} c_\ell (1+\ell t)^2 + c_k (1+2kt)^2.$$

Both have rational coefficients and a root at $t = 0$. On the other hand, the derivatives

$$p'(t) = 2\sum_{\ell=1}^{k} \ell c_\ell (1+\ell t), \qquad q'(t) = 2\sum_{\ell=1}^{k-1} \ell c_\ell (1+\ell t) + 4kc_k(1+2kt)$$

can not both vanish at $t = 0$. Therefore at least one of these polynomials must have a second root at some $t \in \mathbb{Q} \setminus \{0\}$. Assume p has a second root (an analogous argument works in the alternative case). Letting d be the denominator of t and $u_\ell = d(1+\ell t)$ for each $\ell = 1, \dots, k$, we now have pairwise distinct $u_1, \dots, u_k \in \mathbb{Z}$ such that $c_1 u_1^2 + \cdots + c_k u_k^2 = 0$. We can also assume that $c_1 u_1 + \cdots + c_k u_k \neq 0$ by changing some nonzero u_ℓ into $-u_\ell$ if necessary.

Let $b = 2(c_1 u_1 + \cdots + c_k u_k)$. Let $\chi : \mathbb{N} \to \{1, \dots, r\}$ be an arbitrary finite coloring of $\mathbb{N}$ and define a new coloring $\tilde{\chi}$ of $\mathbb{N}$ in $r + b - 1$ colors by:

$$\tilde{\chi}(n) := \begin{cases} \chi\left(\frac{n}{b}\right) & \text{if } n \text{ is divisible by } b \\ r + (n \bmod b) & \text{otherwise} \end{cases}$$

where $n \bmod b \in \{0, 1, \dots, b-1\}$ is the remainder of the division of n by b. Next apply Corollary 6.10 to find $x, y \in \mathbb{N}$ such that the set $\{x, xy, x+y, x+u_1 y, \dots, x+u_k y\}$ is monochromatic with respect to $\tilde{\chi}$.

Observe that, in view of the construction of the coloring $\tilde{\chi}$, all the numbers $x, xy, x+y$ share the same congruence class modulo b, which implies that both x and y are divisible by b. We deduce that the set $\left\{\frac{xy}{b}, \frac{x+u_1 y}{b}, \dots, \frac{x+u_k y}{b}\right\}$ consists of integers and is monochromatic with respect to χ. Letting $a_0 = \frac{xy}{b}$ and $a_\ell = \frac{x+u_\ell y}{b}$ for $\ell = 1, \dots, k$, we have the desired relation (6.11). $\square$

BIBLIOGRAPHY

[Arn70] V. I. Arnautov. “Nondiscrete topologizability of countable rings”. In: *Dokl. Akad. Nauk SSSR* 191 (1970), pp. 747–750. ISSN: 0002-3264.

[Arv76] W. Arveson. *An invitation to C^*-algebras*. Graduate Texts in Mathematics, No. 39. New York: Springer-Verlag, 1976, pp. x+106.

[Aus88] J. Auslander. *Minimal flows and their extensions*. Vol. 153. North-Holland Mathematics Studies. Notas de Matemática [Mathematical Notes], 122. North-Holland Publishing Co., Amsterdam, 1988, pp. xii+265. ISBN: 0-444-70453-1.

[BBHS06] M. Beiglböck, V. Bergelson, N. Hindman, and D. Strauss. “Multiplicative structures in additively large sets”. In: *J. Combin. Theory Ser. A* 113.7 (2006), pp. 1219–1242.

[BBHS08] M. Beiglböck, V. Bergelson, N. Hindman, and D. Strauss. “Some new results in multiplicative and additive Ramsey theory”. In: *Trans. Amer. Math. Soc.* 360.2 (2008), pp. 819–847. ISSN: 0002-9947.

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

[Ber03] V. Bergelson. “Minimal Idempotents and Ergodic Ramsey Theory”. In: *Topics in Dynamics and Ergodic Theory*. Vol. 310. London Math Soc. Lecture Note Ser. Cambridge: Cambridge Univ. Press, 2003, pp. 8–39.

[Ber05] V. Bergelson. “Multiplicatively large sets and ergodic Ramsey theory”. In: *Israel J. Math.* 148 (2005), pp. 23–40. ISSN: 0021-2172.

[Ber06] V. Bergelson. “Combinatorial and Diophantine Applications of Ergodic Theory”. In: *Handbook of Dynamical Systems*. Ed. by B. Hasselblatt and A. Katok. Vol. 1B. Elsevier, 2006, pp. 745–841.

[Ber10] V. Bergelson. “Ultrafilters, IP sets, dynamics, and combinatorial number theory”. In: *Ultrafilters across mathematics*. Vol. 530. Contemp. Math. Providence, RI: Amer. Math. Soc., 2010, pp. 23–47.

[Ber86] V. Bergelson. “A density statement generalizing Schur’s theorem”. In: *J. Combin. Theory Ser. A* 43.2 (1986), pp. 338–343. ISSN: 0097-3165.

[Ber87] V. Bergelson. “Ergodic Ramsey theory”. In: *Logic and combinatorics (Arcata, Calif., 1985)*. Vol. 65. Contemp. Math. Amer. Math. Soc., Providence, RI, 1987, pp. 63–87.

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

[BFM96] V. Bergelson, H. Furstenberg, and R. McCutcheon. “IP-sets and polynomial recurrence”. In: *Ergodic Theory Dynam. Systems* 16.5 (1996), pp. 963–974. ISSN: 0143-3857.

[BH01] V. Bergelson and N. Hindman. “Partition regular structures contained in large sets are abundant”. In: *J. Combin. Theory Ser. A* 93.1 (2001), pp. 18–36. ISSN: 0097-3165.

[BH90] V. Bergelson and N. Hindman. “Nonmetrizable topological dynamics and Ramsey theory”. In: *Trans. Amer. Math. Soc.* 320.1 (1990), pp. 293–320. ISSN: 0002-9947.

[BH94] V. Bergelson and N. Hindman. “On IP^* sets and central sets”. In: *Combinatorica* 14.3 (1994), pp. 269–277. ISSN: 0209-9683.

[BJM] V. Bergelson, J. Johnson, and J. Moreira. “New polynomial and multidimensional extensions of classical partition results”. Submitted, available online at `http://arxiv.org/abs/1501.02408`.

[BL96] V. Bergelson and A. Leibman. “Polynomial extensions of van der Waerden’s and Szemerédi’s theorems”. In: *J. Amer. Math. Soc.* 9.3 (1996), pp. 725–753. ISSN: 0894-0347.

[BL99] V. Bergelson and A. Leibman. “Set-polynomials and polynomial extension of the Hales-Jewett theorem”. In: *Ann. of Math. (2)* 150.1 (1999), pp. 33–75. ISSN: 0003-486X.

[BLL08] V. Bergelson, A. Leibman, and E. Lesigne. “Intersective polynomials and the polynomial Szemerédi theorem”. In: *Adv. Math.* 219.1 (2008), pp. 369–388. ISSN: 0001-8708.

[BLM05] V. Bergelson, A. Leibman, and R. McCutcheon. “Polynomial Szemerédi theorems for countable modules over integral domains and finite fields”. In: *J. d’Analyse Math.* 95 (2005), pp. 243–296. ISSN: 0021-7670.

[BM07] V. Bergelson and R. McCutcheon. “Central sets and a non-commutative Roth theorem”. In: *Amer. J. Math.* 129.5 (2007), pp. 1251–1275. ISSN: 0002-9327.

[BM16a] V. Bergelson and J. Moreira. “Ergodic Theorem involving additive and multiplicative groups of a field and $\{x + y, xy\}$ patterns”. To appear in *Ergodic Theory Dynam. Systems*, available online doi:10.1017/etds.2015.68. 2016.

[BM16b] V. Bergelson and J. Moreira. “Measure preserving actions of affine semigroups and $\{x+y, xy\}$ patterns”. To appear in *Ergodic Theory Dynam. Systems*, available at `http://arxiv.org/abs/1509.07574`. 2016.

[BM16c] V. Bergelson and J. Moreira. “Van der Corput’s difference theorem: some modern developments”. In: *Indag. Math. (N.S.)* 27.2 (2016), pp. 437–479. ISSN: 0019-3577.

[BM96] V. Bergelson and R. McCutcheon. “Uniformity in the polynomial Szemerédi theorem”. In: *Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993–1994)*. Vol. 228. London Math. Soc. Lecture Note Ser. Cambridge: Cambridge Univ. Press, 1996, pp. 273–296.

[Bra28] A. Brauer. “Über Sequenzen von Potenzresten”. In: *Sitzungsberichte de Preussischen Akademie der Wissenschaften, Physicalish-Mathematische Klasse* (1928), pp. 9–16.

[Bro68] T. Brown. “Locally finite semigroups”. In: *Ukrain. Mat. Ž.* 20 (1968), pp. 732–738. ISSN: 0041-6053.

[CGS12] P. Csikvári, K. Gyarmati, and A. Sárközy. “Density and Ramsey type results on algebraic equations with restricted solution sets”. In: *Combinatorica* 32.4 (2012), pp. 425–449. ISSN: 0209-9683.

[Cil12] J. Cilleruelo. “Combinatorial problems in finite fields and Sidon sets”. In: *Combinatorica* 32.5 (2012), pp. 497–511. ISSN: 0209-9683.

[Deu73] W. Deuber. “Partitionen und lineare Gleichungssysteme”. In: *Math. Z.* 133 (1973), pp. 109–123. ISSN: 0025-5874.

[Ell58] R. Ellis. “Distal transformation groups”. In: *Pacific J. Math.* 8 (1958), pp. 401–405. ISSN: 0030-8730.

[ET36] P. Erdős and P. Turán. “On some sequences of integers”. In: *J. London Math. Soc.* 11 (1936), pp. 261–264.

[FH] N. Frantzikinakis and B. Host. “Higher order Fourier analysis of multiplicative functions and applications”. To appear in J. Amer. Math. Soc.; available online at http://arxiv.org/abs/1403.0945.

[FKO79] H. Furstenberg, Y. Katznelson, and D. Orstein. “The ergodic theoretical proof of Szemerédi’s theorem”. In: *Bull. Amer. Math. Soc.* 7 (1979), pp. 427–552.

[Fur77] H. Furstenberg. “Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions”. In: *J. d’Analyse Math.* 31 (1977), pp. 204–256. ISSN: 0021-7670.

[Fur81] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton, N.J.: Princeton University Press, 1981, pp. xi+203.

[FW78] H. Furstenberg and B. Weiss. “Topological dynamics and combinatorial number theory”. In: *J. d’Analyse Math.* 34 (1978), pp. 61–85. ISSN: 0021-7670.

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

[GRS90] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey theory.* Second. John Wiley & Sons, Inc., New York, 1990, pp. xii+196. ISBN: 0-471-50046-1.

[GS16] B. Green and T. Sanders. “Monochromatic sums and products”. In: *Discrete Analysis* (2016:5), pp. 1–43.

[Gun02] D. S. Gunderson. “On Deuber’s partition theorem for (m, p, c)-sets”. In: *Ars Combin.* 63 (2002), pp. 15–31. ISSN: 0381-7032.

[Han13] B. Hanson. “Capturing forms in dense subsets of finite fields”. In: *Acta Arith.* 160.3 (2013), pp. 277–284. ISSN: 0065-1036.

[Hin11] N. Hindman. “Monochromatic sums equal to products in $\mathbb{N}$”. In: *Integers* 11.4 (2011), pp. 431–439. ISSN: 1867-0652.

[Hin74] N. Hindman. “Finite sums from sequences within cells of a partition of N”. In: *J. Combinatorial Theory Ser. A* 17 (1974), pp. 1–11.

[HJ63] A. W. Hales and R. I. Jewett. “Regularity and positional games”. In: *Trans. Amer. Math. Soc.* 106 (1963), pp. 222–229. ISSN: 0002-9947.

[HMS96] N. Hindman, A. Maleki, and D. Strauss. “Central Sets and Their Combinatorial Characterization”. In: *Journal of Combinatorial Theory, Series A* 74.2 (1996), pp. 188 –208. ISSN: 0097-3165.

[HS98] N. Hindman and D. Strauss. *Algebra in the Stone-Čech compactification.* Berlin: Walter de Gruyter & Co., 1998, pp. xiv+485.

[Khi34] A. Khintchine. "Korrelationstheorie der stationären stochastischen Prozesse". In: *Math. Ann.* 109.1 (1934), pp. 604–615. ISSN: 0025-5831.

[McC10] R. McCutcheon. "A variant of the density Hales-Jewett theorem". In: *Bull. Lond. Math. Soc.* 42.6 (2010), pp. 974–980. ISSN: 0024-6093.

[McC99] R. McCutcheon. "An infinitary polynomial van der Waerden theorem". In: *J. Combin. Theory Ser. A* 86.2 (1999), pp. 214–231. ISSN: 0097-3165.

[Mor] J. Moreira. "Monochromatic sums and products in $\mathbb{N}$". Submitted, available at https://arxiv.org/abs/1605.01469.

[Pat88] A. L. T. Paterson. *Amenability.* Vol. 29. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1988, pp. xx+452. ISBN: 0-8218-1529-6.

[Rad33] R. Rado. "Studien zur Kombinatorik". In: *Math. Zeit.* 36 (1933), pp. 242–280.

[San68] J. H. Sanders. *A generalization of Schur's theorem.* Thesis (Ph.D.)–Yale University. ProQuest LLC, Ann Arbor, MI, 1968, p. 41.

[Sch16] I. Schur. "Über die Kongruenz $x^m + y^m \equiv z^m (\mathrm{mod}\ p)$". In: *Jahresbericht der Deutschen Math. Verein.* 25 (1916), pp. 114–117.

[Shk10] I. D. Shkredov. "On monochromatic solutions of some nonlinear equations in $\mathbb{Z}/p\mathbb{Z}$". In: *Mat. Zametki* 88.4 (2010), pp. 625–634. ISSN: 0025-567X.

[Sto37] M. H. Stone. "Applications of the theory of Boolean rings to general topology". In: *Trans. Amer. Math. Soc.* 41.3 (1937), pp. 375–481. ISSN: 0002-9947.

[Sze75] E. Szemerédi. "On the sets of integers containing no k elements in arithmetic progressions". In: *Acta Arith.* 27 (1975), pp. 299–345.

[Sár78] A. Sárközy. "On difference sets of sequences of integers. I". In: *Acta Math. Acad. Sci. Hungar.* 31.1–2 (1978), pp. 125–149. ISSN: 0001-5954.

[Vin14] L. A. Vinh. “Monochromatic sum and product in $\mathbb{Z}/m\mathbb{Z}$”. In: *J. Number Theory* 143 (2014), pp. 162–169. ISSN: 0022-314X.

[Wae27] B. L. van der Waerden. “Beweis einer Baudetschen Vermutung”. In: *Nieuw. Arch. Wisk.* 15 (1927), pp. 212–216.