On partition regular systems

by

Colin Desmarais

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Mathematics

University of Manitoba

Winnipeg, Manitoba

Abstract

An equation or system of equations is called "partition regular in a set S" if and only if for any finite colouring of S a solution to the system is guaranteed to be contained in some colour class. This thesis is a survey of partition regular systems, starting with early results in arithmetic Ramsey theory, including Hilbert's cube lemma, Schur's theorem, and van der Waerden's theorem. A proof is given of Rado's characterization of all finite partition regular systems of homogeneous linear equations, and results concerning infinite and nonlinear partition regular systems are also proved. Several tools, including linear algebra and topology, are used in the proofs in this thesis.

I would like to extend my sincere gratitude to my advisor Dr. David Gunderson for his encouragement, his support, and for pushing me to be the best that I can be. This thesis would not be the same without his wealth of knowledge and the time he devoted to my research.

I would like to thank my advisory committee, Dr. Robert Craigen and Dr. William Kocay, as well Dr. Karen Gunderson, Andrii Arman, and Sergii Tsaturian for their useful notes on this thesis.

I am grateful for the support I received from the Government of Manitoba in the form of a Research Manitoba Scholarship.

To my closest friends Abdellah, Dan, Max, Michael, Stefan, and Thomas, I would like to thank you for making sure I had some fun while I was completing my master's degree.

Finally, I would like to thank my family, especially my mother Jeannette and my father Guy, for their love and encouragement.

Contents

List of Figures

Chapter 1

Introduction and organization of the thesis

For a set S and a positive integer r, an *r-colouring of S* is a function $\Delta : S \to \{1, 2, \ldots, r\}$ and the integers $1, 2, \ldots, r$ are called *colours* (in fact any r-set can be used for colourings, for example, $\Delta : S \to \{\text{red}, \text{blue}\}$ is a 2-colouring of S). For a colour i, the i-th colour class is the set $C_i = \Delta^{-1}(i)$. If S is r-coloured, then a subset $T \subseteq S$ is called *monochromatic* if and only if T is contained in some colour class. Equivalently, an r-colouring is a partition $S = C_1 \cup \cdots \cup C_r$ where the C_i's are the colour classes (some C_i's may be empty). A system of equations $A\boldsymbol{x} = \boldsymbol{b}$ is called *partition regular in S* if and only if for every finite colouring of S there is a solution $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ to the system so that $\{x_1, x_2, \ldots, x_n\}$ is monochromatic.

As a first example, assign each of the elements of $S = \{1, 2, 3, 4, 5\}$ one of two colours, say red or blue. A triple x, y, z, not necessarily all different, that satisfies $x+y-z=0$ is called a *Schur triple*. The following partial 2-colouring of S avoids monochromatic Schur triples $1, 1, 2$; $2, 2, 4$; and $1, 3, 4$:

$$1\ 2\ 3\ 4\ 5$$

If 5 is assigned the colour blue, then $1, 4, 5$ is a blue Schur triple, and if 5 is assigned the colour red, then $2, 3, 5$ is a red Schur triple. A monochromatic Schur triple is unavoidable in any 2-colouring of $\{1, 2, 3, 4, 5\}$.

In fact, in 1916 Schur [112] proved that for any number of colours r, there is a positive integer $n = S(r)$ so that any r-colouring of $\{1, 2, \ldots, n\}$ contains a Schur triple in the same colour class. Two proofs of this fact (known as Schur's theorem) are presented in Section 3.3.

The pigeonhole principle can be stated in terms of colourings: let r and m be positive integers and let S be a set with more than rm elements. Then for any r-colouring of S there are $m+1$ elements in the same colour class. In 1930, Ramsey [101] proved a generalization of the pigeonhole principle: for all positive integers k, m, r there exists a least positive integer $n = R_k(m; r)$ so that for every r-colouring of the k-subsets of $\{1, 2, \ldots, n\}$, there exists an m-subset $\{i_1, i_2, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$ so that all of the k-subsets of $\{i_1, i_2, \ldots, i_m\}$ are monochromatic. Much of Ramsey theory is devoted to finding the "Ramsey number" $R_k(m; r)$. A solution to the "party problem", which is a standard introduction to the topic of Ramsey theory, is given in Chapter 2, as well as a proof of Ramsey's theorem (Theorem 2.1.1).

Some early famous results of Hilbert, Schur, and van der Waerden are now seen as "Ramsey-type" theorems. Hilbert's cube lemma, which was proved by Hilbert [64] in 1892, is considered to be the first result in Ramsey theory. For positive integers $a_0, a_1, \ldots, a_m$, a Hilbert cube $H(a_0, a_1, \ldots, a_m)$ contains for every subset $\{a_{i_1}, a_{i_2} \ldots, a_{i_j}\}$ of $\{a_1, a_2, \ldots, a_m\}$ the sum $a_0 + a_{i_1} + a_{i_2} + \cdots + a_{i_j}$. Hilbert showed that for any positive integers r, m, there is a least positive integer n so that in any r-colouring of $\{1, 2, \ldots, n\}$, there is always a monochromatic Hilbert cube $H(a_0, a_1, \ldots, a_m)$. A very well-studied theorem is van der Warden's theorem. For a positive integer k and positive integers a and d, an arithmetic progression of length k is a set of the form $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$. In 1927, van der Waerden [123] proved that for any positive integers k, r there is a least positive integer $n = W(k; r)$ so that in any r-colouring of $\{1, 2, \ldots, n\}$, there is a monochromatic arithmetic progression of length k. Hilbert's cube lemma, Schur's theorem, and van der Waerden's theorem are examined in Chapter 3, as well as other results of what is called "arithmetic Ramsey theory".

The results of Chapter 3 can be stated in terms of linear systems. For example, consider the matrices

$$A = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & k-1 \end{bmatrix}.$$

Schur's theorem guarantees that for every r-colouring of $\{1, 2, \ldots, S(r)\}$, there is a monochromatic vector $\boldsymbol{x} = (x, y, x + y)$ so that

$$A\boldsymbol{x} = \mathbf{0}.$$

The matrix A is an example of what is called a *partition regular matrix* (see Definition 4.1.1), Rado (a student of Schur's) characterized such matrices as those having what is called the *columns property* (see Definition 4.1.2; Rado's characterizations of partition regular matrices are presented in Chapter 4.

Van der Waerden's theorem (mentioned above) guarantees that for $n = W(k; r)$ and every r-colouring of $\{1, 2, \ldots, n\}$, there is a vector $\boldsymbol{x} = (a, d) \in (\mathbb{Z}^+)^2$ and a vector $\boldsymbol{y} \in \{1, 2, \ldots, n\}^k$ so that the entries of $\boldsymbol{y}$ have all the same colours and

$$B\boldsymbol{x} = \boldsymbol{y}.$$

The matrix B is an example of what is called an *image partition regular matrix* (see Definition 5.1.1), which were studied by Hindman and Leader [72] in 1993. Chapter 5 contains characterizations of image partition regular matrices.

If a subset of the positive integers contains a solution to *every* partition regular system, this subset is called *large*. To characterize large sets, Deuber [23] introduced (m, p, c)-sets in 1973. For positive integers m, p, c, let

$$[-p, p] = \{-p, -(p-1), \ldots, -1, 0, 1, \ldots, p-1, p\},$$

and let $x_0, x_1, \ldots, x_m$ be any positive integers so that for any $i = 0, 1, \ldots, m$, the set

$$R_i = \{cx_i + \lambda_{i+1}x_{i+1} + \lambda_{i+2}x_{i+2} + \cdots + \lambda_m x_m : \lambda_{i+1}, \lambda_{i+2}, \ldots, \lambda_m \in [-p, p]\}$$

is contained in the set of positive integers. Each set R_i contains multiple arithmetic progressions. The set $R_0 \cup R_1 \cup \ldots \cup R_m$ is an example of a (m, p, c)-set. Each λ_i is a "parameter" that varies in P. Section 6.2 is a discussion of the Hales–Jewett theorem, a result for the "Ramsey-type" properties of parameter words introduced by Hales and Jewett [60] in 1963. Section 6.3 includes a proof of a partition regularity result of Deuber using the Hales–Jewett theorem, that for any positive integers m, p, c, r, there are positive integers n, q, d so that any r-colouring of any (n, q, d)-set produces a monochromatic (m, p, c)-set. Using (m, p, c)-sets, Deuber proved a result (first conjectured to be true by Rado) that for every finite colouring of a large set there is a monochromatic large set.

In Chapter 7, the focus is switched to infinite results in partition regularity. The first of these is the "finite sums theorem", first proved by Hindman in 1974 [65]. For a set A, the *finite sums set* of A, denoted $FS(A)$, consists of all possible sums of finitely many distinct elements of A. Hindman proved that under any finite colouring of the positive integers, there is an infinite set A whose finite sum set $FS(A)$ is monochromatic. A proof of Hindman's theorem, first developed by Galvin and Glazer in 1975 (see [76]) using ultrafilters, is presented in Section 7.4. In the last section of Chapter 7, a lack of characterizations for infinite partition regular matrices analogous to those for finite partition regular matrices is discussed.

Chapter 8 contains results of nonlinear partition regular systems. The first of these results presented was proved independently by Sárközy in 1978 [109] and Furstenberg in 1977 [42]. The Sárközy–Furstenberg theorem guarantees that in any finite colouring of the positive integers two elements x, y whose difference is a square have the same colour. The precise statement of the theorem is that for any $\delta > 0$ there is a positive integer n large enough so that any set A with density $|A| > \delta n$ contains two elements whose difference is a square. Chapter 8 includes a proof of the Sárközy–Furstenberg theorem, and statements of other nonlinear partition regularity results.

The field of partition regularity has a long history arising from the broader field of Ramsey theory. A wide range of tools are used to prove results of partition regularity; including the pigeonhole principle, linear algebra, number theory, Fourier analysis, probability, and topology.

The following notation is used throughout this thesis. For two sets A and B, the notation $A \subseteq B$ is used to denote "A is a subset of B", and $A \subsetneq B$ denotes "A is a strict subset of B". Let $\mathbb{Z}$, $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ denote the sets of integers, rational numbers, real numbers, and complex numbers respectively. The sets of positive integers, positive rationals, and positive reals are denoted by $\mathbb{Z}^+$, $\mathbb{Q}^+$, and $\mathbb{R}^+$ respectively.

For $n \in \mathbb{Z}^+$, define $[n] = \{1, 2, \ldots, n\}$. For any set S and $k \in \mathbb{Z}^+$ define $[S]^k = \{T \subseteq S : |T| = k\}$, the set of all of the k- subsets of S (note that $[S]^k$ is not a Cartesian product). The k-subsets of $[n]$ are denoted by $[n]^k$ instead of using double brackets $[[n]]^k$. The set $\{a, a+1, a+2, \ldots, b-1, b\} \subseteq \mathbb{Z}$ is denoted by $[a, b]$.

A *graph* is a pair $G = (V, E)$, where V is non-empty set and $E \subseteq [V]^2$. The elements of V are called *vertices of* G and the elements of E are called *edges of* G. The definition of a graph above does not allow for multiple edges or loops. The *complete graph* on n

vertices V is the graph $K_n = (V, [V]^2)$.

Unless otherwise noted, vectors are assumed to be column vectors. Matrices are denoted by capital letters, and the corresponding lower case letter is used to denote the entries. For example, for an $m \times n$ matrix $A = [a_{i,j}]$, the entry in the i-th row and the j-th entry is a_{ij}.

Chapter 2

Ramsey theory

2.1 Introduction

The study of partition regular systems is closely related to Ramsey theory, which gets its name from Frank Plumpton Ramsey who proved the following theorem:

Theorem 2.1.1 (Ramsey, 1930 [101]). *For every $k, m, r \in \mathbb{Z}^+$, there exists a least integer $n = R_k(m; r)$ such that for every r-colouring $\Delta : [n]^k \to [r]$, there exists a set $M \in [n]^m$ so that $[M]^k$ is monochromatic.*

Notice that when $k = 1$, Ramsey's theorem states that there is $n \in \mathbb{Z}^+$ large enough so that whenever $[n]$ is r-coloured, there are m points with the same colour. By the pigeonhole principle, $R_1(m; r) = r(m - 1) + 1$

A standard example to illustrate Ramsey's theorem is called the "party problem"; what is the smallest number of people attending a party needed to guarantee that at least three people are all friends or three people are all strangers? The attendees can be thought of as the n vertices of a graph and the pairs of people are the edges of the complete graph K_n which are coloured "blue"' for friend and "red" for stranger. The party problem asks for the smallest n so that for any 2-colouring of the edges of K_n, there is a monochromatic triangle K_3.

The answer to the party problem is $n = 6$. To see that $n = 5$ does not suffice, consider the counterexample in Figure 2.1 of a 2-colouring of the edges of K_5 with no monochromatic triangle.

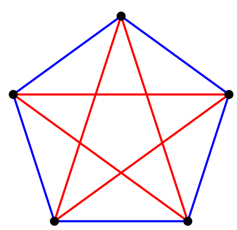

Figure 2.1: K_5 does not guarantee a monochromatic K_3

Figure 2.2 below is an example of a complete graph on 6 vertices. Choose any vertex, say v_1, which is incident with five edges. By the pigeonhole principle, three of the edges, say $\{v_1, v_2\}, \{v_1, v_3\}, \{v_1, v_5\}$, have the same colour, say "red". Then to avoid a monochromatic triangle, the edges $\{v_2, v_3\}, \{v_2, v_5\}$ and $\{v_3, v_5\}$ have to be coloured "blue", which forces a monochromatic triangle.

The party problem is an example of a problem in graph Ramsey theory which, in its simplest form, asks: given $r, m \in \mathbb{Z}^+$, what is the smallest $n \in \mathbb{Z}^+$ so that for any r-colouring of the edges of K_n there is a monochromatic copy of K_m? This number n is

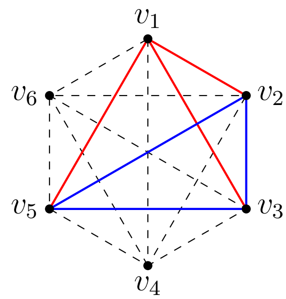

Figure 2.2: K_6 guarantees a monochromatic triangle

precisely the number $R_2(m;r)$ from Ramsey's theorem.

Ramsey proved his theorem while studying a problem in logic. Theorem 2.1.1 was reproved by Erdős and Szekeres [36] by bounding the Ramsey number $R_k(m;r)$ with a recursive process. The Erdős-Szekeres recursion was established to prove a theorem in combinatorial geometry: for any $m \in \mathbb{Z}^+$, there is a least number $n \in \mathbb{Z}^+$ so that if n points are in general position (no three collinear) in the plane, there is a convex m-gon.

In arithmetic Ramsey theory (see Section 3), given $r \in \mathbb{Z}^+$ and a specific arithmetic structure (for example, an arithmetic progression of length three), the goal is to find a smallest $n \in \mathbb{Z}^+$, if it exists, so that every r-colouring of $[n]$ has a monochromatic subset with the desired arithmetic structure. For example Schur's theorem (Theorem 3.3.1) asserts that for any $r \in \mathbb{Z}^+$ there is a smallest $n = S(r)$ so that any r-colouring of $[n]$ contains x, y so that $\{x, y, x+y\}$ is monochromatic. A proof of Schur's theorem using the graph Ramsey number for triangles $R_2(3;r)$ is presented in Section 3.3.

A standard reference for an overview of Ramsey theory is the book by Graham, Rothschild and Spencer [48], [49].

2.2 Ramsey's theorem

Ramsey proved an infinite version of Theorem 2.1.1 and then proved the finite version as a consequence. To go from the infinite version to the finite version, the proof here uses König's lemma. A *tree* is a graph that contains no cycles, and a *rooted tree* is a tree with a vertex v_0 identified as the root. Every vertex adjacent to the root is called a *child* of v_0, for every other vertex $v \neq v_0$ the vertex w adjacent to v along the path from v_0 to v is called the *parent* of v, and all remaining vertices adjacent to v are called its *children*. If every vertex in a tree has finite degree, then the tree is called *locally finite*.

Lemma 2.2.1 (König, 1927 [81])**.** *Every infinite locally finite tree has an infinite path.*

Proof. Let $T = (V, E)$ be an infinite tree with root v_0 and neighbours $w_{0,1}, \ldots, w_{0,n_0}$. Let $T_{0,1}, \ldots, T_{0,n_0}$ be the components of $T \setminus \{v_0\}$. For each $i \in [n_0]$, $T_{0,i}$ is a subtree of T rooted at w_i. Since $V(T) = \{v_0\} \cup V(T_{0,1}) \cup \cdots \cup V(T_{0,n_0})$ and n_0 is finite, by the infinite pigeonhole principle, one of the trees, say $T_{0,i}$, is infinite. Set $T_1 = T_{0,i}$ with root $v_1 = w_{0,i}$.

Let $w_{1,1}, \ldots, w_{1,n_1}$ be the children of v_1 and let $\{T_{1,1}, \ldots, T_{1,n_1}\}$ be the components of the subtree $T_1 \setminus \{v_1\}$. By a similar argument as above, one of the trees $T_{1,j}$ is infinite. Set $T_2 = T_{1,j}$ with root $v_2 = w_{1,j}$.

Repeat the argument above, at every step $k \geq 1$ finding an infinite subtree T_k of T with root v_k. If the sequence of roots $V = \{v_0, v_1, v_2, \ldots, v_k, \ldots\}$ is finite, then there is some $l < \infty$ so that for the infinite subtree T_l, the components of $T_l \setminus \{v_l\}$ are all finite, which contradicts the infinite pigeonhole principle. Therefore, the sequence of roots V is infinite, and since each v_i is the parent of v_{i+1}, the sequence $v_0, v_1, v_2, \ldots, v_k, \ldots$ is an infinite path in T. $\square$

Theorem 2.2.2 (Ramsey, 1930 [101]). *Let A be a countably infinite set. Then for every $k, r \in \mathbb{Z}^+$, and every r-colouring $\Delta : [A]^k \to [r]$, there exists an infinite set $M \subseteq A$ so that $[M]^k$ is monochromatic.*

Proof. Fix $r \geq 1$. The proof is by induction on $k \geq 1$.

Base Case: The base case when $k = 1$ is simply the infinite pigeonhole principle.

Inductive Step: For a fixed $k \geq 1$, suppose the theorem holds for a fixed $k \geq 1$. Let $\Delta : [A]^{k+1} \to [r]$ be any r-colouring. For any element $a_1 \in A$, the colouring Δ induces an r-colouring $\Delta_1 : [A \setminus \{a_1\}]^k \to [r]$ defined for each $K \in [A \setminus \{a_1\}]^k$ by $\Delta_1(K) = \Delta(K \cup \{a_1\})$. By the induction hypothesis, there is an infinite set $A_1 \subseteq A \setminus \{a_1\}$ so that $[A_1]^k$ is monochromatic with respect to Δ_1, and therefore $\{K \cup \{a_1\} : K \in [A_1]^k\}$ is monochromatic with respect to Δ, say with colour r_1. Assign the colour r_1 to the element a_1. For any element $a_2 \in A_1$, the colouring Δ induces an r-colouring $\Delta_2 : [A_1 \setminus \{a_2\}]^k \to [r]$. Then, again by the induction hypothesis, there is an infinite set $A_2 \subseteq A_1 \setminus \{a_2\}$ so that $[A_2]^k$ is monochromatic with respect to Δ_2, and so $\{K \cup \{a_2\} : K \in [A_2]^k\}$ is monochromatic with respect to Δ, say with colour r_2 (note, r_2 may or may not be the same colour as r_1). Assign the colour r_2 to the element a_2.

Continue the argument above for $j \geq 2$: at each step choose $a_j \in A_{j-1}$ and, by the induction hypothesis, let $A_j \subseteq A_{j-1} \setminus \{a_j\}$ be an infinite set for which $\{K \cup \{a_j\} : K \in$

$[A_j]^k\}$ is monochromatic with respect to Δ; say with colour r_j. Assign the colour r_j to the element a_j.

The infinite set $\{a_1, a_2, \ldots, a_j, \ldots\}$ is r-coloured, and so by the infinite pigeonhole principle, there is a monochromatic infinite subset $M = \{a_{i_1}, a_{i_2}, \ldots, a_{i_j}, \ldots\}$, say with the colour s. For any $K \subseteq [M]^{k+1}$, let j_i be the smallest index of the elements in K. Then $K \setminus \{a_{j_i}\} \subseteq A_{j_i}$, and so $\Delta(K) = r_{i_j} = s$. Therefore, the set $[M]^{k+1}$ is monochromatic, concluding the inductive step.

By the principle of mathematical induction, Ramsey's theorem is proved. □

The proof of the finite version of Ramsey's theorem now follows by König's lemma.

Proof of Theorem 2.1.1. Assume, in hopes of contradiction, that the theorem is false. That is, for some k, m, r and every $n \in \mathbb{Z}^+$, there is a colouring of $[n]^k$ for which there is no set $M \in [n]^m$ so that $[M]^k$ is monochromatic. Call such a colouring a "bad" colouring. For any $i \in [n-1]$, a "bad" colouring of $[n]$ restricted to $[n-i]$ is a "bad" colouring of $[n-i]$. Construct a tree T as follows: let 0 be the root with r children representing all possible colours of 1. At every generation n, the children of each vertex represent a different colour of the element $n+1$ so that the colouring of $[n+1]$ represented by the vertices along the path from the root to $n+1$ is a "bad" colouring. The tree T, by assumption, is infinite. Each vertex has degree at most r, and so by Lemma 2.2.1, there is an infinite path. This infinite path represents a "bad" colouring for $\mathbb{Z}^+$, contradicting Theorem 2.2.2. □

Chapter 3

Arithmetic Ramsey theory

3.1 Introduction

Although Ramsey published Theorem 2.1.1 and Theorem 2.2.2 in 1930, many "Ramsey-type" results predate Ramsey's theorem. These include what is known as "Hilbert's cube lemma" (Theorem 3.2.2) published in 1892, Schur's theorem (Theorem 3.3.1) published in 1916, van der Waerden's theorem (Theorem 3.4.4) published in 1927, and the Schur–Brauer theorem (Theorem 3.4.9) published in 1928.

These theorems guarantee a monochromatic set with a specific arithmetic structure under every finite colouring of $\mathbb{Z}^+$. For example, Schur's theorem guarantees a monochromatic set $\{x, y, x + y\}$, while van der Waerden's theorem guarantees a monochromatic arithmetic progression.

While today these theorems are often discussed in the context of Ramsey theory, they were often proved for other reasons; Hilbert proved his cube lemma while investigating the irreducibility of rational functions with integer entries. Schur proved his lemma while reproving a modular version of Fermat's last theorem, first proved by Dickson [22] in 1909. Van der Waerden's theorem and its extension to the Schur–Brauer theorem arose from a conjecture of Schur (see Theorem 3.4.1) that for every $k \in \mathbb{Z}^+$ and every prime p large enough, there are k consecutive integers that are quadratic residues modulo p and k consecutive integers that are non-quadratic residues. See [97] for more details on the motivations of early results in arithmetic Ramsey theory.

3.2 Hilbert's cube lemma

Hilbert's cube lemma, which is the first "Ramsey-type" result, was published 38 years before Ramsey's theorem. The lemma is a Ramsey statement about the following sets:

Definition 3.2.1. *For $a_0 \in \mathbb{Z}$ (usually, $a_0 \in \mathbb{Z}^+$) and $a_1, a_2, \ldots, a_m \in \mathbb{Z}^+$, not necessarily all different, define the m-*dimensional Hilbert cube

$$H(a_0, a_1, \ldots, a_m) = \left\{ a + \sum_{i \in I} a_i : I \subseteq [m] \right\}.$$

The set $H(a_0, a_1, \ldots, a_m)$ is sometimes called an affine m-cube

Theorem 3.2.2 (Hilbert, 1892 [64])**.** *For every $m, r \in \mathbb{Z}^+$, there exists a least integer $n = h(m; r)$ so that for every r-colouring $\Delta : [n] \to [r]$, there is a monochromatic m-dimensional Hilbert cube.*

The proof presented here is standard (see, for example, the solution to problem 12 in Section 14 of [89]).

Proof. Fix $r \in \mathbb{Z}^+$. The proof is by induction on $m \geq 1$. For $m \in \mathbb{Z}^+$, let $P(m)$ be the statement that there is a least $h(m;r) \in \mathbb{Z}^+$ satisfying the statement of the theorem for m and r.

BASE CASE: For $m = 1$, no matter how $[r+1]$ is r-coloured, the pigeonhole principle guarantees that two elements $a < b$ are the same colour. Then $H(a, b-a) = \{a, b\}$ is monochromatic, and so $h(1;r) \leq r+1$ satisfies $P(1)$.

INDUCTIVE STEP: Let $k \geq 1$, assume $P(k)$ is true, and set $h = h(k;r)$. Let

$$\Delta : [(r^h + 1)h] \to [r]$$

be any r-colouring. For $i \in [r^h + 1]$, define the *block* $B_i = [(i-1)h + 1, ih]$. Then Δ induces an r^h-colouring

$$\Delta' : \{B_1, B_2, \ldots, B_{r^h+1}\} \to [r^h]$$

of the blocks by

$$\Delta'(B_i) = \big(\Delta((i-1)h+1), \Delta((i-1)h+2), \ldots, \Delta(ih)\big).$$

Since there are $r^h + 1$ blocks and r^h colours, by the pigeonhole principle, two of these blocks, say B_i, B_j with $i < j$, have the same colour pattern under Δ'. Since B_i has h elements, there is a k-dimensional Hilbert cube $H(a_0, a_1, \ldots, a_k) \subseteq B_i$ that is monochromatic. Since $\Delta'(B_i) = \Delta'(B_j)$, then $H(a_0 + (j-i)h, a_1, \ldots, a_k) \subseteq B_j$ is also monochromatic with the same colour. Therefore, the $k+1$-dimensional Hilbert cube

$$H(a_0, a_1, \ldots, a_k, a_{k+1}) = H(a_0, a_1, \ldots, a_k) \cup H(a_0 + a_{k+1}, a_1, \ldots, a_k)$$

is monochromatic, and so $h(k+1; r) \leq (r^h + 1)h$ satisfies the statement of the theorem.

By the principle of mathematical induction, $P(m)$ holds for all $m \in \mathbb{Z}^+$. □

Hilbert's original proof [64], as pointed out by Brown *et al.* [16], made use of the Fibonacci numbers $F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, \ldots$ to give an upper bound of

$$h(m; r) < (r+1)^{F_{2m}}.$$

The original proof in German can be found in [97], while an English version of Hilbert's argument can be found in [113]. Gunderson and Rödl [58] note that this implies that for $c_1 \sim 2.6$,

$$h(m; r) < r^{c_1^m}.$$

In 1998, Gunderson and Rödl [58] give the following bounds on $h(m; r)$ when $m \geq 3$ and $r \geq 2$:

$$r^{(1-o(1))(2^m-1)/m} \leq h(m; r) \leq (2r)^{2^{m-1}},$$

where $o(1)$ denotes a function f for which $\lim_{n\to\infty} \frac{f(n)}{n} = 0$.

3.3 Schur's theorem

Theorem 3.3.1 (Schur, 1916 [112])**.** *For any $r \in \mathbb{Z}^+$, there exists a least positive integer $n = S(r)$ so that for any r-colouring $\Delta : [n] \to [r]$, there exists a monochromatic set*

$\{x, y, z\}$ *so that* $x + y = z$.

A set $\{x, y, z\}$ that satisfies $x + y = z$ is called a *Schur triple*. A Schur triple is contained in a 2-dimensional Hilbert cube $H(0, x, y)$. As discussed in [97], Schur's main goal in proving Theorem 3.3.1 was to reprove the following result in number theory;

Theorem 3.3.2 (Dickson, 1909 [22]). *For all* $m \in \mathbb{Z}^+$, *there is a number* n *so that for every prime* $p > n$, *the equation*

$$x^m + y^m \equiv z^m \pmod p$$

has a non-trivial solution.

Schur's proof of Dickson's theorem can be found in [48].

Two proofs of Theorem 3.3.1 are presented here. The first is Schur's original proof [112], which was written in German. An English translation is available in [97].

First proof of Theorem 3.3.1: Fix $r \in \mathbb{Z}^+$. The goal of the proof is to show that if for some $n \in \mathbb{Z}^+$ there is an r-colouring that is free of a monochromatic Schur triple $\{x, y, x+y\}$ then $n < r!e$, where e is the base of the natural logarithm, and so $S(r) \leq r!e$.

Let $\Delta : [n] \to [r]$ be an r-colouring that is free of monochromatic Schur triples. Let c_1 be the colour most used, let C_1 be all elements with colour c_1, and $n_1 = |C_1|$. Then for every other colour, at most n_1 elements have that colour, and so

$$n \leq rn_1. \tag{3.1}$$

Let m_1 be the smallest element in C_1 and

$$N_1 = \big\{m - m_1 : m \in C_1 \setminus \{m_1\}\big\}.$$

No element $(m - m_1) \in N_1$ is assigned the colour c_1, otherwise $\{m_1, (m - m_1), m\}$ is a monochromatic Schur triple.

For $k \geq 2$, define recursively c_k to be the colour most used in N_{k-1}, C_k to be all elements of N_{k-1} assigned the colour c_k, and $n_k = |C_k|$. Stop the recursion if $n_k = 1$, otherwise let m_k be the smallest element in C_k and

$$N_k = \big\{m - m_k : m \in C_k \setminus \{m_k\}\big\}.$$

If any element $(m - m_{k-1}) \in N_{k-1}$ has the colour c_{k-1}, then $\{m_{k-1}, (m - m_{k-1}), m\}$ is a monochromatic Schur triple. For $i < k - 1$, suppose some element $m \in N_{k-1}$ is assigned the colour c_i. Then for some $m', m'' \in C_i$,

$$m = m' - m_i - m_{i-1} - \cdots - m_{k-2} - m_{k-1}$$

and

$$m_{k-1} = m'' - m_i - m_{i-1} - \cdots - m_{k-2}.$$

Therefore, since m' and m'' are assigned the colour c_i and

$$m'' + m = (m_{k-1} + m_i + \cdots + m_{k-2}) + (m' - m_i - \cdots - m_{k-2} - m_{k-1}) = m',$$

the set $\{m'', m, m'\}$ is a monochromatic Schur triple. Therefore, at most $r - k + 1$ colours are used to colour the elements of N_{k-1}, and so C_k contains at least $\frac{|N_{k-1}|}{r-k+1}$ of the elements of N_{k-1}. Since $|N_{k-1}| = n_{k-1} - 1$,

$$n_{k-1} \leq (r - k + 1)n_k + 1. \tag{3.2}$$

Note that if $k = r$, then there are no longer any colours to colour the elements of N_r, and so C_r contains a single element m_r. Hence if it exists, $n_r = 1$. In any case, $n_t = 1$ for some $t \leq r$. Then by applying the inequalities (3.1) and (3.2),

$$\begin{aligned}
n &\leq rn_1 \\
&\leq r((r-1)n_2 + 1) \\
&\leq r((r-1)((r-2)(\dots(r-t+2)((r-t+1)n_t + 1) + 1) + \dots + 1) + 1) \\
&= \frac{r!}{(r-t)!}\left(n_t + \sum_{i=1}^{t-1} \frac{(r-t)!}{(r-i)!}\right) \\
&\leq r!\left(1 + \sum_{i=1}^{r-1} \frac{1}{(r-i)!}\right) \\
&= r! \sum_{i=0}^{r-1} \frac{1}{i!}.
\end{aligned}$$

Since $e = \sum_{i=0}^{\infty} \frac{1}{i!}$, then $n \leq r! \sum_{i=0}^{r-1} \frac{1}{i!} < r!e$, which proves the theorem. □

A modern proof of Schur's theorem published by Abbott and Moser [2] uses Ramsey's theorem, and such a proof provides a link between Theorem 3.3.1 and graph Ramsey numbers for monochromatic K_3's.

Second proof of Theorem 3.3.1. Let n_r be the Ramsey number $R_2(3, r)$ guaranteed to exist by Theorem 2.1.1, and let $\Delta : [n_r - 1] \to [r]$ be an r-colouring. Colour the edges of the complete graph $K_{n_r} = ([n_r], [n_r]^2)$ by $\Delta(\{i, j\}) = \Delta(|i-j|)$. By the choice of n_r, there is a monochromatic K_3, that is, there is $i < j < k$, so that $\Delta(j-i) = \Delta(k-j) = \Delta(k-i)$. Letting $x = j - i, y = k - j$, and $z = k - i$, then $x + y = j - i + k - j = k - i = z$. □

It is trivial that $S(1) = 2$, and as seen in Chapter 1, $S(2) = 5$. With a little more work, it can be shown that $S(3) = 14$ (see [5]). The number $S(4) = 45$ was proved using a computer search by Baumert and Golomb in 1965 [5]. Computer searches were also used to show that $S(5) \geq 161$ [37], $S(6) \geq 537$ and $S(7) \geq 1681$ [41].

For general lower bounds, in 1979 Fredricksen [40] gave the following lower bound: for an appropriate constant c and for all $r \geq 5$,

$$S(r) \geq c(315)^{r/5} > c(3.1598)^r.$$

Fredricksen improved Schur's [112] bound $S(r) \geq \frac{3^r+1}{2}$. A proof of Schur's lower bound is presented here (and can also be found in [84]).

Theorem 3.3.3 (Schur, 1916 [112])**.** *For* $r \in \mathbb{Z}^+$, $S(r) \geq \frac{3^r+1}{2}$.

Proof. The proof is by induction on $r \geq 1$.

Base Case: For $r = 1$, setting $x = 1, y = 1$ gives $\{x, y, x + y\}$ monochromatic, so $S(1) = 2 \geq \frac{3+1}{2}$.

Inductive Step: Let $r \in \mathbb{Z}^+$, and suppose $n = S(r) - 1 \geq \frac{3^r+1}{2} - 1$. Let $\Delta : [n] \to [r]$ be an r-colouring that avoids a monochromatic set $\{x, y, x + y\}$. Define a new colouring

$\Delta' : [3n+1] \to [r+1]$ by

$$\Delta'(i) = \begin{cases} \Delta(i) & i \in [n] \\ r+1 & i \in [n+1, 2n+1] \\ \Delta(i-(2n+1)) & i \in [2n+2, 3n+1]. \end{cases}$$

Let $x, y \in [3n+1]$ be elements with the same colour r' under Δ' and set $z = x + y$. If $x, y \in [n]$, then $z \leq 2n$, and so either $z \in [n]$ and $\Delta'(z) = \Delta(z) \neq r'$ by the induction hypothesis, or $z \in [n+1, 2n+1]$ and $\Delta'(z) = r+1 \neq r'$. If both $x, y \in [2n+1, 3n+1]$, then $z \geq 4n+4 > 3n+1$, and so $z \notin [3n+1]$. Suppose either x or y is in $[n+1, 2n+1]$, without loss of generality $x \in [n+1, 2n+1]$. If $y \in [n]$ or $[2n+2, 3n+1]$, then $\Delta'(y) \neq r+1 = \Delta'(x)$. If $y \in [n+1, 2n+1]$, then $z \geq 2n+2$, and so $\Delta'(x+y) \neq r+1$. The last case is when one of x, y is in $[n]$ and the other is in $[2n+2, 3n+1]$. Without loss of generality, $x \in [n]$ and $y \in [2n+2, 3n+1]$. Since $\Delta'(x) = \Delta'(y) = r'$, then $\Delta'(y-(2n+1)) = r'$. But by the induction hypothesis, $\Delta'(x+y-(2n+1)) = \Delta(x+y-(2n+1)) \neq r'$, therefore, $\Delta'(z) = \Delta'(x+y-(2n+1)) \neq r'$. In all cases, there is no monochromatic $\{x, y, x+y\}$. Therefore,

$$S(r+1) \geq 3n+2 = 3(S(r)-1)+2 = 3\left(\frac{3^r+1}{2}\right) - 1 = \frac{3^{r+1}+1}{2},$$

which concludes the inductive step, and so the result follows from the principle of mathematic induction. □

In 1973, Irving [79] gave the upper bound

$$S(r) < r!\left(e - \frac{1}{24}\right). \tag{3.3}$$

Irving gives credit for (3.3) to Whitehead [125] who showed that $49 \leq R_2(3;4) \leq 65$ and as a consequence that $R_2(3;r) < r!\left(e - \frac{1}{24}\right) + 1$. The bound (3.3) is derived from $S(r) \leq R_2(3;r) - 1$, which is the second proof of Theorem 3.3.1. Whitehead credits Folkman for part of the proof of the upper bound $R_2(3;4) \leq 65$; Folkman's work was posthumously published in 1974 [38].

3.4 Van der Waerden's theorem

3.4.1 Van der Waerden numbers

Schur conjectured (see [97]) the following about the distribution of quadratic residues and non-quadratic residues modulo p, which was proved by Brauer:

Theorem 3.4.1 (Brauer 1928, [15]). *For every $k \in \mathbb{Z}^+$, there exists a positive integer $n = n(k)$ so that for every prime $p > n$, there are k consecutive integers that are quadratic residues modulo p, and k consecutive integers that are non-quadratic residues modulo p.*

While working on this conjecture, Schur also conjectured what is now known as van der Waerden's theorem, which is a statement about arithmetic progressions.

Definition 3.4.2. *For $k \in \mathbb{Z}^+$, and any $a, d \in \mathbb{Z}^+$, the set*

$$\{a, a+d, a+2d, \ldots, a+(k-1)d\}$$

is called an arithmetic progression of length k, *denoted AP_k.*

Remark 3.4.3. *Note that when $k = 1$, and AP_1 is a single element $\{a\}$, and when $k = 2$, any pair of elements a, b with $a < b$ forms an AP_2 $\{a, a + (b - a)\}$.*

Theorem 3.4.4 (van der Waerden, 1927 [123])**.** *For any $k, r \in \mathbb{Z}^+$, there exists a least positive integer $n = W(k; r)$ such that for any r-colouring $\Delta : [n] \to [r]$, there is a monochromatic AP_k.*

The first non-trivial case is when $k = 3$. It is a short exercise to see that $W(3; 2) = 9$ (see the introduction to [80] for example). Chvátal [18] showed that $W(3; 3) = 27$. In 1979, Beeler and O'Neil [7] showed that $W(3; 4) = 76$.

Bounds on $W(3; r)$ can be achieved by density arguments (see Section 9.2). For example, In 1946, Behrend [12] proved that there exists a constant c so that for n sufficiently large, there exists an AP_3-free set $B \subseteq [n]$ with $|B| \geq ne^{-c\sqrt{\ln n}}$. Behrend's proof is reworked by Gunderson and Rödl [58] to give a lower bound $r^{\frac{1}{9}\ln r} < W(3; r)$. Upper bounds are also reached by using density results for AP_3's, namely Roth's theorem [102, 103], which states that there exists a constant c so that for n sufficiently large, if $B \subseteq [n]$ has $\frac{cn}{\ln \ln n}$ elements, then B contains an AP_3. Roth's theorem was improved by Bourgain [14], which then gives for an appropriate constant c, the upper bound for the van der Waerden number

$$W(3; r) < r^{cr^2}.$$

See [55] for a discussion of bounds on $W(k; r)$, as well as a slight improvement on a weaker upper bound of

$$W(3;r) < 2^{2^{cr}}$$

achieved by combinatorial methods by Huang and Yang [78].

Exact values for $W(k;r)$ are known for only a few values of k and r. Chvàtal [18] showed that $W(4;2) = 35$. Stevens and Shantaram [115] in 1978 showed that $W(5;2) = 178$, and Kouril [82] showed that $W(4;3) = 293$ in 2012. Many computer searches have been performed to prove lower bounds of other van der Waerden numbers (see for example [27], [83], [63], and [98]; note that some of the authors denote van der Waerden numbers by $W(r;k)$ instead of $W(k;r)$).

For lower bounds, Erdős and Rado proved the following in 1952:

Theorem 3.4.5 (Erdős and Rado, 1952 [35])**.** *For any $k, r \in \mathbb{Z}^+$,*

$$W(k+1;r) > \sqrt{2kr^k}.$$

Proof. For $k, r \in \mathbb{Z}^+$, suppose $n \in \mathbb{Z}^+$ is a number so that for every r-colouring $\Delta : [n] \to [r]$, there is a monochromatic AP_{k+1}. To count the number of AP_{k+1}'s in $[n]$, there are at most $\frac{n-1}{k}$ possible values for d, and with d chosen, $n - kd$ possible values for a. Therefore the total number of AP_{k+1}'s is at most

$$\begin{aligned} A &= \sum_{d=1}^{\lfloor (n-1)/k \rfloor} (n - kd) \\ &< \frac{n(n-1)}{k} - \frac{(n-1)^2}{2k} \\ &< \frac{n^2}{2k}. \end{aligned}$$

There are r^n possible r-colourings of $[n]$. For each AP_{k+1}, there are $r \cdot r^{n-k-1} = r^{n-k}$ colourings with that AP_{k+1} monochromatic. Since, by assumption, for every colouring there is a monochromatic AP_{k+1}, then the number of colourings of $[n]$ is less than the number of AP_{k+1}'s times the number of colourings where a specific AP_{k+1} is monochromatic, that is, $r^n \leq Ar^{n-k}$. Solving for A,

$$r^k \leq A < \frac{n^2}{2k}.$$

Therefore, $n > \sqrt{2kr^k}$. □

For $r = 2$, the lower bound of Theorem 3.4.5 has been improved. Szabó [117] proved in 1990 that for every $\epsilon > 0$ and k large enough, $W(k;2) \geq \frac{2^k}{k^\epsilon}$. For any prime p, in 1968 Berlekamp [10] proved that $W(p+1;2) \geq p2^p$. For more colours, Schmidt showed that there exists a constant c so that $W(k+1;r) > r^{k-c\sqrt{k \ln k}}$ [111], which is improved by Moser [94] when k is relatively small compared to r to $W(k+1;r) > kr^{c \ln t}$, for another constant c. Abbott and Liu [1] showed that for $s \in \mathbb{Z}^+$, there is $c = c(s)$ so that for any k such that $2^s \leq k+1 < 2^{s+1}$, $W(k+1;r) > r^{c(\ln r)^s}$.

For upper bounds on $W(k;r)$, the proof of Theorem 3.4.4 (see Section 3.4.2) gives an upper bound that is so large, it is not even primitive recursive (see [19] for a definition of primitive recursive). In 1988, Shelah [114] was able to provide an upper bound that is primitive recursive. In 2001, Gowers [45] proved that

$$W(k;r) \leq 2^{2^{r^{2^{2^{k+9}}}}},$$

which improves Shelah's upper bound. Gowers' bound was achieved by a density argument, see Section 9.2 for a discussion on density results.

3.4.2 Proof of van der Waerden's theorem

The proof of van der Waerden's theorem (Theorem 3.4.4) presented here follows a proof given by Graham and Rothschild [47], which can also be found in [48] and [97]. Theorem 3.4.4 is also proved in Chapter 6 using the Hales–Jewett theorem (see Section 6.2). For a survey of several proofs as well as generalizations of van der Waerden's theorem, see [80].

To introduce how the proof below of van der Waerden's theorem is achieved, a proof that $W(3;2) \leq 325$ is given first.

Theorem 3.4.6 (Graham and Rothschild, 1980 [48])**.** *For every 2-colouring* $\Delta : [325] \to \{red, blue\}$*, there is a monochromatic* AP_3*.*

Proof. Let $\Delta : [325] \to \{red, blue\}$ be any 2-colouring. Consider the sets $B_1, \ldots, B_{65}$ where $B_i = [5i-4, 5i]$. Then Δ imposes a 32-colouring Δ' on $\{B_1, B_2, \ldots, B_{65}\}$ by colouring each block with a 5-tuple:

$$\Delta'(B_i) = \big(\Delta(5i-4), \Delta(5i-3), \Delta(5i-2), \Delta(5i-1), \Delta(5i)\big).$$

Then by the pigeonhole principle, there are i, j so that $i < j \leq 33$ and $\Delta'(B_i) = \Delta'(B_j)$. Let $d_1 = 5(j-i)$. By the pigeonhole principle, two of the first three elements of B_i are assigned the same colour under Δ, say x and $x + d_2$, where $d_2 \in \{1, 2\}$. If $x + 2d_2$ is also assigned the same colour, then $\{x, x+d_2, x+2d_2\}$ is a monochromatic AP_3. Otherwise, since B_i and B_{i+d_1} are assigned the same colour under Δ', then x and $x+d_1+d_2$ are assigned the same colour under Δ, say red, and $x+2d_2$ and $x+d_1+2d_2$ are assigned the other colour, blue. Consider $y = x+2d_1+2d_2 \leq 325$. If y is red, then

$$\{x, x+d_1+d_2, x+2d_1+2d_2\}$$

is a monochromatic AP_3, and if y is blue, then

$$\{x+2d_2, x+d_1+2d_2, x+2d_1+2d_2\}$$

is a monochromatic AP_3, completing the proof. □

Before starting the general proof of van der Waerden's theorem, some notation is given. For fixed $m, k \in \mathbb{Z}^+$, define an equivalence relation $\sim$ on the Cartesian product $[0,k]^m$ in the following way: let $\boldsymbol{g} = (g_1, g_2, \ldots, g_m)$ and $\boldsymbol{h} = (h_1, h_2, \ldots, h_m)$ be any two elements in $[0,k]^m$. Let $i_{\boldsymbol{g}} = 0$ if k does not appear in $\boldsymbol{g}$ and set $i_{\boldsymbol{g}} = \max\{i : g_i = k\}$ otherwise; similarly define $i_{\boldsymbol{h}}$. Then $\boldsymbol{g} \sim \boldsymbol{h}$ if and only if $i_{\boldsymbol{g}} = i_{\boldsymbol{h}}$ and for all $i \leq i_{\boldsymbol{g}}$, $g_i = h_i$. For example, here are all the equivalence classes $[\boldsymbol{g}]_\sim$ when $m = k = 2$:

$$[(0,0)]_\sim = \big\{(0,0),(0,1),(1,0),(1,1)\big\}$$

$$[(2,0)]_\sim = \big\{(2,0),(2,1)\big\}$$

$$[(0,2)]_\sim = \big\{(0,2)\big\}$$

$$[(1,2)]_\sim = \big\{(1,2)\big\}$$

$$[(2,2)]_\sim = \big\{(2,2)\big\}$$

It is straightforward to see that $\sim$ is reflexive, symmetric and transitive, and so is indeed an equivalence relation.

To prove van der Waerden's theorem, the following stronger theorem is proved.

Theorem 3.4.7 (Graham and Rothschild, 1974 [47]). *For any $k, m, r \in \mathbb{Z}^+$, there exists a least integer $n = S(k, m; r)$ so that for any r-colouring $\Delta : [n] \to [r]$, there exists $a, d_1, \ldots, d_m \in \mathbb{Z}^+$ so that $a + k \sum_{i=1}^{m} d_i \leq n$, and for any $\boldsymbol{g}, \boldsymbol{h} \in [0, k]^m$ with $\boldsymbol{g} \sim \boldsymbol{h}$,*

$$\Delta\left(a + \sum_{i=1}^{m} g_i d_i\right) = \Delta\left(a + \sum_{i=1}^{m} h_i d_i\right).$$

Remark 3.4.8. *When $m = 1$, one of the equivalence classes is the set*

$$[(0)]_\sim = \{(0), (1), \ldots, (k-1)\},$$

and so Theorem 3.4.7 states that for any r-colouring of $[S(k, 1; r)]$, there exists a, d so that $a, a + d, a + 2d, \ldots, a + (k-1)d$ have the same colour and so is a monochromatic AP_k. The other equivalence class is $[(k)]_\sim = \{(k)\}$, and $\{a + kd\}$ is monochromatic. Therefore, $W(k; r) \leq S(k, 1; r)$. The inequality is not necessarily an equality, since van der Waerden's theorem does not require $a + k \cdot d_i \leq n$.

Proof of Theorem 3.4.7. Fix $r \in \mathbb{Z}^+$. The proof is by double induction on k and m. For $k = m = 1$, $S(1, 1; r) = 2$ satisfies the theorem with $a = d = 1$, since both $\{1\}$ and $\{2\}$ form AP_1's. ($S(1, 1; r) = 2$ is needed, since the condition that $a + d \leq S(1, 1; r)$ must be satisfied.) This proves the base step.

The inductive steps are achieved by applying the following two inequalities:

$$S(k, m+1; r) \leq S(k, m; r) \cdot S(k, 1; r^{S(k,m;r)}), \tag{3.4}$$

$$S(k+1, 1; r) \leq S(k, r; r). \tag{3.5}$$

To prove (3.4), set $M = S(k, m; r)$ and $N = S(k, 1; r^{S(k,m;r)})$ and let

$$\Delta : [MN] \to [r]$$

be any r-colouring. For $i \in [N]$, define the N blocks $B_i = [(i-1)M + 1, iM]$. Then Δ imposes a r^M-colouring

$$\Delta' : [N] \to [r^M]$$

on the indices of the blocks B_i by

$$\Delta'(i) = \big(\Delta((i-1)M + 1), \Delta((i-1)M + 2), \ldots, \Delta(iM)\big).$$

By the choice of N, there is a monochromatic AP_k in the colouring Δ' of $[N]$. Therefore, there is b, d so that for every $j \in [M]$, the j-th entry of each of the blocks

$$B_b, B_{b+d}, \ldots, B_{b+(k-1)d}$$

have the same colour under Δ. The block B_b has M elements, and so by the choice of M, there exists $a, d_1, \ldots, d_m$ satisfying the statement of the theorem for k and m. Consider any $\boldsymbol{g}, \boldsymbol{h} \in [0, k]^{m+1}$ with $\boldsymbol{g} \sim \boldsymbol{h}$. If k appears in the last entry of $\boldsymbol{g}$ and $\boldsymbol{h}$, then $\boldsymbol{g} = \boldsymbol{h}$, and so suppose this is not the case. Let $\boldsymbol{g'}$ and $\boldsymbol{h'}$ be the vectors $\boldsymbol{g}$ and $\boldsymbol{h}$ with the last entry removed. Then again by choice of M,

$$\Delta\left(a + \sum_{i=1}^{m} g_i d_i\right) = \Delta\left(a + \sum_{i=1}^{m} h_i d_i\right), \tag{3.6}$$

and these sums are in the block B_b. Let j be the index of $a + \sum_{i=1}^{m} g_i d_i$ as it appears in B_b. Letting $d_{m+1} = dM$, then

$$a + \sum_{i=1}^{m+1} g_i d_i = a + \sum_{i=1}^{m} g_i d_i + g_{m+1} dM$$

is the j-th entry in the block $B_{b+g_{m+1}dM}$. Since $g_{m+1} < k$, both b and $b + g_{m+1}dM$ have the same colouring under Δ'. Therefore,

$$\Delta\left(a + \sum_{i=1}^{m} g_i d_i\right) = \Delta\left(a + \sum_{i=1}^{m+1} g_i d_i\right).$$

By a similar argument,

$$\Delta\left(a + \sum_{i=1}^{m} h_i d_i\right) = \Delta\left(a + \sum_{i=1}^{m+1} h_i d_i\right),$$

and so by (3.6),

$$\Delta\left(a + \sum_{i=1}^{m+1} g_i d_i\right) = \Delta\left(a + \sum_{i=1}^{m+1} h_i d_i\right),$$

completing the proof of (3.4).

To prove (3.5), set $n = S(k, r; r)$, and let $\Delta : [n] \to [r]$ be any r-colouring. By Remark 3.4.8, it suffices to show that [n] has a monochromatic AP_{k+1}. By the choice of n, there exists $a, d_1, \ldots, d_r$ so that or every $\boldsymbol{g}, \boldsymbol{h} \in [0, k]^r$ with $\boldsymbol{g} = (g_1, \ldots, g_r) \sim (h_1, \ldots, h_r) = \boldsymbol{h}$,

$$\Delta\left(a \sum_{i=1}^{r} g_i d_i\right) = \Delta\left(a + \sum_{i=1}^{r} h_i d_i\right).$$

Look at the vectors

$$\boldsymbol{g}_0 = (0, 0, \ldots, 0),$$

$$\boldsymbol{g}_1 = (k, 0, \ldots, 0),$$

$$\vdots$$

$$\boldsymbol{g}_r = (k, k, \ldots, k).$$

There are $r+1$ of these vectors, and so by the pigeonhole principle there are two of them, say $\boldsymbol{g}_\alpha$ and $\boldsymbol{g}_\beta$ for which

$$\Delta\left(a + k\sum_{i=1}^{\alpha} d_i\right) = \Delta\left(a + k\sum_{i=1}^{\beta} d_i\right).$$

For any $j \in [k-1]$, let $\boldsymbol{h}_j \in [0,k]^r$ be the vector with k in the first α entries, j in the $(\alpha+1)$-th to the β-th entry, and 0 afterwards. Then $\boldsymbol{h}_j \sim \boldsymbol{g}_\alpha$, and so

$$\Delta\left(a + k\sum_{i=1}^{\alpha} d_i + j\sum_{i=\alpha+1}^{\beta} d_i\right) = \Delta\left(a + k\sum_{i=1}^{\alpha} d_i\right).$$

Let $a' = a + k\sum_{i=1}^{\alpha} d_i$ and $d = \sum_{i=\alpha+1}^{\beta} d_i$. Then

$$a', a'+d, a'+2d, \ldots, a'+kd$$

forms a monochromatic AP_{k+1}. Therefore $S(k+1,1;r) \leq n$, completing the proof of (3.5). $\square$

3.4.3 The Schur–Brauer theorem

The extension of van der Waerden's theorem to include the difference d was published by Brauer [15]. Brauer, however, credits Schur with its proof (see [97]), and so today

Theorem 3.4.9 below is often called the Schur–Brauer theorem.

Theorem 3.4.9 (Brauer and Schur, 1928 [15]). *For any $k, r \in \mathbb{Z}^+$, there exists a least positive integer $n = SB(k; r)$ so that for any r-colouring $\Delta : [n] \to [r]$, there exists $a, d \in \mathbb{Z}^+$ so that $\{d, a, a + d, a + 2d, \ldots, a + (k-1)d\}$ is monochromatic.*

Proof. Fix $k \in \mathbb{Z}^+$. The proof is by induction on $r \geq 1$. For $r \in \mathbb{Z}^+$, let $P_k(r)$ be the statement of the theorem.

Base Case: When $r = 1$, $n = k$ suffices, since $1, 2, 3, \ldots, k$ is a monochromatic AP_k, and the difference $d = 1$ has the same colour.

Inductive Step: Let $r \geq 1$, and assume $P_k(r)$ is true. Let $n = W(k \cdot SB(k; r) + 1; r + 1)$ and let $\Delta : [n] \to [r + 1]$ be any $(r + 1)$-colouring. By the choice of n, there is a monochromatic $AP_{k \cdot SB(k;r)+1}$

$$B = \{a, a + d, a + 2d, \ldots, a + SB(k; r) \cdot d\}.$$

If for any $j \in [SB(k; r)]$ the element jd is assigned the same colour as B, then

$$\{jd, a, a + jd, a + 2jd, \ldots, a + (k-1)jd\}$$

is monochromatic. Otherwise, the set

$$\{d, 2d, 3d, \ldots, k \cdot SB(k; r) \cdot d\}$$

is coloured with only r colours, and so by $P_k(r)$, there is a monochromatic AP_k with the difference having the same colour. In both cases, $P_k(r + 1)$ holds.

Therefore, by mathematical induction, the statement of the theorem is true for all $r \in \mathbb{Z}^+$. □

Note that when $k = 2$, the Schur–Brauer theorem states that there exists a, d so that $\{a, d, a + d\}$ is monochromatic, and so Theorem 3.4.9 provides another proof of Schur's theorem (Theorem 3.3.1).

With a similar proof, it can be proved that any multiple s of d can also be guaranteed to be the same colour as the AP_k (see [48, Thm. 2, Section 3]).

Brauer [15] used Theorem 3.4.9 to prove Theorem 3.4.1 (see [97]).

3.5 The Folkman–Rado–Sanders theorem

The Folkman–Rado–Sanders theorem (Theorem 3.5.2) is a statement about a finite set A along with all possible sums of distinct elements of A.

Definition 3.5.1. *For any additive semigroup X and any subset $A = \{a_i : i \in I\}$, define the* finite sums set *(also called a finite sum set)*

$$FS(A) = \left\{ \sum_{i \in J} a_i : J \subseteq I, J \neq \emptyset, |J| < \infty \right\}.$$

For a set $A = \{a_1, a_2, \ldots, a_m\} \subseteq \mathbb{Z}^+$, the finite sums set $FS(A)$ is the m-dimensional Hilbert cube $H(0, a_1, a_2, \ldots, a_m)$, and can be considered the "projective versions" of Hilbert cubes. Also, the finite sums set $FS(\{x, y\}) = \{x, y, x + y\}$ is a Schur triple.

The Folkman–Rado–Sanders theorem (Theorem 3.5.2 below) can be proved using Theorem 4.1.3 from Chapter 4, which is a result of Rado's [99] from 1933. In 1968, Sanders [105] gave a direct proof of the Folkman–Rado–Sanders theorem. Folkman's proof was presented by Graham, Rothschild, and Spencer [48]. Theorem 3.5.2 is sometimes called Folkman's theorem in honour of Jon Folkman who passed away in 1969.

Theorem 3.5.2 (Folkman (see [48]); Rado, 1933 [99]; Sanders, 1968 [105]). *For every $m, r \in \mathbb{Z}^+$, there exists a least positive integer n so that for every r-colouring $\Delta : [n] \to [r]$, there exists $A = \{a_1, \ldots, a_m\} \subseteq [n]$ so that $FS(A) \subseteq [n]$ is monochromatic.*

Proof. For a fixed $r \in \mathbb{Z}^+$ and any $m \in \mathbb{Z}^+$, let $P(m)$ be the following statement: there exists a least positive integer $n = n(m; r)$ so that for all r-colourings $\Delta : [n] \to [r]$, there exists $A = \{a_1, \ldots, a_m\}$ so that $FS(A) \subseteq [n]$ and for all $I \subseteq [m]$,

$$\Delta\left(\sum_{i \in I} a_i\right) = \Delta\left(\max_{i \in I} a_i\right).$$

The proof that for all $m \geq 1$, $P(m)$ is true, is by induction on m.

Base Case: For $m = 1$, $n(m; 1) = 1$ suffices since $FS(\{1\})$ satisfies $P(1)$.

Inductive Step: Let $m \geq 1$, and suppose the claim is true for $n(m; r)$. Let

$$n = 2 \cdot W(n(m; r) + 1; r),$$

where $W(n(m, r) + 1; r)$ is the van der Waerden number guaranteed to exist by Theorem 3.4.4, and let $\Delta : [n] \to [r]$ be any r-colouring. By the choice of n, there is a monochromatic $AP_{n(m;r)+1}$

$$B = \{a_{m+1}, a_{m+1} + d, \ldots, a_{m+1} + n(m; r) \cdot d\} \subseteq \left[\frac{1}{2} + 1, n\right].$$

Since $a_{m+1} + n(m;r)d \leq n$ and $a_{m+1} > \frac{n}{2}$, the difference d satisfies $n(m;r)d \leq \frac{n}{2}$. Define a new colouring $\Delta' : [n(m;r)] \to [r]$ by

$$\Delta'(i) = \Delta(id).$$

By the induction hypothesis, there is $A' = \{a'_1, \dots, a'_m\}$ so that $FS(A') \subseteq [n(m,r)]$ and for all $I \subseteq [m]$,

$$\Delta'\left(\sum_{i \in I} a'_i\right) = \Delta'\left(\max_{i \in I} a'_i\right).$$

Consider the set $A = \{a'_i d : a'_i \in A'\} \cup \{a_{m+1}\}$. For any $I \subseteq [m]$, if $\max_{i \in I} a_i \neq a_{m+1}$, then $P(m+1)$ follows by the definition of the colouring Δ'. So consider the case where $\max_{i \in I} a_i = a_{m+1}$. Then

$$\sum_{i \in I} a_i = a_{m+1} + d \sum_{\substack{i \in I \\ i \neq m+1}} a'_i$$

is an element of B, and so has the same colour as a_{m+1}. This concludes the inductive step.

For any $m \in \mathbb{Z}^+$, let $n = n((m-1)r + 1; r)$, and let $\Delta : [n] \to [r]$ be any r-colouring. By $P((m-1)r+1)$, there exists $S = \{s_1, \dots, s_{(m-1)r+1}\}$ so that $FS(S) \subseteq [n]$ and for all $I \subseteq [(m-1)r + 1]$,

$$\Delta\left(\sum_{i \in I} a_i\right) = \Delta\left(\max_{i \in I} a_i\right).$$

By the pigeonhole principle, there are m elements of S that are monochromatic, call this set $A = \{a_1, \dots, a_m\}$. Then any sum of these elements has the same colour as maximum element in A, and so $FS(A)$ is monochromatic. □

While the numbers $h(m;r)$ of Hilbert's cube lemma are bounded above by $(2r)^{2^{m-1}}$ [58], the best known upper bound for the number n that guarantees a monochromatic $FS(\{a_1, \ldots, a_m\})$ under any r-colouring of $[n]$, due to Taylor [121], is

$$n \leq 2^{r^{3^{\cdot^{\cdot^{\cdot^{r^3}}}}}} \Big\}_{2r(m-1)}.$$

Theorem 3.5.2 has an analogous infinite version, that is, for every finite colouring of $\mathbb{Z}^+$, there exist an infinite set A so that $FS(A)$ is monochromatic. This was proved by Hindman in 1974 [65], and the proof of Hindman's theorem is given in Section 7.4.

Chapter 4

Partition regular matrices

4.1 Introduction

The sets with arithmetic properties studied in Chapter 3 also form solutions to certain homogeneous systems. For example, consider the matrices

$$S = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$F = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Any solution to $S\boldsymbol{x} = \mathbf{0}$ satisfies $x_1 + x_2 = x_3$. Any solution to $B\boldsymbol{x} = \mathbf{0}$ has the form

$$x_1, x_6$$

$$x_2 = x_1 + x_6$$

$$x_3 = x_1 + 2x_6$$

$$x_4 = x_1 + 3x_6$$

$$x_5 = x_1 + 4x_6,$$

and any solution to $F\boldsymbol{x} = \mathbf{0}$ consists of x_1, x_2, x_3 along with all sums $x_i + x_j$ for $i, j \in [3]$. Therefore for any finite colouring of $\mathbb{Z}^+$, monochromatic solutions to $S\boldsymbol{x} = \mathbf{0}$, $B\boldsymbol{x} = \mathbf{0}$, and $F\boldsymbol{x} = \mathbf{0}$ are guaranteed by Schur's theorem, the Schur–Brauer theorem, and the Folkman–Rado–Sanders theorem respectively.

Definition 4.1.1. *An $m \times n$ matrix A is called* partition regular in $\mathbb{Z}^+$, *denoted $PR/\mathbb{Z}^+$, if and only if for every finite colouring of $\mathbb{Z}^+$ there is a monochromatic solution $\boldsymbol{x}$ to the system $A\boldsymbol{x} = \mathbf{0}$.*

The matrices S, B and F are partition regular in $\mathbb{Z}^+$. To see that the matrix $C = \begin{bmatrix} 1 & 1 & -3 \end{bmatrix}$ is not partition regular, colour $\mathbb{Z}^+$ in the following way: for each integer

$z \in \mathbb{Z}^+$ let a, b and $r = r(z) \in \{1, 2, 3, 4\}$ be nonnegative integers so that $z = 5^a(5b + r)$. Define $\Delta : \mathbb{Z}^+ \to [4]$ by $\Delta(z) = r(z)$. Since $r + r \not\equiv 3r \pmod 5$ for every $r \in \{1, 2, 3, 4\}$, there is no monochromatic solution to $C\boldsymbol{x} = \mathbf{0}$.

Rado characterized the matrices that are partition regular in 1933 [99] and 1939 [100] as those having the following property:

Definition 4.1.2. *Let A be a matrix with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Then A satisfies the* columns property over $\mathbb{Q}$, *denoted $CP(\mathbb{Q})$, if and only if there exists a partition $[n] = I_1 \cup \cdots \cup I_l$ so that $\sum_{i \in I_1} \boldsymbol{a}_i = \mathbf{0}$ and for all $j = 1, 2, \ldots, l-1$ and every $i \in I_1 \cup \cdots \cup I_j$, there exists $\alpha_{i,j} \in \mathbb{Q}$ so that*

$$\sum_{i \in I_1 \cup \cdots . \cup I_j} \alpha_{i,j} \boldsymbol{a}_i = \sum_{i' \in I_{j+1}} \boldsymbol{a}_{i'},$$

that is, the sum of the vectors with indices in I_{j+1} is a linear combination in $\mathbb{Q}$ of the vectors with indices in $I_1 \cup \cdots \cup I_j$.

Theorem 4.1.3 (Rado, 1933 [99])**.** *Let A be a matrix with integer entries. A is partition regular in $\mathbb{Z}^+$ if and only if A satisfies the columns property over $\mathbb{Q}$.*

A matrix A with a single row satisfies the columns property if and only if there is a nonempty subset of the entries A that sum to zero. For example, the matrix $S = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$ has its first and second entry sum to zero, while the matrix $C = \begin{bmatrix} 1 & 1 & -3 \end{bmatrix}$ has no subset of its entries that sum to zero. The special case of Rado's characterization for these matrices, known as Rado's single equation theorem, is studied in Section 4.2. Rado's full characterization is partially proved in Section 4.4, as well as proofs of equivalences between partition regularity in $\mathbb{Z}^+, \mathbb{Z}$, and $\mathbb{Q}$. One direction of the proof of Rado's characterization is given in Chapter 6, where specific sets that contain solutions to the system $A\boldsymbol{x} = \mathbf{0}$ for matrices A satisfying the columns property are studied.

4.2 Rado's single equation theorem

Rado's single equation theorem is stated as follows:

Theorem 4.2.1 (Rado, 1933 [99]). *Let*

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

be a $1 \times n$ *matrix with nonzero integer entries.* A *is partition regular in* $\mathbb{Z}^+$ *if and only if for some* $I \subseteq [n]$, $\sum_{i \in I} a_i = 0$.

This special case of Rado's characterization for partition regular matrices has received considerable attention. By a compactness argument using König's lemma (Lemma 2.2.1), if A has a monochromatic solution in every r-colouring of $\mathbb{Z}^+$, then there is a least $n \in \mathbb{Z}^+$ so for every r-colouring of $[n]$ there is a monochromatic solution to $A\boldsymbol{x} = \mathbf{0}$. Such a number n is called the *r-colour Rado number* for A. The Rado numbers of several single row partition regular matrices have been studied.

In 1982, Beutelspacher and Brestovansky [11] showed that for every $m \geq 3$, the 2-colour Rado number for the equation

$$x_1 + x_2 + \cdots + x_{m-1} = x_m$$

is $m^2 - m - 1$. In 2008, Schaal and Vestal [110] showed that for $m \geq 6$, the 2-colour Rado number for the equation

$$x_1 + x_2 + \cdots + x_{m-1} = 2x_m$$

is

$$\left\lceil \frac{m-1}{2} \left\lceil \frac{m-1}{2} \right\rceil \right\rceil,$$

and in the same year, Guo and Sun [59] proved the following: for $m \geq 3$ and positive integers $a_1, \ldots, a_{m-1}$, let $a = \min\{a_1, \ldots, a_{m-1}\}$ and $w = a_1 + \cdots + a_{m-1}$. Then the 2-colour Rado number for the equation

$$a_1x_1 + a_2x_2 + \cdots + a_{m-1}x_{m-1} = x_m$$

is $aw^2 + w - a$. For $a_1, \ldots, a_k \in \mathbb{Z}^+$, $m \geq 2$ and $m \geq a_1 + \cdots + a_l$, Saracino in 2016 [108] determined the 2-colour Rado number of the equation

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = x_{k+1} + x_{k+2} + \cdots + x_{k+n}.$$

For $a, b \in \mathbb{Z}^+$ with $gcd(a, b) = 1$, Harborth and Maasberg [61] [62] determined the 2-colour Rado number for all equations

$$a(x + y) = bz.$$

A result of Elsholtz and Gunderson from 2015 [30] states that for $a, b, c \in \mathbb{Z} \setminus \{0\}$ so that $a + b + c = 0$, if $n = W(a, b, c) \in \mathbb{Z}^+$ is the least integer so that every 2-colouring of $[n]$ has a monochromatic solution to $ax + by + cz = 0$, then

$$W(a, b, c) \leq 32 \max\{|a|, |b|, |c|\} + 1.$$

To prove Rado's single equation theorem, first the necessity is proved, that is, A is partition regular in $\mathbb{Z}^+$ if A has a subset of its entries summing to zero. The proof, which can be found in English in [84], uses van der Waerden's theorem.

Proof of necessity of Theorem 4.2.1. Suppose without loss of generality that $a_1 > 0$,

$$a_1 + a_2 + \cdots + a_m = 0,$$

and m is maximal in the sense that the sum of any $m+1$ entries of A is nonzero. Let

$$s = a_{m+1} + a_{m+1} + \cdots + a_n.$$

Suppose that for every finite colouring of $\mathbb{Z}^+$ there are monochromatic $x_1, x_m, x_n \in \mathbb{Z}^+$ so that

$$a_1(x_1 - x_m) + sx_n = 0.$$

Then setting $x_2 = x_3 = \cdots = x_m$, $x_{m+1} = x_{m+1} = \cdots = x_n$, and $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$,

$$\begin{aligned} A\boldsymbol{x} &= a_1x_1 + a_2x_2 + \cdots + a_mx_m + a_{m+1}x_{m+1} + \cdots + a_nx_n \\ &= a_1x_1 - a_1x_m + a_1x_m + a_2x_m + \cdots + a_mx_m + a_{m+1}x_n + \cdots + a_nx_n \\ &= a_1(x_1 - x_m) + x_m(a_1 + a_2 + \cdots + a_m) + x_n(a_{m+1} + \cdots + a_n) \\ &= a_1(x_1 - x_m) + sx_n = 0, \end{aligned}$$

and so $\boldsymbol{x}$ is a monochromatic solution to $A\boldsymbol{x} = \mathbf{0}$. Therefore, the result is proved if the system $a_1(x_1 - x_m) + sx_n$ is partition regular in $\mathbb{Z}^+$.

For every $r \in \mathbb{Z}^+$, let $P(r)$ be the statement that there exists a least $n_r \in \mathbb{Z}^+$ so that for every r-colouring $\Delta : [n_r] \to [r]$, there are monochromatic $x_1, x_m, x_n \in [n_r]$ so that $a_1(x_1 - x_m) + sx_n = 0$. The proof of $P(r)$ is by induction on $r \geq 1$.

Base Case: Let $n_1 = \max\{a_1, s+1\}$. Setting $x_n = a_1$ and choosing x_1 and x_m so that $x_m - x_1 = s$, then $a_1(x_1 - x_m) + sx_n = -a_1 s + sa_1 = 0$, and so the base case is true.

Inductive Step: Let $r \geq 1$ and suppose $P(r)$ is true, that is, there is a least n_r so that for any r-colouring of $[n_r]$, there is a monochromatic solution to $a_1(x_1 - x_m) + sx_n$. Set $n_{r+1} = |s|W(n_r + 1; r + 1)$ and let $\Delta : [n_{r+1}] \to [r + 1]$ be any $(r + 1)$-colouring. Define a new colouring $\Delta' : [W(n_r + 1; r + 1)] \to [r + 1]$ by

$$\Delta'(i) = \Delta(|s|i).$$

By van der Waerden's theorem (Theorem 3.4.4), there is an AP_{n_r+1} in $[W(n_r + 1; r + 1)]$ monochromatic with respect to Δ', and so $[n_{r+1}]$ contains a monochromatic AP_{n_r+1} of the form

$$B = \{a|s|, a|s| + d|s|, a|s| + 2d|s|, \ldots, a|s| + n_r d|s|\},$$

with respect to Δ, say with colour $r + 1$. If for some $j \in [n_r]$, $\Delta(jda_1) = r + 1$, then let x_1, x_m be elements of B with $x_m - x_1 = jd|s|$ and set $x_n = jda_1$. Then $\Delta(x_1) = \Delta(x_m) = \Delta(x_n) = r + 1$ and

$$a_1(x_1 - x_m) + sx_n = -a_1 jd|s| + sjda_1 = 0.$$

Otherwise, define the r-colouring $\Delta'' : [n_r] \to [r]$ by

$$\Delta''(j) = \Delta(jda_1).$$

By the induction hypothesis, there are x_1'', x_m'', x_n'' monochromatic with respect to Δ'' so that $a_1(x_1'' - x_m'') + sx_n'' = 0$. Then $x_1 = x_1'' da_1, x_m = x_m'' da_1$ and $x_n = x_n'' da_1$ are monochromatic with respect to Δ, and

$$a_1(x_1 - x_m) + sx_n = da_1(a_1(x_1'' - x_m'') + sx_n'') = 0.$$

Therefore, $P(r+1)$ holds.

By mathematical induction, the system $a_1(x_1 - x_m) + sx_n$ is partition regular in $\mathbb{Z}^+$. □

In the other direction a more general result is proved to include partition regularity in sets other than $\mathbb{Z}^+$.

Definition 4.2.2. *Let $\mathbb{S}$ be any one of $\mathbb{Z}^+, \mathbb{Z}, \mathbb{Q}^+, \mathbb{Q}, \mathbb{R}^+, \mathbb{R}$ or $\mathbb{C}$. An $m \times n$ matrix A is called* partition regular in $\mathbb{S}$, *denoted $PR/\mathbb{S}$, if and only if for every finite colouring of $\mathbb{S} \setminus \{0\}$ there is a monochromatic solution $\boldsymbol{x}$ to the system $A\boldsymbol{x} = \mathbf{0}$.*

Lemma 4.2.3 (Rado, 1939 [100])**.** *Let*

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

be a $1 \times n$ matrix with nonzero complex entries. If A is partition regular in $\mathbb{C}$, then for some $I \subseteq [n]$, $\sum_{i \in I} a_i = 0$.

Proof. Let A be partition regular in $\mathbb{C}$ and that in hopes of a contradiction, for all $I \subseteq [n]$, $\sum_{i \in I} a_i \neq 0$. Let $a \in \mathbb{R}^+$ be so that for all $I \subseteq [n]$,

$$\left| \sum_{i \in I} a_i \right| \geq \frac{1}{a} \left(\sum_{i=1}^{n} |a_i| \right), \tag{4.1}$$

and choose $q, \delta \in \mathbb{R}$ so that $q > 1, \delta > 0$ and $a^{-1} + 1 - q^{1+\delta} - 2\pi\delta > 0$. Let $m \in \mathbb{Z}^+$ be so that

$$m > 1 + \frac{1}{\delta} - \frac{\ln(a^{-1} + 1 - q^{1+\delta} - 2\pi\delta)}{\delta \ln q}. \tag{4.2}$$

Every $z \in \mathbb{C}$ has a unique representation as $z = q^r e^{2\pi i t}$ for some real numbers $r = r(z) \in \mathbb{R}$ and $t = t(z) \in \mathbb{R}$ satisfying $0 \leq r - t < 1$. Let $\Delta : \mathbb{C} \setminus \{0\} \to [0, m-1]$ be the m-colouring of $\mathbb{C}$ defined by

$$\Delta(z) = \left\lfloor \frac{t}{\delta} \right\rfloor \pmod{m}.$$

Since A is partition regular in $\mathbb{C}$, there exists a monochromatic $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in (\mathbb{C} \setminus \{0\})^n$ so that $A\boldsymbol{x} = \mathbf{0}$. For each $i \in [n]$, let $r_i = r(x_i)$ and $t_i = t(x_i)$ be the values defined above so that $x_i = q^{r_i} e^{2\pi i t_i}$, and let $m_i = \lfloor t_i/\delta \rfloor$. Then since $\boldsymbol{x}$ is monochromatic,

$$m_1 \equiv m_2 \equiv \cdots \equiv m_n \pmod{m}.$$

Assume without loss of generality that $m_1 = m_2 = \cdots = m_k > m_{k+1} \geq \cdots \geq m_n$ and also that $r_1 \leq r_2 \leq \cdots \leq r_k$. Since $0 \leq r_i - t_i < 1$ for every i,

$$t_i \leq r_i < t_i + 1. \tag{4.3}$$

Also,

$$m_i\delta = \left\lfloor \frac{t_i}{\delta} \right\rfloor \delta \leq t_i < \left(\left\lfloor \frac{t_i}{\delta} \right\rfloor + 1 \right) \delta = m_i\delta + \delta. \tag{4.4}$$

Putting (4.3) and (4.4) together,

$$m_1\delta \leq t_1 \leq r_1 \leq r_2 \leq \cdots \leq r_k < t_k + 1 < m_k\delta + \delta + 1 = m_1\delta + \delta + 1,$$

and in particular, for any $i \in [k]$,

$$r_i - r_1 < m_1\delta + \delta + 1 - t_1 \leq m_1\delta + \delta + 1 - m_1\delta = \delta + 1, \tag{4.5}$$

and since $\lfloor t_i/\delta \rfloor = m_i = m_1 = \lfloor t_1/\delta \rfloor$,

$$|t_i - t_1| < \delta. \tag{4.6}$$

Therefore, using the fact that for any $x, y \in \mathbb{R}$, $|e^{2\pi i x}| = 1$ and $|e^{2\pi i x} - e^{2\pi i y}| \leq 2\pi|x - y|$, for $i \in [k]$,

$$\begin{aligned}|x_i - x_1| &= |q^{r_i}e^{2\pi i t_i} - q^{r_1}e^{2\pi i t_1}| \\ &= |(q^{r_i}e^{2\pi i t_i} - q^{r_1}e^{2\pi i t_i}) + (q^{r_1}e^{2\pi i t_i} - q^{r_1}e^{2\pi i t_1})| \\ &\leq (q^{r_i} - q^{r_1})|e^{2\pi i t_i}| + q^{r_1}|e^{2\pi i t_i} - e^{2\pi i t_1}| \\ &\leq q^{r_1}(q^{r_i - r_1} - 1) + q^{r_1}2\pi|t_i - t_1|,\end{aligned}$$

and so by (4.5) and (4.6),

$$\begin{aligned}|x_i - x_1| &\leq q^{r_1}(q^{r_i - r_1} - 1) + q^{r_1}2\pi|t_i - t_1| \\ &< q^{r_1}(q^{\delta+1} - 1) + q^{r_1}2\pi\delta \\ &= q^{r_1}(q^{1+\delta} - 1 + 2\pi\delta).\end{aligned} \tag{4.7}$$

For $i \in [k+1, n]$, since $m_1 > m_i$ and $m_1 \equiv m_i \pmod{m}$, then $m_i \leq m_1 - m$, and so

$$\begin{aligned}|x_i| &= q^{r_i} < q^{t_i+1} < q^{m_i\delta+\delta+1} \\ &\leq q^{(m_1-m)\delta+\delta+1} \leq q^{t_1-m\delta+\delta+1} \\ &\leq q^{r_1-(m-1)\delta+1}.\end{aligned} \tag{4.8}$$

Since $\boldsymbol{x}$ is a solution to $A\boldsymbol{x} = 0$, the triangle inequality along with (4.7), (4.8), and (4.1) give

$$\begin{aligned}
0 = \left|\sum_{i=1}^{n} a_i x_i\right| &= \left|\sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} a_i(x_i - x_1) + \sum_{i=k+1}^{n} a_i x_i\right| \\
&\geq |x_1| \left|\sum_{i=1}^{k} a_k\right| - \sum_{i=1}^{k} |a_i||x_i - x_1| - \sum_{i=k+1}^{n} |a_i||x_i| \\
&> q^{r_1} \left|\sum_{i=1}^{k} a_k\right| - q^{r_1}(q^{1+\delta} - 1 + 2\pi\delta) \sum_{i=1}^{k} |a_i| - q^{r_1-(m-1)\delta+1} \sum_{i=k+1}^{n} |a_i| \\
&> \frac{q^{r_1}}{a} \sum_{i=1}^{n} |a_i| - q^{r_1}(q^{1+\delta} - 1 + 2\pi\delta) \sum_{i=1}^{n} |a_i| - q^{r_1-(m-1)\delta+1} \sum_{i=1}^{n} |a_i|.
\end{aligned}$$

After dividing out $q^{r_1} \sum_{i=1}^{n} |a_i|$ and rearranging,

$$q^{-(m-1)\delta+1} > a^{-1} - q^{1+\delta} + 1 - 2\pi\delta,$$

which after taking logarithms and rearranging gives

$$m < 1 + \frac{1}{\delta} - \frac{\ln(a^{-1} + 1 - q^{1+\delta} - 2\pi\delta)}{\delta \ln q},$$

a contradiction to the choice of m in (4.2). □

4.3 Partition regularity implies the columns property

The proof of one direction of Theorem 4.1.3 is presented here (see Section 6.4 for the other direction), and can be found in [97].

Theorem 4.3.1. *Let A be a matrix with integer entries. If A is partition regular in $\mathbb{Z}^+$ then A satisfies the columns property over $\mathbb{Q}$.*

Proof. Suppose A is an $m \times n$ matrix that is partition regular in $\mathbb{Z}^+$ with columns $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$. For any $I \subseteq [n]$ and any $J \subseteq [n]$ with $I \cap J = \emptyset$, if $\boldsymbol{a}_J = \sum_{i \in J} \boldsymbol{a}_i$ is not a linear combination of the vectors $\{\boldsymbol{a}_i : i \in I\}$, let $P_{I,J}$ be the set of all primes p for which there is an $m \in \mathbb{Z}^+$ such that $p^m \boldsymbol{a}_J$ is a linear combination of the vectors $\{\boldsymbol{a}_i : i \in I\}$ modulo p^{m+1}, that is, there are $x_i \in \mathbb{Q}$ for which

$$\sum_{i \in I} x_i \boldsymbol{a}_i = p^m \sum_{i \in J} \boldsymbol{a_i} \pmod{p^{m+1}}.$$

If $\boldsymbol{a}_J$ is a linear combination of the vectors indexed by I, then define $P_{I,J} = \emptyset$.

First to show that for every $I, J \subseteq [n]$, $I \cap J = \emptyset$, $P_{I,J}$ is finite. Assume $\boldsymbol{a}_J$ is not a linear combination of the vectors indexed by I, otherwise $P_{I,J} = \emptyset$ is finite. Choose a vector $\boldsymbol{v} \notin \text{span}(\{\boldsymbol{a}_i : i \in I\})$ so that $(\boldsymbol{v} \cdot \boldsymbol{a}_J) \neq \mathbf{0}$ and $(\boldsymbol{v} \cdot \boldsymbol{a}_i) = 0$ for all $i \in I$. Without loss of generality, it can be assumed that $\boldsymbol{v} \in \mathbb{Z}^m$ (simply multiply $\boldsymbol{v}$ by a common multiple of the denominators of the entries of $\boldsymbol{v}$). Then for every $p \in P_{I,J}$, there is an $m \in \mathbb{Z}^+$ so that

$$p^m \boldsymbol{a}_J = \sum_{i \in I} x_i \boldsymbol{a}_i \pmod{p^{m+1}},$$

and so

$$p^m (\boldsymbol{v} \cdot \boldsymbol{a}_J) = \sum_{i \in I} (\boldsymbol{v} \cdot \boldsymbol{a}_i) \equiv 0 \pmod{p^{m+1}},$$

that is there is some $t \in \mathbb{Z}$ so that $p^m (\boldsymbol{v} \cdot \boldsymbol{a}_J) = t p^{m+1}$. Therefore, $p | (\boldsymbol{v} \cdot \boldsymbol{a}_J)$, which is only possible for finitely many primes p, and so $P_{I,J}$ is finite.

Let p be a prime so that

$$p \notin \bigcup_{\substack{I,J \subseteq [n] \\ I \cap J = \emptyset}} P_{I,J}$$

and for every $I \subseteq [n]$, if $\sum_{i \in I} \boldsymbol{a}_i \neq 0$, p is not one of the finitely many primes for which $\sum_{i \in I} \boldsymbol{a}_i \equiv \boldsymbol{0} \pmod{p}$. Every integer $x \in \mathbb{Z}^+$ can be expressed uniquely as $x = yp^z$, where $y \not\equiv 0 \pmod{p}$. Let $\Delta : \mathbb{Z}^+ \to [p-1]$ be the $(p-1)$-colouring of $\mathbb{Z}^+$ defined by

$$\Delta(x) = y \pmod{p}.$$

Since A is partition regular in $\mathbb{Z}^+$, there is a monochromatic vector $\boldsymbol{x} = (x_1, \dots, x_n)$ so that $A\boldsymbol{x} = \boldsymbol{0}$. Let $\boldsymbol{x}$ be such a vector. Then there is an $r \in [p-1]$ so that for every $x_i = y_i p^{z_i}$, y_i can be expressed as $y_i = pk_i + r$. Say there are l different values of $z_1, \dots, z_n$, call them $m_1 < m_2 < \cdots < m_l$. For $i = 1, \dots, l$, define the set $I_i = \{k \in [n] : z_k = m_i\}$.

With the partition $[n] = I_1 \cup \cdots \cup I_l$, A satisfies the columns property over $\mathbb{Q}$. To see this, first look at I_1. By the choice of $\boldsymbol{x}$, $\sum_{i \in [n]} x_i \boldsymbol{a}_i = \boldsymbol{0}$, and so in particular,

$$\sum_{i \in I_1} x_i \boldsymbol{a}_i + \sum_{i \in I_2 \cup \cdots \cup I_l} x_i \boldsymbol{a}_i \equiv \boldsymbol{0} \pmod{p^{1+m_1}}. \tag{4.9}$$

For any $i \in I_2 \cup \cdots \cup I_k$, $z_i \geq m_1 + 1$, and so $x_i = y_i p^{z_i} \equiv 0 \pmod{p^{1+m_1}}$. Therefore, the right summand in (4.9) is congruent to $\boldsymbol{0}$ modulo p^{1+m_1}. As for $i \in I_1$, $z_i = m_1$, and so

$$x_i = (pk_i + r)p^{m_1} = p^{1+m_1} k_i + p^{m_i} r \equiv p^{m_i} r \pmod{p^{1+m_1}}.$$

Therefore (4.9) reduces to

$$p^{m_1} \Big(r \sum_{i \in I_1} \boldsymbol{a}_i \Big) \equiv \boldsymbol{0} \pmod{p^{1+m_1}},$$

and so $r\sum_{i\in I_1}\boldsymbol{a}_i \equiv \boldsymbol{0} \pmod p$. Since p was chosen so that $\sum_{i\in I_1}\boldsymbol{a}_i \not\equiv \boldsymbol{0} \pmod p$ if $\sum_{i\in I_1}\boldsymbol{a}_i \neq \boldsymbol{0}$, this implies that $\sum_{i\in I_1}\boldsymbol{a}_i = \boldsymbol{0}$.

For any $j = 1, \ldots, k-1$ assume, in hopes of a contradiction, that $\sum_{i\in I_{j+1}}\boldsymbol{a}_i$ is not a linear combination of the vectors indexed by $I_1 \cup \cdots \cup I_j$. Again, by the choice of $\boldsymbol{x}$,

$$\sum_{i\in I_1\cup\cdots\cup I_j} x_i\boldsymbol{a}_i + \sum_{i\in I_{j+1}} x_i\boldsymbol{a}_i + \sum_{i\in I_{j+2}\cup\cdots\cup I_k} x_i\boldsymbol{a}_i \equiv \boldsymbol{0} \pmod{p^{1+m_{j+1}}}. \tag{4.10}$$

For $i \in I_{j+2}\cup\cdots\cup I_k$, then $z_i \geq m_{j+1}+1$, and so $x_i = y_i p^{z_i} \equiv 0 \pmod{p^{1+m_{j+1}}}$, and the rightmost summand in (4.10) is congruent to $\boldsymbol{0}$ modulo $p^{1+m_{j+1}}$. For $i \in I_{j+1}$, $z_i = m_{j+1}$, and so

$$x_i = (pk_i + r)p^{m_{j+1}} = p^{m_{j+1}+1}k_i + p^{m_{j+1}}r \equiv p^{m_{j+1}}r \pmod{p^{1+m_{j+1}}}.$$

Therefore, (4.10) reduces to

$$\sum_{i\in I_1\cup\cdots\cup I_j} x_i\boldsymbol{a}_i + p^{m_{j+1}}\Big(r\sum_{i\in I_{j+1}}\Big) \equiv \boldsymbol{0} \pmod{p^{1+m_{j+1}}}. \tag{4.11}$$

Let $J_j = I_1 \cup \ldots \cup I_j$. By the choice of $p \notin P_{J_j,I_{j+1}}$, then $p^{m_{j+1}}\boldsymbol{a}_{I_{j+1}} = p^{m_{j+1}}\sum_{i\in I_{j+1}}\boldsymbol{a}$ is not a linear combination of the vectors indexed by $I_1\cup\cdots\cup I_j$ modulo $p^{1+m_{j+1}}$, in particular,

$$\sum_{i\in I_1\cup\cdots\cup I_j}\Big(p^{1+m_{j+1}} - \frac{x_i}{r}\Big)\boldsymbol{a}_i \not\equiv p^{m_{j+1}}\sum_{i\in I_{j+1}}\boldsymbol{a}_i \pmod{p^{1+m_{j+1}}},$$

which contradicts (4.11). Therefore $\sum_{i\in I_{j+1}}\boldsymbol{a}_i$ must in fact be a linear combination of the vectors indexed by $I_1\cup\cdots\cup I_j$. The second part of the columns property is therefore satisfied, concluding the proof. □

In Section 4.4, characterizations of partition regular matrices with entries in $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are given and a more general definition of the columns property is needed.

Definition 4.3.2. *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, and let A be a matrix with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Then A satisfies the* columns property over $\mathbb{S}$, *denoted $CP(\mathbb{S})$ if and only if there exists a partition $[n] = I_1 \cup \cdots \cup I_l$ so that $\sum_{i \in I_1} \boldsymbol{a}_i = \boldsymbol{0}$ and for all $j = 1, 2, \ldots, l-1$ and every $i \in I_1 \cup \cdots \cup I_j$, there exists $\alpha_{i,j} \in \mathbb{S}$ so that*

$$\sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \boldsymbol{a}_i = \sum_{i' \in I_{j+1}} \boldsymbol{a}_{i'},$$

that is, the sum of the vectors with indices in I_{j+1} is a linear combination in $\mathbb{S}$ of the vectors with indices in $I_1 \cup \cdots \cup I_j$.

The following results were proved by Rado in 1939 [100]. First, Rado proved the following lemma, which he calls "geometrically obvious".

Lemma 4.3.3 (Rado, 1939 [100])**.** *For $r \geq 1$ and $s \geq 1$, let $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r, \boldsymbol{m}_1, \ldots, \boldsymbol{m}_s \in \mathbb{C}^n$ be vectors. Suppose that for all $\boldsymbol{t} \in \mathbb{C}^n$, $(\boldsymbol{l}_1 \cdot \boldsymbol{t}) = \cdots = (\boldsymbol{l}_r \cdot \boldsymbol{t}) = 0$ implies that for some $j \in [n]$, $(\boldsymbol{m}_j \cdot \boldsymbol{t}) = 0$. Then at least one of the $\boldsymbol{m}_j$'s is a linear combination of $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$.*

Remark 4.3.4. *When $\mathbb{S} = \mathbb{Q}$ or $\mathbb{R}$, if $\boldsymbol{m}_j, \boldsymbol{l}_1, \ldots, \boldsymbol{l}_r \in \mathbb{S}^n$ and $\boldsymbol{m}_j$ is a linear combination in $\mathbb{C}$ of $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_r$, then it is also a linear combination in $\mathbb{S}$.*

One of Rado's theorems [100, p.142, Thm 5] states that for any subfield $\mathbb{S}$ of $\mathbb{C}$, if A has entries in $\mathbb{S}$ and is partition regular in $\mathbb{S}$ then it satisfies the columns property over $\mathbb{S}$. Here $\mathbb{S}$ is restricted to $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, and with Remark 4.3.4 in mind, it is shown that partition regularity in $\mathbb{C}$ is enough.

Theorem 4.3.5 (Rado, 1939 [100]). *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$, and let A be a matrix with entries in $\mathbb{S}$. If A is partition regular in $\mathbb{C}$, then A satisfies the columns property over $\mathbb{S}$.*

Proof. Suppose A is an $m \times n$ matrix with entries in $\mathbb{S}$ that is partition regular in $\mathbb{C}$, with columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$. For every nonempty subset $I \subseteq [n]$, define $\boldsymbol{m}_I = \sum_{i \in I} \boldsymbol{a}_i$.

Let $\boldsymbol{t} = (t_1, t_2, \ldots, t_m)$ be any vector with complex entries. Then $\boldsymbol{t} \cdot \boldsymbol{0} = 0$, and any solution to $A\boldsymbol{x} = \boldsymbol{0}$ is also a solution to $(\boldsymbol{a}_1 \cdot \boldsymbol{t})x_1 + \cdots + (\boldsymbol{a}_n \cdot \boldsymbol{t})x_n = 0$. Therefore,

$$(\boldsymbol{a}_1 \cdot \boldsymbol{t})x_1 + \cdots + (\boldsymbol{a}_n \cdot \boldsymbol{t})x_n = 0$$

is a single equation that is partition regular in $\mathbb{C}$, and so by Lemma 4.2.3, for some nonempty $I \subseteq [n]$, $\sum_{i \in I}(\boldsymbol{a}_i \cdot \boldsymbol{t}) = \boldsymbol{m}_I \cdot \boldsymbol{t} = 0$. This is true for every vector $\boldsymbol{t}$, and so by Lemma 4.3.3, some $\boldsymbol{m}_{I_1}$ is a linear combination of $\boldsymbol{0}$, that is, $\boldsymbol{m}_{I_1} = \sum_{i \in I_1} \boldsymbol{a}_i = \boldsymbol{0}$. This satisfies the first part of the columns property.

Let $\boldsymbol{t}$ be any vector so that for every $i \in I_1$, $(\boldsymbol{a}_i \cdot \boldsymbol{t}) = 0$. Then since

$$\sum_{i \in [n]} (\boldsymbol{a}_i \cdot \boldsymbol{t})x_i = \sum_{i \in [n] \setminus I_1} (\boldsymbol{a}_i \cdot \boldsymbol{t})x_i$$

is partition regular in $\mathbb{C}$, again by Lemma 4.2.3, for some nonempty $I \subseteq [n] \setminus I_1$,

$$\sum_{i \in I} (\boldsymbol{a}_i \cdot \boldsymbol{t}) = \boldsymbol{m}_I \cdot \boldsymbol{t} = 0.$$

By Lemma 4.3.3 there is a nonempty subset $I_2 \subseteq [n] \setminus I_1$ so that $\boldsymbol{m}_{I_2}$ is a linear combination (in $\mathbb{S}$) of the vectors indexed by I_1.

Continue this process, at each step j, so long as $I_1 \cup \cdots \cup I_j \neq [n]$. Let $\boldsymbol{t}$ be any vector so that for every $i \in I_1 \cup \cdots \cup I_j$, $(\boldsymbol{a}_i \cdot \boldsymbol{t}) = 0$. Then since

$$\sum_{i \in [n]} (\boldsymbol{a}_i \cdot \boldsymbol{t}) x_i = \sum_{i \in [n] \setminus (I_1 \cup \cdots \cup I_j)} (\boldsymbol{a}_i \cdot \boldsymbol{t}) x_i$$

is partition regular in $\mathbb{C}$, Lemma 4.2.3 guarantees that there is some nonempty subset $I \subseteq [n] \setminus (I_1 \cup \cdots \cup I_j)$ for which

$$\sum_{i \in I} (\boldsymbol{a}_i \cdot \boldsymbol{t}) = \boldsymbol{m}_I \cdot \boldsymbol{t} = 0.$$

By Lemma 4.3.3, there is some nonempty subset $I_{j+1} \subseteq [n] \setminus (I_1 \cup \cdots \cup I_j)$ so that $\boldsymbol{m}_{I_{j+1}}$ is a linear combination (in $\mathbb{S}$) of the columns indexed by $I_1 \cup \cdots \cup I_j$.

Since A has finitely many columns, this process terminates with a partition $[n] = I_1 \cup \cdots \cup I_k$ that satisfies the second part of the columns property. Since each of the linear combinations were in $\mathbb{S}$, A satisfies the columns property over $\mathbb{S}$. □

4.4 Characterizations of partition regular matrices

Rado's characterization of partition regular matrices in terms of the columns property gives a deterministic process to check if a matrix A is partition regular, in that an algorithm that tests every partition of the columns to see if it satisfies the conditions of the columns property will terminate. Rado [99] also showed that it is equivalent for a matrix with integer entries to be partition regular in $\mathbb{Z}^+, \mathbb{Z}$ and $\mathbb{Q}$. The first of these equivalences is generalized in the following lemma:

Lemma 4.4.1 (Rado, 1933 [99])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$. Then the $m \times n$ matrix A is partition regular in $\mathbb{S}$ if and only if A is partition regular in $\mathbb{S}^+$.*

Proof. Since $\mathbb{S}^+ \subseteq \mathbb{S}\backslash\{0\}$, if A is partition regular in $\mathbb{S}^+$ then A is also partition regular in $\mathbb{S}$.

In the other direction, assume A is partition regular in $\mathbb{S}$, and let $\Delta : \mathbb{S}^+ \to [r]$ be any r-colouring. Define a new colouring $\Delta' : \mathbb{S}\backslash\{0\} \to [2r]$ by

$$\Delta'(s) = \begin{cases} \Delta(s) & s > 0 \\ \Delta(-s) + r & s < 0. \end{cases}$$

Since A is partition regular in $\mathbb{S}$, there exists a monochromatic $\boldsymbol{x} \in (\mathbb{S} \setminus \{0\})^n$ for which $A\boldsymbol{x} = \mathbf{0}$. If $\boldsymbol{x}$ is assigned a colour in $[r]$, then the entries of $\boldsymbol{x}$ are in $\mathbb{S}^+$. If $\boldsymbol{x}$ is assigned a colour in $[r+1, 2r]$, then $\boldsymbol{z} = -\boldsymbol{x}$ has entries in $\mathbb{S}^+$ and $A\boldsymbol{z} = -A\boldsymbol{x} = \mathbf{0}$. Therefore, A is partition regular in $\mathbb{S}^+$. □

The equivalence between partition regularity in $\mathbb{Z}$ and partition regularity in $\mathbb{Q}$ can be proved using König's lemma, an argument given in [97].

Lemma 4.4.2 (Rado, 1933 [99])**.** *Let A be a $m \times n$ matrix with integer entries. Then A is partition regular in $\mathbb{Z}$ if and only if A is partition regular in $\mathbb{Q}$.*

Proof. Since $\mathbb{Z} \subseteq \mathbb{Q}$, if A is partition regular in $\mathbb{Z}$ then A is also partition regular in $\mathbb{Q}$.

In the other direction, suppose A is partition regular in $\mathbb{Q}$ and fix an ordering $q_1, q_2, q_3, \ldots$ on $\mathbb{Q}$. The next step in the proof is to show that for every $r \in \mathbb{Z}^+$, there is some i_r for which every r-colouring of $\{q_1, q_2, \ldots, q_{i_r}\}$ has a monochromatic vector $\boldsymbol{x} \in \{q_1, q_2, \ldots, q_{i_r}\}^n$ so that $A\boldsymbol{x} = \mathbf{0}$.

Suppose for some $r \in \mathbb{Z}^+$, there is no such i_r, that is, for every $i \in \mathbb{Z}^+$, there is a colouring of $\{q_1, q_2, \ldots, q_i\}$ with no monochromatic solution to $A\boldsymbol{x} = \mathbf{0}$. Call such a colouring a "bad" colouring. Such a "bad" colouring restricted to $\{q_1, \ldots, q_{i-1}\}$ is also a "bad" colouring. Construct a tree as follows: Let 0 be the root with r children representing all possible colours for q_1. At every generation i, the children of each vertex at this generation represents a colour for $i + 1$ so that the colouring of $\{q_1, \ldots, q_i, q_{i+1}\}$ represented by the colours of each vertex along the path from the root to q_{i+1} is a "bad" colouring. This tree, by assumption, is infinite, and each vertex has finite degree. By Lemma 2.2.1 there is an infinite path, but this infinite path represents a "bad" colouring of $\mathbb{Q}$, which is a contradiction to A being partition regular over $\mathbb{Q}$. Therefore, $i_r \in \mathbb{Z}^+$ exists.

Let c be a common multiple of the denominators of $q_1, q_2, \ldots, q_{i_r}$. Then for $\boldsymbol{x} \in \{q_1, q_2, \ldots, q_{i_r}\}^n$, $A\boldsymbol{x} = \mathbf{0}$ if and only if $cA\boldsymbol{x} = \mathbf{0}$, and so A is also partition regular over $\{cq_1, cq_2, \ldots, cq_i\} \subseteq \mathbb{Z}$. Therefore, A is partition regular in $\mathbb{Z}$. □

A characterization is given of integer-valued matrices that are partition regular in $\mathbb{Z}^+$.

Theorem 4.4.3 (Rado, 1933 [99]). *For a set $\mathbb{S}$, let $PR/\mathbb{S}$ denote partition regularity in $\mathbb{S}$ and $CP(\mathbb{S})$ denote the columns property over $\mathbb{S}$. For a matrix A with integer entries, the following are equivalent:*

1. *A is* $PR/\mathbb{Z}^+$*;*

2. *A is* $PR/\mathbb{Z}$*;*

3. *A is* $PR/\mathbb{Q}$;

4. *A satisfies* $CP(\mathbb{Q})$*.*

The equivalences

$$A \text{ is } PR/\mathbb{Z}^+ \quad \longleftrightarrow \quad A \text{ is } PR/\mathbb{Z} \quad \longleftrightarrow A \text{ is } PR/\mathbb{Q}$$

are established by Lemmas 4.4.1 and 4.4.2. The implication

$$A \text{ is } PR/\mathbb{Z}^+ \quad \longrightarrow \quad A \text{ satisfies } CP(\mathbb{Q})$$

is proved in Theorem 4.3.1. The missing implication

$$A \text{ satisfies } CP(\mathbb{Q}) \quad \longrightarrow \quad A \text{ is } PR/\mathbb{Z}^+$$

is proved in Section 6.4 (see Lemma 6.4.2.)

Remark 4.4.4. *Suppose that A is a matrix with rational entries that is partition regular in $\mathbb{Q}$, and let c be a common multiple of the denominators of the entries of A. Then cA is a matrix with integer entries, and any $\boldsymbol{x}$ is a solution to $A\boldsymbol{x} = \mathbf{0}$ if and only if $\boldsymbol{x}$ is also a solution to $cA\boldsymbol{x} = \mathbf{0}$, and so cA is partition regular in $\mathbb{Q}$. By Theorem 4.4.3, cA is partition regular in $\mathbb{Z}^+$, and so A is also partition regular in $\mathbb{Z}^+$. Therefore, integer-valued matrices may be treated as matrices with rational entries for a more general characterization.*

The statement of Theorem 4.4.5 below is derived from the results of Rado [100].

Theorem 4.4.5 (Rado, 1939 [100]). *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. Let $PR/\mathbb{S}$ denote partition regularity in $\mathbb{S}$ and $CP(\mathbb{S})$ denote the columns property over $\mathbb{S}$.*

1. *For a matrix A with entries in $\mathbb{Q}$, the following are equivalent:*

    (a) *A is $PR/\mathbb{Z}^+$;*

    (b) *A is $PR/\mathbb{Z}$;*

    (c) *A is $PR/\mathbb{Q}^+$;*

    (d) *A is $PR/\mathbb{Q}$;*

    (e) *A is $PR/\mathbb{R}^+$;*

    (f) *A is $PR/\mathbb{R}$;*

    (g) *A is $PR/\mathbb{C}$;*

    (h) *A satisfies $CP(\mathbb{Q})$.*

2. *For a matrix A with entries in $\mathbb{R}$, the following are equivalent:*

    (a) *A is $PR/\mathbb{R}^+$;*

    (b) *A is $PR/\mathbb{R}$;*

    (c) *A is $PR/\mathbb{C}$;*

    (d) *A satisfies $CP(\mathbb{R})$.*

3. *For a matrix A with entries in $\mathbb{C}$, the following are equivalent:*

(a) A is $PR/\mathbb{C}$;

(b) A satisfies $CP(\mathbb{C})$.

For a matrix A with entries in $\mathbb{Q}$, by Remark 4.4.4, the following implications are true by Lemmas 4.4.1 and 4.4.2 and by set inclusion.

$$\begin{array}{ccccccc}
A \text{ is } PR/\mathbb{Z}^+ & \longrightarrow & A \text{ is } PR/\mathbb{Q}^+ & \longrightarrow & A \text{ is } PR/\mathbb{R}^+ & & \\
\updownarrow & & \updownarrow & & \updownarrow & & \\
A \text{ is } PR/\mathbb{Z} & \longleftrightarrow & A \text{ is } PR/\mathbb{Q} & \longrightarrow & A \text{ is } PR/\mathbb{R} & \longrightarrow & A \text{ is } PR/\mathbb{C}
\end{array}$$

The implication

$$A \text{ is } PR/\mathbb{C} \;\longrightarrow\; \text{A satisfies } CP(\mathbb{Q})$$

is established by Theorem 4.3.5.

If A has entries in $\mathbb{R}$, then by Lemma 4.4.1, Theorem 4.3.5, and set inclusion, the following implications are true.

$$A \text{ is } PR/\mathbb{R}^+ \longleftrightarrow A \text{ is } PR/\mathbb{R} \longrightarrow A \text{ is } PR/\mathbb{C} \longrightarrow A \text{ satisfies } CP(\mathbb{R})$$

If A has entries in $\mathbb{C}$, the implication

$$A \text{ is } PR/\mathbb{C} \;\longrightarrow\; A \text{ satisfies } CP(\mathbb{C})$$

is established by Theorem 4.3.5.

The remaining implications $CP(\mathbb{Q}) \longrightarrow PR/\mathbb{Q}$ for matrices with rational entries, $CP(\mathbb{R}) \longrightarrow PR/\mathbb{R}$ for matrices with real entries, and $CP(\mathbb{C}) \longrightarrow PR/\mathbb{C}$ for matrices with complex entries are proved in Section 6.4 (see Lemma 6.4.3).

Although some implications are missing, Theorem 4.4.5 is assumed to be true for the rest of this chapter and the next chapter.

In Rado's 1939 paper [100], a slightly different definition of the columns property is given. Letting $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, say that an $m \times n$ matrix A with columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ satisfies $CP^*(\mathbb{S})$ if and only if there is a partition $[n] = I_1 \cup \cdots \cup I_l$ and $c \in \mathbb{S}$ so that $\sum_{i \in I_1} \boldsymbol{a}_i = 0$ and for all $j = 1, 2, \dots, l-1$ and every $i \in I_1 \cup \cdots \cup I_j$, there exists $\beta_{i,j} \in \mathbb{S}$ so that

$$\sum_{i \in I_1 \cup \cdots \cup I_j} \beta_{i,j} \boldsymbol{a}_i = c \sum_{i \in I_{j+1}} \boldsymbol{a}_i.$$

Letting $\beta_{i,j} = c\alpha_{i,j}$, then $CP^*(\mathbb{S})$ is equivalent to $CP(\mathbb{S})$ for $\mathbb{S}$ being any of $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. If A satisfies $CP(\mathbb{Q})$ and c is a common multiple of the denominators of all $\alpha_{i,j}$'s, then A also satisfies $CP^*(\mathbb{Z})$, and so these two notions are equivalent. With Remark 4.4.4 in mind, part 1 of Theorem 4.4.5 applies to matrices with integer and rational entries, and so a corollary would be that for $\mathbb{S}$ being any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, a matrix with entries in $\mathbb{S}$ is $PR/\mathbb{S}$ if and only if it satisfies $CP^*(\mathbb{S})$. In fact, even more can be said. Rado stated the following as Theorem VII in 1939 [100].

Theorem 4.4.6 (Rado, 1939 [100]). *Let $\mathbb{S}$ be any subring of $\mathbb{C}$. Then a matrix with entries in $\mathbb{S}$ is $PR/\mathbb{S}$ if and only if it satisfies $CP^*(\mathbb{S})$.*

A corollary to Theorem VII of [100] states that a matrix A is $PR/\mathbb{C}$ if and only if A is partition regular over the ring generated by the entries of A.

4.5 Non-homogeneous partition regularity

4.5.1 Rado's characterizations

Rado considered the partition regularity of non-homogeneous linear systems [99], [100]. The theorems in this section are given without proof.

Definition 4.5.1. *Let $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$. Let A be an $m \times n$ matrix with complex entries and $\boldsymbol{b} \in \mathbb{C}^m \setminus \{\boldsymbol{0}\}$. The pair $(A, \boldsymbol{b})$ is partition regular in $\mathbb{S}$ if and only if for every finite colouring of $\mathbb{S}$, there is a monochromatic vector $\boldsymbol{x} \in \mathbb{S}^n$ so that $A\boldsymbol{x} = \boldsymbol{b}$.*

Rado's characterization from 1933 of non-homogeneous partition regular matrices and its proof is translated from its original German to English by Hindman [71].

Theorem 4.5.2 (Rado, 1933 [99])**.** *Let A be an $m \times n$ matrix with rational entries and $\boldsymbol{b} \in \mathbb{Q}^m \setminus \{\boldsymbol{0}\}$.*

1. *The pair $(A, \boldsymbol{b})$ is partition regular in $\mathbb{Z}^+$ if and only if either*

   (a) There exists $\boldsymbol{k} = (k, k, \ldots, k) \in (\mathbb{Z}^+)^m$ so that $A\boldsymbol{k} = \boldsymbol{b}$, or

   (b) A satisfies the columns property over $\mathbb{Q}$ and there exists $\boldsymbol{k} = (k, k, \ldots, k) \in (\mathbb{Z}^+)^m$ so that $-A\boldsymbol{k} = \boldsymbol{b}$.

2. *The pair $(A, \boldsymbol{b})$ is partition regular in $\mathbb{Z}$ if and only if there exists $\boldsymbol{k} = (k, k, \ldots, k) \in \mathbb{Z}^m$ so that $A\boldsymbol{k} = \boldsymbol{b}$.*

3. *The pair* $(A, \boldsymbol{b})$ *is partition regular in* $\mathbb{Q}$ *is and only if there exists* $\boldsymbol{k} = (k, k, \ldots, k) \in \mathbb{Q}^m$ *so that* $A\boldsymbol{k} = \boldsymbol{b}$

In Rado's paper from 1939 [100], another similar result is given. Let $\mathcal{A} \subseteq \mathcal{C}$ be the set of all algebraic numbers.

Theorem 4.5.3 (Rado, 1939 [100])**.** *Let* A *be an* $m \times n$ *matrix with entries in* $\mathbb{C}$ *and* $\boldsymbol{b} \in \mathbb{C}^m \setminus \{\boldsymbol{0}\}$. *The pair* $(A, \boldsymbol{b})$ *is partition regular in* $\mathcal{A}$ *if and only if there exists* $\boldsymbol{k} = (k, k, \ldots, k) \in \mathcal{A}^m$ *so that* $A\boldsymbol{k} = \boldsymbol{b}$

Note that Theorem 4.5.3 does not include non-homogeneous partition regularity in $\mathbb{R}$ or $\mathbb{C}$. This fact was proved later (see Corollary 4.5.8 below). Rado presents a proof in 1939 [100], using Lemma 4.3.3, that only the cases for single row matrices need to be considered:

Lemma 4.5.4 (Rado, 1939 [100])**.** *Let* $\mathbb{F}$ *be a field,* $M \subseteq \mathbb{F}$, A *an* $m \times n$ *matrix with entries in* $\mathbb{F}$ *and columns* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, *and* $\boldsymbol{b} \in \mathbb{F}^n \setminus \{\boldsymbol{0}\}$. *If for every* $\boldsymbol{t} \in \mathbb{F}^m$, *there exists* $k' \in M$ *so that*

$$(\boldsymbol{t} \cdot \boldsymbol{a}_1)k' + (\boldsymbol{t} \cdot \boldsymbol{a}_2)k' + \cdots + (\boldsymbol{t} \cdot \boldsymbol{a}_1)k' = (\boldsymbol{t} \cdot \boldsymbol{b}),$$

then there exists $k \in M$ *so that for* $\boldsymbol{k} = (k, k, \ldots, k) \in M^m$,

$$A\boldsymbol{k} = \boldsymbol{b}.$$

Since any $\boldsymbol{k} = (k, k, \ldots, k)$ always has monochromatic entries, as a consequence of Theorems 4.5.2 and 4.5.3, Ramsey properties of sets with arithmetic structures cannot be extracted from proving the non-homogeneous partition regularity of a pais $(A, \boldsymbol{b})$ in the same way that Ramsey properties are extracted from partition regular matrices.

4.5.2 A result in Euclidean Ramsey theory

In Euclidean Ramsey theory, a geometric configuration H is given along with $r \in \mathbb{Z}^+$, and the question is asked whether there is a large enough metric space X so that whenever the points of X are r-coloured there is a congruent, homothetic, similar, or translation of a copy of H that is monochromatic. An extensive catalogue of results in Euclidean Ramsey theory was researched by Erdős, Graham, Montgomery, Rothschild, Spencer and Straus [31], [32], [33].

Erdős *et al.* [31] proved a strengthening of Theorem 4.5.3 to then prove a result on spherical sets (sets that can be embedded in a sphere) and Euclidean Ramsey theory.

Definition 4.5.5. *Let $\mathbb{E}^n$ denote the n-dimensional Euclidean space. A configuration of points $K = \{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{E}^n$ is* spherical *if and only if there exists $r \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{E}^n$ so that for all $i = 0, 1, \ldots, k$, $||\boldsymbol{v}_i - \boldsymbol{x}|| = r$, where $||\cdot||$ denotes the euclidean norm.*

Two sets $K, K' \in \mathbb{E}^n$ are congruent if and only if one can be transformed into the other by a combination of translations, rotations and reflections.

Theorem 4.5.6 (Erdős *et al.*, 1973 [31])**.** *Let $K = \{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{E}^n$ be a non-spherical set. Then for every $n \in \mathbb{Z}^+$, there is a finite colouring of $\mathbb{E}^n$ so that no set K' congruent to K is monochromatic.*

As of yet, there is no proof of a converse of Theorem 4.5.6. A reward of \$1000 dollars is available for the proof of the converse (see [46]).

The idea of the proof of Theorem 4.5.6 is to first show that K is non-spherical if and only if there exists $b, c_1, \ldots, c_k \in \mathbb{R}$ not all zero so that $\sum_{i=1}^{k} c_i(\boldsymbol{v}_i - \boldsymbol{v}_0) = \mathbf{0}$ and

$$\sum_{i=1}^{k} c_i \left(||\boldsymbol{v}_i||^2 - ||\boldsymbol{v}_0||^2\right) = b,$$

and the same values of $b, c_1, \ldots, c_k$ also hold for any congruent copy $K' = \{\boldsymbol{v}_0', \boldsymbol{v}_1' \ldots, \boldsymbol{v}_n'\}$. For an r-colouring $\Delta : \mathbb{R} \to [r]$, define a new colouring $\Delta' : \mathbb{E}^{n'} \to [r]$ by

$$\Delta'(\boldsymbol{v}) = \Delta(||\boldsymbol{v}||).$$

In hopes of a contradiction, if for every $r \in \mathbb{Z}^+$ and every r-colouring of $\Delta : \mathbb{R} \to [r]$, there is K' monochromatic with respect to Δ', then letting $x_0 = ||\boldsymbol{v}_0'||^2, x_1 = ||\boldsymbol{v}_1'||^2, \ldots, x_k = ||\boldsymbol{v}_k'||^2$ there is a monochromatic solution to

$$c_1(x_1 - x_0) + c_2(x_2 - x_0) + \cdots + c(x_k - x_0) = b. \tag{4.12}$$

If there is a colouring $\mathbb{R}$ with no monochromatic solutions to (4.12), then a contradiction is reached.

Theorem 4.5.7 (Erdős *et. al*, 1973 [31])**.** *Let $\mathbb{F}$ be a field and $b, a_1, \ldots, a_n \in \mathbb{F}$ with $b \neq 0$. There exists a finite colouring Δ of $\mathbb{F}$ so that*

$$a_1(x_1 - x_i') + a_2(x_2 - x_2') + \cdots + a_n(x_n - x_n') = b$$

has no solution $x_1, x_1', x_2, x_2', \ldots, x_n, x_n'$ with the property that for all $i = 1, 2, \ldots, n$, $\Delta(x_i) = \Delta(x_i')$.

With $x_0 = x'_1 = x'_2 = \cdots = x'_n$, there is a finite colouring of $\mathbb{R}$ with no monochromatic solution to (4.12). In particular, if for any $b, a_1, \ldots, a_n \in \mathbb{F}$ with $a_1 + a_2 + \cdots + a_n = 0$, then there is a finite colouring with no monochromatic solution to

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b + x_0(a_1 + \cdots + a_n) = b, \tag{4.13}$$

and so if $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is partition regular, then $s = a_1 + \cdots + a_m \neq 0$, and in particular with $k = \frac{b}{s}$, then $x_1 = k, x_2 = k, \ldots, x_n = k$ is a solution to 4.13. By using Lemma 4.5.4, the general characterization of non-homogeneous partition regularity is true:

Corollary 4.5.8. *Let $\mathbb{F}$ be a field, A be an $m \times n$ matrix with entries in $\mathbb{F}$ and $\boldsymbol{b} \in \mathbb{F}^m \backslash \{0\}$. The pair $(A, \boldsymbol{b})$ is non-homogeneous partition regular in $\mathbb{F}$ if and only if there exists $\boldsymbol{k} = (k, k, \ldots, k) \in \mathbb{F}^m$ so that $A\boldsymbol{k} = \boldsymbol{0}$.*

Chapter 5

Image partition regular matrices

5.1 Introduction

The matrices

$$H = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$$

and

$$W = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

are partition regular in $\mathbb{Z}^+$ by Hilbert's cube lemma and van der Waerden's theorem respectively, since any Hilbert cube $H(a_0, a_1, a_2)$ is a solution to $H\boldsymbol{x} = \mathbf{0}$ and any AP_5 is a solution to $W\boldsymbol{x} = \mathbf{0}$. However, these systems also have the solutions $\boldsymbol{x} = (1, 1, 1, 1)^T$ and $(1, 1, 1, 1, 1)^T$ respectively, and so Hilbert's cube lemma and van der Waerden's theorem do not follow from the matrices H and W having the columns property. Any Hilbert cube $H(a_0, a_1, a_2)$ is a subset of $FS(\{a, a_1, a_2\})$ and so the partition regularity of matrices for finite sums set guarantee a non-constant monochromatic solution to $H\boldsymbol{x} = \boldsymbol{0}$. The partition regularity of matrices for $\{a + rd : r \in [k]\} \cup \{d\}$ guarantee a non-constant monochromatic solution to $W\boldsymbol{x} = \boldsymbol{0}$.

Hindman and Leader considered *image partition regular matrices* in 1993 [72].

Definition 5.1.1. *Let $\mathbb{S}$ be any one of $\mathbb{Z}^+, \mathbb{Z}, \mathbb{Q}^+, \mathbb{Q}, \mathbb{R}^+, \mathbb{R}$ or $\mathbb{C}$. An $m \times n$ matrix A is* image partition regular in $\mathbb{S}$*, denoted $IPR/\mathbb{S}$, if and only if for every finite colouring of $\mathbb{S} \setminus \{0\}$ there is a vector $\boldsymbol{x} \in (\mathbb{S} \setminus \{0\})^n$ and a monochromatic vector $\boldsymbol{y} \in (\mathbb{S} \setminus \{0\})^m$ so that $A\boldsymbol{x} = \boldsymbol{y}$.*

To differentiate image partition regular from the matrices of Chapter 4.4, partition regular matrices are sometimes called *kernel partition regular matrices*.

Consider the matrices

$$H' = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad W' = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}.$$

Then $H'(a_0, a_1, a_2)^T$ has $H(a_0, a_1, a_2)$ as its image, and $W'(a, d)^T$ has $(a, a+d, a+2d, a+3d, a+4d)$, which is an AP_5, as its image, and so the image partition regularity of H' and W' imply Hilbert's cube lemma and van der Waerden's theorem respectively.

Hindman and Leader first considered image partition regularity of matrices with rational entries in 1993 [72]. Since then, those characterizations have expanded to include real-valued matrices by Hindman [70], and all the different notions mentioned above for matrices with rational entries were explored in 2005 by Hindman and Strauss [75]. The results presented in this chapter are proved in the three sources mentioned, although some have been slightly reworked to include real-valued and complex-valued matrices.

Unlike partition regularity, image partition regularity in $\mathbb{Z}^+$ is not equivalent to image partition regularity in $\mathbb{Z}$ (similarly for $\mathbb{Q}$ and $\mathbb{R}$). In this chapter, characterizations of image partition regular matrices are presented, but only results for image partition regularity in $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are proved. These are proved by first studying a weaker condition on image partition regularity, where $\boldsymbol{x}$ is allowed to have zero entries (not all zero). The proofs of characterizations of image partition regularity in $\mathbb{Z}^+, \mathbb{Q}^+$ and $\mathbb{R}^+$ require more sophisticated tools.

5.2 Weakly image partition regular matrices

A formal definition of a weaker version of image partition regularity is given.

Definition 5.2.1. *Let $\mathbb{S}$ be any one of $\mathbb{Z}^+, \mathbb{Z}, \mathbb{Q}^+, \mathbb{Q}, \mathbb{R}^+, \mathbb{R}$ or $\mathbb{C}$. Let $\mathbb{T} = \mathbb{S} \cup \{-s : s \in \mathbb{S}\}$. An $m \times n$ matrix A is* weakly image partition regular in $\mathbb{S}$, *denoted $WIPR/\mathbb{S}$, if and only if for every finite colouring of $\mathbb{S} \setminus \{0\}$, there is a vector $\boldsymbol{x} \in \mathbb{T}^n \setminus \{\mathbf{0}\}$ and a monochromatic vector $\boldsymbol{y} \in (\mathbb{S} \setminus \{0\})^m$ so that $A\boldsymbol{x} = \boldsymbol{y}$.*

When $\mathbb{S}$ is $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$, the definition of weakly image partition regular allows the case when $\boldsymbol{y}$ has entries in $\mathbb{S}^+$ and $\boldsymbol{x}$ has entries in $\mathbb{S}$. This extra generalization, however, is equivalent to allowing $\boldsymbol{y}$ to range over $\mathbb{S} \setminus \{0\}$.

Lemma 5.2.2 (Hindman and Strauss, 2005 [75])**.** *For $\mathbb{S}$ being any one of $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$, A is weakly image partition regular in $\mathbb{S}$ ($WIPR/\mathbb{S}$) if and only if A is weakly image partition regular in $\mathbb{S}^+$ ($WIPR/\mathbb{S}^+$).*

Proof. The proof is a similar argument to the proof of Lemma 4.4.1. Since $\mathbb{S}^+ \subseteq \mathbb{S}$, then $WIPR/\mathbb{S}^+ \longrightarrow WIPR/\mathbb{S}$ is immediate. In the other direction, suppose A is $WIPR/\mathbb{S}$. Then for any colouring $\Delta : \mathbb{S}^+ \to [r]$, define a new colouring $\Delta' : \mathbb{S} \setminus \{0\} \to [2r]$ by

$$\Delta'(s) = \begin{cases} \Delta(s) & s > 0 \\ \\ \Delta(-s) + r & s < 0. \end{cases}$$

Since A is $WIPR/\mathbb{S}$, there is a monochromatic $\boldsymbol{y}$ with colour r' and a vector $\boldsymbol{x}$ so that $A\boldsymbol{x} = \boldsymbol{y}$. If $r' \leq r$, then $\boldsymbol{y}$ has positive entries. Otherwise, let $\boldsymbol{x}' = -\boldsymbol{x}$ and $\boldsymbol{y}' = -\boldsymbol{y}$. Then $A\boldsymbol{x}' = \boldsymbol{y}'$ and $\boldsymbol{y}$ has monochromatic entries in $\mathbb{S}^+$ with colour $r' - r$, and so A is $WIPR/\mathbb{S}^+$. □

The goal of the next lemma is to establish an equivalence between weakly image partition regularity of A and partition regularity of the following matrix:

Definition 5.2.3. *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. For an $m \times n$ matrix A with entries in $\mathbb{S}$ and $rank(A) = l < m$, let $I_0 \subseteq [m]$ be the set of indices of l linearly independent rows* *of A, and relabel the rows of A so that $I_0 = \{1, 2, \ldots, l\}$. Then for each $t \in [m]\setminus I_0 = \{l+1, \ldots, m\}$ and $i \in I_0$, let $\gamma_{t,i} \in \mathbb{S}$ be so that $\boldsymbol{r}_t = \sum_{i\in I_0} \gamma_{t,i}\boldsymbol{r}_i$. Define the $(m-l)\times m$ matrix $D(A)$ as follows:*

$$D(A) = \begin{bmatrix} \gamma_{l+1,1} & \gamma_{l+1,2} & \cdots & \gamma_{l+1,l} & -1 & 0 & \cdots & 0 \\ \gamma_{l+2,1} & \gamma_{l+2,2} & \cdots & \gamma_{l+2,l} & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{m,1} & \gamma_{m,2} & \cdots & \gamma_{m,l} & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

Remark 5.2.4. *Note that if A has integer entries, the matrix $D(A)$ need not have integer entries.*

A first characterization of weakly image partition regular matrices is given.

Lemma 5.2.5 (Hindman and Leader, 1993 [72]; Hindman, 2003 [70])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$ and A be an $m \times n$ matrix with entries in $\mathbb{S}$. Then A is weakly image partition regular in $\mathbb{S}$ ($WIPR/\mathbb{S}$) if and only if $rank(A) = m$ or $D = D(A)$ is partition regular in $\mathbb{S}$.*

Proof. In the first direction, suppose A is $WIPR/\mathbb{S}$ and $rank(A) = l < m$. For a finite colouring of $\mathbb{S} \setminus \{0\}$, there is $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{S}^n \setminus \{\boldsymbol{0}\}$ and a monochromatic $\boldsymbol{y} = (y_1, \ldots, y_m) \in (\mathbb{S} \setminus \{0\})^m$ so that $A\boldsymbol{x} = \boldsymbol{y}$.

By the definition of D, for any $t \in \{l+1, \ldots, m\}$,

$$\left(\sum_{i=1}^{l} \gamma_{t,i}\boldsymbol{r}_i\right) - \boldsymbol{r}_t = \boldsymbol{0},$$

hence for any column $j = 1, \dots, n$,

$$\left(\sum_{i=1}^{l} \gamma_{t,i} a_{i,j}\right) - a_{t,j} = 0.$$

Looking at the k-th entry of $\boldsymbol{z} = D\boldsymbol{y}$,

$$\begin{aligned} z_k &= \sum_{i=1}^{l} \gamma_{l+k,i} y_i - y_{l+k} \\ &= \sum_{i=1}^{l} \gamma_{l+k,i} \left(\sum_{j=1}^{n} a_{i,j} x_j\right) - \sum_{j=1}^{n} a_{l+k,j} x_j \\ &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{l} \gamma_{l+k,i} a_{i,j} - a_{l+k,j}\right) = 0, \end{aligned}$$

and so $D\boldsymbol{y} = \boldsymbol{0}$. Since $\boldsymbol{y}$ is monochromatic, D is partition regular in $\mathbb{S}$.

In the other direction, for any $r \in \mathbb{Z}^+$, let $\Delta : \mathbb{S} \setminus \{0\} \to [r]$ be an r-colouring of $\mathbb{S} \setminus \{0\}$. There are two cases to look at: either $rank(A) = m$, or $rank(A) < m$.

Case I: Let A' be an $m \times m$ submatrix of A with m linearly independent columns, suppose without loss of generality that these columns are indexed from 1 to l in A. Since A' is nonsingular, there is $\boldsymbol{x}' = (x_1, \dots, x_m)$ with entries in $\mathbb{S}$ (in $\mathbb{Q}$ if $\mathbb{S} = \mathbb{Z}$) so that $A'\boldsymbol{x}' = (1, 1, \dots, 1)^T$, which is always monochromatic. Define $\boldsymbol{x}$ with x_i in the first $i = 1, \dots, m$ entries, and 0 everywhere else. Then $A\boldsymbol{x} = (1, 1, \dots, 1)^T$. If $\mathbb{S} = \mathbb{Z}$, Let d be a common multiple of the denominators of $x_1, \dots, x_m$. Then $Ad\boldsymbol{x} = (d, d, \dots, d)$ is monochromatic, and $d\boldsymbol{x}$ has integer entries.

Case II: Let A' denote the nonsingular $l \times l$ submatrix of A consisting of the l linearly independent rows indexed by I_0 used to make $D = D(A)$ in Definition 5.2.3 and l linearly independent columns of these rows. Since D is partition regular in $\mathbb{S}$, there is a monochromatic vector $\boldsymbol{y} = (y_1, \dots, y_m)$ with entries in $\mathbb{S} \setminus \{0\}$ so that $D\boldsymbol{y} = \mathbf{0}$.

Since A' is non-singular, there is $\boldsymbol{x}' = (x_1, \dots, x_l)$ with entries in $\mathbb{S}$ (in $\mathbb{Q}$ if $\mathbb{S} = \mathbb{Z}$) so that

$$A' \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_l \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{bmatrix}.$$

Define the vector $\boldsymbol{x}$ whose first l entries are $x_1, \dots, x_l$, and 0 for the rest of the entries. All that is left is to show that $A\boldsymbol{x} = \boldsymbol{y}$.

For any $k = 1, \dots, l$, along the k-th row of A,

$$\sum_{j=1}^{n} a_{k,j} x_j = \sum_{j=1}^{l} a_{k,j} x_j = y_k.$$

For any $k \in l+1, \dots, m$, then from the definition of $D(A)$, $\boldsymbol{r}_k = \sum_{i=1}^{l} \gamma_{k,i} \boldsymbol{r}_i$, and so for any $j = 1, \dots, n$, $a_{k,j} = \sum_{i=1}^{l} \gamma_{k,i} a_{i,j}$. Since $D\boldsymbol{y} = \mathbf{0}$,

$$0 = \left(\sum_{i=1}^{l} \gamma_{k,i} y_i \right) - y_k,$$

and so

$$y_k = \sum_{i=1}^{l} \gamma_{k,i} y_i.$$

Along the k-th row of A,

$$\begin{aligned}\sum_{j=1}^{n} a_{k,j}x_j &= \sum_{j=1}^{l}\left(\sum_{i=1}^{l}\gamma_{k,i}a_{i,j}\right)x_j \\ &= \sum_{i=1}^{l}\gamma_{k,i}\left(\sum_{j=1}^{l}a_{i,j}x_j\right) \\ &= \sum_{i=1}^{l}\gamma_{k,i}y_i = y_k.\end{aligned}$$

Therefore, $A\boldsymbol{x} = \boldsymbol{y}$.

For $\mathbb{S} = \mathbb{Z}$, let $c = |A'| \neq 0$ be the determinant of A' and define a new colouring $\Delta' : \mathbb{Z} \to [r]$ by

$$\Delta'(z) = \Delta(cz).$$

Then repeat the argument to find a monochromatic $\boldsymbol{y} = (y_1, \ldots, y_m)$ with respect to Δ' with integer entries and $\boldsymbol{x}' = (x_1, \ldots, x_l)$ with rational entries so that

$$A'\boldsymbol{x}' = (y_1, \ldots, y_l)^T,$$

and $A\boldsymbol{x} = \boldsymbol{y}$, where $\boldsymbol{x}$ has x_i in its first l entries and 0 for the rest of its entries. Letting $adj(A')$ denote the adjoint of A', since A' is nonsingular,

$$\boldsymbol{x}' = (A')^{-1}(y_1, \ldots, y_l)^T = \frac{1}{|A'|}adj(A')(y_1, \ldots, y_l)^T.$$

Since $adj(A')$ has integer entries, so does $|A'|\boldsymbol{x}' = c\boldsymbol{x}'$. Letting $\boldsymbol{z} = c\boldsymbol{x}$, then $\boldsymbol{z}$ has integer entries, and $A\boldsymbol{z} = c\boldsymbol{y}$. By definition of Δ', the vector $c\boldsymbol{y}$ is monochromatic with respect to Δ with the same colour as $\boldsymbol{y}$. $\square$

The introduction of the matrix $D(A)$ allows for easy proofs of certain results, for example a result similar to that of Lemma 4.4.2.

Lemma 5.2.6 (Hindman and Strauss, 2005 [75])**.** *Let A be a $m \times n$ matrix with rational entries. Then A is weakly image partition regular in $\mathbb{Q}$ if and only if A is weakly image partition regular in $\mathbb{Z}$.*

Proof. In the first direction, since $\mathbb{Z} \subseteq \mathbb{Q}$, then if A is $WIPR/\mathbb{Z}$ then it is $WIPR/\mathbb{Q}$.

In the other direction, first suppose A has integer entries. Then by Lemma 5.2.5, either $rank(A) = m$ or $D(A)$ is partition regular over $\mathbb{Q}$. If $rank(A) = m$, then again by Lemma 5.2.5 (since A has entries in $\mathbb{Z}$), A is $WIPR/\mathbb{Z}$. If instead $rank(A) < m$ and $D(A)$ is partition regular $\mathbb{Q}$, then by Lemma 4.4.2, $D(A)$ is partition regular in $\mathbb{Z}$. Again by Lemma 5.2.5, A is $WIPR/\mathbb{Z}$.

If A has non-integer rational entries, let c be a common multiple of the denominators of the entries of A. Then cA has integer entries, and so by the previous argument it is partition regular in $\mathbb{Z}$ (if $A\boldsymbol{x} = \boldsymbol{y}$, then $cA(\frac{\boldsymbol{x}}{c}) = \boldsymbol{y}$, and so cA is $WIPR/\mathbb{Q}$). Then for any $r \in \mathbb{Z}^+$ and any r-colouring $\Delta : \mathbb{Z} \setminus \{0\} \to [r]$, there is a vector $\boldsymbol{x} \in \mathbb{Z}^n \setminus \{\boldsymbol{0}\}$ and a monochromatic vector $\boldsymbol{y} \in (\mathbb{Z} \setminus \{0\})^m$, so that $cA\boldsymbol{x} = \boldsymbol{y}$. Of course $\boldsymbol{z} = c\boldsymbol{x}$ has integer entries, and $A\boldsymbol{z} = \boldsymbol{y}$, and so A is $WIPR/\mathbb{Z}$. $\square$

The following lemma is useful for proving an equivalence between weakly image partition regularity and certain cases of image partition regularity.

Lemma 5.2.7 (Hindman and Leader, 1993 [72], Hindman, 2003 [70])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. Then A is weakly image partition regular in $\mathbb{S}$ ($WIPR/\mathbb{S}$) if and only if for every $\boldsymbol{p} \in \mathbb{S}^n \setminus \{\mathbf{0}\}$, there exists $b \in \mathbb{S} \setminus \{0\}$ so that*

$$\begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}$$

is weakly image partition regular in $\mathbb{S}$.

Proof. For $\boldsymbol{p} \in \mathbb{S}^n \setminus \{\mathbf{0}\}$ and $b \in \mathbb{S} \setminus \{0\}$, define the matrix

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}.$$

In the first direction, a weakly image partition regular matrix with a row removed is also weakly image partition regular.

In the other direction, assume A with rows $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m$ is $WIPR/\mathbb{S}$. For any $\boldsymbol{p} \in \mathbb{S}^n \setminus \{0\}$, the goal is to find b so that $B_{b\boldsymbol{p}}$ has rank $m+1$ or $D(B_{b\boldsymbol{p}})$ satisfies the columns property over $\mathbb{S}$, which by Theorem 4.4.5 is equivalent to partition regularity in $\mathbb{S}$. If the columns property is satisfied, then by Lemma 5.2.5, $B_{b\boldsymbol{p}}$ is $WIPR/\mathbb{S}$.

First assume $rank(A) = m$. If $\boldsymbol{p}$ is not a linear combination of the rows of A, let $b = 1$. Then $B_{\boldsymbol{p}}$ has rank $m + 1$. If $\boldsymbol{p}$ is a linear combination of the rows of A, choose $\beta_1, \ldots, \beta_m \in \mathbb{S}$ so that $\boldsymbol{p} = \beta_1\boldsymbol{r}_1 + \cdots + \beta_m\boldsymbol{r}_m$. Choose any $\beta_j \neq 0$, and let $b = 1/\beta_j$. Then

$$D(B_{b\boldsymbol{p}}) = \begin{bmatrix} b\beta_1 & b\beta_2 & \cdots & b\beta_m & -1 \end{bmatrix}$$

has the columns property with $I_1 = \{j, m+1\}$ and $I_2 = [m] \setminus \{j\}$, with $\alpha_{j,1} = 0$ and $\alpha_{m+1,1} = -\sum_{i \in I_2} b\beta_i$.

Now assume $rank(A) = l < m$. Then $D(A)$ with columns $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_m$ is partition regular in $\mathbb{S}$ and satisfies $CP(\mathbb{S})$. Let $[n] = I_1 \cup \cdots \cup I_k$ be a partition of the columns of $D(A)$ and for $j = 1, 2, \ldots, k-1$ and $i \in I_1 \cup \cdots \cup I_k$, let $\alpha_{i,j} \in \mathbb{S}$ be guaranteed by the columns property so that $\sum_{i \in I_1} \boldsymbol{d}_i = \boldsymbol{0}$ and $\sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \boldsymbol{d}_i = \sum_{i \in I_{j+1}} \boldsymbol{d}_i$. Assume $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_l$ are the l linearly independent rows of A used to make $D(A)$ (and so $I_0 = \{1, 2, \ldots, l\}$).

If $\boldsymbol{p}$ is not a linear combination of $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_l$, let $b = 1$, and rearrange $B_{b\boldsymbol{p}}$ so that the $(l+1)$-th row is $\boldsymbol{p}$. Add $l+1$ to I_0, and let $D(B_{\boldsymbol{p}})$ be the matrix $D(A)$ with an added empty column $\boldsymbol{c}_{l+1}$. Since $D(A)$ satisfies the columns property over $\mathbb{S}$, so does the matrix $D(B_{\boldsymbol{p}})$ with $l+1$ added to I_1.

If $\boldsymbol{p}$ is a linear combination of $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_l$, choose $\beta_1, \beta_2, \ldots, \beta_l$ so that $\boldsymbol{p} = \sum_{i \in I_0} \beta_i \boldsymbol{r}_1$. For a nonzero b, let the matrix $D(B_{b\boldsymbol{p}})$ be the matrix with $D(A)$ in the $m - l \times l$ upper left corner, $d_{i,m+1} = 0$ for $i = 1, 2, \ldots, m-l$, and the $m-l+1$ row being

$$\begin{pmatrix} b\beta_1 & b\beta_2 & \cdots & b\beta_l & 0 & \cdots & 0 & -1 \end{pmatrix}.$$

All that is left to finish the proof is to find a value for b so that $D(B_{b\boldsymbol{p}})$ satisfies $CP(\mathbb{S})$. The new $(m+1)$-th column needs to be added to some I_t.

If $\sum_{i \in I_1} \beta_i \neq 0$ let $b = 1/\sum_{i \in I_1} \beta_i$. The columns property is satisfied by adding $m+1$

to I_1 and for $j = 1, 2, \ldots, k-1$ defining recursively

$$\alpha_{m+1,j} = \sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} b\beta_i - \sum_{i \in I_{j+1}} b\beta_i.$$

If $\sum_{i \in I_1} \beta_i = 0$ and for some $t = 1, 2, \ldots, k-1$,

$$\sum_{i \in I_1 \cup \cdots \cup I_t} \alpha_{i,j} b\beta_i \neq \sum_{i \in I_{t+1}} b\beta_i,$$

choose the least such t and let $b = 1/(\sum_{i \in I_1 \cup \cdots \cup I_t} \alpha_{i,j} b\beta_i - \sum_{i \in I_{t+1}} b\beta_i)$. The columns property is satisfied by adding $m+1$ to I_t and for $j = t, t+1, \ldots, k-1$ defining recursively

$$\alpha_{m+1,j} = \sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} b\beta_i - \sum_{i \in I_{j+1}} b\beta_i.$$

If no such t exists, then let $b = 1$, and let $I_{k+1} = \{m+1\}$. Choose $s \in \{1, 2, \ldots, l\}$ so that $\beta_s \neq 0$. The columns property is satisfied by defining, $\alpha_{s,k} = \frac{-1}{\beta_s}$, for $i \in \{1, 2, \ldots, l\} \setminus \{s\}$ defining $\alpha_{i,k} = 0$, and for $i \in \{l+1, \ldots, m\}$ defining $\alpha_{i,k} = \frac{-\gamma_{i,s}}{\beta_s}$. □

5.3 Characterizations of image partition regular matrices

Having proved several characterizations of matrices that are weakly image partition regular, the next lemma states that these are equivalent to image partition regularity when $\mathbb{S}$ is $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. Note that for $\mathbb{S} \neq \mathbb{C}$ image partition regularity over $\mathbb{S}$ is not equivalent to image partition regularity over $\mathbb{S}^+$, even though the equivalence is true for weakly image partition regularity.

Lemma 5.3.1 (Hindman and Strauss, 2005 [75])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. The $m \times n$ matrix A is weakly image partition regular in $\mathbb{S}$ ($WIPR/\mathbb{S}$) if and only if A is image partition regular in $\mathbb{S}$ ($IPR/\mathbb{S}$).*

Proof. The implication $IPR/\mathbb{S} \longrightarrow WIPR/\mathbb{S}$ is immediate, since $(\mathbb{S} \setminus \{0\})^n \subseteq \mathbb{S}^n \setminus \{\mathbf{0}\}$.

In the other direction, if A is $WIPR/\mathbb{S}$, then by consecutively applying Lemma 5.2.7 with $\boldsymbol{p}_1 = (1, 0, \ldots, 0, 0)$, $\boldsymbol{p}_2 = (0, 1, \ldots, 0, 0)$, $\ldots$, $\boldsymbol{p}_n = (0, 0, \ldots, 0, 1)$, there exists nonzero $b_1, \ldots, b_n$ so that

$$B = \begin{bmatrix} & A & & \\ b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{bmatrix}$$

is $WIPR/\mathbb{S}$. By the definition of $WIPR/\mathbb{S}$, for every finite colouring of $\mathbb{S} \setminus \{0\}$, there is $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{S}^n \setminus \{\mathbf{0}\}$ and a monochromatic $\boldsymbol{y} \in (\mathbb{S} \setminus \{0\})^{m+n}$ so that $B\boldsymbol{x} = \boldsymbol{y}$. Let $\boldsymbol{y}' \in (\mathbb{S} \setminus \{0\})^m$ consist of the first m entries of $\boldsymbol{y}$. Then $A\boldsymbol{x} = \boldsymbol{y}'$, and $\boldsymbol{y}'$ is monochromatic. Looking at the last n entries of $\boldsymbol{y} = (y_1, \ldots, y_{m+n})$,

$$\begin{aligned} y_{m+1} &= b_1 x_1, \\ y_{m+2} &= b_2 x_2, \\ &\vdots \\ y_{m+n} &= b_n x_n. \end{aligned}$$

Since none of $b_1, \ldots, b_n$ are zero and none of $y_1, \ldots, y_n$ are zero, then none of $x_1, \ldots, x_n$ are zero. Therefore, $\boldsymbol{x} \in (\mathbb{S} \setminus \{0\})^n$, and so A is $IPR/\mathbb{S}$. □

Using Lemma 5.3.1 along with the characterizations of Section 5.2, a characterization of image partition regular matrices in $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ can be proved. Recall from Definition 5.2.3 the definition of the matrix $D(A)$.

Theorem 5.3.2 (Hindman and Leader, 1993 [72]; Hindman, 2003 [70]; Hindman and Strauss, 2005 [75])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, let $IPR/\mathbb{S}$ denote image partition regularity in $\mathbb{S}$ and $PR/\mathbb{S}$ denote partition regular in $\mathbb{S}$.*

1. *Let A be a matrix with entries in $\mathbb{Q}$. Then the following are equivalent:*

   (a) A is $IPR/\mathbb{Z}$;

   (b) A is $IPR/\mathbb{Q}$;

   (c) A is $IPR/\mathbb{R}$;

   (d) A is $IPR/\mathbb{C}$;

   (e) $rank(A) = m$ or $D(A)$ is $PR/\mathbb{Q}$;

   (f) For every $\boldsymbol{p} \in \mathbb{Q}^n$, there exists $b \in \mathbb{Q} \setminus \{0\}$ so that

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}$$

   is $IPR/\mathbb{Q}$.

2. *Let A be a matrix with entries in $\mathbb{R}$. Then the following are equivalent:*

(a) A is $IPR/\mathbb{R}$*;*

(b) A is $IPR/\mathbb{C}$*;*

(c) $rank(A) = m$ *or* $D(A)$ *is* $PR/\mathbb{R}$*;*

(d) For every $\boldsymbol{p} \in \mathbb{R}^n$*, there exists* $b \in \mathbb{R} \setminus \{0\}$ *so that*

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}$$

is $IPR/\mathbb{R}$*.*

3. Let A be a matrix with entries in $\mathbb{C}$*. Then the following are equivalent:*

(a) A is $IPR/\mathbb{C}$*;*

(b) $rank(A) = m$ *or* $D(A)$ *is* $PR/\mathbb{C}$*;*

(c) For every $\boldsymbol{p} \in \mathbb{C}^n$*, there exists* $b \in \mathbb{C} \setminus \{0\}$ *so that*

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}$$

is $IPR/\mathbb{C}$*.*

Proof. Let $\mathbb{S}$ be any one of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. By Lemma 5.3.1, A is $IPR/\mathbb{S}$ if and only if A is $WIPR/\mathbb{S}$. Then by Lemma 5.2.7, A is $WIPR/\mathbb{S}$ if and only if for every $\boldsymbol{p} \in (\mathbb{S})^n$, there exists a nonzero $b \in \mathbb{S} \setminus \{0\}$ so that $B_{b\boldsymbol{p}}$ is $WIPR/\mathbb{S}$. Then again by Lemma 5.3.1, $B_{b\boldsymbol{p}}$ is $WIPR/\mathbb{S}$ if and only if $B_{b\boldsymbol{p}}$ is $IPR/\mathbb{S}$.

The rest of the equivalences are achieved by applying Lemmas 5.3.1 and 5.2.5, and the characterization theorem for partition regularity Theorem 4.4.5.

If A has rational entries, then the following equivalencies hold:

$$\begin{array}{ccccccc} A \text{ is } IPR/\mathbb{Z} & & A \text{ is } IPR/\mathbb{Q} & & A \text{ is } IPR/\mathbb{R} & & A \text{ is } IPR/\mathbb{C} \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ A \text{ is } WIPR/\mathbb{Z} & & A \text{ is } WIPR/\mathbb{Q} & & A \text{ is } WIPR/\mathbb{R} & & A \text{ is } WIPR/\mathbb{C} \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ D(A) \text{ is } PR/\mathbb{Z} & \longleftrightarrow & D(A) \text{ is } PR/\mathbb{Q} & \longleftrightarrow & D(A) \text{ is } PR/\mathbb{R} & \longleftrightarrow & D(A) \text{ is } PR/\mathbb{C}. \end{array}$$

If A has real entries, then the following equivalencies hold:

$$\begin{array}{ccc} A \text{ is } IPR/\mathbb{R} & & A \text{ is } IPR/\mathbb{C} \\ \updownarrow & & \updownarrow \\ A \text{ is } WIPR/\mathbb{R} & & A \text{ is } WIPR/\mathbb{C} \\ \updownarrow & & \updownarrow \\ D(A) \text{ is } PR/\mathbb{R} & \longleftrightarrow & D(A) \text{ is } PR/\mathbb{C}. \end{array}$$

Finally, if A has complex entries, the following equivalencies hold:

$$A \text{ is } IPR/\mathbb{C} \longleftrightarrow A \text{ is } WIPR/\mathbb{C} \longleftrightarrow D(A) \text{ is } PR/\mathbb{C}.$$

□

Characterizations of image partition regular matrices in $\mathbb{Z}^+, \mathbb{Q}^+$ and $\mathbb{R}^+$ are very similar. A notable exception is the lack of partition regularity of $D(A)$. If $D(A)$ is partition regular in $\mathbb{Z}^+$ for example, then by Lemma 4.4.1, it is also partition regular in

$\mathbb{Z}$, and so only image partition regularity in $\mathbb{Z}$ can be guaranteed. In Section 5.2, the matrix $D(A)$ was instrumental in many of the proofs of weakly image partition regularity, and so many of the proofs of characterizations of image partition regularity in positive sets do not follow from a simple reworking of the proof of the analogous statement for weakly image partition regularity.

Only a short list of characterizations presented in [70],[72], and [75] are given here. For example, the characterization of image partition regular matrices over the reals in [70] has 13 equivalencies.

Theorem 5.3.3 (Hindman and Leader, 1993 [72]; Hindman, 2003 [70]; Hindman and Strauss, 2005 [75])**.** *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ and let $IPR/\mathbb{S}$ denote image partition regularity in $\mathbb{S}$.*

1. *Let A be a matrix with entries in $\mathbb{Q}$. Then the following are equivalent:*

   (a) A is $IPR/\mathbb{Z}^+$;

   (b) A is $IPR/\mathbb{Q}^+$;

   (c) A is $IPR/\mathbb{R}^+$;

   (d) For every $\boldsymbol{p} \in (\mathbb{Q} \setminus \{0\})^n$, there exists $b \in \mathbb{Q}^+$ so that

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ b\boldsymbol{p} \end{bmatrix}$$

   is $IPR/\mathbb{Q}^+$.

2. Let A be a matrix with entries in $\mathbb{R}$. Then the following are equivalent:

(a) A is $IPR/\mathbb{R}^+$;

(b) For every $\boldsymbol{p} \in (\mathbb{R} \setminus \{0\})^n$, there exists $b \in \mathbb{R}^+$ so that

$$B_{b\boldsymbol{p}} = \begin{bmatrix} A \\ \\ b\boldsymbol{p} \end{bmatrix}$$

is $IPR/\mathbb{R}^+$.

Chapter 6

Deuber's (m, p, c)-sets

6.1 Introduction

Rado conjectured that for every finite colouring of $\mathbb{Z}^+$, one of the colour classes contain a solution to every partition regular system $A\boldsymbol{x} = \mathbf{0}$ in $\mathbb{Z}^+$.

Definition 6.1.1. *The set $X \subseteq \mathbb{Z}^+$ is* large *if and only if X contains a solution to every partition regular system. That is, for every $m \times n$ partition regular matrix A, there is a vector $\boldsymbol{v} \in X^n$ for which $A\boldsymbol{v} = \mathbf{0}$.*

In 1973, Deuber [23] proved Rado's conjecture, by proving that large sets are partition regular using what are called (m, p, c)-sets, what are sometimes called *Deuber sets* (see, for example, [57]). The proof is presented in Section 6.5.

Definition 6.1.2. *For $m, p, c \in \mathbb{Z}^+$, the set $M \subseteq \mathbb{Z}^+$ is an* (m, p, c)-set *if and only if* *there exists $x_0, x_1 \ldots, x_m \in \mathbb{Z}^+$ and sets*

$$R_0(M) = \Big\{cx_0 + \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m : \lambda_1, \lambda_2, \ldots, \lambda_m \in [-p,p]\Big\} \subseteq \mathbb{Z}^+$$

$$R_1(M) = \Big\{cx_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m : \lambda_2, \ldots, \lambda_m \in [-p,p]\Big\} \subseteq \mathbb{Z}^+$$

$$\vdots$$

$$R_m(M) = \Big\{cx_m\Big\} \subseteq \mathbb{Z}^+$$

such that $M = \cup_{k=0}^{m} R_k(M)$. The set $R_k(M)$ is called the k-th row of M *and such an (m,p,c)-set M is denoted $M = (x_0, x_1, \ldots, x_m)_{p,c}$.*

The Hilbert cube $H(a_0, a_1, \ldots, a_m)$ is a subset of the 0-th row of the $(m,1,1)$-set $(a_0, a_1, \ldots, a_n)_{1,1}$. Any set $\{x, y, x+y\}$ is a subset of the $(1,1,1)$-set $(x,y)_{1,1}$. An arithmetic progression $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$ is a subset of the 0-th row of the $(1, k-1, 1)$-set $(a,d)_{k-1,1}$, and the difference $\{d\}$ forms row 1 of $(a,d)_{k-1,1}$. Finally, a finite sums set $FS(\{a_1, \ldots, a_m\})$ is a subset of the $(m-1,1,1)$-set $(a_1, a_2, \ldots, a_m)_{1,1}$. Deuber [23] proved that (m,p,c)-sets are partition regular. In fact, he proved that for every $m, p, c \in \mathbb{Z}^+$, there exists $n, q, d \in \mathbb{Z}^+$ so that for every r-colouring of any (n,q,d)-set there is a monochromatic (m,p,c)-set. The proof of Deuber's result given here, Theorem 6.3.1, is due to Leeb [85] and uses the Hales–Jewett theorem. Section 6.2 is devoted to the Hales–Jewett theorem.

Section 6.4 contains an exploration of the connection between (m,p,c)-sets and matrices with the columns property, as well as the missing direction of the proof of Theorem 4.4.3. To complete the missing directions of the proof of Theorem 4.4.5, a new generalization of (m, p, c)-sets to $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ is introduced:

Definition 6.1.3. *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. For $m \in \mathbb{Z}^+$ and $P \subseteq \mathbb{S}$ with $|P| < \infty$ and $0 \in P$, the set $M \subseteq \mathbb{S}$ is a* generalized Deuber set $GD(m, P, 1)$ *if and only if there exists $x_0, \ldots, x_m \in \mathbb{Z}^+$ and sets*

$$R_0(M) = \Big\{x_0 + \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m : \lambda_1, \lambda_2, \ldots, \lambda_m \in P\Big\} \subseteq \mathbb{S}$$
$$R_1(M) = \Big\{x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m : \lambda_2, \ldots, \lambda_m \in P\Big\} \subseteq \mathbb{S}$$
$$\vdots$$
$$R_m(M) = \Big\{x_m\Big\} \subseteq \mathbb{S}$$

such that $M = \cup_{k=0}^{m} R_k(M)$. The set $R_k(M)$ is called the k-th row of M.

The idea of (m, p, c)-sets is generalized by Hindman and Leader [72] as the images of what are called first entries matrices (see [70], [72], and [75]). Both (m, p, c)-sets and $GD(m, P, 1)$ are images of first entries matrices.

Deuber's original proofs are in German, but English versions of the proofs were published by Gunderson in 2002 [53].

6.2 The Hales–Jewett Theorem

It is well known that every game of Tic-Tac-Toe can end in a tie. If the game is played in more dimensions, however, is there is a guarantee that someone wins, that is, there are always three **X**'s or three **O**'s in a line? This is different from asking whether for some dimension n the first player has a strategy to always win (a strategy can be formulated for 3-dimensional Tic-Tac-Toe, which is left as an exercise in [6]), but rather, if there exists an n so that for any 2-coloring of the entries of an n-dimensional Tic-Tac-Toe grid with **O** and **X**, there are 3 **O**'s or 3 **X**'s in a line.

Before continuing the example of Tic-Tac-Toe, a few definitions for parameter words are introduced.

Definition 6.2.1. *Let $A = \{a_1, \ldots, a_k\}$ be a finite alphabet, and $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$ be a set of elements called* parameters. *Let*

$$(A \cup \Lambda)^n = \{f = (f(1), \ldots, f(n)) : f(1), \ldots, f(n) \in A \cup \Lambda\},$$

and for $f \in (A \cup \Lambda)$ and $j \in [m]$, let $f^{-1}(\lambda_j) = \{i \in [n] : f(i) = \lambda_j\}$. Then define the set of parameter words of length n over A *to be*

$$[A]\binom{n}{m} = \Big\{f \in (A \cup \Lambda)^n : \forall j \in [m], f^{-1}(\lambda_j) \neq \emptyset \text{ and } \forall j_1 < j_2,$$

$$\min\{i \in f^{-1}(\lambda_{j_1})\} < \min\{i \in f^{-1}(\lambda_{j_2})\}\Big\}.$$

The set $[A]\binom{n}{0}$ is simply A^n.

These parameter words have a composition similar to the composition of functions.

Definition 6.2.2. *For $f \in [A]\binom{n}{m}$ and $g \in [A]\binom{m}{k}$, define $f \circ g \in [A]\binom{n}{k}$ to be the* composition of f and g *by*

$$(f \circ g)(i) = \begin{cases} f(i) & f(i) \in A \\ g(j) & f(i) = \lambda_j. \end{cases}$$

Definition 6.2.3. *For $f \in [A]\binom{n}{m}$, define the* space of f *to be*

$$sp(f) = \left\{ f \circ g : g \in [A]\binom{m}{0} \right\}.$$

The space of f comprises all the ways of replacing the parameters of f with elements of A, where the parameter λ_j is replaced with the same element of A wherever it appears in f. *Combinatorial lines* are the spaces of parameter words with a single parameter.

The positions in the n-dimensional Tic-Tac-Toe grid can be identified by coordinates in an n-dimensional space. For example, the classic Tic-Tac-Toe board has the following identifications:

(1,3)	(2,3)	(3,3)
(1,2)	(2,2)	(3,2)
(1,1)	(2,1)	(3,1)

Then placing **O**'s and **X**'s in the grid is the same as a 2-colouring of

$$\{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}.$$

In the simple example for classic Tic-Tac-Toe the set

$$[\{1,2,3\}]\binom{2}{1} = \{(1,\lambda), (2,\lambda), (3,\lambda), (\lambda,1), (\lambda,2), (\lambda,3), (\lambda,\lambda)\}$$

is the set of all parameter words of length two over $\{1,2,3\}$. For any $f \in [\{1,2,3\}]\binom{2}{1}$, the set $sp(f)$ is a combinatorial line. For example, if $f = (2,\lambda)$, then $sp(f)$ is the combinatorial line $\{(2,1), (2,2), (2,3)\}$. All such combinatorial lines are also geometric lines. The only geometric line that is not a combinatorial line is

$$\{(1,3), (2,2), (3,1)\}.$$

A special case of the Hales–Jewett theorem states that for dimensions n large enough, and any two colouring of the coordinates of the n-dimensional Tic-Tac-Toe grid

$$\big\{(a_1, a_2, \ldots, a_n) : a_i \in \{1, 2, 3\}\big\}$$

as either **O** or **X**, there is a monochromatic combinatorial line. While the Hales–Jewett theorem was proved in 1963, it was only in 2014 that Hindman and Tessler [77] proved that for this specific case, $n \geq 4$ will do. Since every combinatorial line is also a geometric line, as a consequence in every game of 4-dimensional Tic-Tac-Toe, there is always a winner. The number of dimensions needed to guarantee a geometric line, however, is not always the same as the number of dimensions needed to guarantee a combinatorial line. In the case of 3-dimensional Tic-Tac-Toe for example, there is always a geometric line (see [6], again given as an exercise), but Figure 6.1 is an example of a game with no combinatorial lines and 9 geometric lines.

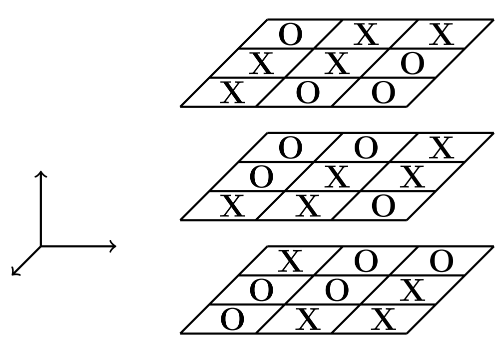

Figure 6.1: A game of 3-dimensional Tic-Tac-Toe with no combinatorial line

For two parameter words f and g, it is useful to 'add' g to the end of f. Formally, this concatenation is defined below.

Definition 6.2.4. *For $f \in [A]\binom{m}{n}$ and $g \in [A]\binom{k}{l}$, define $f \wedge g \in [A]\binom{n+k}{m+l}$ to be the* concatenation of f and g *by*

$$(f \wedge g)(i) = \begin{cases} f(i) & i \in [n] \\ g(i-n) & i \in [n+1, n+m], \ \ g(i-n) \in A \\ \lambda_{j-m} & i \in [n+1, n+m], \ \ g(i-n) = \lambda_j. \end{cases}$$

The general statement of the Hales–Jewett theorem allows for more than 2 colours to be used, for A to have any number of symbols, and for spaces of words with more than one parameter to be monochromatic.

Theorem 6.2.5 (Hales and Jewett, 1963 [60])**.** *For every finite alphabet A and $m, r \in \mathbb{Z}^+$, there exists a least $n = HJ(|A|, m, r) \in \mathbb{Z}^+$ so that for every r-colouring $\Delta : [A]\binom{n}{0} \to [r]$, there exists $f \in [A]\binom{n}{m}$ for which $sp(f)$ is monochromatic.*

The proof given here can be found in [97].

Proof. For $t, m, r \in \mathbb{Z}^+$, the Hales–Jewett theorem is proved by double induction using the following two inequalities:

$$HJ(t, m+1, r) \leq HJ(t, 1, r) + HJ(t, m, r^{t^{HJ(t,1,r)}}), \tag{6.1}$$

$$HJ(t+1, 1, r+1) \leq HJ(t, 1 + HJ(t+1, 1, r), r+1), \tag{6.2}$$

and the base case $HJ(2, 1, r) \leq r$.

Base Case: For any r-colouring of $\Delta : [\{0,1\}]\binom{r}{0} \to [r]$, look at the $r+1$ words

$$f_0 = (0, 0, 0, \dots, 0, 0)$$

$$f_1 = (1, 0, 0, \dots, 0, 0)$$

$$f_2 = (1, 1, 0, \dots, 0, 0)$$

$$\vdots$$

$$f_{r-1} = (1, 1, 1, \dots, 1, 0)$$

$$f_r = (1, 1, 1, \dots, 1, 1)$$

By the pigeonhole principle, two of these words have the same colour, say f_i and f_j with $i < j$. Then let $f \in [\{0,1\}]\binom{r}{1}$ consist of 1 in the first i coordinates, λ in coordinates $i+1, i+2, \dots, j$, and 0 in the remaining coordinates. Then $sp(f) = \{f_1, f_2\}$ is monochromatic, completing the base case.

To prove (6.1), let A be a set with $|A| = t$. Set $M = HJ(t, 1, r)$, $N = HJ(t, m, r^{t^M})$ and fix an indexing $g_1, \dots, g_{t^M}$ on $[A]\binom{M}{0}$. Notice that for every $f \in [A]\binom{N}{0}$ and $g \in [A]\binom{M}{0}$, $f \wedge g \in [A]\binom{M+N}{0}$. Then for any r-colouring $\Delta : [A]\binom{M+N}{0} \to [r]$, this colouring imposes a r^{t^M}-colouring $\Delta_N : [A]\binom{N}{0} \to [r^{t^M}]$ defined by

$$\Delta_N(f) = (\Delta(g_1 \wedge f), \dots, \Delta(g_{t^M} \wedge f)).$$

Then by the choice of N, there exists $f' \in [A]\binom{N}{m}$ so that $sp(f')$ is monochromatic with respect to Δ_N. That is, for any fixed $g \in [A]\binom{M}{0}$, $sp(g \wedge f')$ is monochromatic with respect to Δ, not necessarily the same colour for every g. See Figure 6.2 for an illustration of $g \wedge f'$.

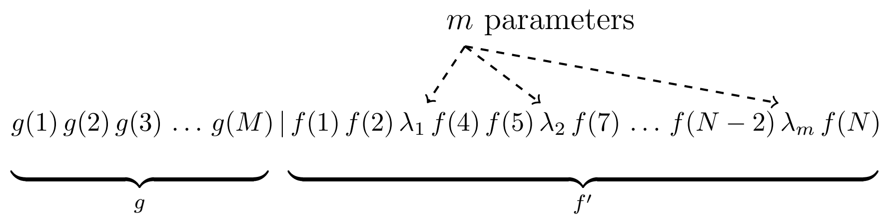

Figure 6.2: For any fixed g, $sp(g \,^\wedge f')$ is monochromatic.

For any $h \in [A]\binom{m}{0}$, define an r-colouring $\Delta_M : [A]\binom{M}{0} \to [r]$ by

$$\Delta_M(g) = \Delta(g \,^\wedge (f' \circ h)).$$

The choice of h does not matter, since $sp(f')$ is monochromatic. Then by the choice of M, there exists $g' \in [A]\binom{M}{1}$ so that $sp(g')$ is monochromatic. And so, with $g' \,^\wedge f' \in [A]\binom{M+N}{m+1}$, $sp(g' \,^\wedge f')$ is monochromatic with respect to Δ, completing the proof of (6.1). Figure 6.3 illustrates $g' \,^\wedge f'$.

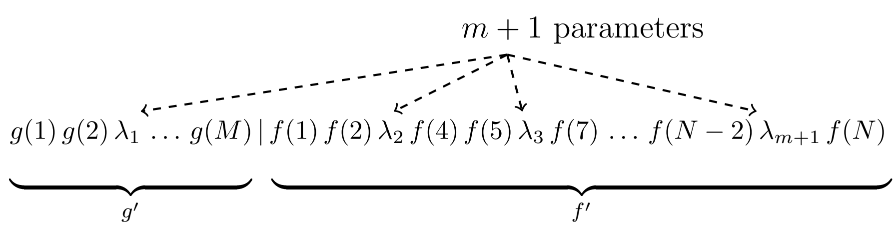

Figure 6.3: $sp(g' \,^\wedge f')$ is monochromatic.

To prove (6.2), set $M = HJ(t+1, 1, r)$ and $N = HJ(t, 1+M, r+1)$. Let B be an alphabet with $|B| = t+1$. For any $b \in B$, let $A = B \setminus \{b\}$. Fix an r-colouring $\Delta : [B]\binom{N}{0} \to [r+1]$. Define an $r+1$-colouring

$$\Delta_A : [A]\binom{N}{0} \to [r+1]$$

by restricting Δ to $[A]\binom{N}{0}$. By the choice of N, there exists $f' \in [A]\binom{N}{1+M}$ so that $sp(f')$ is monochromatic, say with colour $r+1$.

Look at the following two cases:

Case I: There exists $g' \in [B]\binom{M}{0}$ so that for $f' \circ (b \,^\wedge g') \in [B]\binom{N}{0}$,

$$\Delta(f' \circ (b \,^\wedge g')) = r + 1.$$

Then define $h \in [B]\binom{N}{1}$ to be $h = f' \circ (\lambda_1 \,^\wedge g')$. Then for any $a \in A$, the composition $h \circ a \in sp(f')$, and so has colour $r+1$. For b, the composition $h \circ b = f' \circ (b \,^\wedge g')$, and also has colour $r+1$. And so $sp(h)$ is monochromatic.

Case II: No such g' exists. Then for every $g \in [B]\binom{M}{0}$,

$$\Delta(f' \circ (b \,^\wedge g)) \neq r + 1.$$

Then define an r-colouring $\Delta_M : [B]\binom{M}{0} \to [r]$ by

$$\Delta_M(g) = \Delta(f' \circ (b \,^\wedge g)).$$

By the choice of M, there exists $g' \in [B]\binom{M}{1}$ so that $sp(g')$ is monochromatic with respect to Δ_M, say with colour r. Define $h \in [B]\binom{N}{1}$ to be $h = f' \circ (b \,^\wedge g')$. Then for any $x \in B$,

$$h \circ x = (f' \circ (b \,^\wedge g')) \circ x,$$

and since the parameter λ appears in indices of h corresponding to g'',

$$(f' \circ (b \,^\wedge g')) \circ x = f' \circ (b \,^\wedge (g' \circ x)).$$

Therefore, for any $x \in B$,

$$\Delta(h \circ x) = \Delta(f' \circ (b^\wedge (g' \circ x))) = \Delta_M(g' \circ x) = r,$$

and so $sp(h)$ is monochromatic, completing the proof of (6.2). □

The Hales–Jewett theorem gives a simple proof of van der Waerden's theorem (Theorem 3.4.4); the proof presented here is from [48]. The result of Shelah's that the van der Waerden numbers $W(k;r)$ are primitive recursive actually stems from proving that the Hales–Jewett numbers $HJ(k,m,r)$ are primitive recursive [114], then bounding the van der Waerden numbers by a function of $HJ(k,m,r)$.

Second proof of Theorem 3.4.4. For any $k, r \in \mathbb{Z}^+$, let $m = HJ(k-1, 1, r)$, and let $n = k^m$. Then every $x \in [n]$ can be written as $1 + x_1 + x_2 k + x_3 k^2 + \cdots + x_m k^{m-1}$, where $x_0, x_1, \ldots, x_{m-1} \in [0, k-1] = A$. For any r-colouring $\Delta : [n] \to [r]$, define a new r-colouring $\Delta' : [A]\binom{m}{0} \to [r]$ by

$$\Delta'(f') = \Delta(1 + f'(1) + f'(2)k + \cdots + f'(m)k^{m-1}).$$

By the choice of m, there exists $f \in [A]\binom{m}{1}$ so that $sp(f)$ is monochromatic with respect to Δ'. Then setting

$$d = \sum_{i \in f^{-1}(\lambda_1)} k^{i-1}$$

and

$$a = \sum_{i \notin f^{-1}(\lambda_1)} f(i) k^{i-1},$$

the AP_k $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$ is monochromatic. Therefore $W(k;r) \le k^m$. □

6.3 Partition regularity of (m,p,c)-sets

In the next theorem, the partition regularity of (m,p,c)-sets is proved.

Theorem 6.3.1 (Deuber, 1973 [23]). *For every $m,p,c,r \in \mathbb{Z}^+$, there exists $n,q,d \in \mathbb{Z}^+$ so that for every (n,q,d)-set N and every r-colouring $\Delta : N \to [r]$, N contains a monochromatic (m,p,c)-set M.*

Proof. Without loss of generality, assume $p > c$. Otherwise, simply prove the theorem for $p' > c \geq p$, then notice that for any m, $(x_0,\dots,x_m)_{p,c} \subseteq (x_0,\dots,x_m)_{p',c}$, and so any n,q,d satisfying the theorem for m,p',c also suffices for m,p,c.

Let $r \in \mathbb{Z}^+$, and for $k = 0,1,\dots,rm$, set $q_k = p^{2^{rm-k}}$, $d_k = c^{2^{rm-k}}$, and $Q_k = [-q_k, q_k]$. Note that $d_k = d_{k+1}^2$, $q_k = q_{k+1}^2$, and $q_k > d_{k+1}q_{k+1}$. Let $n_{rm} = rm+1$, and for $k = rm-1,\dots,2,1,0$, define recursively

$$n_k = HJ(2q_{k+1}+1, n_{k+1}-k, r) + k.$$

Let $n = n_0, q = q_0$ and $d = d_0$. Let $N = N_0$ be an (n,q,d)-set. Let $\Delta : N_0 \to [r]$ be any r-colouring.

The idea of the proof is to find (n_k,q_k,d_k)-sets $N_k \subseteq N$ whose first k rows

$$R_0(N_k), R_1(N_k), \dots, R_{k-1}(N_k)$$

are monochromatic (not necessarily the same colour for each row). See Figure 6.4 for an illustration.

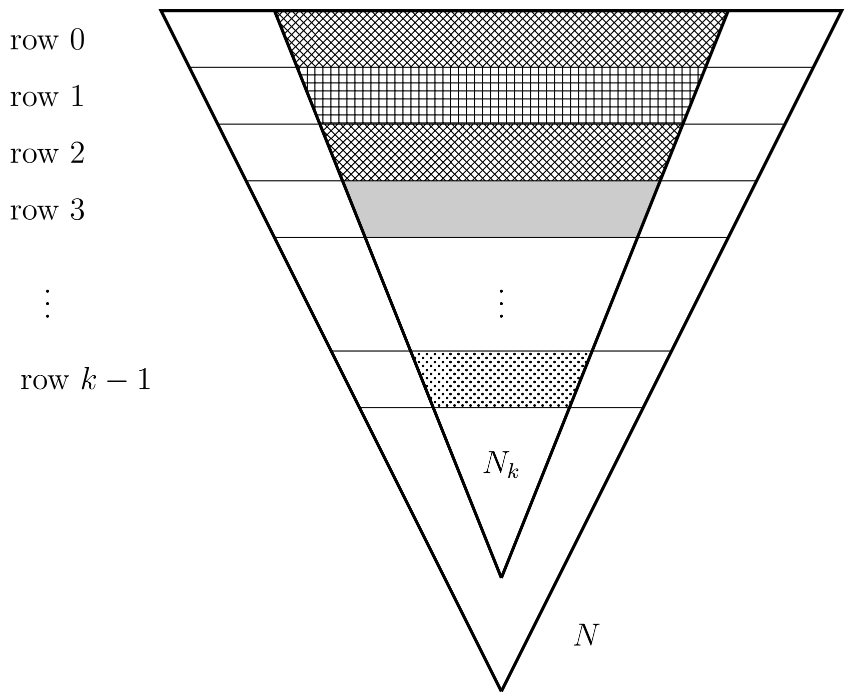

Figure 6.4: The first k rows of $N_k \subseteq N$ are monochromatic.

Suppose that for some $k \geq 0$, there is a (n_k, q_k, d_k)-set $(y_0, y_1, \ldots, y_{n_k})_{q_k,d_k} = N_k \subseteq N$ whose first k rows are monochromatic (for $k = 0$, this is trivial, since $N_0 = N$, and none of its rows need be monochromatic). Look at the row $R_k(N_k)$ of N_k. The colouring Δ restricted to this row induces another colouring $\Delta_k : [Q_k]\binom{n_k-k}{0} \to [r]$ defined by

$$\Delta_k((f(k+1), \ldots, f(n_k))) = \Delta(d_k y_k + d_{k+1} f(k+1) y_{k+1} + \cdots + d_{k+1} f(n_k) y_{n_k}).$$

By the choice of n_k, the Hales–Jewett theorem guarantees an $h \in [Q_k]\binom{n_k-k}{n_{k+1}-k}$ so that $sp(h)$ is monochromatic. Fix such an $h = (h(k+1), \ldots, h(n_k))$. Define the set of fixed coordinates by $C_k = \{i : h(i) \in Q_k\}$ and for each $j = [k+1, n_{k+1}]$ the set $C_j = \{i : h(i) = \lambda_j\}$. By Definition 6.2.1 of $h \in [Q_k]\binom{n_k-k}{n_{k+1}-k}$, if $j < j'$, then

$\min\{i \in C_j\} < \min\{i \in C_{j'}\}$. Since $sp(h)$ is monochromatic with respect to Δ_k, the set

$$\begin{aligned} S_k = &\left\{ d_k y_k + d_{k+1} \sum_{i \in C_k} h(i) y_i + \lambda_{k+1} \sum_{i \in C_{k+1}} d_{k+1} y_i + \cdots \right. \\ &\qquad \left. + \lambda_{n_{k+1}} \sum_{i \in C_{n_{k+1}}} d_{k+1} y_i : \lambda_{k+1}, \ldots, \lambda_{n_{k+1}} \in Q_k \right\} \\ = &\left\{ d_{k+1} \left(d_{k+1} y_k + \sum_{i \in C_k} h(i) y_i \right) + \lambda_{k+1} \sum_{i \in C_{k+1}} d_{k+1} y_i + \cdots \right. \\ &\qquad \left. + \lambda_{n_{k+1}} \sum_{i \in C_{n_{k+1}}} d_{k+1} y_i : \lambda_{k+1}, \ldots, \lambda_{n_{k+1}} \in Q_k \right\} \end{aligned}$$

is monochromatic with respect to Δ. Set $z_0 = d_{k+1} y_0$, $z_1 = d_{k+1} y_1$, $\ldots$, $z_{k-1} = d_{k+1} y_{k-1}$, set $z_k = d_{k+1} + \sum_{i \in C_k} h(i) y_i$, and for $j = k+1, \ldots, n_{k+1}$, set $z_j = \sum_{i \in C_j} d_{k+1} y_i$. Consider the $(n_{k+1}, q_{k+1}, d_{k+1})$-set $N_{k+1} = (z_0, \ldots, z_{n_{k+1}})_{q_{k+1}, d_{k+1}}$.

To verify that $N_{k+1} \subseteq N$ and its first $k+1$ rows are monochromatic, first look at any $j = 0, \ldots, k-1$. Then

$$\begin{aligned} R_j(N_{k+1}) &= \{ d_{k+1} z_j + \lambda_{j+1} z_{j+1} + \cdots + \lambda_{n_{k+1}} z_{n_{k+1}} : \lambda_{j+1}, \ldots, \lambda_{n_{k+1}} \in Q_{k+1} \} \\ &= \left\{ d_{k+1}^2 y_j + \lambda_{j+1} d_{k+1} y_{j+1} + \cdots + \lambda_{k-1} d_{k+1} y_{k-1} + \lambda_k \left(d_{k+1} + \sum_{i \in C_k} h(i) y_i \right) \right. \\ &\qquad \left. + \sum_{i=k+1}^{n_{k+1}} \left(\lambda_i \sum_{u \in C_i} d_{k+1} y_u \right) : \lambda_{j+1}, \ldots, \lambda_{n_{k+1}} \in Q_{k+1} \right\} \\ &\subseteq \{ d_k y_j + \mu_{j+1} y_{j+1} + \cdots + \mu_{n_k} y_{n_k} : \mu_{j+1}, \ldots, \mu_{n_k} \in Q_k \} = R_j(N_k). \end{aligned}$$

Since $R_j(N_k)$ is monochromatic, so is $R_j(N_{k+1})$.

For $j = k$, then $R_k(N_{k+1}) = S_k \subseteq R_k(N)$, which is monochromatic, and so the first $k+1$ of N_{k+1} are monochromatic.

Finally, for $j = k+1, \ldots, n_k$, let $l = \min\{i \in C_j\}$. Then

$$\begin{aligned} R_j(N_{k+1}) &= \{d_{k+1}z_j + \lambda_{j+1}z_{j+1} + \cdots + \lambda_{n_{k+1}}z_{n_{k+1}} : \lambda_{j+1}, \ldots, \lambda_{n_{k+1}} \in Q_{k+1}\} \\ &= \left\{ d_{k+1}^2 \sum_{i \in C_j} y_i + \sum_{i=j+1}^{n_{k+1}} \left(\lambda_i \sum_{u \in C_i} d_{k+1} y_u \right) : \lambda_j + 1, \ldots, \lambda_{n_{k+1}} \in Q_{k+1} \right\} \\ &\subseteq \{d_k y_l + \mu_{l+1} y_{l+1} + \cdots + \mu_{n_k} y_{n_k} : \mu_{l+1}, \ldots, \mu_{n_k} \in Q_k\} = R_l(N_k). \end{aligned}$$

Therefore $N_{k+1} \subseteq N_k \subseteq N$.

The (rm, p, c)-set $N_{rm} \subseteq N$ has its first rm rows monochromatic. The row $R_{rm}(N_{rm})$ is a single element, and so is trivially monochromatic. By the pigeonhole principle, there are $m+1$ rows $R_{i_0}(N_{rm}), R_{i_1}(N_{rm}), \ldots, R_{i_m}(N_{rm})$ of N_{rm} with the same colour. Then

$$M = (y_{i_0}, y_{i_1}, \ldots, y_{i_m})_{p,c} \subseteq N_{rm}$$

is a monochromatic (m, p, c)-set. □

To generalize Theorem 6.3.1, recall $GD(m, P, 1)$-sets from Definition 6.1.3, where $\mathbb{S}$ is any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, and $P \subseteq \mathbb{S}$. Set $d_k = 1$ for $k = 0, 1, \ldots, rm$; let $Q_{rm} = P$ and $n_{rm} = rm + 1$, and define recursively for $k = rm - 1, \ldots, 2, 1, 0$

$$Q_k = Q_{k+1} \cup \{ab : a, b \in Q_{k+1}\}$$

and

$$n_k = HJ(|Q_{k+1}|, n_{k+1} - k, r) + k.$$

Then with $n = n_0$ and $Q = Q_0$, the proof of Theorem 6.3.1 gives that any r-colouring of any $GD(n, Q, 1)$-set $N \subseteq \mathbb{S}$ yields a monochromatic $GD(m, P, 1)$-set.

6.4 Partition regular matrices and (m, p, c)-sets

In the introduction to this chapter, it is shown that the sought after sets of Chapter 3 are subsets of (m, p, c)-sets. The next theorem proves that every solution to a partition regular system is a subset of a (m, p, c)-set.

Theorem 6.4.1 (Deuber, 1973 [23]). *If a matrix A satisfies the columns property over $\mathbb{Q}$, then there exists $m, p, c \in \mathbb{Z}^+$ so that any (m, p, c)-set $M \subseteq \mathbb{Z}^+$ contains a solution to $A\boldsymbol{x} = \mathbf{0}$.*

Proof. Let A be a matrix with n columns satisfying the columns property with partition $[n] = I_1 \cup \cdots \cup I_m$, and for every $j = 1, 2, \ldots, m-1$ and $i \in I_1 \cup \cdots \cup I_j$, fix $\alpha_{i,j} \in \mathbb{Q}$ so that

$$\sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \boldsymbol{a}_i = \sum_{i \in I_{j+1}} \boldsymbol{a}_i$$

guaranteed to exist by the columns property. Let c be a common multiple of the denominators of all the $\alpha_{i,j}$'s and $p = \max_{i,j} |c\alpha_{i,j}|$. The values of m, p, c are the values that satisfy the theorem. For every $k = 1, 2, \ldots, m$, let A_k be the submatrix of A with columns indexed by $I_1 \cup \cdots \cup I_k$, and let $S(k)$ be the statement that every (k, p, c)-set contains a solution to $A_k\boldsymbol{v} = \mathbf{0}$. The proof that $S(k)$ is true for all $k = 1, 2, \ldots, m$ is by induction.

Base Case: For $k = 1$, S(1) holds since $\sum_{i=I_1} \boldsymbol{a}_i = \mathbf{0}$ by the columns property. So any constant multiple of $(1, 1, \ldots, 1)$ is a solution to the system $A_1\boldsymbol{x} = \mathbf{0}$.

Inductive Step: Fix $1 \leq k \leq m-1$ and assume that $S(k)$ is true. Let $M =$

$(x_0, x_1, \ldots, x_{k+1})_{p,c}$ be a $(k{+}1, p, c)$-set, and M' be the (k, p, c)-set $(x_0, x_1, \ldots, x_k)_{p,c} \subseteq M$. By the induction hypothesis, for every $i \in I_1 \cup \cdots I_k$, there is $v_i \in M'$ so that

$$\sum_{i \in I_1 \cup \cdots \cup I_k} v_i \boldsymbol{a}_i = \mathbf{0}. \tag{6.3}$$

For every $i \in I_1 \cup \cdots I_{k+1}$, define u_i by

$$u_i = \begin{cases} v_i - c\alpha_{i,k} x_{k+1} & i \in I_1 \cup \cdots \cup I_k \\ \\ c x_{k+1} & i \in I_{k+1}. \end{cases}$$

Of course $cx_{k+1} \in M$. By the choice of p, $-c\alpha_{i,k} \in [-p, p]$, and so $v_i - c\alpha_{i,k}x_{k+1} \in M$, and so all $i \in I_1 \cup \cdots I_{k+1}$, each u_i is an element of M. By the columns property,

$$\sum_{i \in I_1 \cup \cdots \cup I_k} \alpha_{i,k} \boldsymbol{a}_i - \sum_{i \in I_{k+1}} \boldsymbol{a}_i = \mathbf{0}. \tag{6.4}$$

Multiplying equation (6.4) by $-cx_{k+1}$ and adding Equation (6.3) produces

$$\mathbf{0} = \sum_{i \in I_1 \cup \cdots \cup I_k} (v_i - c\alpha_{i,k} x_{k+1}) \boldsymbol{a}_i + \sum_{i \in I_{k+1}} c x_{k+1} \boldsymbol{a}_i = \sum_{i \in I_1 \cup \cdots \cup I_{k+1}} \boldsymbol{a}_i u_i,$$

and so $A_{k+1}\boldsymbol{x} = \mathbf{0}$ has a solution contained in the $(k+1, p, c)$-set M.

By mathematical induction, m, p, c satisfy the statement of the theorem. □

The missing direction in the proof of Theorem 4.4.3 is presented.

Lemma 6.4.2. *If A be a matrix with integer entries that satisfies the columns property over $\mathbb{Q}$, then A is partition regular in $\mathbb{Z}^+$.*

Proof. If A satisfies the columns property over $\mathbb{Q}$, then by Theorem 6.4.1, there exists $m, p, c \in \mathbb{Z}^+$ so that every (m, p, c)-set $M \subseteq \mathbb{Z}^+$ has a solution to $A\boldsymbol{x} = \mathbf{0}$. By Theorem

6.3.1, for every $r \in \mathbb{Z}^+$ there exists $n, q, d \in \mathbb{Z}^+$ so that every r-colouring of an (n, q, d)-set N yields a monochromatic (m, p, c)-set M. Every r-colouring of $\mathbb{Z}^+$ induces an r-colouring of N, and so there is a monochromatic (m, p, c)-set M that contains a solution to $A\boldsymbol{x} = \mathbf{0}$. □

Letting $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, the proof of Theorem 6.4.1 can be reworked for matrices with entries in $\mathbb{S}$. Let A be a matrix with entries in $\mathbb{S}$ satisfying the columns property over $\mathbb{S}$ with partition $[n] = I_1 \cup \cdots \cup I_m$ and for every $j = 1, 2, \ldots, m-1$ and $i \in I_1 \cup \cdots \cup I_j$, fix $\alpha_{i,j} \in \mathbb{S}$ to be the values guaranteed by the columns property. Then setting $P = \{-\alpha_{i,j} : j \in [m-1], i \in I_1 \cup \cdots \cup I_j\} \cup \{0\}$, the proof of Theorem 6.4.1 shows that every $GD(m, P, 1)$-set $M \subseteq \mathbb{S}$ contains a solution to $A\boldsymbol{x} = \mathbf{0}$. The missing direction in the proofs of Theorem 4.4.5 is completed in the same manner as Lemma 6.4.2.

Lemma 6.4.3. *Let $\mathbb{S}$ be any one of $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$, and A a matrix with entries in $\mathbb{S}$. If A satisfies $CP(\mathbb{S})$, then A is $PR/\mathbb{S}$.*

Finally, the following theorem is needed to prove Rado's conjecture.

Theorem 6.4.4 (Deuber, 1973 [23])**.** *For every $m, p, c \in \mathbb{Z}^+$, there exists a partition regular matrix A for which every solution to $A\boldsymbol{x} = \mathbf{0}$ contains an (m, p, c)-set.*

Proof. For $p, c \in \mathbb{Z}^+$, let $s = 2p$ and $[-p, p] \setminus \{0\} = \{p_1, p_2, \ldots, p_s\}$, and for $m \geq 0$, let $S(m)$ be the statement of the theorem for m, p, c. The proof of the theorem is by induction on m.

Base Case: For $m = 0$, the matrix $A = \begin{bmatrix} c & -c & -1 \end{bmatrix}$ has the column property with the partition $I_1 = \{1, 2\}$ and $I_2 = \{3\}$, and $\alpha_{1,1} = 0, \alpha_{2,1} = 1/c$. By Theorem 4.4.3, A is partition regular. If $\boldsymbol{y} = (y_1, y_2, y_3)$ is a solution to $A\boldsymbol{x} = 0$, then $cy_1 - cy_2 - y_3 = 0$, so $y_3 = cx_0$ for some x_0. The set $\{cx_0\}$ is a $(0, p, c)$-set, concluding the base step.

Inductive Step: Let $k \geq 0$ and assume $S(k)$, that is, there exists a matrix A with the columns property so that every solution to $A\boldsymbol{x} = \boldsymbol{0}$ contains an (k, p, c)-set. Suppose A is an $r \times n$ matrix. Let $[n] = I_1 \cup \cdots \cup I_l$ with $\alpha_{i,j} \in \mathbb{Q}$ guaranteed by the columns property, that is $\sum_{i \in I_1} \boldsymbol{a_i} = \boldsymbol{0}$ and for $j \in [l-1]$,

$$\sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \boldsymbol{a_i} = \sum_{i \in I_{j+1}} \boldsymbol{a_i}.$$

Let B be the $(r + sn) \times (n + sn + 1)$ matrix in Figure 6.5 with columns $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n+sn+1})$.

For each $i \in [n]$, let

$$C_i = [n + s(i-1) + 1, n + si].$$

Then $\boldsymbol{a}'_i = \boldsymbol{b}_i + \sum_{i' \in C_i} \boldsymbol{b}_{i'}$ simply has the entries of $\boldsymbol{a}_i$ for the first r entries, and 0 for the rest of the entries (the c's negate each other). For each $1 \leq j \leq l$, let

$$I'_j = I_j \bigcup \left(\cup_{i \in I_j} C_i \right),$$

and define $I'_{l+1} = \{n + sn + 1\}$.

For $j \neq l$ and for each $i' \in C_i$, let $\beta_{i',j} = \alpha_{i,j}$. Then the columns property for A guarantees that

$$\sum_{i' \in I'_1} \boldsymbol{b}_{i'} = \sum_{i \in I_1} \left(\boldsymbol{b}_i + \sum_{i' \in C_i} \boldsymbol{b}_{i'} \right) = \sum_{i \in I_1} \boldsymbol{a}'_i = \boldsymbol{0},$$

$$
B = \begin{array}{c}
\begin{array}{cccccccccccccc}
 & & & & \multicolumn{4}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{C_1}} & & \multicolumn{4}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{C_n}} & \\
\end{array} \\
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
a_{21} & a_{22} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{r1} & a_{r2} & \cdots & a_{rn} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
 & & & & & & & & & & & & & \\
c & 0 & \cdots & 0 & -c & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & p_1 \\
c & 0 & \cdots & 0 & 0 & -c & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & p_2 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
c & 0 & \cdots & 0 & 0 & 0 & \cdots & -c & \cdots & 0 & 0 & \cdots & 0 & p_s \\
 & & & & & & & & & & & & & \\
 & & \vdots & & & & \vdots & & \ddots & & & \vdots & & \vdots \\
 & & & & & & & & & & & & & \\
0 & 0 & \cdots & c & 0 & 0 & \cdots & 0 & \cdots & -c & 0 & \cdots & 0 & p_1 \\
0 & 0 & \cdots & c & 0 & 0 & \cdots & 0 & \cdots & 0 & -c & \cdots & 0 & p_2 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & c & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & -c & p_s
\end{bmatrix}
\end{array}
$$

Figure 6.5: The matrix B

and for $1 \le j \le l-1$,

$$\begin{aligned}
\sum_{i' \in I'_1 \cup \cdots \cup I'_j} \beta_{i',j} \boldsymbol{b}_{i'} &= \sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \left(\boldsymbol{b}_i + \sum_{i' \in C_i} \boldsymbol{b}_{i'} \right) \\
&= \sum_{i \in I_1 \cup \cdots \cup I_j} \alpha_{i,j} \boldsymbol{a}'_i \\
&= \sum_{i \in I_{j+1}} \boldsymbol{a}'_i \\
&= \sum_{i \in I_{j+1}} \alpha_{i,j} \left(\boldsymbol{b}_i + \sum_{i' \in C_i} \boldsymbol{b}_{i'} \right) \\
&= \sum_{i' \in I'_{j+1}} \boldsymbol{b}_{i'}
\end{aligned}$$

and so the columns property is satisfied for $1 \le j \le l-1$.

All that is left is to verify that the columns property holds when $j = l$. By the columns property for A, $\sum_{i \in I_l} \boldsymbol{a}_i$ is a linear combination of the columns $\boldsymbol{a}_i$ indexed by $I_1 \cup \cdots \cup I_{l-1}$; that is, $\sum_{i \in I_l} \boldsymbol{a}_i$ and the columns indexed by $I_1 \cup \cdots \cup I_{l-1}$ are linearly dependent. Therefore, for $i \in [n]$, $\beta_{i,l}$ can be chosen so that the first r entries of $\sum_{i \in [n]} \beta_{i,l} \boldsymbol{b}_i$ are all zero, which is exactly the first r entries of $\boldsymbol{b}_{n+sn+1}$. For every $i \in [n]$, and every $i' \in C_i$, suppose the only non-zero entry of $\boldsymbol{b}'_i$ is in row t. Let $p_{i'}$ be the t-th entry in $\boldsymbol{b}_{n+sn+1}$, and define $\beta_{i',l} = (\beta_{i,l}c - p_{i'})/c$. Then the t-th entry of $\sum_{k \in I'_1 \cup \cdots \cup I'_l} \beta_{k,l} \boldsymbol{b}_k$ is

$$\beta_{i,l}c - \beta_{i',l}c = c(\beta_{i,l} + (\beta_{i,l}c - p_{i'})/c) = p_{i'},$$

which is indeed the t-th entry of $\boldsymbol{b}_{n+sn+1}$. The columns property is satisfied for $j = l$.

The next step is to show that every solution to $B\boldsymbol{x} = \boldsymbol{0}$ contains a $(k+1,p,c)$-set. Let $\boldsymbol{y}$ be a solution, and let $x_{k+1} = y_{n+sn+1}/c$. Then the first n entries of $\boldsymbol{y}$ is a solution to $A\boldsymbol{x} = \mathbf{0}$, and so by the induction hypothesis, there is a (k, p, c)-set Y contained in the first n entries of $\boldsymbol{y}$. Let $i \in [r]$ and $i' \in C_i$, and suppose the only non-zero entry of $\boldsymbol{b}_{i'}$ is in row t. Then along the t-th row of B, $B\boldsymbol{y} = \mathbf{0}$ gives that

$$cy_i - cy_{i'} = p_{i'}y_{n+sn+1} = p_{i'}cx_{k+1},$$

where $p_{i'}$ is the t-th entry in the last column of B. Divide by c and rearrange to get

$$y_{i'} = y_i + p_{i'}x_{k+1}.$$

Since p' varies in $[-p, p] \setminus \{0\}$, this means that for every $i \in [n]$, and every $\lambda \in [-p, p]$, $y_i + \lambda x_{k+1}$ is in $\boldsymbol{y}$ ($y_i + 0x_{k+1}$ is in $\boldsymbol{y}$ since y_i is in $\boldsymbol{y}$.) Therefore,

$$X = Y \cup \{y + \lambda x_{k+1} : y \in Y, \lambda \in [-p, p]\}$$

is contained in $\boldsymbol{y}$. Adding the last entry of $\boldsymbol{y}$ to X, the $(k+1, p, c)$-set $X \cup \{cx_{k+1}\}$ is contained in $\boldsymbol{y}$.

Thus by mathematical induction, for every m, p, c, there exists a matrix A with the columns property so that every solution to $A\boldsymbol{x} = \mathbf{0}$ contains an (m, p, c)-set. □

6.5 Proof of Rado's conjecture

Recall from Definition 6.1.1 that a set is large if and only if it contains a solution to every partition regular system.

Theorem 6.5.1 (Deuber, 1973 [23]). *A set $X \subseteq \mathbb{Z}^+$ is large if and only if for every $m, p, c \in \mathbb{Z}^+$, there is an (m, p, c)-set contained in X.*

Proof. Assume $X \subseteq \mathbb{Z}^+$ is large. By Theorem 6.4.4, find a matrix A satisfying the columns property so that every solution of $A\boldsymbol{x} = \mathbf{0}$ contains an (m, p, c)-set. Then since X is large and A is partition regular, there is a $\boldsymbol{v} \in X^n$ so that $A\boldsymbol{v} = \mathbf{0}$. The entries of $\boldsymbol{v}$ contain an (m, p, c)-set, and so X contains an (m, p, c)-set.

In the other direction, assume that for every $m, p, c \in \mathbb{Z}^+$ the set $X \subseteq \mathbb{Z}^+$ contains an (m, p, c)-set. Let A be any partition regular matrix. Since A satisfies the columns property, fix by Theorem 6.4.1 $m', p', c' \in \mathbb{Z}^+$ so that every (m', p', c')-set contains a solution to to $A\boldsymbol{x} = \mathbf{0}$. Then X contains a solution to the partition regular system. Since this is true for every partition regular matrix, X is large. □

The proof of Rado's conjecture now follows from Theorem 6.5.1 and Theorem 6.3.1.

Theorem 6.5.2 (Deuber, 1973 [23])**.** *Let $X \subseteq \mathbb{Z}^+$ be a large set finitely partitioned into $X = X_1 \cup \cdots \cup X_r$. Then some X_i is large.*

Proof. The proof is by induction on $r \geq 2$.

Base Case: Let $X \subseteq \mathbb{Z}^+$ be a large set and assume, contrary to the theorem, that there is $X = X_1 \cup X_2$ so that neither X_1 nor X_2 is large. By Theorem 6.5.1, there exists $m_1, p_1, c_1 \in \mathbb{Z}^+$ and $m_2, p_2, c_2 \in \mathbb{Z}^+$ so that for every (m_1, p_1, c_1)-set M_1 and every (m_2, p_2, c_2)-set M_2, $M_1 \not\subseteq X_1$ and $M_2 \not\subseteq X_2$. Let $m = \max\{m_1, m_2\}$, $c = c_1c_2$ and $p = \max\{p_1c_2, p_2c_1\}$. Every (m, p, c)-set contains both an (m_1, p_1, c_1)-set and an (m_2, p_2, c_2)-set. By Theorem 6.3.1, there exists n, q, d so that for every finite partition of any (n, q, d)-set, some partition set contains an (m, p, c)-set. Since X is large, X contains an (n, q, d)-set, call it N. Let $X_1' = N \cap X_1$, and $X_2' = N \cap X_2$. Then $N = X_1' \cup X_2'$ is a partition of N, and so one of X_1' or X_2' contains an (m, p, c)-set, and so contains both a (m_1, p_1, c_1)-set M_1 and an (m_2, p_2, c_2)-set M_2. This is a contradiction to $M_1 \not\subseteq X_1$ and $M_2 \not\subseteq X_2$.

INDUCTIVE STEP: Let $r \geq 2$ and assume that for any partition of a large set $X \subseteq \mathbb{Z}^+$ into r-sets, one of the sets is large. Consider a partition

$$X = (X_1 \cup \cdots \cup X_r) \cup X_{r+1} = Y \cup X_{r+1}.$$

By the base case, either Y or X_{r+1} is large. If Y is large, then by the induction hypothesis, one of $X_1, X_2, \ldots, X_r$ is large, concluding the inductive step. □

Chapter 7

Infinite partition regular systems

7.1 Introduction

In 1974, Hindman proved what is called the *finite sums theorem*, or *Hindman's theorem*, which extends the Folkman–Rado–Sanders theorem (Theorem 3.5.2). Recall the definition of finite sums sets from Definition 3.5.1. For any semigroup additive X and any subset $A = \{a_i \in X : i \in I\}$, the *finite sums set* of A is

$$FS(A) = \left\{ \sum_{i \in J} a_i : J \subseteq I, J \neq \emptyset, |J| < \infty \right\}.$$

Hindman [65] proved that in any finite colouring of $\mathbb{Z}^+$, there is an infinite set A so that $FS(A)$ is monochromatic. Hindman's proof "was one of enormous complexity" [76, p. vi], and is not presented here. The proof of the finite sums theorem given in Section 7.4 was introduced by Galvin and Glazer in 1975 [76, p. 123]. Galvin and Glazer never published their proof, although it appears in several works (for example, see [20], [69],

[76], and [122]).

For a set S, denote the *power set* of S by $\mathcal{P}(S) = \{T : T \subseteq S\}$. Many proofs in this chapter use filters and ultrafilters. A standard reference for ultrafilters is the book by Comfort and Negrepontis [21].

Definition 7.1.1. *For a set S, a* filter on S *is a collection $\mathcal{F} \subseteq \mathcal{P}(S)$ satisfying*

1. *for all $A \in \mathcal{F}$, if $A \subsetneq B \subseteq S$ then $B \in \mathcal{F}$ (upward closed);*

2. *for all $A, B \in \mathcal{F}$, $A \cap B \in \mathcal{F}$ ($\mathcal{F}$ is closed under intersections).*

Since every filter $\mathcal{F}$ is upward closed, $S \in \mathcal{F}$.

Definition 7.1.2. *A filter $\mathcal{F}$ on S is called a* proper *filter if and only if $\mathcal{F} \neq \mathcal{P}(S)$.*

Notice that by upward closure of filter, a filter $\mathcal{F}$ is a proper filter if and only if $\emptyset \notin \mathcal{F}$.

Definition 7.1.3. *A filter $\mathcal{F}$ on S is an* ultrafilter on S *if and only if for every $A \subseteq S$, either $A \in \mathcal{F}$ or $S \setminus A \in \mathcal{F}$.*

In Section 7.3, using ultrafilters, S is extended to a compact Hausdorff topological space βS, called the *Stone–Čech compactification of S.* In Section 7.4, an addition operator on βS is imposed, giving a topological semigroup. Hindman's theorem (Theorem 7.4.9) for finite sums then follows from a known result of Ellis [29].

Recall that a set $X \subseteq \mathbb{Z}^+$ is large if and only if X contains a solution to every finite partition regular system (Definition 6.1.1). By Theorem 6.5.2, every finite colouring of $\mathbb{Z}^+$ yields a monochromatic large set, which by Theorem 6.5.1, means that for every finite colouring of $\mathbb{Z}^+$, there is a colour class that contains for every $m, p, c \in \mathbb{Z}^+$ an (m, p, c)-set M. In 1987, Deuber and Hindman [25] proved that in fact some colour class contains what is called an (M, P, C)-system. Throughout the rest of this chapter, fix an ordering $\{(m_i, p_i, c_i)\}_{i\geq 1}$ of $(\mathbb{Z}^+)^3$.

Definition 7.1.4. *The set S is an (M, P, C)-system if and only if for every $i \in \mathbb{Z}^+$ there is an (m_i, p_i, c_i)-set M_i so that the sets in $\{M_i\}_{i\geq 1}$ are pairwise disjoint, and for*

$$\mathcal{S} = \left\{ X \subseteq \bigcup_{i\in\mathbb{Z}^+} M_i : \text{ for all } i \in \mathbb{Z}^+, |X \cap M_i| \leq 1 \right\},$$

the set $S = \bigcup_{X\in\mathcal{S}} FS(X)$.

An (M, P, C)-system S contains for every $m, p, c \in \mathbb{Z}^+$ an (m, p, c)-set, with all (m, p, c)-sets disjoint, along with the finite sums set of any set X that intersects every (m, p, c)-set contained in S in at most one element.

A proof of Deuber and Hindman's theorem (Theorem 7.5.5) is contained in Section 7.5.

In Section 7.6, possible generalizations of characterizations of partition regular matrices to infinite partition regular matrices are examined. In 1995, Deuber, Hindman, Leader, and Lefmann [26] found two infinite partition regular matrices A and B and a finite colouring $\mathbb{Z}^+$ so that no colour class contains a solution to both $A\boldsymbol{x} = \mathbf{0}$ and $B\boldsymbol{x} = \mathbf{0}$. In 2015, Barber, Hindman, Leader, and Strauss [4] considered an infinite columns property, and showed that the infinite columns property is neither sufficient, nor necessary, for infinite partition regularity.

The four main references used for Sections 7.3 and 7.4 are [54], [69], [76], and [80]. The axiom of choice is assumed throughout this chapter in the form of Zorn's lemma.

Lemma 7.1.5 (Zorn's lemma, see [28])**.** *Let $(P, \leq)$ be a partially ordered set. If every chain in P has an upper (lower) bound, then P contains a maximal (minimal) element.*

7.2 Topology review

Basic results in topology are used in many of the proofs of this chapter, and so a review of topology is included. Theorems and lemmas in this section are included without proof. The statements in this section are from Dugundji's *Topology* [28].

Definition 7.2.1. *Let X be a set. A* topology *in X is a family $\mathcal{T} \subseteq \mathcal{P}(X)$ that satisfies*

1. *$\emptyset, X \in \mathcal{T}$;*

2. *for every collection $\{U_i\}_{i \in I} \subseteq \mathcal{T}$, the union $\displaystyle\bigcup_{i \in I} U_i$ is an element of $\mathcal{T}$.*

3. *for every finite collection $\{U_i\}_{i=1}^{n} \subseteq \mathcal{T}$, the intersection $\displaystyle\bigcap_{i=1}^{n} U_i$ is an element of $\mathcal{T}$.*

Definition 7.2.2. *For a set X and a topology $\mathcal{T}$ in X, the pair $(X, \mathcal{T})$ is called a* topological space*.*

The elements of $\mathcal{T}$ are called the *open sets* of the topological space $(X, \mathcal{T})$.

Definition 7.2.3. *Let $(X, \mathcal{T})$ be a topological space. A collection $\mathcal{B} \subseteq \mathcal{T}$ is called a* basis *for $\mathcal{T}$ if and only if every open set $U \in \mathcal{T}$ is the union of elements of $\mathcal{B}$.*

Lemma 7.2.4 (see [28, Thm. 2.3, Ch. 3])**.** *Let* $(X, \mathcal{T})$ *be a topological space, and* $\mathcal{B}$ *a basis for* $\mathcal{T}$*. Then* $U \subseteq X$ *is open if and only if for every* $x \in U$ *there exists* $B \in \mathcal{B}$ *so that* $x \in B \subseteq U$*.*

Lemma 7.2.5 (see [28, Thm. 3.2, Ch. 3])**.** *Let* X *be a set and* $\mathcal{B} \subseteq \mathcal{P}(X)$ *a collection of subsets of* X *with the condition that for every* $B_1, B_2 \in \mathcal{B}$ *and every* $x \in B_1 \cap B_2$ *there exists* $B \in \mathcal{B}$ *so that* $x \in B \subseteq B_1 \cap B_2$*. Then* $\mathcal{B} \cup \{\emptyset, X\}$ *is a basis for some topology in* X*.*

Definition 7.2.6. *Let* $(X, \mathcal{T})$ *be a topological space. A set* $C \subset X$ *is* closed *if and only if the complement* $X \setminus C$ *of* C *is an open set.*

By De Morgan's law for sets, any intersection of closed sets is also closed, and any finite union of closed sets is closed.

Definition 7.2.7. *Let* $(X, \mathcal{T})$ *be a topological space. For any set* $A \subseteq X$*, the set*

$$\overline{A} = \{x \in X : \text{ for all } U \in \mathcal{T} \text{ with } x \in U,\, U \cap A \neq \emptyset\}$$

is called the closure of A.

Definition 7.2.8. *Let* $(X, \mathcal{T})$ *be a topological space. A set* $D \subseteq X$ *is* dense in X *if and only if* $\overline{D} = X$*.*

Lemma 7.2.9 (see [28, 4.13, Ch. 3])**.** *Let* $(X, \mathcal{T})$ *be a topological space, and let* $\mathcal{B}$ *be a basis for* $\mathcal{T}$*. Then* $D \subseteq X$ *is dense in* X *if and only if for every nonempty* $B \in \mathcal{B}$*, the intersection* $B \cap D$ *is nonempty.*

Definition 7.2.10. *Let $(X, \mathcal{T})$ be a topological space, and let $Y \subseteq X$. The* induced topology *on Y is $\mathcal{T}_Y = \{Y \cap U : U \in \mathcal{T}\}$. The pair $(Y, \mathcal{T}_Y)$ is called a* subspace *of $(X, \mathcal{T})$.*

The induced topology $\mathcal{T}_Y$ is a topology in Y, and $(Y, \mathcal{T}_Y)$ is a topological space.

Definition 7.2.11. *For two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, the function $f : X \to Y$ is* continuous *if and only if for every open set $V \in \mathcal{T}_Y$, the preimage*

$$f^{-1}(V) = \{x \in X : f(x) \in V\} = U$$

is open in X, that is, $U \in \mathcal{T}_X$.

Lemma 7.2.12 (see [28, Thm. 8.3, Ch. 3])**.** *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces, and $f : X \to Y$ a function from X to Y.*

1. *The function f is continuous if and only if for all closed set $C \subseteq Y$, the preimage $f^{-1}(C)$ is a closed set in X.*

2. *Let $\mathcal{B}$ be a basis for $\mathcal{T}_Y$. The function f is continuous if and only if for all $B \in \mathcal{B}$, the set $f^{-1}(B)$ is open in X.*

Definition 7.2.13. *A topological space $(X, \mathcal{T})$ is* Hausdorff *if and only if for every distinct $x, y \in X$, there are disjoint open sets $U, V \in \mathcal{T}$ so that $x \in U$ and $y \in V$.*

Lemma 7.2.14 (see [28, Thm. 1.3, Ch. 7])**.** *Let $(X, \mathcal{T})$ be a Hausdorff space. Then for any set $Y \subseteq X$, the topological space $(Y, \mathcal{T}_Y)$ is also a Hausdorff space.*

Lemma 7.2.15 (see [28, 1.4, Ch. 7])**.** *Let $(X, \mathcal{T})$ be a Hausdorff space. Then any finite set C is a closed set.*

For a topological space $(X, \mathcal{T})$, a collection $\mathcal{A} = \{U_i\}_{i \in I} \subseteq \mathcal{T}$ is called an *open covering* of X if and only if $\bigcup_{i \in I} U_i = X$. A *subcover of* $\mathcal{A}$ is a collection $\mathcal{B} \subseteq \mathcal{A}$ that is also an open covering of X.

Definition 7.2.16. *A Hausdorff space* $(X, \mathcal{T})$ *is* compact *if and only if every open covering of* X *has a finite subcover.*

Instead of using open coverings, another formulation of compact spaces using the finite intersection property is used throughout this chapter.

Definition 7.2.17. *For a set* X*, the collection* $\mathcal{F} \subseteq \mathcal{P}(X)$ *has the* finite intersection property *if and only if every finite subcollection* $\{F_i\}_{i=1}^{n} \subseteq \mathcal{F}$ *has nonempty intersection, that is,* $\bigcap_{i=1}^{n} F_i \neq \emptyset$.

Lemma 7.2.18 (see [28, Thm. 1.3, Ch. 11])**.** *The topological space* $(X, \mathcal{T})$ *is compact if and only if for every collection* $\mathcal{F}$ *of closed sets of* X *with the finite intersection property,* $\bigcap_{F \in \mathcal{F}} F \neq \emptyset$.

Lemma 7.2.19 (see [28, Thm. 1.4, Ch. 11])**.** *Let* $(X, \mathcal{T})$ *be a compact Hausdorff space.*

1. *Any set* $Y \subseteq X$ *is closed in* X *if and only if* Y *is compact.*

2. *Let* $(Z, \mathcal{T}_Z)$ *be a topological space and* $f : X \to Z$ *a continuous function. If* $Y \subseteq X$ *is compact in* X *then* $f(Y)$ *is compact in* Z.

Definition 7.2.20. *A continuous bijective function* $f : X \to Y$ *is a* homeomorphism *if and only if* $f^{-1} : Y \to X$ *is continuous.*

Definition 7.2.21. *For a topological space $(X, \mathcal{T}_X)$, the pair (X', f) is a* compactification *of X if and only if X' is compact and Hausdorff and there is a dense subset $D \subseteq X'$ so that $f : X \to D$ is a homeomorphism.*

It is common to refer to X' as the compactification of X.

7.3 Ultrafilters and the Stone–Čech compactification

The goal of this section is to introduce ultrafilters and the Stone–Čech compactification of a discrete set.

Definition 7.3.1. *A filter $\mathcal{F}$ is* maximal *if and only if $\mathcal{F}$ is a proper filter and whenever $\mathcal{F}' \supseteq \mathcal{F}$ is a filter, then $\mathcal{F}' = \mathcal{P}(S)$.*

Theorem 7.3.2. *Let $\mathcal{F}$ be a filter on S. Then $\mathcal{F}$ is maximal if and only if $\mathcal{F}$ is an ultrafilter.*

Proof. Suppose $\mathcal{F}$ is an ultrafilter and $\mathcal{F}'$ is a proper filter so that $\mathcal{F} \subsetneq \mathcal{F}'$. Then for any $A \in \mathcal{F}' \setminus \mathcal{F}$, since $A \notin \mathcal{F}$, by definition of ultrafilters, $S \setminus A \in \mathcal{F} \subsetneq \mathcal{F}'$. Therefore, both A and $S \setminus A$ are in $\mathcal{F}'$, and so $A \cap (S \setminus A) = \emptyset \in \mathcal{F}'$ and $\mathcal{F}' = \mathcal{P}(S)$.

Let $\mathcal{F}$ be a maximal filter and suppose $\mathcal{F}$ is not an ultrafilter. For any $A \in S$, if both A and $S \setminus A$ are contained in $\mathcal{F}$, then so is their intersection $\emptyset$. Since $\mathcal{F}$ is proper, $\emptyset \notin \mathcal{F}$ and so if $\mathcal{F}$ is not an ultrafilter, there is some A for which $A \notin \mathcal{F}$ and $S \setminus A \notin \mathcal{F}$. Then either $\mathcal{F} \cup \{A\}$ or $\mathcal{F} \cup \{S \setminus A\}$ has the finite intersection property:

otherwise there is $C_1, \dots, C_{n_1} \in \mathcal{F}$ and $D_1, \dots, D_{n_2} \in \mathcal{F}$ so that $A \cap C_1 \cap \cdots \cap C_{n_1} = \emptyset$ and $(S \setminus A) \cap D_1 \cap \cdots \cap D_{n_2} = \emptyset$. Since $\mathcal{F}$ is closed under intersections, both $C = C_1 \cap \cdots \cap C_{n_1} \neq \emptyset$ and $D = D_1 \cap \cdots \cap D_{n_2} \neq \emptyset$ are elements of $\mathcal{F}$. Since $\mathcal{F}$ is proper, $C \cap D \neq \emptyset$. But $A \cap (C \cap D) = (S \setminus A) \cap (C \cap D) = \emptyset$, which is impossible. Without loss of generality, $\mathcal{F} \cup \{A\}$ has the finite intersection property. Then consider

$$\mathcal{F}' = \big\{B \subseteq S : \text{ there exists } B_1, \dots, B_n \in \mathcal{F} \cup \{A\} \text{ so that } B_1 \cap \cdots \cap B_n \subseteq B\big\}.$$

The set $\mathcal{F}'$ is upward closed, and also closed under intersections. The intersection of any two elements of $\mathcal{F}'$ contains the intersection of finitely many elements of $\mathcal{F} \cup \{A\}$, and so is nonempty. Therefore, $\mathcal{F}'$ is a proper filter and $\mathcal{F} \subsetneq \mathcal{F}'$, contradicting $\mathcal{F}$ being maximal. □

Lemma 7.3.3. *Let $\mathcal{F}$ be an ultrafilter on a set S. For $A_1, A_2, \dots, A_r \subseteq S$, if*

$$A_1 \cup \cdots \cup A_r \in \mathcal{F}$$

then there exists $i \in [r]$ so that $A_i \in \mathcal{F}$.

Proof. Suppose for all $i \in [r]$ that $A_i \notin \mathcal{F}$. Since $\mathcal{F}$ is an ultrafilter, then for all $i \in [r]$, $S \setminus A_i \in \mathcal{F}$. Filters are closed under intersection, and so $(S \setminus A_1) \cap (S \setminus A_2) \cap \cdots \cap (S \setminus A_r) = \emptyset \in \mathcal{F}$, which contradicts $\mathcal{F}$ is a proper filter. Therefore, there is some $i \in [r]$ so that $A_i \in \mathcal{F}$. □

Theorem 7.3.4. *Every proper filter $\mathcal{F}$ can be extended to an ultrafilter.*

Proof. Let $\mathcal{H}$ be the set of all proper filters containing $\mathcal{F}$ along with the partial order by the subset relation. For any chain $\mathcal{C} = \{\mathcal{F}_i\}_{i\in I}$ in $\mathcal{H}$, let

$$\mathcal{F}_{\mathcal{C}} = \cup_{i\in I}\mathcal{F}_i.$$

For any $A \in \mathcal{F}_{\mathcal{C}}$, let $\mathcal{F}_i$ be a filter with $A \in \mathcal{F}_i$. For any $A \subsetneq B \subseteq S$ since $\mathcal{F}_i$ is upward closed, $B \in \mathcal{F}_i \subseteq \mathcal{F}_{\mathcal{C}}$, and so $\mathcal{F}_{\mathcal{C}}$ is upward closed. For any $A, B \in \mathcal{F}_{\mathcal{C}}$, let $\mathcal{F}_i \in \mathcal{C}$ be so that $A \in \mathcal{F}_i$ and $\mathcal{F}_j \in \mathcal{C}$ so that $B \in \mathcal{F}_j$. Since $\mathcal{C}$ is a chain, one filter is contained in the other, say $\mathcal{F}_i \subsetneq \mathcal{F}_j$. Then $A \cap B \in \mathcal{F}_j$, and $A \cap B \in \mathcal{F}_{\mathcal{C}}$, and so $\mathcal{F}_{\mathcal{C}}$ is a filter. Therefore $\mathcal{F}$ is an upper bound of $\mathcal{C}$.

By Zorn's lemma, $\mathcal{H}$ has a maximal element $\mathcal{F}$'. The filter $\mathcal{F}$' is a proper filter, and so by Theorem 7.3.2, $\mathcal{F}$' is an ultrafilter that contains $\mathcal{F}$. □

For a discrete space D, let βD be the set of all ultrafilters on D. It is standard to denote the elements of βD as lower case letters. For $A \subseteq D$, let

$$\widehat{A} = \{p \in \beta D : A \in p\}.$$

Notice that $p \in \beta D \setminus \widehat{A}$ if and only if $A \notin p$. Since p is an ultrafilter, $A \notin p$ if and only if $D \setminus A \in p$, and so $\beta D \setminus \widehat{A} = \widehat{D \setminus A}$. For $A, B \subseteq D$ and an ultrafilter $p \in \beta D$, since p is upward closed and closed under intersection, $A \cap B \in p$ if and only if $A, B \in p$. Therefore, $\widehat{A} \cap \widehat{B} = \widehat{A \cap B}$.

Theorem 7.3.5. *The collection $\mathcal{B} = \{\widehat{A} : A \subseteq D\} \cup \{\emptyset, \beta D\}$ is the basis for a compact Hausdorff topology $\mathcal{T}$ on βD.*

Proof. For any $\widehat{A_1}, \widehat{A_2} \in \mathcal{B}$ and $p \in \widehat{A_1} \cap \widehat{A_2}$, the set $\widehat{A_1 \cap A_2} \in \mathcal{B}$ contains p, and $\widehat{A_1 \cap A_2} = \widehat{A_1} \cap \widehat{A_2}$. By Lemma 7.2.5, the collection $\mathcal{B}$ is the basis for a topology on βD.

To show that $(\beta D, \mathcal{T})$ is Hausdorff, let $p, q \in \beta D$ be distinct, and let $A \in p \setminus q$. Since q is an ultrafilter, $D \setminus A \in q$. Therefore, $p \in \widehat{A}$ and $q \in \widehat{D \setminus A} = \beta D \setminus \widehat{A}$, and $\widehat{A} \cap (\beta D \setminus \widehat{A}) = \emptyset$.

To show that $(\beta D, \mathcal{T})$ is a compact topological space, it is enough to show that any collection of closed sets $\mathcal{C} = \{C_i\}_{i \in I}$ with the finite intersection property has nonempty intersection (by Lemma 7.2.18). For such a collection $\mathcal{C}$ of closed sets, let

$$\mathcal{D} = \left\{ A \subseteq D : C_{i_1} \cap \cdots \cap C_{i_n} \subseteq \widehat{A} \text{ for some } i_1, \ldots, i_n \in I \right\}.$$

Then $\bigcap_{i \in I} C_i \subseteq \bigcap_{A \in \mathcal{D}} \widehat{A}$, but the other inclusion is also true. Recall that closed sets are complements of open sets, and that open sets are unions of basis elements. Then for any $C \in \mathcal{C}$, by De Morgan's law for sets, there is a collection of basis elements $\left\{ \widehat{B_j} \right\}_{j \in J_i} \subseteq \mathcal{B}$ so that $C_i = \bigcap_{j \in J_i} \left(\beta D \setminus \widehat{B}_j \right)$. But each $\beta D \setminus \widehat{B_j}$ is precisely $\widehat{D \setminus B_j}$ and so letting $A_j = D \setminus B_j$,

$$C_i = \bigcap_{j \in J_i} \widehat{A_j}.$$

Each $\widehat{A_j}$ contains C_i, and so is an element of $\mathcal{D}$. Therefore,

$$\bigcap_{A \in \mathcal{D}} \widehat{A} \subseteq \bigcap_{i \in I} \left(\bigcap_{j \in J_i} \widehat{A_j} \right) = \bigcap_{i \in I} C_i.$$

Since the intersection of any finite subcollection of elements of $\mathcal{C}$ is nonempty, the collection $\mathcal{D}$ does not contain $\emptyset$. The collection $\mathcal{D}$ is upward closed, and also closed under intersections. Therefore, $\mathcal{D}$ is a proper filter, and can be extended to an ultrafilter p by Theorem 7.3.4. Each $A \in \mathcal{D}$ is an element of p, and so $p \in \widehat{A}$. Therefore,

$$p \in \bigcap_{A \in \mathcal{D}} \widehat{A} = \bigcap_{i \in I} C_i,$$

proving that $\mathcal{C}$ has nonempty intersection. The topological space $(\beta D, \mathcal{T})$ is compact. $\square$

For the rest of this chapter, set $\mathcal{B} = \{\widehat{A} : A \subseteq D\} \cup \{\emptyset, \beta D\}$ and $\mathcal{T}$ the topology generated by $\mathcal{B}$.

Definition 7.3.6. *For $C \subseteq S, C \neq \emptyset$, the* principal filter generated by C *is $\mathcal{F}_C = \{A \subseteq S : C \subseteq A\}$.*

For $d \in D$, let $p_d = \{A \subseteq D : d \in A\}$ denote the principal filter generated by $\{d\}$. Any filter $\mathcal{F}$ that contains p_d contains a set A with $d \notin A$. Then by closure under intersection, $\mathcal{F}$ also contains $\{d\} \cap A = \emptyset$, and so $\mathcal{F} = \mathcal{P}(D)$. The proper filter p_d is maximal, and so by 7.3.2, p_d is an ultrafilter. Let $\mathcal{D} = \{p_d : d \in D\}$ and define the function $e : D \to \mathcal{D}$ by $e(d) = p_d$.

Lemma 7.3.7. *For a discrete set D, The pair $(\beta D, e)$ is a compactification of the topological space $(D, \mathcal{P}(D))$.*

Proof. By Theorem 7.3.5, $(\beta D, \mathcal{T})$ is compact. For any distinct $c, d \in D$, the ultrafilters p_c and p_d are distinct since $\{d\} \in p_d$ and $D \setminus \{d\} \in p_c$ because $c \in D \setminus \{d\}$. Therefore e is a bijection. Since every set $A \in \mathcal{P}(D)$ is open in D the preimage of any set $U \subseteq \mathcal{D}$ is open, and so e is continuous. Any set $A \subseteq D$ is mapped to $e(A) = \{p_a \in \mathcal{D} : a \in$

$A\} = \widehat{A} \cap \mathcal{D}$, which is open in $(\mathcal{D}, \mathcal{T}_{\mathcal{D}})$. Therefore $e^{-1} : \mathcal{D} \to D$ is continuous, and e is a homeomorphism.

To show that $\mathcal{D}$ is dense in βD, let $\widehat{A} \in \mathcal{B}$ be any basis element of $\mathcal{T}$. For any $a \in A$, the ultrafilter p_a contains A, and so $p_a \in \widehat{A}$. Therefore $\widehat{A} \cap \mathcal{D} \neq \emptyset$, and so by Lemma 7.2.9, $\mathcal{D}$ is dense in βD. □

The set βD is called the Stone–Čech compactification of D, which was developed by Stone [116] and Čech [17] in 1937.

7.4 Hindman's theorem

The goal of this section is to extend the operation $'+'$ from $\mathbb{Z}^+$ to an operation on $\beta\mathbb{Z}^+$, then to use a result of Ellis concerning idempotents in right topological semigroups (Definition 7.4.3 below) to extract a proof of the finite sums theorem.

Definition 7.4.1. *Let $(X, \mathcal{T})$ be a topological space, and let $* : X \times X \to X$ be a binary operation on X. The operation $*$ is called* right continuous *if and only if for every $x \in X$, the function $\rho_x : X \to X$ defined by $\rho_x(y) = y * x$ is continuous.*

Recall that a *semigroup* is a set along with a closed associative binary operation.

Definition 7.4.2. *Let $(X, *)$ be a semigroup. An element $x \in X$ is an* idempotent *if and only if $x * x = x$.*

Definition 7.4.3. *For a set X, $\mathcal{T} \subseteq \mathcal{P}(X)$ and a binary operation $*$ on X, the triple $(X, *, \mathcal{T})$ is called a* right topological semigroup *if and only if $(X, *)$ is a semigroup, $(X, \mathcal{T})$ is a topological space, and the operation $*$ is right continuous.*

Theorem 7.4.4 (Ellis, 1969 [29])**.** *Let $(X, *, \mathcal{T})$ be a compact Hausdorff right topological semigroup. Then X contains an idempotent.*

For a subset $A \subseteq X$, let $A * A = \{a_1 * a_2 : a_1, a_2 \in A\}$, and for $x \in X$, let $A * x = \{a * x : a \in A\}$.

Proof. Define the set

$$\mathcal{A} = \{A \subseteq X : A * A \subseteq A, A \neq \emptyset, \text{ and } A \text{ is compact}\},$$

which is partially ordered by the subset relation. The set $\mathcal{A}$ is nonempty, since X is compact and $X * X \subseteq X$. Let $\mathcal{C} = \{C_i\}_{i \in I} \subseteq \mathcal{A}$ be any chain. Each C_i is closed (by (1) of Lemma 7.2.19) and so $C = \bigcap_{i \in I} C_i$ is also closed and compact (recall that any intersection of closed sets is also closed). For any $c \in C$, since for all $i \in I$, $c * c \in C_i$, then $c * c \in C$. Therefore $C \in \mathcal{A}$, and C is a minimal element of $\mathcal{C}$. By Zorn's lemma, the set $\mathcal{A}$ has a minimal element, call it A.

For any $x \in A$, since $*$ is right continuous, the function ρ_x is continous. By (2) of Lemma 7.2.19, $A * x = \rho_x(A)$ is also compact. Also, since $A \in \mathcal{A}$, then $A * x \subseteq A * A \subseteq A$, and so by associativity of $*$,

$$(A * x) * (A * x) \subseteq (A * A) * (A * x) \subseteq (A * A) * x \subseteq A * x.$$

Therefore, $A * x \in \mathcal{A}$, but since A is minimal, $A * x = A$.

Define the set $B = \{y \in A : y * x = x\}$. It was just shown that $A * x = A$, and so $x \in B \neq \emptyset$. The singleton $\{x\}$ is a closed set (by Lemma 7.2.15), and since ρ_x is continuous, the set $\rho_x^{-1}(\{x\})$ is a closed set. Therefore by (1) of Lemma 7.2.12, $B = \rho_x^{-1}(\{x\}) \cap A$ is closed and compact. For any $y, z \in B$, by associativity of $*$,

$$(y * z) * x = y * (z * x) = y * x = x,$$

and so $y * z \in B$. Then $B * B \subseteq B$, and so $B \in \mathcal{A}$. But $B \subseteq A$, and since A is minimal, $A = B$. Therefore, $x * x = x$. □

The Stone–Čech compactification of $\mathbb{Z}^+$ provides a topological space $\mathcal{T}$ that is compact and Hausdorff. A right continuous associative binary operation $\oplus$ is added to $(\beta\mathbb{Z}^+, \mathcal{T})$ to make a compact Hausdorff right topological semigroup.

Definition 7.4.5. *For any set $A \subseteq \mathbb{Z}^+$ and $x \in \mathbb{Z}^+$, define*

$$A - x = \{y \in \mathbb{Z}^+ : y + x \in A\}.$$

For ultrafilters $p, q \in \beta\mathbb{Z}^+$, define

$$p \oplus q = \{A \subseteq \mathbb{Z}^+ : \{x \in \mathbb{Z}^+ : A - x \in q\} \in p\}.$$

Lemma 7.4.6. *For any $p, q \in \beta\mathbb{Z}^+$, $p \oplus q \in \beta\mathbb{Z}^+$, and so $\oplus : \beta\mathbb{Z}^+ \times \beta\mathbb{Z}^+ \to \mathbb{Z}^+$ is a well defined binary operation on $\beta\mathbb{Z}^+$.*

Proof. For any $x \in \mathbb{Z}^+$, $\mathbb{Z}^+ - x = \mathbb{Z}^+$. Both p and q contain $\mathbb{Z}^+$, and so $\{x : \mathbb{Z}^+ - x \in q\} = \mathbb{Z}^+ \in p$. Therefore, $\mathbb{Z}^+ \in p \oplus q$.

Fix any $A \in p \oplus q$ and any $B \subseteq \mathbb{Z}^+$ with $A \subseteq B$. For all $x \in \mathbb{Z}^+$, $A - x \subseteq B - x$, and so if $A - x \in q$, then $B - x \in q$ since q is upward closed. Therefore, $\{x \in \mathbb{Z}^+ : A - x \in q\} \subseteq \{x \in \mathbb{Z}^+ : B - x \in q\}$, and since p is upward closed, $B \in p \oplus q$.

Look at any $A, B \in p \oplus q$ and let $C = A \cap B$. For all $x \in \mathbb{Z}^+$, $C - x = (A - x) \cap (B - x)$. Therefore, if $A - x \in q$ and $B - x \in q$, since q is closed under intersections, $C - x \in q$. Also,

$$\begin{aligned}\{x \in \mathbb{Z}^+ : C - x \in q\} &= \{x \in \mathbb{Z}^+ : (A - x) \cap (B - x) \in q\} \\ &= \{x \in \mathbb{Z}^+ : A - x \in q\} \cap \{x \in \mathbb{Z}^+ : B - x \in q\},\end{aligned}$$

and since p is closed under intersections, $C \in p \oplus q$.

The collection $p \oplus q$ is a filter. To show that $p \oplus q$ is proper, notice for any $x \in \mathbb{Z}^+$ that $\emptyset - x = \emptyset$ and so since q is proper, $\emptyset - x \notin q$. Then $\{x \in \mathbb{Z}^+ : \emptyset - x \in q\} = \emptyset$ and since p is proper, $\emptyset \notin p \oplus q$.

The final step in the proof is to show that $p \oplus q$ is an ultrafilter. Fix any $A \notin p \oplus q$. Notice that for any $x \in \mathbb{Z}^+$, $\mathbb{Z}^+ \setminus (A - x) = (\mathbb{Z}^+ \setminus A) - x$. Since p is an ultrafilter and $\{x \in \mathbb{Z}^+ : A - x \in q\} \notin p$, then $\{x \in \mathbb{Z}^+ : A - x \notin q\} \in p$. But since q is an ultrafilter, $A - x \notin q$ if and only if $\mathbb{Z}^+ \setminus (A - x) \in q$. Therefore, $\{x \in \mathbb{Z}^+ : (\mathbb{Z}^+ \setminus A) - x \in q\} \in p$, and so $\mathbb{Z}^+ \setminus A \in p \oplus q$. □

Lemma 7.4.7. *The operation $\oplus$ is associative, and so $(\beta\mathbb{Z}^+, \oplus)$ is a semigroup.*

Proof. For any $A \subseteq \mathbb{Z}^+$, define $X = \{x \in \mathbb{Z}^+ : A - x \in r\}$ and $Y = \{y \in \mathbb{Z}^+ : X - y \in q\}$. Then $A \in (p \oplus q) \oplus r$ if and only if $X \in p \oplus q$ if and only if $Y \in p$. Suppose $A \in (p \oplus q) \oplus r$.

For any $y \in Y$, by definition of Y, $X - y \in q$. And so

$$X - y = \{z \in \mathbb{Z}^+ : z + y \in X\} = \{z \in \mathbb{Z}^+ : (A - y) - z \in r\} \in q,$$

therefore $A - y \in q \oplus r$. This is true for every $y \in Y$, and so $Y = \{y \in \mathbb{Z}^+ : A - y \in q \oplus r\}$. Then $Y \in p$ if and only if $A \in p \oplus (q \oplus r)$. □

Lemma 7.4.8. *The operation $\oplus$ is right continuous.*

Proof. For any $q \in \beta\mathbb{Z}^+$, look at ρ_q (defined by $\rho_q(p) = p \oplus q$). By (2) of Lemma 7.2.12, it is enough to show that the preimage under ρ_q of any element of $\mathcal{B}$ is open. For any $\widehat{A} \in \mathcal{B}$, then

$$\oplus_q^{-1}(\widehat{A}) = \{p \in \beta\mathbb{Z}^+ : p \oplus q \in \widehat{A}\}.$$

By definition of $\widehat{A}$, $p \oplus q \in \widehat{A}$ if and only if $A \in p \oplus q$, and so for every $p \in \widehat{A}$, $X = \{x \in \mathbb{Z}^+ : A - x \in q\} \in p$. Therefore,

$$\oplus_q^{-1}(\widehat{A}) = \{p \in \beta\mathbb{Z}^+ : \{x \in \mathbb{Z}^+ : A - x \in q\} \in p\} = \{p \in \beta\mathbb{Z}^+ : X \in p\},$$

which is precisely $\widehat{X}$. The set $\widehat{X}$ is a basis element, and so is open. □

The topological space $(\beta\mathbb{Z}^+, \mathcal{T})$ is compact and Hausdorff, $(\beta\mathbb{Z}^+, \oplus)$ is a semigroup, and $\oplus$ is right continuous. Therefore $(\beta\mathbb{Z}^+, \oplus, \mathcal{T})$ is a compact Hausdorff right topological semigroup and so Ellis' theorem can be applied to find an idempotent.

Theorem 7.4.9 (Hindman, 1974 [65])**.** *For every finite colouring of $\mathbb{Z}^+$, there is an infinite set $A = \{a_1, a_2, \ldots\} \subseteq \mathbb{Z}^+$ so that $FS(A)$ is monochromatic.*

Proof. Since $(\beta\mathbb{Z}^+, \oplus, \mathcal{T})$ is a compact Hausdorff right topological semigroup, by Theorem 7.4.4, $\beta\mathbb{Z}^+$ has an idempotent p. For any $A \in p$ let $X = \{x \in \mathbb{Z}^+ : A - x \in p\} \in p$. Since p is closed under intersection, $A \cap X \in p$, and since p is proper, $A \cap X \neq \emptyset$.

The first part of the proof is to show that any set $A \in p$ contains a finite sums set. Let $A_1 = A$ and $X_1 = \{x \in \mathbb{Z}^+ : A_1 - x \in p\}$, and choose any $a_1 \in A_1 \cap X_1$. For $k > 1$, define recursively $A_k = A_{k-1} \cap (A_{k-1} - a_{k-1})$ and $X_k = \{x \in \mathbb{Z}^+ : A_k - x \in p\}$, and choose $a_k \in A_k \cap X_k$. For $A_{k-1} \in p$, then $X_{k-1} \in p$, and since $a_{k-1} \in X_{k-1}$, then $A_{k-1} - a_{k-1} \in p$. Since p is closed under intersections $A_k \in p$. Also, $A_k \subseteq A_{k-1} \subseteq A_{k-2} \subseteq \cdots \subseteq A_1 = A$.

For $k \geq 1$, let $S(k)$ be the following statement: for any $I \subseteq \mathbb{Z}^+$ with $|I| = k$ and $m = \min\{i \in I\}$, then $\sum_{i \in I} a_i \in A_m$. The statement is proved by induction on k.

Base Case: For $k = 1$, then $I = \{m\}$ and $a_m \in A_m$.

Inductive Step: Let $l \geq 1$ and suppose $S(l)$ holds. For any $|I| = l + 1$ with $\min\{i \in I\} = m$, let $J = I \setminus \{m\}$ and $\min\{j \in J\} = n$. By the induction hypothesis, $\sum_{j \in J} a_j \in A_n$. Since $n > m$, then

$$A_n \subseteq A_{m+1} = A_m \cap (A_m - a_m) \subseteq A_m - a_m.$$

Then for any $a \in A_m$, $a + a_m \in A_m$. In particular,

$$\sum_{i \in I} a_i = a_m + \sum_{j \in J} a_j \in A_m,$$

concluding the inductive step. By mathematical induction, for every $I \subseteq \mathbb{Z}^+$ with $|I| < \infty$,

$$\sum_{i \in I} a_i \in A.$$

To finish the proof, let $\Delta : \mathbb{Z}^+ \to [r]$ be an r-colouring of $\mathbb{Z}^+$. Let $A_1, A_2, \ldots, A_r$ be the colour classes of the colouring. By Lemma 7.3.3, since $A_1 \cup \cdots A_r = \mathbb{Z}^+ \in p$, then one of the colour classes A_i is contained in p. Since every set of p contains a finite sum set, the colour class $A_i \in p$ also contains a finite sums set. □

In the proof of Theorem 7.4.9, it is shown that there is an ultrafilter p so that every $A \in p$ contains a finite sums set. In fact, Hindman characterized all Ramsey statements in a similar way.

Theorem 7.4.10 (Hindman, 1979 [67]). *Let X be a set and $\mathcal{H} \subseteq \mathcal{P}(X)$. Then the following are equivalent.*

1. *For every $r \in \mathbb{Z}^+$ and every r-colouring of X, there exists $H \in \mathcal{H}$ so that H is monochromatic.*

2. *There is an ultrafilter $\mathcal{F}$ on X so that if $A \in \mathcal{F}$, then there exists an $H \in \mathcal{H}$ so that $H \subseteq A$.*

Proof. In the first direction, suppose (1) is true. If $\emptyset \in \mathcal{H}$, then any ultrafilter satisfies (2) with $H = \emptyset$. Otherwise, suppose $\emptyset \notin \mathcal{H}$. Let

$$\mathcal{A} = \{A \subseteq X : \text{ for all } H \in \mathcal{H}, A \cap H \neq \emptyset\}.$$

The collection $\mathcal{A}$ has the finite intersection property. Assume otherwise, that for some $A_1, A_2, \ldots, A_k \in \mathcal{A}$, $\cap_{i=1}^{k} A_i = \emptyset$. Then by De Morgan's law, $X = \cup_{i=1}^{k}(X \setminus A_i)$. Define a k-colouring $\Delta : X \to [k]$ as follows: $\Delta(x) = i$ if and only if $x \in (X \setminus A_i)$ and for all

$j < i$, $x \notin A_j$. Then by (1), there is a monochromatic $H \in \mathcal{H}$, that is, $H \subseteq (X \setminus A_i)$ for some i, and so $H \cap A_i = \emptyset$, which contradicts the definition of $\mathcal{A}$. Therefore, $\mathcal{A}$ has the finite intersection property. Let

$$\mathcal{F}' = \{A_1 \cap \cdots \cap A_k : A_1, \ldots, A_k \in \mathcal{A}\}.$$

Then $\mathcal{F}'$ is closed under intersections, and since $\mathcal{A}$ is upward closed so is $\mathcal{F}'$. Also, $\emptyset \notin \mathcal{F}'$ since $\mathcal{A}$ has the finite intersections property. The collection $\mathcal{F}'$ is then a proper filter, which can be extended to an ultrafilter $\mathcal{F}$ by Theorem 7.3.4. For any $A \in \mathcal{F}$, if it were the case that for all $H \in \mathcal{H}$, $H \cap (X \setminus A) \neq \emptyset$, then $(X \setminus A) \in \mathcal{A} \subseteq \mathcal{F}$, which cannot happen since $\mathcal{F}$ is an ultrafilter and so does not contain $(X \setminus A)$. Therefore, there is some $H \in \mathcal{H}$ so that $H \subseteq A$.

In the other direction, suppose (2) is true for some ultrafilter $\mathcal{F}$. By Lemma 7.3.3, for every finite partition $X = X_1 \cup X_2 \cup \cdots \cup X_r$, some partition class X_i is contained in $\mathcal{F}$. By (2), there is some $H \subseteq X_i$. □

7.5 (M,P,C)-systems

Throughout this section, set an ordering $(m_i, p_i, c_i)_{i \in \mathbb{Z}^+}$ on $(\mathbb{Z}^+)^3$. Recall the definition of (M, P, C)-systems:

Definition 7.5.1. *The set S is an (M, P, C)-system if and only if for every $i \in \mathbb{Z}^+$ there* *is an (m_i, p_i, c_i)-set M_i so that the sets in $\{M_i\}_{i\geq 1}$ are pairwise disjoint, and for*

$$\mathcal{S} = \left\{ X \subseteq \bigcup_{i\in\mathbb{Z}^+} M_i \;\middle|\; \text{for all } i \in \mathbb{Z}^+, |X \cap M_i| \leq 1 \right\},$$

the set $S = \bigcup_{X\in\mathcal{S}} FS(X)$.

A similar technique to the proof of the finite sums theorem is used to show that any finite colouring of $\mathbb{Z}^+$ contains a monochromatic (M, P, C)-system. However the idempotent q has the property that every set $A \in q$ contains an (M, P, C)-system. To find such an idempotent, compact subspaces of $(\beta\mathbb{Z}^+, \mathcal{T})$ are studied. The proofs here are from [25].

Definition 7.5.2. *For all $i \in \mathbb{Z}^+$, let*

$$T(i) = \left\{ q \in \beta\mathbb{Z}^+ : \text{ for all } A \in q, \text{ there exists an } (m_i, p_i, c_i)\text{-set } M_i \in A \right\}.$$

Also, let

$$\mathcal{U} = \bigcap_{i\in\mathbb{Z}^+} T(i).$$

Lemma 7.5.3 (Deuber and Hindman, 1987 [25])**.** *For every $i \in \mathbb{Z}^+$, the subset $T(i) \subseteq \beta\mathbb{Z}^+$ is nonempty and closed.*

Proof. For any $i \in \mathbb{Z}^+$, then by Theorem 6.3.1, any finite colouring of $\mathbb{Z}^+$ yields a monochromatic (m_i, p_i, c_i)-set. By Theorem 7.4.10, there is an ultrafilter q so that every $A \in q$ contains an (m_i, p_i, c_i)-set, and so $T(i) \neq \emptyset$. The set $T(i)$ is closed if and only if its complement is open. Choose any $q \notin T(i)$ and $B \in q$ so that for all (m_i, p_i, c_i)- sets

M_i, $M_i \notin B$. Every ultrafilter in $\widehat{B}$ is not in $T(i)$, and so $q \in \widehat{B} \subseteq \beta\mathbb{Z}^+ \setminus T(i)$. Therefore, $T(i)$ is closed. □

Lemma 7.5.4 (Deuber and Hindman, 1987 [25]). *The triple $(\mathcal{U}, \oplus, \mathcal{T}_{\mathcal{U}})$ is a nonempty compact Hausdorff right topological semigroup.*

Proof. The first step is to show that $\mathcal{U}$ is non-empty. For any finite number of sets

$$T(i_1), T(i_2), \ldots, T(i_n),$$

let $m = \max\{m_{i_1}, m_{i_2}, \ldots, m_{i_n}\}$, $c = c_{i_1} c_{i_2} \cdots c_{i_n}$ and $p = \max\{cp_{i_1}, cp_{i_2}, \ldots, cp_{i_n}\}$. Then every (m, p, c)-set contains for every $i_k = i_1, \ldots, i_n$ an $(m_{i_k}, p_{i_k}, c_{i_k})$-set. Let $i \in \mathbb{Z}^+$ be such that $(m, p, c) = (m_i, p_i, c_i)$, then for any $q \in T(i)$ and every $A \in q$, A contains an (m, p, c)-set, and so

$$T(i) \subseteq \bigcap_{k=1}^{n} T(i_k).$$

Therefore $\{T(i)\}_{i \in \mathbb{Z}^+}$ has the finite intersection property, and since $(\beta\mathbb{Z}^+, \mathcal{T})$ is compact, $\mathcal{U} = \bigcap_{i \in \mathbb{Z}^+} T(i)$ is nonempty.

Since $\mathcal{U}$ is the intersection of closed sets, then $\mathcal{U}$ is also closed, and compact since $\beta\mathbb{Z}^+$ is compact. The topological space $(\mathcal{U}, \mathcal{T}_{\mathcal{U}})$ is then compact by (1) of Lemma 7.2.19 and Hausdorff by Lemma 7.2.14.

The next step is to show that $(\mathcal{U}, \oplus)$ is a subsemigroup of $(\beta\mathbb{Z}^+, \oplus)$. For any $q, s \in \mathcal{U}$, look at any $A \in q \oplus s$, and any $i \in \mathbb{Z}^+$. Then $B = \{x \in \mathbb{Z}^+ : A - x \in s\} \in q$, and so B contains an (m_i, p_i, c_i)-set $M_i = (x_0, x_1, \ldots, x_{m_i})_{p_i, c_i}$. Let

$$C = \bigcap_{x \in M_i} (A - x).$$

Each $A - x \in s$, and M_i is a finite set, and so since s is closed under intersections, $C \in s$. Therefore C contains an (m_i, p_i, c_i)-set $N_i = (y_0, y_1, \ldots, y_{m_i})_{p_i, c_i}$. For $j = 0, 1, \ldots, m_i$ define $z_j = x_j + y_j$. Then for any

$$y = c_i y_j + \lambda_{j+1} y_{j+1} + \cdots + \lambda_{m_i} y_{m_i} \in N_i$$

and

$$x = c_i x_j + \lambda_{j+1} x_{j+1} + \cdots + \lambda_{m_i} x_{m_i} \in M_i,$$

since $y \in C$, then

$$x + y = c_i z_j + \lambda_{j+1} z_{j+1} + \cdots + \lambda_{m_i} z_{m_i} \in A.$$

Therefore A contains the (m_i, p_i, c_i)-set $L_i = (z_0, z_1, \ldots, z_{m_i})_{p_i, c_i}$. This is true for any $i \in \mathbb{Z}^+$ and any $A \in q \oplus s$, and so $q \oplus s \in \mathcal{U}$. Therefore, $(\mathcal{U}, \oplus)$ is a subsemigroup of $(\beta\mathbb{Z}^+, \oplus)$.

Any continuous function restricted to a subspace topology is also continuous. Therefore, $\rho_s : \mathcal{U} \times \mathcal{U} \to \mathcal{U}$ defined by $\rho_s(q) = q \oplus s$ is continuous for any q, and so $(\mathcal{U}, \oplus, \mathcal{T}_{\mathcal{U}})$ is a compact Hausdorff right topological semigroup. □

Theorem 7.5.5 (Deuber and Hindman, 1987 [25])**.** *For every $r \in \mathbb{Z}^+$ and every r-colouring of $\mathbb{Z}^+$, there is monochromatic (M, P, C)-system.*

Proof. By Theorem 7.4.4, the set $\mathcal{U}$ has an idempotent q. For any set $A \in p$, let $A_1 = A$ and $X_1 = \{x \in \mathbb{Z}^+ : A_1 - x \in q\} \in q$. Since $q \in T(1)$, there is an (m_1, p_1, c_1)-set $M_1 \subseteq X_1$. Recursively for $k > 1$, suppose $A_{k-1} \in q$ has been defined and $M_{k-1} \subseteq \{x \in$

$\mathbb{Z}^+ : A_{k-1} - x \in q\} \in q$ is an $(m_{k-1}, p_{k-1}, c_{k-1})$-set. Let

$$A_k = A_{k-1} \bigcap_{x \in M_{k-1}} (A_{k-1} - x).$$

Since M_{k-1} is finite and $A_{k-1} - x \in q$ for every $x \in M_{k-1}$, and since q is closed under intersections, $A_k \in q$. Let $a_k = \max\left\{x \in \cup_{i=1}^{k-1} M_i\right\}$, and define $X_k = \{x \in \mathbb{Z}^+ : x > a_k$ and $A_k - x \in q\}$. Since $q \oplus q = q$, then $\{x \in \mathbb{Z}^+ : A - x \in q\} \in q$. Also, the set $\{1, 2, \ldots, a_k\}$ is finite, does not contain an (m, p, c)-set for every $m, p, c \in \mathbb{Z}^+$, and therefore is not an element of q. Since q is an ultrafilter the complement $\{a_k{+}1, a_k{+}2, \ldots\}$ is an element of q, and since q is closed under intersection,

$$X_k = \{x \in \mathbb{Z}^+ : A - x \in q\} \cap \{a_k + 1, a_k + 2, \ldots\} \in q.$$

Choose an (m_k, p_k, c_k)-set $M_k \subseteq X_k \subseteq \{x \in \mathbb{Z}^+ : A_k - x \in q\}$. Notice that $A_k \subseteq A_{k-1} \subseteq \cdots \subseteq A_1 = A$.

Let $\{x_i\}_{i \in \mathbb{Z}^+} \subseteq A$ be any sequence so that for every $i \in \mathbb{Z}^+$, $x_i \in M_i$. For $k \geq 1$ let $S(k)$ be the following statement: for any $I \subseteq \mathbb{Z}^+$ with $|I| = k$ and $n = \min\{i \in I\}$, then $\sum_{i \in I} x_i \in A_n$. The statement is proved by induction on $k \geq 1$.

Base Case: For $k = 1$, then $I = \{n\}$ and so $x_n \in M_n \subseteq X_n \subseteq A_n$.

Inductive Step: Let $l > 1$ and suppose $S(l)$ holds. For any set I with $|I| = l + 1$ and $n = \min\{i \in I\}$, let $J = I \setminus \{n\}$ and $n' = \min\{j \in J\}$. By the induction hypothesis, $\sum_{j \in J} x_j \in A_{n'}$. Since $n' > n$, then $x_{n'} > x_n$ by the construction of $X_{n'}$, and so

$$A_{n'} \subseteq A_{n+1} = A_n \bigcap_{x \in M_n} (A_n - x) \subseteq (A_n - x_n).$$

Therefore $a + x_n \in A_n$ for any $a \in A$, and in particular,

$$\sum_{i\in I} x_i = x_n + \sum_{j\in J} x_j \in A_n,$$

concluding the inductive step. By mathematical induction, for every $I \subseteq \mathbb{Z}^+$ with $|I| < \infty$,

$$\sum_{i\in I} x_i \in A.$$

The same is true for any sequence $\{x_i\}_{i\in\mathbb{Z}^+}$ with $x_i \in M_i$, and so A contains an (M, P, C)-system. This is true for any $A \in q$, and so q is an ultrafilter whose every set contains an (M, P, C)-system. By Theorem 7.4.10, any finite colouring of $\mathbb{Z}^+$ has a monochromatic (M, P, C)-system. □

Hindman and Deuber point out in [25] that this theorem cannot be generalized to include sums with more than one element from an (m, p, c)-set. For example, if $m \geq 2$ and $p = 2$, then for any $X = (x_0, x_1, \ldots, x_m)_{p,c}$, all of $cx_1 + x_2, cx_1$ and $cx_1 + 2x_2$ are in X, but in any colouring for which $z, 2z$ are different colours for every $z \in \mathbb{Z}^+$, then $cx_1 + 2x_2$ and cx_1 are a different colour then $(cx_1 + 2x_2) + cx_1 = 2(cx_1 + x_2)$.

In 1993, Hindman and Lefmann [74] showed that a monochromatic (M, P, C)-system can be found in any finite colouring of any (M, P, C)-system.

7.6 A lack of characterizations

The finite sums sets and (M, P, C)-systems form solutions to certain infinite homogeneous linear systems. For example, consider the matrix F of Figure 7.1.

$$F = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \\ & & & & & & & \vdots & & & & & & & & \ddots \end{bmatrix}$$

Figure 7.1: The matrix H.

Every $(2^k - k - 1) \times (2^k - 1)$ submatrix F_k of F consisting of the first $2^k - k - 1$ rows and $2^k - 1$ columns is a partition regular matrix whose solution to $F_k\boldsymbol{x} = \mathbf{0}$ is a finite sums set $FS(\{a_1, a_2, \ldots, a_k\})$. Any solution to $F\boldsymbol{x} = \mathbf{0}$, where $\boldsymbol{x}$ and $\mathbf{0}$ are infinite vectors, is a finite sums set $FS(\{a_1, a_2, a_3, \ldots.\})$.

Definition 7.6.1. *An infinite matrix A is* infinite partition regular *in $\mathbb{Z}^+$ if and only if for every finite colouring of $\mathbb{Z}^+$, there is a infinite monochromatic vector $\boldsymbol{x}$ so that*

$A\boldsymbol{x} = \mathbf{0}$.

Infinite partition regular matrices and variations of these matrices are studied in [3], [4], [26], and [75]. A common theme of the results concerning infinite partition regular matrices is that characterizations of finite partition regular matrices do not generalize to the infinite case. The infinite matrices studied in the papers mentioned above are restricted to those with a finite number of nonzero entries in each row.

The results of Chapter 6 show that every finite colouring of $\mathbb{Z}^+$ admits a colour class (a large set) which contains a solution to every partition regular system (Theorem 6.5.2). This result is proved by characterizing large sets as those which contain an (m, p, c)-set for every $m, p, c \in \mathbb{Z}^+$. It was the hope of the authors of [26] that a similar characterization is true for infinite partition regular systems, and suggested that (M, P, C)-systems would be good candidates: every (M, P, C)-system contains for every $m, p, c \in \mathbb{Z}^+$ an (m, p, c)-set, and so contain a solution to every finite partition regular system. Also, every (M, P, C)-system contains the finite sums set of several infinite sets. Every finite colouring of an (M, P, C)-system has a monochromatic (M, P, C)-system [74]. Also, Lefmann and Kaddour [87] showed that (M, P, C)-systems are maximal in the following sense: let M be an infinite partition regular matrix whose every solution to $M\boldsymbol{x} = \mathbf{0}$ contains an (M, P, C)-system, then if adding a row $\boldsymbol{r}$ to M is still infinite partition regular, then $\boldsymbol{r}$ is a linear combination of some rows of M.

However, it was discovered that such a characterization of "large" sets for infinite partition regular matrices does not exist: in fact that it cannot be guaranteed that a colour class contains a solution to every infinite partition regular system.

Theorem 7.6.2 (Deuber, Hindman, Leader, and Lefmann, 1995 [26])**.** *There exists infinite partition regular matrices A and B and a finite colouring of $\mathbb{Z}^+$, so that no colour class contains a solution to both $A\boldsymbol{x} = \mathbf{0}$ and $B\boldsymbol{y} = \mathbf{0}$.*

The matrices used to prove Theorem 7.6.2 were produced using systems based on results of Milliken [92] and Taylor [120].

Barber, Hindman, Leader and Strauss [4] studied an "infinite columns property":

Definition 7.6.3. *An infinite matrix A with columns $\{\boldsymbol{a}_i\}_{i\in\mathbb{Z}^+}$ satisfies the* infinite columns property *if and only there is a partition partition of $\mathbb{Z} = I_1 \cup I_2 \cup I_3 \cup \cdots$ so that $\sum_{i\in I_1} \boldsymbol{a}_i = \mathbf{0}$ and for all $j \geq 1$ and every $i \in I_1 \cup \cdots \cup I_j$ there exists $\alpha_{i,j} \in \mathbb{Q}$ so that*

$$\sum_{i\in I_1\cup\cdots\cup I_j} \alpha_{i,j}\boldsymbol{a}_i = \sum_{i\in I_{j+1}} \boldsymbol{a}_i.$$

Notice that the definition allows for $\mathbb{Z}^+$ to be partitioned into finitely or infinitely many sets.

In [4], the following matrix is considered:

$$A = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & \\ & & & & \vdots & & & & & \ddots \end{bmatrix}.$$

This matrix satisfies the infinite columns property with $I_1 = \{1, 3, 5, \ldots\}$ and $I_2 = \{2, 4, 6, \ldots\}$. However, the system $A\boldsymbol{x} = \mathbf{0}$ has no solution in $\mathbb{Z}^+$; any solution $\boldsymbol{x} = \{x_1, y_1, x_2, y_2, \ldots\}$ satisfies $x_i - x_{i+1} = y_i$ and for $y_i > 0$, the solution also satisfies $x_i > x_{i+1}$, implying $x_1 > x_2 > \cdots > x_i > \cdots$, an impossibility for $x_1, x_2, \ldots, x_i, \ldots \in \mathbb{Z}^+$. Therefore, a matrix A having the infinite columns property does not imply infinite partition regularity. The authors of [4] also gave an example of a matrix that is infinite partition regular, but does not satisfy the infinite columns property. However, they note that the first part of the columns property can be satisfied under an extra condition on the entries of the matrix A.

Theorem 7.6.4 (Barber, Hindman, Leader, and Strauss [4]). *Let A be an infinite partition regular with columns $\{\boldsymbol{a}_i\}_{i \in \mathbb{Z}^+}$. If A has the property that there exists an $m \in \mathbb{Z}^+$ so that along every row r_j of A, the sum of the absolute values of the entries along r_j is bounded by m, then there exists $I \subseteq \mathbb{Z}^+$ so that $\sum_{i \in I} \boldsymbol{a}_i = \mathbf{0}$.*

Chapter 8

Nonlinear partition regular systems

8.1 Introduction

In a straight forward manner, multiplicative results for partition regularity can be achieved from the additive ones. For example, setting $n = 2^{S(r)}$, where $S(r)$ denotes the Schur number from Section 3.3, then any r-colouring of $[2, n]$ yields x, y so that $\{x, y, xy\}$ is monochromatic: simply define a new colouring $\Delta' : [\log_2(n)] \to [r]$ by

$$\Delta'(z) = \Delta(2^z).$$

Then by Schur's theorem, there exists x', y' so that $\{x', y', x'+y'\}$ is monochromatic with respect to Δ'. Then set $x = 2^{x'}$ and $y = 2^{y'}$, and so x, y and $xy = 2^{x'+y'}$ are assigned the same colour with respect to Δ. Graham (see [66]) showed, using colourings of powers of 2 similar to the argument above, that a multiplicative Hindman's theorem (recall Theorem

7.4.9) is true: for an infinite set $A \subseteq \mathbb{Z}^+$, let

$$FP(A) = \left\{ \prod_{i \in I} a_i : I \subseteq \mathbb{Z}^+, I \neq \emptyset, |I| < \infty \right\}.$$

Then in any finite colouring of $\mathbb{Z}^+$, there is an infinite set A so that $FP(A)$ is monochromatic.

In 1991, Lefmann [86] published the following result (recall the definition of the columns property from Definition 4.1.2):

Theorem 8.1.1 (Lefman, 1991 [86, Thm. 2.6])**.** *Let A be an $m \times n$ matrix with rational entries. The system*

$$A \begin{pmatrix} x_1^{-1} \\ x_2^{-1} \\ \vdots \\ x_n^{-1} \end{pmatrix}$$

has a monochromatic solution $(x_1, x_2, \ldots, x_n)$ under any finite colouring of $\mathbb{Z}^+$ if and only if A satisfies the columns property in $\mathbb{Q}$.

Since $S = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$ satisfies the columns property in $\mathbb{Q}$, Lefmann's theorem implies that the system

$$\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$$

is partition regular in $\mathbb{Z}^+$.

Section 8.2 focuses on a result known as the Sárközy–Furstenberg theorem [109], [42], which can be interpreted to say that for $r \in \mathbb{Z}^+$, there is a least $n \in \mathbb{Z}^+$ so that for any

r-colouring $\Delta : [n] \to [r]$, there exists x, y so that $\{x, x + y^2\}$ is monochromatic. The theorem in fact states that for any $\delta > 0$, there is $n \in \mathbb{Z}^+$ so that any set A with density $|A| > \delta n$ contains two elements whose difference is a square. The proof presented in Section 8.2 is due to Green, Tao, and Ziegler (see [119]).

The Sárközy–Furstenberg has been generalized to include the value y with the same colour and also to specify different powers as the difference. In 1996, Bergelson and Leibman [9] showed that for any polynomials $p_1, p_2, \ldots, p_k$ with integer coefficients so that $p_1(0) = p_2(0) = \cdots = p_k(0)$, there is a least n so that for any r-colouring of $[n]$ there is x, y so that $\{x, y, x + p_1(y), x + p_2(y), \ldots, x + p_k(y)\}$ is monochromatic. Bergelson and Leibman's result is known as the *polynomial van der Waerden theorem.*

8.2 The Sárközy–Furstenberg theorem

Sárközy in 1978 [109] and Furstenberg in 1977 [42] proved independently that for any $\delta > 0$ and n large enough, if $A \subseteq [n]$ has density at least δ in $[n]$, then A contains two elements whose difference is a square. Sárközy applied the Hardy–Littlewood circle method along with a density increment argument similar to Roth's density result for arithmetic progressions [102] [103], while Furstenberg used topological dynamics to achieve this result.

Theorem 8.2.1 (Sárközy, 1978 [109]; Furstenberg, 1977 [42]). *For every $\delta > 0$, there is* *a $N \in \mathbb{Z}^+$ so that for every $n > N$ and every $A \subseteq [n]$ with*

$$\frac{|A|}{n} > \delta,$$

the set A contains $x < y$ so that $y - x$ is a square.

The result for Theorem 8.2.1 is immediate for $\delta > \frac{1}{2}$: If $n \geq 4$ is even, then any set $A \subseteq [n]$ with density greater than $\frac{1}{2}$ contains two elements $x < y$ that are consecutive, and so $y - x = 1^2$. If $n \geq 4$ is odd and a set $A \subseteq [n]$ with density greater than $\frac{1}{2}$ has no consecutive elements, then A consists of all odd numbers in $[n]$, and in particular, $1, 5 \in A$.

To see that the Sárközy–Furstenberg theorem implies a Ramsey result, the pigeonhole principle guarantees some colour class contains at least $\frac{n}{r}$ elements in any r-colouring of $[n]$. And so for $\delta \leq \frac{1}{r}$, Theorem 8.2.1 guarantees that some colour class contains two elements whose difference is a square.

There are many proofs of Theorem 8.2.1 using Fourier analysis (for example [50], [91], and [95]). The proof presented here, which follows the one discovered by Green, Tao and Ziegler and published on Tao's blog [119], does not use any Fourier analysis and relies instead on the Cauchy-Schwarz inequality to guarantee the density increment argument can be applied.

Throughout, for any $n \in \mathbb{Z}^+$, let $M = [n], R = [\lfloor n^{1/3} \rfloor]$, and $H = [\lfloor n^{1/24} \rfloor]$. For a set

$A \subseteq [n]$, the indicator function $\mathbb{1}_A : [n] \to \{0, 1\}$ is defined by

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A \\ \\ 0 & x \notin A. \end{cases}$$

For any finite set X and any function $g : X \to \mathbb{R}$, the *expected value* is denoted

$$\mathop{\mathbb{E}}_{x \in X} g(x) = \frac{1}{|X|} \sum_{x \in X} g(x).$$

The expected value gives the average value or $g(x)$ over all $x \in X$. Therefore, if $\mathbb{E}_{x \in X} g(x) = a$, then there is some $x \in X$ so that $g(x) \geq a$ and some $x' \in X$ so that $g(x') \leq a$.

For a set $A \subseteq [n]$ with density δ, define $f : [n] \to \mathbb{R}$ as

$$f(x) = \mathbb{1}_A(x) - \delta \mathbb{1}_M(x).$$

Then

$$\mathop{\mathbb{E}}_{x \in M} f(x) = \frac{1}{n} \left(\sum_{x \in X} (\mathbb{1}_A(x) - \delta \mathbb{1}_M(x)) \right) = \frac{1}{n}(|A| - \delta n) = 0.$$

If A does not contain elements whose difference is a square, then say A is *square difference free*, or *free of square differences.*

Lemma 8.2.2. *Let $A \subseteq [n]$ have density*

$$\frac{|A|}{n} = \delta$$

and be free of square differences. Then

$$\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} \mathop{\mathbb{E}}_{h \in H} f(m) f(m + (r + h)^2) < -\delta^2 + 2\delta(n^{-1/6} + n^{-11/24})^2.$$

Proof. Suppose A has density greater than δ in $[n]$, and A is free of elements with square difference. For any $m, r, h \in [n]$, since A has no square differences, $\mathbb{1}_A(m)\mathbb{1}_A(m + (r + h)^2) = 0$. Therefore,

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathbb{1}_A(m)\mathbb{1}_A(m+(r+h)^2)=0.$$

The Lemma follows from the three inequalities:

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta^2\mathbb{1}_M(m)\mathbb{1}_M(m+(r+h)^2)<\delta^2 \tag{8.1}$$

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_A(m)\mathbb{1}_M(m+(r+h)^2)>\delta^2-\delta(n^{-1/6}+n^{-11/24})^2 \tag{8.2}$$

$$\mathop{\mathbb{E}}_{n\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_M(m)\mathbb{1}_A(m+(r+h)^2)>\delta^2-\delta(n^{-1/6}+n^{-11/24})^2, \tag{8.3}$$

since by linearity of expectation

$$\begin{aligned}&\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\left[\mathbb{1}_A(m)-\delta\mathbb{1}_M(m)\right]\left[\mathbb{1}_A(m+(r+h)^2)-\delta\mathbb{1}_M(m+(r+h)^2)\right]\\&\quad=\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathbb{1}_A(m)\mathbb{1}_A(m+(r+h)^2)-\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_A(m)\mathbb{1}_M(m+(r+h)^2)\\&\quad-\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_M(m)\mathbb{1}_A(m+(r+h)^2)+\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta^2\mathbb{1}_M(m)\mathbb{1}_M(m+(r+h)^2).\end{aligned}$$

For $r \in R$ and $h \in H$, if $m > n - (r + h)^2$ then $m + (r + h)^2 \notin M$, and so $\mathbb{1}_M(n + (r + h)^2) \leq 1$. Therefore

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta^2\mathbb{1}_M(m)\mathbb{1}_M(m+(r+h)^2)<\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta^2\mathbb{1}_M(m)=\delta^2,$$

which proves (8.1). Letting

$$T=[\lfloor n-(n^{1/3}+n^{1/24})^2\rfloor],$$

then for any $r \in R$ and $h \in H$, if $m \in T$, then $m + (r+h)^2 \in M$. Also, $|A \cap T| > |A| - (n^{1/3} + n^{1/24})^2$. Therefore,

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_A(m)\mathbb{1}_M(m+(r+h)^2) > \mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{m\in M}\delta\mathbb{1}_A(m)\mathbb{1}_T(m)$$
$$> \mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\frac{\delta}{n}(|A| - (n^{1/3}+n^{1/24})^2) = \delta^2 - \delta(n^{-1/6}+n^{-11/24})^2.$$

which proves (8.2). For any $r \in R$ and $h \in H$, let $D = \{m \in M : m + (r+h)^2 \in A\}$. Then $|D| > |A| - (n^{1/3} + n^{1/24})^2$, and since $\mathbb{1}_M(m) = 1$ for every $m \in M$,

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\delta\mathbb{1}_M(m)\mathbb{1}_A(m+(r+h)^2) = \mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{m\in M}\delta\mathbb{1}_D(m)$$
$$> \mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\frac{\delta}{n}(|A| - (n^{1/3}+n^{1/24})^2) = \delta^2 - \delta(n^{-1/6}+n^{-11/24})^2,$$

which proves (8.3).

Putting it all together,

$$\mathop{\mathbb{E}}_{n\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H} f(n)f(n+(r+h)^2)$$
$$= \mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\left[\mathbb{1}_A(m) - \delta\mathbb{1}_M(m)\right]\left[\mathbb{1}_A(m+(r+h)^2) - \delta\mathbb{1}_M(m+(r+h)^2)\right]$$
$$< 0 - 2(\delta^2 - \delta(n^{-1/6}+n^{-11/24})^2) + \delta^2 = -\delta^2 + 2\delta(n^{-1/6}+n^{-11/24})^2.$$

□

Lemma 8.2.3. *Let $\delta \in (0, \frac{1}{2}]$. For any $n \in \mathbb{Z}^+$ large enough so that*

$$\frac{1}{4}\delta^3 > 3n^{-1/24},$$

if $A \subseteq [n]$ is free of square differences and has density

$$\frac{|A|}{n} > \delta,$$

then there is an arithmetic progression

$$B = \{a, a + d^2, a + 2d^2, a + 3d^2, \ldots, a + \lfloor n^{1/4} \rfloor d^2\} \subseteq [n],$$

so that

$$\frac{|A \cap B|}{|B|} > \delta + \frac{1}{4}\delta^3.$$

Proof. Suppose A has density $\delta_1 > \delta$ in $[n]$, and A is free of square differences. The choice of n guarantees $(n^{-1/6} + n^{-11/24})^2 < \frac{1}{8}n^{-1/24}$ and that $-\delta^2 + 2\delta(n^{-1/6} + n^{-11/24})^2 < 0$. By Lemma 8.2.2,

$$\begin{aligned} |\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} \mathop{\mathbb{E}}_{h \in H} f(m) f(m + (r+h)^2)|^2 &> | - \delta_1^2 + 2\delta_1 (m^{-1/6} + m^{-11/24})^2|^2 \\ &> \delta_1^4 - 4\delta_1^3 (m^{-1/6} + m^{-11/24})^2 \\ &> \delta_1^4 - \frac{\delta_1^3}{2} n^{-1/24}. \end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned} &\left(\mathop{\mathbb{E}}_{m \in M} |f(m)|^2\right) \left(\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} |\mathop{\mathbb{E}}_{h \in H} f(m + (r+h)^2)|^2\right) \\ &\quad = \left(\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} |f(m)|^2\right) \left(\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} |\mathop{\mathbb{E}}_{h \in H} f(m + (r+h)^2)|^2\right) \\ &\quad \geq |\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} f(m) \mathop{\mathbb{E}}_{h \in H} f(m + (r+h)^2)|^2 > \delta_1^4 - \frac{\delta_1^3}{2} n^{-1/24}. \end{aligned} \tag{8.4}$$

Since $f(m) = \mathbb{1}_A(m) - \delta_1 \mathbb{1}_M(m)$, then for $|A|$ values of $m \in M$, $|f(m)|^2 = (1 - \delta_1)^2$, and for the other $n - |A|$ values of $m \in M$, $|f(m)|^2 = \delta_1^2$. Therefore,

$$\mathop{\mathbb{E}}_{m \in M} |f(m)|^2 = \frac{1}{n} \left(|A|(1 - \delta_1)^2 + (n - |A|)\delta_1^2\right) = \delta_1 - \delta_1^2 < \delta_1.$$

By dividing out $\mathbb{E}_{m\in M}|f(n)|^2$ from (8.4),

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}|\mathop{\mathbb{E}}_{h\in H}f(m+(r+h)^2)|^2 > \frac{\delta_1^4-\frac{\delta_1^3}{2}n^{-1/24}}{\delta_1} > \delta_1^3-\frac{1}{2}n^{-1/24}.$$

By expanding $|\mathbb{E}_{h\in H}f(m+(r+h)^2)|^2 = \mathbb{E}_{h\in H}\mathbb{E}_{h'\in H}f(m+(r+h)^2)f(m+(r+h')^2)$,

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}f(m+(r+h)^2)f(m+(r+h')^2) > \delta_1^3-\frac{1}{2}n^{-1/24}.$$

The next is to remove all the instances of $h=h'$. Since $|f(x)|\leq 1$,

$$\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}\mathbb{1}_{\{h\}}(h')f(m+(r+h)^2)f(m+(r+h')^2) < \mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}\mathbb{1}_{\{h\}}(h') = n^{-1/24},$$

and so

$$\begin{aligned}
&\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}\mathbb{1}_{H\setminus\{h\}}(h')f(m+(r+h)^2)f(m+(r+h')^2)\\
&\quad= \mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}f(m+(r+h)^2)f(m+(r+h')^2)\\
&\quad- \mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}\mathbb{1}_{\{h\}}(h')f(m+(r+h)^2)f(m+(r+h')^2)\\
&\quad> \mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}\mathop{\mathbb{E}}_{h\in H}\mathop{\mathbb{E}}_{h'\in H}f(m+(r+h)^2)f(m+(r+h')^2)-n^{-1/24}\\
&\quad> \delta_1^3-\frac{3}{2}n^{-1/24},
\end{aligned}$$

therefore there are $h\neq h'$ so that

$$\mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R}f(m+(r+h)^2)f(m+(r+h')^2) > \delta_1^3-\frac{3}{2}n^{-1/24}.$$

Without loss of generality, assume $h'>h$. Let $a=(h'-h)(h'+h)>0$ and $d=2(h'-h)>0$. Then

$$m-(r+h)^2+(r+h')^2 = m+(h'-h)(2r+h+h') = m+a+dr.$$

For each $r \in R$,

$$\sum_{m \in M} f(m + (r+h)^2) f(m + (r+h')^2) = \sum_{m \in M} f(m) f(m + a + dr)^2)$$
$$+ \sum_{m \in [n+1, n+(r+h)^2]} f(m) f(m + a + dr) - \sum_{m \in [(r+h)^2]} f(m) f(m + a + dr).$$

If $m > n$, then $f(m) = \mathbb{1}_A(m) - \delta_1 \mathbb{1}_M(m) = 0$, and so

$$\sum_{m \in [n+1, n+(r+h)^2]} f(m) f(m + a + dr) = 0.$$

Since $|f(x)| \leq 1$ for any x, then

$$\begin{aligned}\sum_{m \in [(r+h)^2]} f(m) f(m + a + dr) &< \sum_{m \in [(r+h)^2]} |f(m)||f(m + a + dr)| \\ &< (r+h)^2 \\ &< (n^{1/3} + n^{1/24})^2,\end{aligned}$$

and so

$$\begin{aligned}&\mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} f(m) f(m + a + dr) \\ &\qquad = \mathop{\mathbb{E}}_{r \in R} \mathop{\mathbb{E}}_{m \in M} f(m) f(m + a + dr) \\ &\qquad = \mathop{\mathbb{E}}_{r \in R} \frac{1}{n} \left(\sum_{m \in M} f(m) f(m + a + dr) - \sum_{m \in [(r+h)^2]} f(m) f(m + a + dr) \right) \\ &\qquad > \mathop{\mathbb{E}}_{m \in M} \mathop{\mathbb{E}}_{r \in R} f(m + (r+h)^2) f(m + (r+h')^2) - (n^{-1/6} + n^{-11/24})^2 \\ &\qquad > \delta_1^3 - \frac{3}{2} n^{-1/24} - \frac{1}{8} n^{-1/24} = \delta_1^3 - \frac{13}{8} n^{-1/24}.\end{aligned}$$

Since $|f(x)| \leq 1$, by the triangle inequality

$$\begin{aligned}\mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} f(m+a+dr)| &\geq \mathop{\mathbb{E}}_{m\in M} |f(m)||\mathop{\mathbb{E}}_{r\in R} f(m+a+dr)| \\ &\geq \mathop{\mathbb{E}}_{m\in M}\mathop{\mathbb{E}}_{r\in R} f(m)f(m+a+dr),\end{aligned}$$

and so

$$\mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} f(m+a+dr)| > \delta_1^3 - \frac{13}{8}n^{-1/24}.$$

With $T = [n - (\lfloor n^{1/3}\rfloor + \lfloor n^{1/24}\rfloor)^2]$, then

$$\begin{aligned}&\mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} \mathbb{1}_T(m) f(m+a+dr)| \\ &\qquad \geq \mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} f(m+a+dr)| - \mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} \mathbb{1}_{M\setminus T}(m) f(m+a+dr)| \\ &\qquad > \mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} f(m+a+dr)| - (n^{-1/6}+n^{-11/24})^2.\end{aligned}$$

and so

$$\mathop{\mathbb{E}}_{m\in M} |\mathop{\mathbb{E}}_{r\in R} \mathbb{1}_T(m) f(m+a+dr)| > \delta_1^3 - \frac{13}{8}n^{-1/24} - (n^{-1/6}+n^{-11/24})^2 > \delta_1^3 - \frac{14}{8}n^{-1/24}.$$

For any $r \in R$, the expected value $\mathbb{E}_{m\in M} \mathbb{1}_T(m) f(m+a+dr)$ counts $|T|$ values of $f(m)$.

Since $\mathbb{E}_{m\in M} f(m) = 0$ and $|f(m)| < 1$,

$$\mathop{\mathbb{E}}_{m\in M} \mathbb{1}_T(m) f(m+a+dr) < \mathop{\mathbb{E}}_{m\in M} f(m) + (n^{-1/6}+n^{-11/24})^2 = (n^{-1/6}+n^{-11/24})^2.$$

Looking at the positive instances of $\mathbb{E}_{r\in R} \mathbb{1}_T(m) f(m+a+dr)$,

$$\begin{aligned}\mathop{\mathbb{E}}_{m\in M} \max\{0, \mathop{\mathbb{E}}_{r\in R} \mathbb{1}_T(m) f(n+a+dr)\} &> \frac{1}{2}\left(\delta_1^3 - \frac{14}{8}n^{-1/24}\right) - (n^{-1/6}+n^{-11/24})^2 \\ &> \frac{1}{2}\delta_1^3 - \frac{7}{8}n^{-1/24} - \frac{1}{8}n^{-1/24} \\ &= \frac{1}{2}\delta_1^3 - n^{-1/24},\end{aligned}$$

and so there is some $m \in T$ so that

$$\mathop{\mathbb{E}}_{r \in R} f(m+a+dr) > \frac{1}{2}\delta_1^3 - n^{-1/24}.$$

For $m \in T$ and every $r \in R$, $m+a+dr \in M$. By definition, $f(x) = \mathbb{1}_A(x) - \delta_1 \mathbb{1}_M(x)$, and so

$$\begin{aligned}\mathop{\mathbb{E}}_{r \in R} \mathbb{1}_A(m+a+dr) &> \mathop{\mathbb{E}}_{r \in R} \delta_1 \mathbb{1}_M(m+a+dr) + \frac{1}{2}\delta_1^3 - n^{-1/24} \\ &= \delta_1 + \frac{1}{2}\delta_1^3 - n^{-1/24},\end{aligned}$$

that is, A has density greater than δ_1 on the arithmetic progression of length $\lfloor n^{1/3} \rfloor$

$$B = \{m+a+dr : r \in R\}.$$

The arithmetic progression B can be partitioned into arithmetic progressions $B_1, B_2, \ldots,$ $B_i, \ldots$ of length $\lfloor n^{1/4} \rfloor$ and difference d^2 with at most $dn^{1/4} < 2n^{1/24}n^{1/4} = 2n^{7/24}$ terms in an error set E. Letting $S = [\lfloor n^{1/4} \rfloor]$, with

$$\mathop{\mathbb{E}}_{r \in R} \mathbb{1}_{A \cap E}(m+a+dr) < \frac{2n^{7/24}}{n^{1/3}} = 2n^{-1/24},$$

there is some $B_i = \{m' + d^2 s : s \in S\}$ so that

$$\begin{aligned}\mathop{\mathbb{E}}_{x \in B_i} \mathbb{1}_A(x) &> \delta_1 + \frac{1}{2}\delta_1^3 - n^{-1/24} - 2n^{-1/24} \\ &= \delta_1 + \frac{1}{2}\delta_1^3 - 3n^{-1/24},\end{aligned}$$

which by the choice of n, implies

$$\frac{|A \cap B_i|}{|B_i|} > \delta_1 + \frac{1}{4}\delta_1^3 > \delta + \frac{1}{4}\delta^3.$$

□

Theorem 8.2.4. *For $\delta > 0$ and*

$$n > \left(\frac{12}{\delta^3}\right)^{24\cdot 4^{2/\delta^3}},$$

if $|A|$ has density $\frac{|A|}{n} > \delta$, then A contains two elements x, y whose difference is a square.

Proof. Let $t = \lfloor \frac{2-4\delta}{\delta^3} \rfloor$. For each $i \in [t]$, let $\delta_i = \delta + \frac{i}{4}\delta^3$ and $n_i = n^{1/4^i}$. Then for each i, $\delta_i \in (0, \frac{1}{2}]$,

$$\frac{1}{4}\delta_i^3 > \frac{1}{4}\delta^3 = 3\left(\frac{12}{\delta^3}\right)^{-1} > 3\left(\left(\frac{12}{\delta^3}\right)^{\frac{24\cdot 4^{2/\delta^3}}{4^i}}\right)^{-1/24} > 3n_i^{-1/24}.$$

Therefore, for $i \in [t]$, if $A \subseteq [n_i]$ is free of square differences and has density greater than δ_i in $[n_i]$, Lemma 8.2.3 can be applied.

For any $n \in \mathbb{Z}^+$, if $|A| > \frac{n}{2}$, then A contains two elements whose difference is a square. Therefore assume in hopes of a contradiction that $\delta \in (0, \frac{1}{2}]$ and for $n_0 > \left(\frac{12}{\delta^3}\right)^{24\cdot 4^{2/\delta^3}}$, A is a square difference free subset of $[n_0]$ and has density greater than δ in $[n_0]$. Then by Lemma 8.2.3 there is an arithmetic progression $B_1 = \{a_1 + rd_1^2 : r \in \lfloor n_1 \rfloor\}$ so that $A \cap B$ has density greater than $\delta_1 = \delta + \frac{1}{4}\delta^3$. Let $A_1 = \{r \in n_1 : a + rd^2 \in A\}$. Then $|A_1| = |A \cap B|$, and if A_1 contains elements $x < y$ with $y - x = e^2$, then $(a + yd^2) - (a + xd^2) = (ed)^2$. Therefore A_1 is also square difference free.

Repeat the argument above for $i \in [t]$: Suppose $A_i \subseteq [n_i]$ has density greater than δ_i. Then there is an arithmetic progression $B_{i+1} = \{a_{i+1} + rd_{i+1}^2 : r \in \lfloor n_{i+1} \rfloor\}$ with

$$\frac{|A_i \cap B_{i+1}|}{|B_{i+1}|} > \delta_i + \frac{1}{4}\delta_i^3 > \delta + \frac{i}{4}\delta^3 + \frac{1}{4}\delta^3 = \delta_{i+1}.$$

Then let $A_{i+1} = \{r \in n_{i+1} : a_{i+1} + rd_{i+1}^2 \in A_i\}$. Again, if A_i has no square-differences, then neither does A_{i+1}.

Since A is assumed to be square difference free, there is an arithmetic progression B_{t+1} of length $n_0^{1/4^{t+1}}$ for which $A_t \cap B_{t+1}$ has density greater than $\delta_t + \frac{1}{4}\delta_t^3 > \delta + \frac{t+1}{4}\delta^3 > \frac{1}{2}$ in B_{t+1}. Then the set $A_{t+1} = \{r \in \lfloor n_{t+1}^{1/4} \rfloor : a_{t+1} + rd_{t+1}^2 \in A_t\}$ contains more than half the elements in $\lfloor n_{t+1}^{1/4} \rfloor$, and so it contains 2 elements with a square difference. This contradicts A containing no square differences. $\square$

For $\delta \in (0, \frac{1}{2}]$, then $2^{2^{1/\delta^3}} > \frac{12}{\delta^3}$ and so for $n > \left(2^{2^{1/\delta^3}}\right)^{24 \cdot 4^{2/\delta^3}}$ any set $A \subseteq [n]$ that is free of square differences must have density less than δ. Letting log denote the base 2 logarithm, rearranging for δ this means that for n large enough, if $A \subseteq [n]$ has density

$$\frac{|A|}{n} > \sqrt[3]{\frac{5}{\log \log n - \log 24}},$$

then A contains square differences. Tao (see [119]) points out that the best known bound for densities that guarantee square differences is due to Pintz, Steiger, and Szemeredi [95] who, in 1988, showed that for some absolute constant C, any set with density at least

$$\frac{C}{(\log n)^{\frac{1}{4} \log \log \log \log n}}$$

contains square differences. As for a lower bound, Ruzsa in 1984 [104] gave constructions of sets with densities greater than $\frac{1}{65} n^{-0.266923}$ free of square differences. The exponent was improved to -0.266588 by Lewko [88] in 2015.

In 1985, Bergelson [8] was able to show that for any finite colouring of $\mathbb{Z}^+$, some colour class contains $\{x, y, x + y^2\}$ monochromatic, that is, some colour class contains two elements who difference is a square, along with the square root of that difference.

8.3 The Polynomial van der Waerden theorem

Using topological dynamics and ergodic theory, Bergelson and Leibman prove in 1996 what is known as the polynomial van der Waerden theorem.

Theorem 8.3.1 (Bergelson and Leibman, 1996 [9])**.** *Let $p_1, p_2, \ldots, p_k$ be polynomials with integer coefficients so that $p_1(0) = p_2(0) = \cdots = p_k(0) = 0$. Then for every r-colouring of $\mathbb{Z}^+$, there is $x, y \in \mathbb{Z}^+$ so that*

$$\{x, x + p_1(y), x + p_2(y), \ldots, x + p_k(y)\}$$

is monochromatic.

With $p_1 = y^2$, the polynomial van der Waerden theorem implies the Sárközy–Furstenberg theorem. Setting $p_1 = y, p_2 = 2y, \ldots, p_k = ky$, then Theorem 8.3.1 implies van der Waerden's theorem, but by setting $p_1 = y^2, p_2 = 2y^2, \ldots, p_k = ky^2$, the difference of consecutive elements in a monochromatic AP_k can be guaranteed to be squares. Similarly, the difference between consecutive elements of a monochromatic AP_k can be guaranteed to be any fixed power.

A combinatorial proof of the polynomial van der Waerden theorem was published in 2000 by Walters [124].

Chapter 9

Generalizations and open problems

9.1 Partition regular independent sets in graphs on the integers

For a graph $G = (V, E)$, an *independent set* is a subset $Y \subseteq V$ of the vertices so that $[Y]^2 \cap E = \emptyset$, that is, there are no edges between vertices in Y. Hajnal (see [34]) asked whether every triangle-free graph G with vertex set $\mathbb{Z}^+$ has an infinite set $A \subseteq \mathbb{Z}^+$ for which $FS(A)$ is an independent set. Hajnal's question was later answered in the negative (see [24]), but before it was answered Erdős "retaliated" (see [56]) by asking whether every triangle-free graph with vertex set $\mathbb{Z}^+$ contains an independent Schur triple $Y = \{x, y, x+y\}$. A stronger statement was proved by Łuczak, Rödl, and Schoen.

Theorem 9.1.1 (Łuczak, Rödl, and Schoen, 1998 [90])**.** *Let $k, d \in \mathbb{Z}^+$. Then for any*

K_k-free graph with vertex set $\mathbb{Z}^+$, there exists $a_1, a_2, \ldots, a_d$ so that $FS(\{a_1, a_2, \ldots, a_d\})$ is an independent set.

A few similar results have also been proved; see, for example, [56] for a proof that an arithmetic progression of arbitrary length can be found in an independent set. The work on independent partition regular sets on graphs culminated with the guarantee of an (m, p, c)-set in an independent set.

Theorem 9.1.2 (Gunderson, Leader, Prömel, and Rödl, 2003 [57]). *Given $k, m, p, c \in \mathbb{Z}^+$ with $p < c$, there exists $n, q, d \in \mathbb{Z}^+$ so that any K_k-free graph with vertices labelled by an (n, q, d)-set contains an (m, p, c)-set as an independent set.*

Theorem 6.3.1 states that for every $m, p, c, r \in \mathbb{Z}^+$, there exists $n, q, d \in \mathbb{Z}^+$ so that every r-colouring of any (n, q, d)-set has a monochromatic (m, p, c)-set. To see that Theorem 9.1.2 generalizes Theorem 6.3.1, choose $k, m, p, c \in \mathbb{Z}^+$ and let $n, q, d \in \mathbb{Z}^+$ be guaranteed by Theorem 9.1.2. Then for any $(k-1)$-colouring of any (n, q, d)-set N, create a $(k-1)$-partite graph $G(V_1 \cup \cdots \cup V_{k-1}, E)$, where each partite set V_i consists of the members of N with colour i, and draw an edge between $v \in V_i$ and $u \in V_j$ if and only if $i \neq j$. Then the graph G is K_k-free (there are no k vertices in k different partite sets). Then by Theorem 9.1.2, there is an independent (m, p, c)-set, which is contained completely in one of the colour classes.

9.2 Density results

Several results in partition regularity can be extracted from density results. In Section 8.2, it is discussed how the Sarközy–Furstenberg theorem implies that for every $r \in \mathbb{Z}^+$, there is a least positive integer n so that every r-colouring of $[n]$ yields two monochromatic elements $x < y$ so that $y - x$ is a square.

The upper bound of $h(m; r) \leq (2r)^{2^{m-1}}$ in Section 3.2 is achieved in [58] by first showing that for $m \geq 3$ and any $n > \left(\frac{2^d-2}{\ln 2}\right)^2$, then any set A with density

$$\frac{|A|}{n} \geq 2n^{-\frac{1}{2^{d-1}}}$$

contains a Hilbert cube of dimension m. Then for $r \in \mathbb{Z}^+$, so long as $n \geq (2r)^{2^{m-1}}$, then in any r-colouring of $[n]$, by the pigeonhole principle, one of the colour classes contains at least $\frac{n}{r}$ elements, and so has density large enough so that it contains a Hilbert cube of dimension m. The upper bound on the density of sets A free of Hilbert cubes of dimension m was improved by Sándor [107] to $\frac{|A|}{n} \leq n^{-\frac{1}{2^{m-1}}} + 2n^{-\frac{1}{2^{m-2}}}$.

Upper bounds for the van der Waerden number $W(k; r)$ are also achieved through density results. Let $r_k(n)$ denote the size of the largest set $A \subseteq [n]$ so that A is free of AP$_k$'s. Roth's theorem [102] [103] gave the upper bound $r_3(k) < \frac{c \ln n}{\ln \ln n}$. The bound on $r_3(n)$ has been improved several times (see for example [14] and [106]); to the best of my knowledge the best known bound is by Bloom in 2016 [13], which is $r_3(n) < \frac{cn(\ln \ln n)^6}{\ln n}$. For AP$_4$'s, the first density result is due to Szemereédi [118]. In 1998, Gowers [44] improved the bound to $r_4(n) < \frac{n}{(\log_2 \log_2 n)^c}$. Recently in 2017, Green and Tao [52] further improved the bound to $r_4(n) < \frac{n}{(\log_2 n)^c}$. For arithmetic progressions of length k, in 2001, Gowers [45] proved that $r_k(n) < \frac{n}{(\log_2 \log_2 n)^{c(k)}}$.

Density arguments cannot always be used to derive results in partition regularity, for example, for any $n \in \mathbb{Z}^+$, large subsets of $[n]$ are free of Schur triples: let A consist of all odd numbers in $[n]$, which has density greater than $\frac{1}{2}$. Then for any $x, y \in A$, their sum $x + y$ is even and so is not in A. Also, this set A does not contain a set $\{x, y, x + y, xy\}$ (because A does not contain a Schur triple).

Most of the results mentioned so far use a form of discrete Fourier analysis. Another useful tool in achieving density results is ergodic theory, which was used by Furstenberg [42] to reprove Szemerédi's density result on arithmetic progressions of length four and to prove the Sárközy–Furstenberg theorem. The use of ergodic theory, however, does not usually give quantitative upper bounds for density results, but it does allow to prove theorems like the polynomial van der Waerden theorem (Theorem 8.3.1). Another density result achieved with ergodic theory is the density Hales–Jewett theorem:

Theorem 9.2.1 (Furstenberg and Katznelson, 1991 [43])**.** *Let A be an alphabet with t elements. For every $\epsilon > 0$ there exists a least $R(\epsilon, t) \in \mathbb{Z}^+$ so that for every positive integer $n > R(\epsilon, t)$, if $S \subseteq [A]\binom{n}{0}$ has at least ϵt^n elements, then S contains a combinatorial line.*

A combinatorial proof of the density Hales–Jewett theorem was achieved by the Polymath Project [96].

9.3 Open problems

In 2003, Hindman, Leader, and Strauss [73] published a list of open problems in partition regularity. A few of those problems are presented here.

For any prime p and every positive integer $x \in \mathbb{Z}^+$, there are unique integers $y, z \in \mathbb{Z}^+$ with $y \not\equiv 0 \pmod{p}$ so that $x = yp^z$. The colouring $c_p : \mathbb{Z}^+ \to [p-1]$ is defined by $c_p(x) = y$. In the proof of Theorem 4.3.1, the columns property of a partition regular matrix A was established by considering the colouring c_p for a large enough prime p, and so a consequence is the following: If a matrix A is partition regular with respect to every colouring c_p, then it satisfies the columns property. Thanks to Rado's characterization of partition regular matrices (or Lemma 6.4.2), any matrix with the columns property is partition regular with respect to every colouring.

Question 9.3.1 (see [73, Question 1, Section 1])**.** *Is there a direct proof (without using Rado's characterization of partition regular matrices) that if for every prime p, A is partition regular with respect to c_p, then A is partition regular with respect to every finite colouring of $\mathbb{Z}^+$?*

If A is a finite matrix that is not partition regular, there may still be small numbers $r \in \mathbb{Z}^+$ so that under any r-colouring $\Delta : \mathbb{Z}^+ \to [r]$, there is a monochromatic solution to $A\boldsymbol{x} = \mathbf{0}$. The following open problem was asked by Rado [99] (see [73]).

Question 9.3.2 (Rado's boundedness conjecture, see [73, Question 2, Section 1])**.** *Given m and n, does there exists a least positive integer $k = k(m, n)$ so that if A is an $m \times n$*

matrix with the property that there is a monochromatic solution to $A\boldsymbol{x} = \mathbf{0}$ *under any* k*-colouring of* $\mathbb{Z}^+$ *then* A *is partition regular under any finite colouring of* $\mathbb{Z}^+$*?*

The case for $m = 1$ and $n = 3$ was solved by Fox and Kleitman [39]:

Theorem 9.3.3 (Fox and Kleitman, 2006 [39]). *For every* $a_1, a_2, a_3 \in \mathbb{Z}^+$*, if for every* 24*-colouring of* $\mathbb{Z}^+$ *there is a monochromatic solution to* $a_1x + a_2y + a_3z = 0$*, then for every finite colouring of* $\mathbb{Z}^+$ *there is a monochromatic solution to* $a_1x + a_2y + a_3z = 0$.

In Section 7.6, it is shown that there is a lack of characterizations for infinite partition regular matrices analogous to characterizations of finite partition regular matrices.

Question 9.3.4 (see [73, Question 4, Section 2]). *Which infinite systems of equations are partition regular?*

One of the major open problems in partition regularity is a proof that for every finite colouring of $\mathbb{Z}^+$, there is $x, y \in \mathbb{Z}^+$, not both 1, so that $\{x, y, x+y, xy\}$ is monochromatic. Graham showed using computer assistance that every 2-colouring of [252] contains a monochromatic set $\{x, y, x + y, xy\}$ (see [66]). The problem is still open for more than 2 colours.

Question 9.3.5 (see [73, Question 3, Section 1]). *If* $\mathbb{Z}^+$ *is finitely coloured, is there always* $x, y \in \mathbb{Z}^+$ *so that* $\{x, y, x + y, xy\}$ *is monochromatic?*

An analagous version of the question in finite fields has an affirmative answer, as proved by Green and Sanders:

Theorem 9.3.6 (Green and Sanders, 2016 [51]). *For any $r \in \mathbb{Z}^+$, there is $P \in \mathbb{Z}^+$ so that for every prime $p > P$ and any r-colouring of $\mathbb{F}_p$, there is $x, y \in \mathbb{F}_p$ so that $\{x, y, x+y, xy\} \subseteq \mathbb{F}_p$ is monochromatic.*

Recently, Moreira was able to show the problem has an affirmative answer if y is excluded;

Theorem 9.3.7 (Moreira, 2017 [93]). *For any $r \in \mathbb{Z}^+$ and any r-colouring of $\mathbb{Z}^+$, there is $x, y \in \mathbb{Z}^+$ so that $\{x, x+y, xy\}$ is monochromatic.*

To generalize the problem, for any $k \in \mathbb{Z}^+$ and any finite colouring of $\mathbb{Z}^+$, does there exist a sequence $x_1, x_2, \ldots, x_k$ so that all of its pairwise sums and products are monochromatic? This generalizes the problem in a similar way that the Folkman–Rado–Sanders theorem generalizes Schur's theorem. To take it a step further, is there always an infinite sequence for which all of its finite sums and products are monochromatic?

For an infinite set $A = \{a_1, a_2, a_3, \ldots\}$, let

$$FP(A) = \left\{ \prod_{i \in I} a_i : I \subseteq \mathbb{Z}^+, I \neq \emptyset, |I| < \infty \right\}.$$

In 1979, Hindman proved the following:

Theorem 9.3.8 (Hindman, 1979, [66]). *For every $r \in \mathbb{Z}^+$ and any r-colouring of $\mathbb{Z}^+$, there are infinite sets $A, B \subseteq \mathbb{Z}^+$ so that $FS(A) \cup FS(B)$ is monochromatic.*

However, in 1984, Hindman [68] showed that there is a 7-colouring of $\mathbb{Z}^+$ with no infinite set $A \subseteq \mathbb{Z}^+$ so that $FS(A) \cup FP(A)$ is monochromatic, proving that the analogous version of partition regularity of $\{x, y, x+y, xy\}$ in the infinite is false.

Bibliography

[*] **NOTE:** Numbers in boxes following each entry are page numbers in this thesis where the entry is cited.

[1] H. L. Abbott and A. C. Liu, On partitioning integers into progression free sets, *J. Combin. Theory Ser. A* **13** (1972), 432–436. 26

[2] H. L. Abbott and L. Moser, Sum-free sets of integers, *Acta Arith.* **10** (1966), 393–396. 20

[3] B. Barber, N. Hindman, and I. Leader, Partition regularity in the rationals, *J. Combin. Theory Ser. A* **120** (2013), 1590–1599. 133

[4] B. Barber, N. Hindman, I. Leader, and D. Strauss, Partition regularity without the columns property, *Proc. Amer. Math. Soc.* **143** (2015), 3387–3399. 110, 134, 135, 136

[5] L. D. Baumert and S. W. Golomb, Backtrack Programming, *J. Assoc. Comput. Mach.* **12** (1965), 516–524. 21

[6] J. Beck, *Combinatorial games: Tic-tac-toe theory*, Encyclopedia Math. Appl. **114**, Cambridge Univ. Press, Cambridge, 2008. 87, 89

[7] M. D. Beeler and P. E. O'Neil, Some new van der Waerden numbers, *Discrete Math.* **28** (1979), 135–146. 24

[8] V. Bergelson, Ergodic Ramsey theory, pp. 63–87 in *Logic and combinatorics* (Arcata, Calif., 1985), Contemp. Math. **65**, Amer. Math. Soc., Providence, RI, 1987. 150

[9] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. Amer. Math. Soc.* **9** (1996), 725–753. 139, 151

[10] E. R. Berlekamp, A construction for partitions which avoid long arithmetic progressions, *Canad. Math. Bull.* **11** (1968), 409–414. 26

[11] A. Beutelspacher and W. Brestovansky, Generalized Schur numbers, pp. 30–38 in *Combinatorial Theory*, Lecture Notes in Math. **969**, Springer-Verlag, Berlin, 1982. 41

[12] F. A. Behrend, On sets of integers which contain no three elements in arithmetic progression, *Proc. Nat. Acad. Sci. U. S. A.* **32** (1946), 331–332. 24

[13] T. F. Bloom, A quantitative improvement for Roth's theorem on arithmetic progressions, *J. Lond. Math. Soc. (2)* **93** (2016), 643–663. 154

[14] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.* **9** (1999), 968–984. 24, 154

[15] A. Brauer, Über Sequenzen von Potenzresten, *Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl.* (1928), 9–16. 23, 32, 33, 34

[16] T. C. Brown, P. Erdős, F. Chung, and R. Graham, Quantitative forms of a theorem of Hilbert, *J. Combin. Theory Ser. A* **38** (1985), 210–216. 17

[17] E. Čech, On bicompact spaces, *Ann. of Math. (2)* **38** (1937), 823–844. 120

[18] V. Chvátal, Some unknown van der Waerden numbers, pp. 31–33 in *Combinatorial Structures and Their Applications* (Proc. Calgary Internat. Conf., Calgary, Alta., 1969), Gordon & Breach, New York, 1970. 24, 25

[19] R. Cori and D. Lascar, *Mathematical logic. A course with exercises. Part II* (transl. of 1993 ed. by Pelletier), Oxford University Press, Oxford, 2001. 26

[20] W. Comfort, Ultrafilters: some old and new results, *Bull. Amer. Math. Soc.* **83** (1977), 417–455. 108

[21] W. Comfort and S. Negrepontis, *The theory of ultrafilters*, Springer-Verlag, Berlin, 1974. 109

[22] L. E. Dickson, Lower limit for the number of sets of solutions of $x^e + y^e + z^e \equiv 0 \pmod p$, *J. Reine Angnew. Math.* **135** (1909), 181–188. 15, 18

[23] W. Deuber, Partitionen und lineare Gleichungssyteme, *Math. Z.* **133** (1973), 109–123. 4, 84, 85, 95, 99, 101, 105, 106

[24] W. Deuber, D. S. Gunderson, N. Hindman, and D. Strauss, Independent finite sums for K_m-free graphs, *J. Combin. Theory Ser. A* **78** (1997), 171–198. 152

[25] W. Deuber and N. Hindman, Partitions and sums of (m, p, c)-sets, *J. Combin. Theory Ser. A* **45** (1987), 300–302. 110, 128, 129, 130, 132

[26] W. Deuber, N. Hindman, I. Leader, and H. Lefmann, Infinite partition regular matrices, *Combinatorica* **15** (1995), 333–355. 110, 134

[27] M. R. Dransfield, L. Liu, V. W. Marek, and M. Truszczyński, Satisfiability and computing van der Waerden numbers, *Electron. J. Combin.* **11** (2004), #R41. 25

[28] J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966. 111, 112, 113, 114

[29] R. Ellis, *Lectures on Topological Dynamics*, Benjamin, New York, 1969. 109, 121

[30] C. Elsholtz and D. S. Gunderson, Congruence properties of multiplicative functions on sumsets and monochromatic solutions of linear equations, *Funct. Approx. Comment. Math.* **52** (2015), 263–281. 42

[31] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus, Euclidean Ramsey theorems I, *J. Combin. Theory Ser. A* **14** (1973), 341–363. 63, 64

[32] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus, Euclidean Ramsey Theorems II, pp. 529–557 in *Infinite and finite sets I*

(Colloquia Keszethly, 1973; dedicated to P. Erdős on his 60th birthday), North Holland, Amsterdam, 1975. 63

[33] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus, Euclidean Ramsey Theorems III, pp. 559–583 in *Infinite and finite sets II* (Colloquia Keszethly, 1973; dedicated to P. Erdős on his 60th birthday), North Holland, Amsterdam, 1975. 63

[34] P. Erdős, A. Hajnal, and J. Pach, On a metric generalization of Ramsey's theorem, *Israel J. Math.* **101** (1997), 283–295. 152

[35] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. London Math. Soc. (3)* **2** (1952), 417–439. 25

[36] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470. 10

[37] G. Exoo, A lower bound for Schur numbers and multicolor Ramsey number of K_3, *Electron. J. Combin.* **1** (1994), #R8. 21

[38] J. Folkman, Notes on the Ramsey number $N(3,3,3,3)$, *J. Combin. Theory Ser. A* **16** (1974), 371–379. 23

[39] J. Fox and D. J. Kleitman, On Rado's boundedness conjecture, *J. Combin. Theory Ser. A* **113** (2006), 84–100. 157

[40] H. Fredricksen, Schur numbers and the Ramsey numbers $N(3, 3, \ldots, 3; 2)$, *J. Combin. Theory Ser. A* **27** (1979), 376–377. 21

[41] H. Fredricksen, Symmetric sum-free partitions and lower bounds for Schur numbers, *Electron. J. Combin.* **7** (2000), #R32. 21

[42] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* **31** (1977), 204–256. 5, 138, 139, 155

[43] H. Furstenberg and Y. Katznelson, A density version of the Hales–Jewett theorem, *J. Anal. Math.* **57** (1991), 64–119. 155

[44] W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, *Geom. Funct. Anal.* **8** (1998), 529–551. 154

[45] W. T. Gowers, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.* **11** (2001), 465–588. 26, 155

[46] R. L. Graham, Some of my favourite problems in Ramsey theory, *Integers* **7** (2007), #A15. 63

[47] R. L. Graham and B. L. Rothschild, A short proof of van der Waerden's theorem on arithmetic progressions, *Proc. Amer. Math. Soc.* **42** (1974), 385–386. 27, 29

[48] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, J. Wiley and Sons, 1980. 10, 18, 27, 34, 35, 94

[49] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, 2nd ed., J. Wiley and Sons, 1990. 10

[50] B. Green, On arithmetic structures in dense sets of integers, *Duke Math. J.* **114** (2002), 215–238. 140

[51] B. Green and T. Sanders, Monochromatic sums and products, *Discrete Anal.*, (2016), Paper No. 5, 43 pages. 158

[52] B. Green and T. Tao, New bounds for Szemerédi's theorem, III: A polylogarithmic bound for $r_4(n)$. Preprint, available on arXiv:1705.01703 (accessed 17 June 2017). 154

[53] D. S. Gunderson, On Deuber's partition theorem for (m, p, c)-sets, *Ars Combin.* **63** (2002), 15–31. 86

[54] D. S. Gunderson, "Ultrafilters and Ramsey theory", Chapter 22 of *Ramsey theory notes*, unpublished manuscript, 353 pages, 11 February 2017. 111

[55] D. S. Gunderson and K. Johannson, On combinatorial upper bounds for van der Waerden numbers W(3;r), pp. 33–46 in Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing, *Congr. Numer.* **190** (2008). 24

[56] D. S. Gunderson, I. Leader, H. J. Prömel, and V. Rödl, Independent arithmetic progressions in clique-free graphs on the natural numbers, *J. Combin. Theory Ser. A* **93** (2001), 1–17. 152, 153

[57] D. S. Gunderson, I. Leader, H. Prömel, and V. Rödl, Independent Deuber sets in graphs on the natural numbers, *J. Combin. Theory Ser. A* **103** (2003), 305–322. 84, 153

[58] D. S. Gunderson and V. Rödl, Extremal problems for affine cubes of integers, *Combin. Probab. Comput.* **7** (1998), 65–79. 17, 24, 37, 154

[59] S. Guo and Z-W. Sun, Determination of the two-color Rado number for $a_1x_1 + \cdots + a_mx_m = x_0$, *J. Combin. Theory Ser. A* **115** (2008), 345–353. 42

[60] A. Hales and R. Jewett, Regularity and positional games, *Trans. Amer. Math. Soc.* **106** (1963), 222–229. 5, 90

[61] H. Harborth and S. Maasberg, Rado numbers for $a(x+y) = bz$, *J. Combin. Theory Ser. A* **80** (1997), 356–363. 42

[62] H. Harborth and S. Maasberg, All two-color Rado numbers for $a(x+y) = bz$, pp. 397–407 in Proceedings of the 16th British Combinatorial Conference, (London, 1997), *Discrete Math.* **197/198** (1999). 42

[63] P. R. Herwig, M. J. Heule, P. M. van Lambalgen, and H. van Maaren, A new method to construct lower bounds for van der Waerden numbers, *Electron. J. Combin.* **14** (2007), #R6 [25]

[64] D. Hilbert, Über die Irreducibilität ganzer rationaler Funktionen mit ganzzahligen Coefizienten, *J. Reine Angew. Math.* **110** (1892), 104–129. [3, 15, 17]

[65] N. Hindman, Finite sums from sequences within cells of a partition of N, *J. Combin. Theory Ser. A* **17** (1974), 1–11. [5, 37, 108, 124]

[66] N. Hindman, Partitions and sums and products of integers, *Trans. Amer. Math. Soc.* **247** (1979), 227–245. [137, 157, 158]

[67] N. Hindman, Ultrafilters and Combinatorial Number Theory, pp. 119–184 in *Number Theory Carbondale 1979* (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), Lecture Notes in Math. **751**, Springer, Berlin, 1979. [126]

[68] N. Hindman, Partitions and pairwise sums and products, *J. Combin. Theory Ser. A* **37** (1984), 46–60. [158]

[69] N. Hindman, Algebra in βS and its applications to Ramsey theory, *Math. Japon.* **44** (1996), 581–625. [108, 111]

[70] N. Hindman, Image partition regularity over the reals, *New York J. Math.* **9** (2003), 79–91. [68, 70, 75, 79, 82, 86]

[71] N. Hindman, Partition regularity of matrices, pp. 265–298 in *Combinatorial number theory*, de Gruyter, Berlin, 2007. 61

[72] N. Hindman and I. Leader, Image partition regularity of matrices, *Combin. Probab. Comput.* **2** (1993), 437–463. 4, 67, 68, 70, 75, 79, 82, 86

[73] N. Hindman, I. Leader, and D. Strauss, Open problems in partition regularity, *Combin. Probab. Comput.* **12** (2003), 571–585. 156, 157

[74] N. Hindman, H. Lefmann, Partition regularity of (M, P, C)-systems, *J. Combin. Theory Ser. A* **64** (1993), 1–9. 132, 134

[75] N. Hindman and D. Strauss, Image partition regularity over the integers, rationals and reals, *New York J. Math.* **11** (2005), 519–538. 68, 69, 74, 78, 79, 82, 86, 134

[76] N. Hindman and D. Strauss, *Algebra in the Stone–Čech compactification: theory and applications*, 2nd ed., de Gruyter, Berlin, 2012. 5, 108, 111

[77] N. Hindman and E. Tessler, The first nontrivial Hales–Jewett number is four, *Ars Combin.* **113** (2014), 385–390. 89

[78] Y. R. Huang and J. S. Yang, New upper bounds for van der Waerden numbers W(3,n), (in Chinese), *Chin. Ann. Math. Ser. A* **21** (2000), 631–634. 25

[79] R. W. Irving, An extension of Schur's theorem on sum-free partitions, *Acta Arith.* **25** (1973/74), 55–64. 22

[80] K. Johannson, *Variations on a theorem by van der Waerden*, Master's thesis, University of Manitoba, 2007. 24, 27, 111

[81] D. König, Über eine Schußweise aus dem Endlichen ins Unendliche, *Acta Litt. Acad. Sci. Hungar.* **3** (1927), 121–130. 11

[82] M. Kouril, Computing the van der Waerden number W(3,4) = 293, *Integers* **12** (2012), #A46. 25

[83] M. Kouril and J. Franco, Resolution tunnels for improved SAT solver performance, pp. 143–157 in *Theory and applications of satisfiability testing*, Lecture Notes in Comput. Sci. **3569**, Springer, Berlin, 2005. 25

[84] B. M. Landman and A. Robertson, *Ramsey theory on the integers*, American Mathematical Society, Providence RI, 2004. 21, 43

[85] K. Leeb, A full Ramsey-theorem for the Deuber-category, pp. 1043–1049 in *Infinite and finite sets II* (Colloquia Keszethly, 1973; dedicated to P. Erdős on his 60th birthday), North Holland, Amsterdam (1975), 1043–1049. 85

[86] H. Lefmann, On partition regular systems of equations, *J. Combin. Theory Ser. A* **58** (1991), 35–53. 138

[87] H. Lefmann and H. S. Kaddour, On a maximality property of partition regular systems of equations, *Canad. Math. Bull.* **36** (1993), 96–102. 134

[88] M. Lewko, An improved lower bound related to the Furstenberg–Sárközy theorem, *Electron. J. Combin* **22** (2015) #P1.32. 150

[89] L. Lovász, *Combinatorial problems and exercises*, 2nd ed., North-Holland Publishing Co., Amsterdam, 1993. 16

[90] T. Łuczak, V. Rödl, and T. Schoen, Independent finite sums in graphs defined on the natural numbers, *Discrete Math.* **181** (1998), 289–294. 152

[91] N. Lyall, A new proof of Sárközy's theorem, *Proc. Amer. Math. Soc.* **141** (2013), 2253–2264. 140

[92] K. Milliken, Ramsey's Theorem with sums or unions, *J. Combin. Theory Ser. A* **18** (1975), 276–290. 135

[93] J. Moreira, Monochromatic sums and products in $\mathbb{N}$, *Ann. of Math. (2)* **185** (2017), 1069–1090. 158

[94] L. Moser, Notes on number theory. II. On a theorem of van der Waerden, *Canad. Math. Bull.* **3** (1960), 23–25. 26

[95] J. Pintz, W.L. Steiger, and E. Szemerédi, On sets of natural numbers whose difference sets contain no squares, *J. London Math. Soc. (2)* **37** (1988), 219–231. 140, 150

[96] D. H. J. Polymath, A new proof of the density Hales–Jewett theorem, *Ann. of Math. (2)* **175** (2012), 1283–1327. 155

[97] H. Prömel, *Ramsey Theory for Discrete Structures*, Springer, 2013. 15, 17, 18, 23, 27, 32, 34, 48, 55, 90

[98] J. Rabung and M. Lotts, Improving the use of cyclic zippers in finding lower bounds for van der Waerden numbers, *Electron. J. Combin.* **19** (2012), #P35. 25

[99] R. Rado, Studien zur Kombinatorik, *Math. Z.* **36** (1933), 424–470. 34, 35, 40, 41, 54, 55, 56, 61, 156

[100] R. Rado, Note on combinatorial analysis, *Proc. Lond. Math. Soc. (2)* **48** (1939), 122–160. 40, 45, 52, 53, 58, 60, 61, 62

[101] F. P. Ramsey, On a problem of formal logic. *Proc. London Math. Soc.*, **30** (1930), 735–744. 2, 8, 12

[102] K. F. Roth, Sur quelques ensembles d'entiers, *C. R. Acad. Sci. Paris* **234** (1952), 388–390. 24, 139, 154

[103] K. F. Roth, On certain sets of integers, *J. Lond. Math. Soc.* **28** (1953), 104–109. 24, 139, 154

[104] I. Z. Ruzsa, Difference sets without squares, *Period. Math. Hungar.* **15** (1984), 205–209. 150

[105] J. Sanders, *A generalization of Schur's theorem*, dissertation, Yale University, 1968. 35

[106] T. Sanders, On certain other sets of integers, *J. Anal. Math.* **116** (2012), 53–82. 154

[107] C. Sándor, An upper bound for Hilbert cubes, *J. Combin. Theory Ser. A* **114** (2007), 1157–1159. 154

[108] D. Saracino, The 2-color Rado number of $x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_k$, *Ars. Combinatoria* **129** (2016), 315–321. 42

[109] A. Sarközy, On difference sets of sequences of integers I, *Acta. Math. Acad. Sci. Hungar.* **31** (1978), 125–149. 5, 138, 139

[110] D. Schaal and D. Vestal, Rado numbers for $x_1 + x_2 + \cdots + x_{m-1} = 2x_m$, pp. 105–116 in Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing, *Congr. Numer.* **191** (2008). 42

[111] W. M. Schmidt, Two combinatorial theorems on arithmetic progressions, *Duke Math. J.* **29** (1962), 129–140. 26

[112] I. Schur, Über die Kongruenz $x^m + y^m = z^m \pmod{p}$, *Jahresber. Dtsch. Math.-Ver.* **25** (1916), 114–117. 2, 17, 18, 21

[113] Y. Setyawan, *Combinatorial number theory: results of Hilbert, Schur, Folkman and Hindman*, Master's thesis, Simon Fraser University, 1998. 17

[114] S. Shelah, Primitive recursive bounds for van der Waerden numbers, *J. Amer. Math. Soc.* **1** (1988), 683–697. 26, 94

[115] R. S. Stevens, R. Shantaram, Computer-generated van der Waerden partitions, *Math. Comp.* **32** (1978), 635–636. 25

[116] M. Stone, Applications of the theory of Boolean rings to general topology, *Trans. Amer. Math. Soc.* **41** (1937), 375–481. 120

[117] Z. Szabó, An application of Lovász's local lemma — a new lower bound for the van der Waerden number, *Random Structures Algorithms* **1** (1990), 343–360. 26

[118] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression, *Acta Math. Acad. Sci. Hungar.* **20** (1969), 89–104. 154

[119] T. Tao, A Fourier-free proof of the Furstenberg-Sarkozy theorem, available at `http://terrytao.wordpress.com/2013/02/28/a-fourier-free-proof-of-the-furstenberg-sarkozy-theorem`, accessed 24 April 2017. 139, 140, 150

[120] A. D. Taylor, A canonical partition relation for finite subsets of ω, *J. Combin. Theory Ser. A* **21** (1976), 137–146. 135

[121] A. D. Taylor, Bounds for the disjoint unions theorem, *J. Combin. Theory Ser. A* **30** (1981), 339–344. 37

[122] S. Todorcevic, *Topics in topology*, Lecture notes in mathematics **1652**, Springer, 1997. 108

[123] B. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wiskd.* **19** (1927), 212–216. 3, 24

[124] M. Walters, Combinatorial proofs of the polynomial van der Waerden Theorem and the polynomial Hales–Jewett Theorem, *J. London Math. Soc. (2),* **61** (2000), 1–12. 151

[125] E. G. Whitehead Jr., The Ramsey number $N(3, 3, 3, 3; 2)$, *Discrete Math.* **4** (1973), 389–396. 23