Variations on a theorem by van der Waerden

by

Karen R. Johannson

A thesis submitted to

the Faculty of Graduate Studies

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Mathematics

University of Manitoba

Winnipeg, Manitoba

Abstract

The central result presented in this thesis is van der Waerden's theorem on arithmetic progressions. Van der Waerden's theorem guarantees that for any integers k and r, there is an n so that however the set $\{1, 2, \ldots, n\}$ is split into r disjoint partition classes, at least one partition class will contain a k-term arithmetic progression. Presented here are a number of variations and generalizations of van der Waerden's theorem that utilize a wide range of techniques from areas of mathematics including combinatorics, number theory, algebra, and topology.

I would like to thank my advisor, Dr. David Gunderson, for his guidance, encouragement and the boundless energy and enthusiasm that he devoted to my research. I am greatly appreciative of the tremendous amount of time and effort he spent helping me throughout the process of preparing this thesis.

I would also like to express my thanks to my advisory committee, Dr. R. Craigen, Dr. R. Padmanabhan and Dr. B. Landman as well as Dr. B. Li.

I am grateful for support received from the Natural Sciences and Engineering Research Council of Canada in the form of a Postgraduate Scholarship.

Finally, I am eternally indebted to my family for their unending love and support.

Contents

List of Figures

Chapter 1

Preface

1.1 Introduction

A notion central to this thesis is that of an arithmetic progression. An arithmetic progression is a sequence of at least three numbers of the form $a, a+d, a+2d, \ldots$. For example, $2,\ 5,\ 8,\ 11,\ 14$ is a 5-term arithmetic progression where $a = 2$ and the difference between consecutive terms is $d = 3$. A recurring theme throughout this thesis is the question of whether it is possible to split a particular set of integers into two groups, neither containing an arithmetic progression. Van der Waerden's theorem states that no matter how the set of positive integers is split into finitely many disjoint groups, for any integer k, at least one group will contain a k-term arithmetic progression. Moreover, something similar holds for some finite sets of integers. Consider the possible divisions of the numbers $1, 2, \ldots, 9$ into two groups, say a red group and a blue group. Is it possible to prevent arithmetic progressions in either group? While attempting to avoid arithmetic progressions that are all red or all blue, examine the possible groupings of the middle three numbers: 4, 5, and 6. These three numbers form an arithmetic progression (with difference $d = 1$) and if the numbers 4, 5 and 6 are not all in the same group, there must be one in one group and two in the other. Assuming for the moment that 4 is contained in the red group, the possibilities are:

$$\underset{\text{R}}{4}\,\underset{\text{B}}{5}\,\underset{\text{R}}{6}, \quad \underset{\text{R}}{4}\,\underset{\text{B}}{5}\,\underset{\text{B}}{6}, \quad \text{and} \quad \underset{\text{R}}{4}\,\underset{\text{R}}{5}\,\underset{\text{B}}{6}.$$

In the case where $4, 5, 6$ are grouped RBR, in order to avoid the progression $2, 4, 6$ the number 2 must be in the blue group. Then if 8 is in the red group, so is the progression $4, 6, 8$ but if 8 is in the blue group so is $2, 5, 8$.

In the case where $4, 5, 6$ are grouped RBB, in order to avoid the progression $5, 6, 7$ the number 7 ought to be in the red group but then in order to avoid $1, 4, 7$, the number 1 should be in the blue group. Because of the progression $1, 3, 5$ the number 3 must be in the red group but then to avoid $2, 3, 4$, the number 2 is in the blue group. Finally, since $2, 5, 8$ is an arithmetic progression, 8 is in the red group. The grouping is then as follows:

$$\underset{\text{B}}{1}\,\underset{\text{B}}{2}\,\underset{\text{R}}{3}\,\underset{\text{R}}{4}\,\underset{\text{B}}{5}\,\underset{\text{B}}{6}\,\underset{\text{R}}{7}\,\underset{\text{R}}{8}\,9.$$

Now, if 9 is in the red group, then so is the arithmetic progression $7, 8, 9$ and if 9 is in the blue group, so is $1, 5, 9$.

Finally, in the case where $4, 5, 6$ are grouped RRB, by symmetry with the case where $4, 5, 6$ are grouped RBB, in order to avoid progressions in either group, the numbers must be divided as follows:

$$1\ \underset{\text{B}}{2}\ \underset{\text{B}}{3}\ \underset{\text{R}}{4}\ \underset{\text{R}}{5}\ \underset{\text{B}}{6}\ \underset{\text{B}}{7}\ \underset{\text{R}}{8}\ \underset{\text{R}}{9}.$$

Then, if 1 is in the red group, so is the progression $1, 5, 9$ and if 1 is in the blue group, so is $1, 2, 3$. Repeating this argument while assuming that the number 4 is in the blue group shows that no matter how the numbers from 1 to 9 are split into two groups, one cannot avoid an arithmetic progression with at least three terms in one of the groups. What if we would like to guarantee a 4-term arithmetic progression? Though more difficult to verify, it turns out that no matter how the integers from 1 to 35 are split into two groups, one group will contain a 4-term arithmetic progression. This can be generalized to arithmetic progressions with any finite number of terms. Van der Waerden's theorem states that for any $k \geq 3$ and $r \geq 2$, there is an n so that if the integers $1, 2, \ldots, n$ are split into r disjoint groups, at least one of the groups will contain an arithmetic progression with k terms.

Van der Waerden's theorem is one of what are called “Ramsey-type” results. A grouping of the elements of a set defines a “partition” of the set, where each group is called a *partition class*. Ramsey theory is the study of structures preserved under partition. Frank Plumpton Ramsey proved a theorem known as Ramsey's theorem, which is the central result in Ramsey theory. Ramsey's theorem can be illustrated with the following, often referred to as the "Party Problem": How many people must one invite to a party to be assured that among the guests there are 3 mutual acquaintances or 3 mutual strangers (or both). The answer is that 6 guests suffice. A simple form of Ramsey's theorem states that for any $k \geq 3$, there is an n so that among any collection of n people, there are either k mutual acquaintances or k mutual strangers.

A number of quantitative questions raised by van der Waerden's theorem are studied in Chapter 5. What is the smallest n so that whenever the numbers $1, 2, \ldots, n$ are partitioned into two classes, one contains a k-term arithmetic progression? As k increases, these numbers become difficult to calculate and known proofs of van der Waerden's theorem produce large upper bounds for n that grow incredibly quickly. For example, the argument from the original proof is only able to guarantee that when the integers from 1 to 325 are partitioned into two classes, one will contain a 3-term arithmetic progression — a far cry from the first example where the numbers from 1 to 9 suffice. There are a number of results giving both lower bounds and more reasonable upper bounds for these numbers. In general, however, the known upper and lower bounds remain of different orders of magnitude.

Returning to the numbers from 1 to 9, instead of dividing them into two groups, suppose that one were to select a single group that contained no arithmetic progressions. How large could that group be? For example, the 5-element group $1, 2, 6, 8, 9$ contains no arithmetic progressions and it can be shown that any sequence of 6 numbers, such as $1, 2, 5, 6, 8, 9$, will contain an arithmetic progression (in this case $1, 5, 9$). This idea can be generalized to determining the maximum number of integers that can be chosen from among $1, 2, \ldots, n$ while avoiding arithmetic progressions with k terms. Szemerédi's theorem states that for any $k \geq 3$ and ε with $0 < \varepsilon < 1$, there is an n so that any collection of more than εn numbers from $1, 2, \ldots, n$ will contain a k-term arithmetic progression. Such a result is called a "density result" as opposed to a Ramsey result. In Chapter 4, some density results related to arithmetic progressions are examined.

It is also possible to examine arithmetic progressions in the context of some larger structure that either contains the integers or provides a generalization of arithmetic progressions. One such generalization, given in Chapter 7, involves games of multi-dimensional Tic-Tac-Toe. In the usual game of Tic-Tac-Toe on a 3×3 grid, there are possible stalemates. The Hales-Jewett theorem guarantees that, for each $k \geq 3$, there is an n so that no matter how X's and O's are distributed in an $(k \times k \times \cdots \times k)$ grid of dimension n, there will be either k X's or k O's in a line (in fact in a particular kind of line called a "combinatorial line"). Coordinatizing the grid with the numbers $1, 2, \ldots, k$, each box in the grid can be associated with an integer either by adding all of its coordinates or regarding the coordinates as a base expansion. In this way, the "combinatorial lines" in the grid can be made to correspond to arithmetic progressions. For example, on the usual 3×3 grid, the line in Figure 1.1 could correspond to the arithmetic progression $11, 22, 33$ (base 10 expansion) or to the arithmetic progression $2, 4, 6$ (adding the coordinates). The Hales-Jewett theorem is a purely combinatorial generalization of van der Waerden's theorem and provides another combinatorial proof of van der Waerden's theorem.

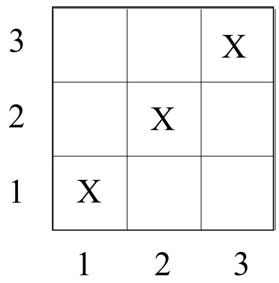

Figure 1.1: Tic-Tac-Toe

Another way to prove van der Waerden's theorem is use structures from set theory and topology called *ultrafilters*. Ultrafilters are a type of "large" set system. In Chapter 8, it is seen how the ultrafilters on the integers can be endowed with a semigroup operation and a topology that are used to prove some Ramsey-type results including van der Waerden's theorem.

A number of other variations of van der Waerden's theorem are presented in Chapter 9. Among these is a polynomial version of van der Waerden's theorem that guarantees that for any partition of the positive integers into two partition classes, at least one partition class will contain an arithmetic progression whose difference is a perfect square, a perfect cube, or in the range of any polynomial with integer coefficients and no constant term. Another variation of van der Waerden's theorem is a "Rainbow Ramsey-type" problem where, if the integers are divided into a certain number of groups, the desired sets are not arithmetic progressions with all terms in one group, but rather those where each term is in a different group from all others. There are a number of other results closely related to van der Waerden's theorem, but beyond the scope of this thesis. Among them is Szemerédi's Regularity Lemma (see for example [77]). The Regularity Lemma (or Uniformity Lemma) is a result in graph theory used by Szemerédi to prove his density theorem for arithmetic progressions and has found a number of applications beyond its original usage.

The results and proofs that are included in this thesis exhibit some of the wide-ranging techniques that have been used to attack some of the variations of van der Waerden's theorem. While some proofs are purely combinatorial or number theoretic, others apply results from areas such as algebra, field theory and topology. Ramsey theory is an appealing subject area for its ability to predict order within chaos and the results related to van der Waerden's theorem illustrate, what I found to be, surprising connections between exceedingly different areas of mathematics.

1.2 Notation

The following standard notation and definitions are used throughout. The integers $\{0, \pm 1, \pm 2, \ldots\}$ are denoted by $\mathbb{Z}$ and the positive integers $\{1, 2, 3, \ldots\}$ are denoted by $\mathbb{Z}^+$.

For any $a, b \in \mathbb{Z}$, set $[a, b] = \{m \in \mathbb{Z} : a \leq m \leq b\}$. In particular, for any $n \in \mathbb{Z}^+$, $[1, n] = \{1, 2, \ldots, n\}$. The notation $[n] = [1, n]$ is also common but avoided here whenever possible. For any sets $S, T \subseteq \mathbb{Z}$ and $x \in \mathbb{Z}$, define

$$S + x = \{s + x : s \in S\} \text{ and } S + T = \{s + t : s \in S \text{ and } t \in T\}.$$

For any $k \in \mathbb{Z}^+$, define $[S]^k = \{X \subseteq S : |X| = k\}$, the collection of *$k$-subsets of S*, and for any $n \in \mathbb{Z}^+$, define $[n]^k = [[1, n]]^k$.

For functions $f, g : \mathbb{Z}^+ \to \mathbb{R}$, if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$, then write $f = o(g)$ and if there is a constant $c > 0$ and $n_0 \in \mathbb{Z}^+$ such that for every $n \geq n_0$, $f(n) \leq cg(n)$, then write $f = O(g)$.

For any set A and a finite set B, any function $\Delta : A \to B$ is called a *finite colouring* of A. In particular, if $|B| = r$, then any function $\Delta : A \to B$ is called an *r-colouring* of A. Any colouring $\Delta : A \to B$ corresponds to a partition of A into sets $\{\Delta^{-1}(b) : b \in B\}$ called *colour classes* and so the terms "colourings" and "partitions" are used interchangeably. The pigeonhole principle says that if $r + 1$ pigeons are placed in r holes, some hole will contain at least two pigeons and more generally, if $nr+1$ pigeons are placed into r holes, some hole will contain at least $n+1$ pigeons. The pigeonhole principle can be stated in terms of colourings.

Pigeonhole Principle Let $n, r \in \mathbb{Z}^+$. For every r-colouring $\Delta : [1, nr+1] \to [1, r]$, there exists an $i \in [1, r]$ so that $|\Delta^{-1}(i)| \geq n+1$.

Definition 1.2.1. A set P of integers is an *arithmetic progression of length k* (denoted by AP_k) iff there are integers a and $d > 0$ so that

$$P = \{a + id : 0 \leq i \leq k-1\}.$$

For any $k \geq 3$, a set of integers A is called *AP_k-free* if A contains no AP_k's. The following type of set is a generalization of an arithmetic progression.

Definition 1.2.2. A set $P \subseteq \mathbb{Z}^+$ is called an *m-fold arithmetic progression of length k* if there are integers $a, d_1, \ldots, d_m$ such that

$$P = \{a + \sum_{i=1}^{m} x_i d_i : x_1, x_2, \ldots x_m \in [0, k-1]\}.$$

Note that a 1-fold arithmetic progression of length k is just an AP_k.

Many proofs throughout require bounds on the number of AP_k's in an interval. The following basic calculation is referred to in numerous proofs to come.

Lemma 1.2.3. For any $n \in \mathbb{Z}^+$ and $k > 1$, the number of AP_k's in $[1, n]$ is less than $\frac{n^2}{2(k-1)}$.

Proof. For any integer d with $1 \leq d \leq \frac{n-1}{k-1}$, the arithmetic progressions in $[1, n]$ with difference d are $\{1, 1+d, \ldots, 1+(k-1)d\}, \{2, 2+d, \ldots, 2+(k-1)d\}, \ldots, \{n-(k-1)d, n-(k-2)d, \ldots, n\}$ and so the number of AP_k's with difference d is $n-(k-1)d$. Therefore, the total number of AP_k's contained in $[1, n]$ is

$$\begin{aligned}
\sum_{d=1}^{\lfloor \frac{n-1}{k-1} \rfloor} (n-(k-1)d) &= \left\lfloor \frac{n-1}{k-1} \right\rfloor n - (k-1)\binom{\lfloor \frac{n-1}{k-1} \rfloor + 1}{2} \\
&= \left\lfloor \frac{n-1}{k-1} \right\rfloor \frac{1}{2} \left(2n - (k-1)\left(\left\lfloor \frac{n-1}{k-1} \right\rfloor + 1 \right) \right) \\
&\leq \frac{n-1}{2(k-1)} (2n - (n-1)) \\
&= \frac{(n-1)(n+1)}{2(k-1)} \\
&< \frac{n^2}{2(k-1)}. \qquad \square
\end{aligned}$$

Chapter 2

Classic Ramsey theorems

2.1 Ramsey's theorem

Van der Waerden's theorem on arithmetic progressions has come to be classified as a "Ramsey-type" theorem. Ramsey proved what has become known as Ramsey's theorem [91] in 1930 as a tool to prove a result in formal logic on propositional sentences.

Theorem 2.1.1 (**Ramsey's theorem**, Ramsey [91])**.** For every $k, m, r \in \mathbb{Z}^+$ there is an integer n so that for any set S with $|S| = n$ and any r-colouring $\Delta : [S]^k \to [1, r]$ there is a set $T \in [S]^m$ such that $[T]^k$ is monochromatic.

A simple case of Ramsey's theorem is often stated as a result in graph theory. A detailed exposition of Ramsey theory and graph theory can be found, for example, in Bollobás's *Modern graph theory* [16, Chapter 6].

Definition 2.1.2. Given a non-empty set V and $E \subseteq [V]^2$, the pair $G = (V, E)$ is called a *graph*. The elements of the set $V = V(G)$ are called the *vertices of* G and the elements of $E = E(G)$ are called the *edges of* G. For each $v \in V$, the set $N(v) = \{x \in V : \{x, v\} \in E\}$ is called the *neighbourhood of* v.

Any graph $G = (V, E)$ with $E = [V]^2$ is called a *complete graph*. For each $n \in \mathbb{Z}^+$, the complete graph on n vertices is denoted by K_n.

Definition 2.1.3. Given graphs $G = (V, E)$ and $G' = (V', E')$, the graph G' is called a *subgraph* (or sometimes a *weak subgraph*) of G iff $V' \subseteq V$ and $E' \subseteq E \cap [V']^2$.

A slightly simpler form of Ramsey's theorem (the case $m = 2$) can be stated in terms of graph theory. A proof of Theorem 2.1.4 is given in this section.

Theorem 2.1.4 (Ramsey [91]). For every $k, r \in \mathbb{Z}^+$ there is a least integer $R(k; r)$ such that for all $n \geq R(k; r)$ and for any r-colouring of the edges of K_n, there is a complete subgraph G on k vertices such that the set of edges $E(G)$ is monochromatic.

Definition 2.1.5. Let $r, k_1, k_2, \ldots, k_r \in \mathbb{Z}^+$. Denote by $R(k_1, k_2, \ldots, k_r; r)$ the least integer N, if it exists, such that for every $n \geq N$ and any r-colouring of the edges of K_n, there is an $i \in [1, r]$ such that K_n contains a subgraph K_{k_i} whose edges are monochromatic in colour i.

Often, the notation $R(k_1, \ldots, k_r; r)$ is abbreviated by $R(k_1, \ldots, k_r)$ when the number of colours is implicit. If for all $r, k_1, \ldots, k_r \in \mathbb{Z}^+$, $R(k_1, \ldots, k_r; r)$ exists, then since $R(k; r) = R(k, \ldots, k; r)$, so does the Ramsey number, $R(k; r)$. Before proceeding to the proof that the number $R(k_1, \ldots, k_r; r)$ will always exist, a few preliminary results will be useful.

Lemma 2.1.6. Let $k_1, \ldots, k_r \in \mathbb{Z}^+$ be such that $R(k_1, \ldots, k_r)$ exists. For any permutation σ of $[1, r]$, $R(k_{\sigma(1)}, \ldots, k_{\sigma(r)}) = R(k_1, \ldots, k_r)$.

Proof. Set $N = R(k_1, \ldots, k_r)$, let σ be a permutation of $[1, r]$ and let Δ be any r-colouring of $E(K_N)$. Define a new r-colouring Δ' of the edges of K_N as follows. For $\{x, y\} \in E(K_N)$, set $\Delta'(\{x, y\}) = \sigma(\Delta(\{x, y\}))$. By the choice of N, for some $i \in [1, r]$, there is a K_{k_i} which is monochromatic, under Δ', in the colour i. That is, for all $x, y \in V(K_{k_i})$, $\sigma(\Delta(\{x, y\})) = i$. Let $j = \sigma^{-1}(i)$. Then $K_{k_i} = K_{k_{\sigma(j)}}$ is a complete graph, monochromatic, under Δ, of colour j. □

In 1935, Erdős and Szekeres [40] gave a new proof of Ramsey's theorem which they used to prove that for every integer $k \geq 4$, there is an $N \in \mathbb{Z}^+$ so that for any collection of N points in the plane with no three collinear, some k points form the vertices of a convex k-gon. Their inductive proof employed a recursion that has come to be known as the "Erdős-Szekeres recursion“. The recursive equation is often given in its simplest form which states that if $k, \ell \geq 3$, then $R(k, \ell) \leq$

$R(k-1,\ell)+R(k,\ell-1)$. The following theorem (see for example [51, p.5]) extends the recursion to Ramsey numbers with any number of colours.

Theorem 2.1.7 (Erdős-Szekeres Recursion [40])**.** For all integers $r \geq 2$ and $k_1, k_2, \ldots, k_r \geq 3$,

$$\begin{aligned} R(k_1, k_2, \ldots, k_r) \leq R(k_1-1, k_2, \ldots, k_r) + R(k_1, k_2-1, \ldots, k_r) \\ + \ldots + R(k_1, k_2, \ldots, k_r-1) - r + 2. \end{aligned}$$

Proof. For each $i \in [1,r]$ define $N_i = R(k_1, \ldots, k_i - 1, \ldots k_r)$ and let $N = N_1 + \ldots + N_r - r + 2$. Let Δ be any r-colouring of the edges of K_N.

Fix an $x \in V(K_N)$ and for each $i \in [1,r]$, define

$$V_i = \{y \in V(K_N) : \Delta(\{x,y\}) = i\}.$$

Then, V_i is the set all vertices in the neighbourhood of x whose edges to x are coloured i by Δ. Then for some $i \in [1,r]$, $|V_i| \geq N_i$, for if not, then,

$$\sum_{i=1}^{r} N_i - r + 2 = 1 + \sum_{i=1}^{r} |V_i| \leq 1 + \sum_{i=1}^{r} (N_i - 1) = \left(\sum_{i=1}^{r} N_i\right) - r + 1$$

which is impossible.

Fix $\ell \in [1,r]$ such that $|V_\ell| \geq N_\ell$. The colouring Δ induces an r-colouring of the complete graph on the vertices of V_ℓ and so by the choice of the number $N_\ell = R(k_1, \ldots, k_\ell - 1, \ldots, k_r)$, either for some $j \in [1,r]\backslash\{\ell\}$, V_ℓ contains a K_{k_j} with edges all colour j, or else V_ℓ contains a $K_{k_\ell - 1}$ with edges all colour ℓ. In the latter case, the vertices of the complete graph $K_{k_\ell - 1}$ together with the vertex x form the vertices of a complete graph on k_ℓ vertices all of whose edges are colour ℓ since all the edges between x and V_ℓ are colour ℓ. □

The Erdős-Szekeres recursion provides a step of a proof by double induction on r and $k_1 + \cdots + k_r$ that for all $r, k_1, \ldots, k_r \in \mathbb{Z}^+$, the Ramsey number $R(k_1, \ldots, k_r; r)$ exists. The necessary base cases are as follows.

For $r = 2$ and $k \in \mathbb{Z}^+$, $R(k, 2; 2) = k$ since for any 2-colouring of the edges of the complete graph K_k, there is either be one edge, a K_2, of the second colour or else all edges of the graph K_k are of the first colour. Similarly, if $k_1, \ldots, k_{r-1} \in \mathbb{Z}^+$, then $R(k_1, \ldots, k_{r-1}, 2; r) = R(k_1, \ldots, k_{r-1}; r - 1)$.

Together, Lemma 2.1.6 and Theorem 2.1.7 show that for any $r, k_1, \ldots, k_r \in \mathbb{Z}^+$, the Ramsey number $R(k_1, \ldots, k_r; r)$ exists.

The Erdős-Szekeres recursion can be used to prove that for every $k, \ell \geq 2$, $R(k, \ell; 2) \leq \binom{k+\ell-2}{k-1}$.

2.2 Hilbert's cube lemma

Definition 2.2.1. For any $m \in \mathbb{Z}^+$, $H \subseteq \mathbb{Z}^+$ is an *affine m-cube* iff there are $a_0, a_1, \ldots a_m \in \mathbb{Z}^+$ such that,

$$H = \left\{ a_0 + \sum_{i \in I} a_i : I \subseteq [1, m] \right\}.$$

Denote such an affine m-cube by $H = H(a_0, a_1, \ldots, a_m)$.

For example, the set $H(1, 2, 3) = \{1, 3, 4, 6\}$ is an affine 2-cube and the set $H(2, 3, 5, 8) = \{2, 5, 7, 10, 13, 15, 18\}$ is an affine 3-cube. For any $a_0, a_1, \ldots, a_m \in \mathbb{Z}^+$, the affine m-cube $H(a_0, a_1, \ldots, a_m)$ can be decomposed as $\{a_0\} + \{0, a_1\} + \ldots + \{0, a_m\}$. Affine m-cubes are related to arithmetic progressions since any arithmetic progression of length $m + 1$ is an affine m-cube. In fact, for any $a, d, m \in \mathbb{Z}^+$,

$$\{a, a + d, \ldots, a + md\} = H(a, \underbrace{d, \ldots, d}_{m-\text{times}}).$$

Also, any m-fold arithmetic progression of length 2 (recall Definition 1.2.2) is an affine m-cube. Even though not all affine cubes are arithmetic progressions, affine cubes are a key component of the proof of Szemerédi's theorem (in Chapter 4 to come) regarding arithmetic progressions. Long arithmetic progressions in a particular set are found by examining shorter arithmetic progressions contained within particular affine cubes. The following colouring theorem about affine m-cubes was published by David Hilbert in 1892. After the pigeonhole principle, Hilbert's result on affine cubes is the earliest non-trivial "Ramsey"-type result.

Theorem 2.2.2 (Hilbert [63]). For every $r, m \in \mathbb{Z}^+$ there exists a positive integer $n = H(m, r)$ such that for every r-colouring of $[1, n]$, there is a monochromatic affine m-cube.

Proof. The proof given here appears in Lovaśz [81, p. 561]. Note that the translation of an affine m-cube remains an affine m-cube. That is, for any $a_0, a_1, \ldots, a_m, t \in \mathbb{Z}^+$,

$$H(a_0, a_1, \ldots a_m) + t = H(a_0 + t, a_1, \ldots, a_m).$$

Thus, if for some $m, r \in \mathbb{Z}^+$, n is such that for any r-colouring of $[1, n]$, there is a monochromatic affine m-cube, then the same is true for any interval of length n.

Fix $r \in \mathbb{Z}^+$. The proof that for all $m \in \mathbb{Z}^+$, $H(m; r)$ exists proceeds by induction on m.

Base Case: If $m = 1$, then by the pigeonhole principle, $H(1, r) = r + 1$ since an affine 1-cube is just a pair of integers.

Inductive Step: Assume that for some $m \geq 1$, the number $H(m, r)$ exists. Set $n = H(m, r)$, $N = r^n + n$, and let $\Delta : [1, N] \to [1, r]$ be any r-colouring. For each $i \in [1, r^n + 1]$ define the interval $I_i = [i, i - 1 + n]$.

In each interval I_i, there are n elements that receive one of r colours and so each of these intervals could be coloured in r^n different ways. Since there are $r^n + 1$ intervals, by the pigeonhole principle, there must be two, say I_j and I_{j+k}, that are coloured in the same way. That is, for any $\ell \in I_j$, $\Delta(\ell) = \Delta(\ell + k)$.

By the choice of $n = H(m, r)$, since I_j is an interval of length n, there is a monochromatic affine m-cube $H(a_0, a_1, \ldots, a_m)$ in I_j.

By the choice of k, for any $h \in H(a_0, a_1, \ldots, a_m)$, $\Delta(h) = \Delta(h + k)$. Therefore, the set

$$H(a_0, a_1, \ldots, a_m) \cup (H(a_0, a_1, \ldots, a_m) + k) = H(a_0, a_1, \ldots, a_m, k)$$

is an affine $(m+1)$-cube in $[1, N]$ which is monochromatic under Δ.

Therefore, for any $m, r \in \mathbb{Z}^+$, the number $H(m, r)$ exists. □

The following extension of Theorem 2.2.2 was given by Szemerédi [107, p. 93] who used it to prove a density result about arithmetic progressions.

Theorem 2.2.3 (Szemerédi [107])**.** For every $\varepsilon > 0$, $m \in \mathbb{Z}^+$ and $n \geq \frac{1}{4}\left(\frac{4}{\varepsilon}\right)^{2^m}$, if $A \subseteq [1, n]$ is such that $|A| \geq \varepsilon n$, then A contains an affine m-cube.

Proof. The proof presented here is different from the original and is due to Lovaśz [81, pp. 561–562]. Throughout this proof, for any set $A \subseteq \mathbb{Z}^+$ and $d \in \mathbb{Z}^+$, let $A_d = A \cap (A + d) = \{x \in A : x + d \in A\}$. It will be useful to have some information on the size of the set A_d. If $A \subseteq [1, n]$ and $|A| \geq 2$, then

$$\begin{aligned}
\sum_{d=1}^{n-1} |A_d| &= \sum_{d=1}^{n-1} |\{x \in A : x + d \in A\}| \\
&= \sum_{\{x,y\} \in [A]^2} |\{d \in [1, n-1] : |x - y| = d\}| \qquad \text{(double counting)} \\
&= \binom{|A|}{2} \\
&= \frac{|A|^2}{2}\left(1 - \frac{1}{|A|}\right) \\
&\geq \frac{|A|^2}{4} \qquad \text{(since } |A| \geq 2\text{).}
\end{aligned}$$

Therefore,

$$\max_{d\in[1,n-1]}|A_d| \geq \operatorname*{avg}_{d\in[1,n-1]}|A_d| \geq \frac{1}{(n-1)}\frac{|A|^2}{4} > \frac{|A|^2}{4n}$$

and so there is at least one $d \in [1, n-1]$ with $|A_d| \geq \frac{|A|^2}{4n}$. For integers d_1, d_2, denote $(A_{d_1})_{d_2} = A_{d_1,d_2}$.

Fix ε with $0 \leq \varepsilon \leq 1$ and $m \in \mathbb{Z}^+$. Let $n \geq \frac{1}{4}\left(\frac{4}{\varepsilon}\right)^{2^m}$ and $A \subseteq [1,n]$ be such that $|A| \geq \varepsilon n$. A sequence of integers $a_1, a_2, \ldots a_m$ will be recursively chosen so that if $k \in [1,m]$, then $|A_{a_1,\ldots,a_k}| \geq 4n\left(\frac{\varepsilon}{4}\right)^{2^k}$.

Since

$$|A| \geq \varepsilon n \geq \frac{\varepsilon}{4}\left(\frac{4}{\varepsilon}\right)^{2^m} = \frac{4^{2^{m-1}}}{\varepsilon} \geq \frac{4}{\varepsilon} > 2,$$

there is an $a_1 \in [1, n-1]$ such that

$$|A_{a_1}| \geq \frac{|A|^2}{4n} \geq \frac{(\varepsilon n)^2}{4n} = 4n\left(\frac{\varepsilon}{4}\right)^2.$$

In general for $k \in [1, m-1]$, having found $a_1, \ldots, a_k$ so that $|A_{a_1,\ldots,a_k}| \geq 4n\left(\frac{\varepsilon}{4}\right)^{2^k}$, since

$$\begin{aligned}|A_{a_1,\ldots,a_k}| &\geq \left(\frac{4}{\varepsilon}\right)^{2^m}\left(\frac{\varepsilon}{4}\right)^{2^k}\\ &= \left(\frac{4}{\varepsilon}\right)^{2^{m-k}}\\ &> \frac{4}{\varepsilon} > 2,\end{aligned}$$

there is an $a_{k+1} \in [1, n-1]$ such that

$$|A_{a_1,\ldots,a_k,a_{k+1}}| \geq \frac{|A_{a_1,\ldots,a_k}|^2}{4n}$$

$$\geq 4n\left(\frac{\varepsilon}{4}\right)^{2^{k+1}}.$$

Therefore, by induction, there are $a_1, \ldots, a_m \in [1, n-1]$ such that

$$|A_{a_1,\ldots,a_m}| \geq 4n\left(\frac{\varepsilon}{4}\right)^{2^m} \geq 4\cdot\frac{1}{4}\left(\frac{4}{\varepsilon}\right)^{2^m}\left(\frac{\varepsilon}{4}\right)^{2^m} = 1.$$

Thus, $A_{a_1,\ldots,a_m} \neq \emptyset$ and so there is at least one element $a_0 \in A_{a_1,\ldots,a_m}$.

Now, $A_{a_1} = \{a \in A : a + a_1 \in A\} = \{a \in A : H(a, a_1) \subseteq A\}$ and $A_{a_1,a_2} = \{a \in A_{a_1} : a + a_2 \in A_{a_1}\} = \{a \in A : H(a, a_1, a_2) \subseteq A\}$. Continuing in this manner, $A_{a_1,\ldots,a_m} = \{a \in A : H(a, a_1, a_2, \ldots, a_m) \subseteq A\}$. Therefore, since $a_0 \in A_{a_1,\ldots,a_m}$, $H(a_0, a_1, \ldots, a_m) \subseteq A$. □

Corollary 2.2.4. For every $\varepsilon > 0$ there is a constant c such that for all $n \in \mathbb{Z}^+$, if $m \leq \log_2\log_2 n + c$, and $A \subseteq [1, n]$ with $|A| \geq \varepsilon n$, then A contains an affine m-cube.

Proof. Fix $\varepsilon > 0$ and set $c = -\log_2\log_2\left(\frac{4}{\varepsilon}\right)$. For any $n, m \in \mathbb{Z}^+$, if

$$m \leq \log_2\log_2 n + c = \log_2\log_2 n - \log_2\log_2\left(\frac{4}{\varepsilon}\right) = \log_2\left(\frac{\log_2 n}{\log_2\left(\frac{4}{\varepsilon}\right)}\right),$$

then,

$$2^m \leq \frac{\log_2 n}{\log_2\left(\frac{4}{\varepsilon}\right)} \Rightarrow 2^m \log_2\left(\frac{4}{\varepsilon}\right) \leq \log_2 n \Rightarrow \left(\frac{4}{\varepsilon}\right)^{2^m} \leq n.$$

Since $n \geq \left(\frac{4}{\varepsilon}\right)^{2^m} > \frac{1}{4}\left(\frac{4}{\varepsilon}\right)^{2^m}$, by Theorem 2.2.3, for any $A \subseteq [1, n]$ with $|A| \geq \varepsilon n$, A must contain an affine m-cube. □

The following lemma gives a further connection between affine cubes and arithmetic progressions.

Definition 2.2.5. An affine m-cube is *replete* iff $|H| = 2^m$.

Lemma 2.2.6 (Gunderson and Rödl [56])**.** For any $m \in \mathbb{Z}^+$, if $A \subseteq \mathbb{Z}^+$ contains no replete affine m-cubes and no AP_3's, then A contains no affine m-cubes.

Proof. Let $A \subseteq \mathbb{Z}^+$ contain no replete affine m-cubes. Suppose that A contains an affine m-cube, $H = H(a, d_1, \ldots, d_m) \subseteq A$. Since H is not replete, there are sets $I, J \subseteq [1, m]$ with $I \neq J$ and $a + \sum_{i \in I} d_i = a + \sum_{j \in J} d_j$. By cancelling common terms, it can be assumed that $I \cap J = \emptyset$. Since $\sum_{i \in I} d_i = \sum_{j \in J} d_j$, the set

$$\left\{a, a + \sum_{i \in I} d_i, a + \sum_{i \in I \cup J} d_i\right\} \subseteq A$$

is an AP_3. □

2.3 Schur's theorem

According to Prömel and Voigt [85], Issai Schur had made a conjecture about arbitrarily long sequences of consecutive quadratic residues and saw that he could achieve the necessary proof given a particular partition result about arithmetic progressions. Although Schur was not successful on this front, he did prove the following arithmetic Ramsey-type theorem in 1916.

Theorem 2.3.1 (Schur [101])**.** For every $r \in \mathbb{Z}^+$ there is a least positive integer $S(r)$ such that for any r-colouring, $\Delta : [1, S(r)] \to [1, r]$, there exist $x, y \in [1, S(r)]$, possibly with $x = y$, such that $\Delta(x) = \Delta(y) = \Delta(x + y)$.

Here x and y need not be distinct, though it is possible to prove the theorem while insisting that $x \neq y$.

Proof. The proof given here uses Ramsey's theorem (Theorem 2.1.4) and shows that $S(r) \leq R(3; r) - 1$ (see [51, p.69]).

Let $r \in \mathbb{Z}^+$ and set $n = R(3; r) - 1$. To see that $S(r) \leq n$, let $\Delta : [1, n] \to [1, r]$ be any r-colouring and consider the graph K_{n+1} on vertices $\{0, 1, \ldots, n\}$ with an edge colouring defined as follows. For $0 \leq i < j \leq n$,

$$\Delta^*(\{i, j\}) = \Delta(j - i).$$

If $0 \leq i < j \leq n$, then $j - i \in [1, n]$ and thus Δ^* is well-defined. By the choice of n, there is a triangle in K_{n+1} which is monochromatic under Δ^*. Thus, in terms of Δ, there are $0 \leq a < b < c \leq n$ such that,

$$\begin{aligned}\Delta(b - a) = \Delta(c - b) &= \Delta(c - a)\\ &= \Delta((b - a) + (c - b)).\end{aligned}$$

Set $x = b - a$, $y = c - b$ and $z = c - a$. Then $\Delta(x) = \Delta(y) = \Delta(z)$ and $x + y = (b - a) + (c - b) = c - a = z$. □

This proof is shorter than Schur's original proof which did not use Ramsey's theorem and gave the slightly better bound: $S(r) \leq er!$, where is e is the natural logarithm base.

2.4 Dickson's theorem

It seems that one of Schur's motivations for Theorem 2.3.1 was to present a new proof of a theorem by L. E. Dickson [30] on congruence equations, Theorem 2.4.5 below. Though unrelated to problems on arithmetic progressions, Dickson's theorem provides an interesting application of Ramsey Theory to a number theoretic problem. At the time, one of the attempts to prove Fermat's Last Theorem focussed on showing that when an integer m has some "nice" properties, there would be only finitely many primes p for which there were non-trivial solutions to the equation $x^m + y^m \equiv z^m \pmod{p}$. Given such a result and x_0, y_0 and z_0 with $x_0^m + y_0^m = z_0^m$, there would be infinitely many primes p for which $x_0^m + y_0^m = z_0^m \pmod{p}$ was a trivial solution. In that case, one of x_0, y_0 or z_0 would be divisible by infinitely many primes and hence 0. This would show that there are no non-trivial solutions to the equation $x^m + y^m = z^m$. Unfortunately, this approach was shown to be futile by Dickson. (More information about the problem can be found for example in [31].)

The proof of Dickson's theorem (Theorem 2.4.5 below) utilizes group theory and a few definitions and results are necessary (see for example Hungerford [68]).

Definition 2.4.1. Let G be a finite group with identity e. The *order* of G, denoted by $|G|$, is the number of elements in G. For each $g \in G$, the *order of* g, denoted by $|g|$, is the least positive integer n such that $g^n = e$.

Lemma 2.4.2. Let G be a finite group and $g \in G$. For any $m \in \mathbb{Z}^+$,

$$|g^m| = \frac{|g|}{\gcd(m, |g|)}.$$

Definition 2.4.3. Let G be a group and H a subgroup of G. For each $g \in G$, the set $gH = \{gh : h \in H\}$ is called a *left coset of* H and the set $Hg = \{hg : h \in H\}$ is called a *right coset of* H.

It is possible to show that, for any group G and subgroup H, the set of left cosets (or the set of right cosets) of H partitions G. It can also be shown that the number of left cosets of H is the same as the number of right cosets of H in G and this number depends only on the orders of G and H. (see, *e.g.*, [68, p.39]).

Theorem 2.4.4 (Lagrange)**.** Let G be a group and H a subgroup of G. The number of cosets of H in G is $\frac{|G|}{|H|}$.

Theorem 2.4.5 (Dickson [30])**.** For every integer m, and for every sufficiently large prime p, the equation

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

has a solution with none of x, y or z divisible by p.

Proof. The following is Schur's proof [101] of Dickson's theorem.

Fix $m \in \mathbb{Z}^+$ and let p be prime with $p \geq S(m) + 1$ and let $G = \{1, 2, \ldots, p-1\}$. Since p is prime, G is a cyclic group under multiplication modulo p. Let a be a generator for G, that is $G = \langle a \rangle$, and set $H = \{x^m : x \in G\}$. Since a is a generator for G, $H = \langle a^m \rangle$ and the order of H is

$$\begin{aligned} |H| = |\langle a^m \rangle| = |a^m| \\ &= \frac{|a|}{\gcd(m, |a|)} \\ &= \frac{p-1}{\gcd(m, p-1)}. \end{aligned}$$

By Theorem 2.4.4, the number of distinct cosets of H in G is

$$\begin{aligned} |G|/|H| &= (p-1)/\frac{p-1}{\gcd(m, p-1)} \\ &= \gcd(m, p-1) \\ &\leq m. \end{aligned}$$

Since the distinct cosets of H partition G and there are no more than m of them, they can be used to define an m-colouring of the elements of G.

By the choice of p, with $p-1 \geq S(m)$, and by Schur's theorem (Theorem 2.3.1), there are $x_0, y_0 \in G$ such that x_0, y_0, and $x_0 + y_0$ are all contained in the same coset of H. That is $y_0 \in x_0 H$ and $x_0 + y_0 \in x_0 H$.

In other words, $x_0^{-1} y_0 \in H$ and $x_0^{-1}(x_0 + y_0) = 1 + x_0^{-1} y_0 \in H$. By the definition of H, there are elements $y_1, z_1 \in G$ such that $y_1^m = x_0^{-1} y_0$ and $z_1^m = 1 + x_0^{-1} y_0$ in G.

That is, $1^m + y_1^m = 1 + x_0^{-1} y_0 = z_1^m \pmod{p}$ and since $1, y_1, z_1 \in G$ none of 1, y_1 or z_1 are divisible by p. □

Chapter 3

Van der Waerden's theorem

3.1 Preliminaries

While working at Hamburg in 1926, Bartel van der Waerden shared a conjecture on arithmetic progressions with Artin and Schrier that he had heard from Baudet in Göttingen. The conjecture, that was likely originally due to Schur, stated that for every partition of $\mathbb{Z}^+$ into two classes, for every $k \geq 3$, one of the partition classes will contain an AP_k. Together with Artin and Schrier, van der Waerden found an equivalent problem that he subsequently proved. (Further details on the history of the problem can be found in [85, 114].)

Theorem 3.1.1 (van der Waerden [113])**.** *For every $k, r \in \mathbb{Z}^+$, there is an integer n such that for every r-colouring of $[1, n]$, there is a monochromatic AP_k.*

There is no similar result for infinite arithmetic progressions. Consider the 2-colouring of $\mathbb{Z}^+$ defined for each $n \in \mathbb{Z}^+$ by $\Delta(n) = \lfloor \log_2 n \rfloor \pmod 2$. That is, for each $i \geq 0$, all of the integers in the interval $[2^i, 2^{i+1} - 1]$ are colour 0 if i is even and colour 1 if i is odd. Both colour classes contain arbitrarily long arithmetic progressions with finitely many terms, but no infinite arithmetic progressions.

The lemmas needed to show that Theorem 3.1.1 is equivalent to the original conjecture are given in this section together with some preliminary results in preparation for the proof of Theorem 3.1.1 in Section 3.2.

Definition 3.1.2. For every $k, r \in \mathbb{Z}^+$ let $W(k; r)$ be the least integer (if it exists) so that for every r-colouring of $[1, W(k; r)]$, there are $a, d \in \mathbb{Z}^+$ so that the arithmetic progression of length k

$$\{a + id : 0 \leq i \leq k - 1\} \subseteq [1, W(k; r)]$$

is monochromatic. The numbers $W(k; r)$ are called *van der Waerden numbers.*

The following lemma, observed by Schrier (see [114]), shows that the problem of proving the existence of the van der Waerden numbers is equivalent to a related problem for the set of all positive integers. The next three lemmas (Lemmas 3.1.3–3.1.5) all have standard proofs (see for example [51], [80] and [114]).

Lemma 3.1.3. Fix $r, k \in \mathbb{Z}^+$. The integer $W(k; r)$ exists iff for every r-colouring of $\mathbb{Z}^+$ there is a monochromatic AP_k.

Proof. If $W(k;r)$ exists, then for every r-colouring of $\mathbb{Z}^+$ there is a monochromatic AP_k since $[1, W(k;r)] \subseteq \mathbb{Z}^+$.

For the converse, suppose that $W(k;r)$ does not exist. Then, for every $n \in \mathbb{Z}^+$, there is an r-colouring Δ_n for which $[1,n]$ contains no monochromatic AP_k's. These r-colourings are used to construct an r-colouring of $\mathbb{Z}^+$ with no monochromatic AP_k's. Recursively build a sequence of colours $\{c_n\}_{n \in \mathbb{Z}^+}$ and a sequence of infinite sets $A_1 \supseteq A_2 \supseteq \ldots$ as follows.

Since there are only finitely many colours, one colour must occur infinitely many times in the sequence $\{\Delta_1(1), \Delta_2(1), \ldots\}$. Let $c_1 \in [1,r]$ be such that $A_1 = \{i \in \mathbb{Z}^+ : \Delta_i(1) = c_1\}$ is infinite.

In general, for $t \geq 1$, having defined the infinite set A_t, there must be one colour, call it c_{t+1}, that occurs infinitely many times in the sequence $\{\Delta_i(t+1) : i \in A_t\}$. Set $A_{t+1} = \{i \in A_t : \Delta_i(t+1) = c_{t+1}\}$.

Define a new colouring $\Delta : \mathbb{Z}^+ \rightarrow [1,r]$ as follows. For each $n \in \mathbb{Z}^+$, set $\Delta(n) = c_n$. Note that for $m, n \in \mathbb{Z}^+$, if $n \in A_m$, then $\Delta_n|_{[1,m]} = \Delta|_{[1,m]}$ by the definition of Δ and the choice of the set A_m. For each $n \in \mathbb{Z}^+$, there are no AP_k's which are monochromatic under Δ_n and so $\mathbb{Z}^+$ also does not contain any AP_k's which are monochromatic under Δ. □

It is worth noting that, in the proof of Lemma 3.1.3, the only property of arithmetic progressions used was that a k-term arithmetic progression is a finite set. A variation of this proof gives a more general result: Let $\mathcal{F}$ be any collection of finite sets of integers and for each $r \geq 2$, let $F(r)$ be the least integer (if it exists) such that for every r-colouring of $[1, F(r)]$, there is a monochromatic element of $\mathcal{F}$. A proof similar to that of Theorem 3.1.3 shows that for any $r \geq 2$, $F(r)$ exists iff for every r-colouring of $\mathbb{Z}^+$, there is an $F \in \mathcal{F}$ that is monochromatic. Further details on problems known as "compactness" results can be found, for example, in [27] and [90].

Lemma 3.1.4. Fix $k, r \in \mathbb{Z}^+$, suppose that the number $n = W(k; r)$ exists and let $P = \{a, a + d, \ldots a + (n-1)d\}$ be any AP_n. Then for any r-colouring of P there is a monochromatic AP_k.

Proof. Let $\Delta : P \to [1, r]$ be any r-colouring. Define an induced r-colouring $\Delta^* : [1, n] \to [1, r]$ by $\Delta^*(i) = \Delta(a + d(i-1))$. By the choice of n, there is an AP_k, $\{c, c + b, \ldots, c + (k-1)b\}$ that is monochromatic under Δ^*. In terms of Δ, for each $0 \leq i \leq k-1$, $\Delta^*(c + ib) = \Delta(a + d(c + ib - 1))$. Therefore the AP_k, $\{a + d(c-1), (a + d(c-1)) + db, \ldots, (a + d(c-1)) + (k-1)db\}$ is monochromatic under Δ. □

The same argument also shows that if $k, r \in \mathbb{Z}^+$ are such that for every r-colouring of $\mathbb{Z}^+$, there is a monochromatic AP_k, then for any infinite arithmetic progression $P = \{a + id : i \geq 0\}$ and any r-colouring of P, there is a monochromatic AP_k contained in P.

The next lemma, observed by Artin (see [114]), shows that the problem for arbitrary colourings is equivalent to that for 2-colourings.

Lemma 3.1.5. If for all $k \in \mathbb{Z}^+$, $W(k;2)$ exists, then for all $k \in \mathbb{Z}^+$, $r \geq 2$, the van der Waerden number $W(k;r)$ exists.

Proof. Assume that for all $k \in \mathbb{Z}^+$, the van der Waerden number $W(k;2)$ exists. The proof is by induction on r. The base case $r = 2$ is trivially true. Fix $k \in \mathbb{Z}^+$ and suppose that $r \geq 3$ is such that $W(k;r-1)$ exists. Set $m = W(k;r-1)$ and $n = W(m;2)$. In order to show that $W(k;r) \leq n$ let $\Delta : [1,n] \rightarrow \{\text{red}_1, \ldots, \text{red}_{(r-1)}, \text{blue}\}$ be any r-colouring. Define a new 2-colouring $\Delta^* : [1,n] \rightarrow \{\text{red}, \text{blue}\}$ by,

$$\Delta^*(i) = \begin{cases} \text{red}, & \text{if for some } j,\ \Delta(i) = \text{red}_j; \\ \text{blue}, & \text{if } \Delta(i) = \text{blue}. \end{cases}$$

If $[1,n]$ contains a blue AP_m under Δ^*, then since the AP_m will also be blue under Δ, $[1,n]$ contains a monochromatic AP_k since $k \leq m$.

Otherwise, by the choice of n, there is a AP_m, call it P, which is red under Δ^*. Thus, Δ restricted to P is an $(r-1)$-colouring and since $m = W(k;r-1)$, by Lemma 3.2.3, P contains an AP_k which is monochromatic under Δ.

Therefore, for every r-colouring of $[1,n]$, there is a monochromatic AP_k and hence $W(k;r)$ exists. □

The original problem, as it came to van der Waerden, asked if it was possible to guarantee that in any 2-colouring of $\mathbb{Z}^+$, one colour class contained arbitrarily long arithmetic progressions. As it turns out, the modifications given by Lemmas 3.1.3 and 3.1.5 led to a seemingly more complicated conjecture that was easier to solve.

The following variation of van der Waerden's theorem (Theorem 3.1.1) states that any sequence for which the difference between consecutive terms is bounded will contain arbitrarily long arithmetic progressions.

Theorem 3.1.6 (Brown [20] and Rabung [86])**.** Let $M \in \mathbb{Z}^+$ be such that for every $(M-1)$-colouring of $\mathbb{Z}^+$, at least one colour class contains arbitrarily long arithmetic progressions. Let $S = \{s_i\}_{i \geq 0}$ be a strictly increasing sequence such that for all $i \geq 0$, $|s_{i+1} - s_i| \leq M$. Then S contains arbitrarily long arithmetic progressions.

Proof. Define a partition of $\mathbb{Z}^+$ into M disjoint sets as follows. Set $A_0 = S$ and for each $n \in [1, M-1]$, having previously defined $A_0, A_1, \ldots, A_{n-1}$, set

$$A_n = \{s_i + n : i \geq 0\} \backslash \left(\bigcup_{j=0}^{n-1} A_j \right).$$

Since for all $i \geq 0$, $|s_{i+1} - s_i| \leq M$, the sets $A_0, \ldots, A_{M-1}$ define a partition of $\mathbb{Z}^+$. By assumption, there is one $n \in [1, M-1]$ so that A_{n_0} contains arbitrarily long arithmetic progressions. That is, for each $k \in \mathbb{Z}^+$, there are $a, d \in \mathbb{Z}^+$ so that

$\{a, a+d, \ldots, a+(k-1)d\} \subseteq A_{n_0} \subseteq S + n_0$. Therefore, the k-term arithmetic progression $\{a - n_0, a - n_0 + d, \ldots, a - n_0 + (k-1)d\}$ is contained in S. □

A consequence of Theorem 3.1.6 is the following result about the existence of arithmetic progressions and strings of consecutive integers in any partition of $\mathbb{Z}^+$.

Corollary 3.1.7 (Rabung [86])**.** If, for every finite colouring of $\mathbb{Z}^+$, one colour class contains arbitrarily long arithmetic progressions, then for any partition of $\mathbb{Z}^+$ into 2 classes, either one class contains arbitrarily long strings of consecutive numbers or else both classes contain arbitrarily long arithmetic progressions.

Proof. Let $\mathbb{Z}^+ = A_1 \cup A_2$ be any partition. If for some $M \in \mathbb{Z}^+$, the longest string of consecutive integers in A_1 is of length M, then for any two consecutive entries $a < b$ in A_2, $|b - a| \leq M + 1$. Therefore, if neither A_1 nor A_2 contain arbitrarily long strings of consecutive numbers, then both A_1 and A_2 satisfy the conditions of Theorem 3.1.6 and hence contain arbitrarily long arithmetic progressions. □

3.2 Block proof

In this section, a combinatorial proof of van der Waerden's theorem (Theorem 3.1.1) is given. Monochromatic arithmetic progressions are found by examining, not just the elements coloured, but also the effect of the colourings on an interval (also called a block). It will be useful to define some terminology and make a few preliminary notes.

Definition 3.2.1. Given two blocks of equal length, $I_1 = [i, i+k]$ and $I_2 = [j, j+k]$ with $I_1, I_2 \subseteq [1, n]$ and an r-colouring Δ of $[1, n]$, I_1 and I_2 are said to have the same *colour pattern* under Δ if for each $0 \leq m \leq k$, $\Delta(i+m) = \Delta(j+m)$.

Definition 3.2.2. A sequence of blocks $B_1, B_2, \ldots, B_k$ form an AP_k *of blocks* iff there is an integer $d > 0$ so that for each $i \in [2, k]$, $B_i = B_1 + (i-1)d$.

Note that the blocks B_i may or may not overlap.

Lemma 3.2.3. Suppose that for some $k, r, n \in \mathbb{Z}^+$, $W(k; r^n) = N$ exists. For any r-colouring $\Delta : [1, nN] \to [1, r]$, there exists an AP_k of blocks each block of length n, all with the same colour pattern under Δ.

Proof. For each $1 \leq i \leq N$, set $B_i = [1 + (i-1)n, in]$. Let $\Delta : [1, nN] \to [1, r]$ be any r-colouring and define the induced r^n-colouring $\Delta^* : [1, N] \to ([1, r])^n$ by

$$\Delta^*(x) = (\Delta(1 + (x-1)n),\ \Delta(2 + (x-1)n),\ \ldots,\ \Delta(xn))$$

where each colour under Δ^* is an n-tuple. Since $N = W(k; r^n)$, $[1, N]$ contains an AP_k, $\{a, a+d, \ldots, a+(k-1)d\}$ which is monochromatic under Δ^*. In terms of the colouring Δ, this means that the blocks $B_a, B_{a+d}, \ldots, B_{a+(k-1)d}$ form an AP_k of blocks of length n, with difference nd, for which all members have the same colour pattern under Δ. $\square$

The following proof of van der Waerden's theorem is a variation of the original proof and can be found in Khinchin [75], where it is attributed to M. A. Lukomskaya.

Theorem 3.2.4 (van der Waerden [113])**.** For each $k, r \geq 2$, $W(k; r)$ exists.

The proof is by induction on k, with a recursive construction in r steps. In order to demonstrate the idea of the proof, the following is a sketch of the inductive step for 2 colours to construct a monochromatic AP_{k+1} from AP_k's.

Let N be large and $\Delta : [1, N] \to \{\text{red }, \text{blue}\}$ be any 2-colouring. Find an AP_k of blocks $\{B(1), B(2), \ldots, B(k)\}$ with difference d_1 in the first half of the interval $[1, N]$, all with the same colour pattern under Δ, and set $B(k+1) = B(k) + d_1$. Let the blocks be large enough so that the first half of the block $B(1)$ is guaranteed to contain a monochromatic AP_k, $\{B(1, 1), B(1, 2), \ldots, B(1, k)\}$ with some difference d_2. Set $B(1, k + 1) = B(1, k) + d_2$. Then $B(1, k + 1) \in B(1)$ since the AP_k $\{B(1, 1), B(1, 2), \ldots, B(1, k)\}$ is contained in the first half of $B(1)$. For each $i \in [2, k + 1]$ and $j \in [1, k + 1]$, define

$$B(i, j) = B(1, j) + (i - 1)d_1.$$

Then, for each $i \in [1, k + 1]$, the block B_i contains an AP_{k+1} of blocks:

$$\{B(i, 1), B(i, 2), \ldots, B(i, k + 1)\}.$$

Since the blocks $B(1), \ldots, B(k)$ all have the same colour pattern, the sets $\{B(i,j) : 1 \le i,j \le k\}$ and $\{B(i,k+1) : 1 \le i \le k\}$ are both monochromatic, though not necessarily the same colour. If these two sets are the same colour, then $\{B(1,j) : 1 \le j \le k+1\}$ is a monochromatic AP_{k+1}.

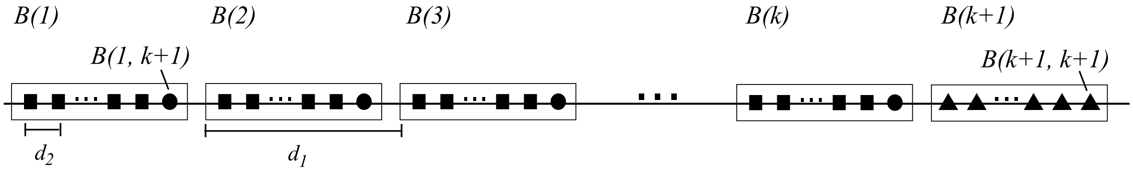

Figure 3.1: Arithmetic progression of blocks

If not, say $\{B(i,j) : 1 \le i,j \le k\}$ is red and $\{B(i,k+1) : 1 \le i \le k\}$ is blue, consider the element $B(k+1,k+1)$. If $B(k+1,k+1)$ is red, then the set $\{B(1,1), B(2,2), \ldots, B(k+1,k+1)\}$ is a red AP_{k+1} with difference $d_1 + d_2$. If $B(k+1,k+1)$ is blue, then $\{B(1,k+1), B(2,k+1), \ldots, B(k+1,k+1)\}$ is a blue AP_{k+1} with difference d_1.

The following proof shows the existence of all van der Waerden numbers. The general idea is similar to that for the previous sketch, though each block is subdivided into many smaller blocks in steps corresponding to the number of colours in question.

Proof of Theorem 3.1.1. The proof shows that for all integers $k, r \ge 2$, $W(k;r)$

exists using an induction on k with a recursive construction in r steps. Fix $r \geq 2$.

Base Case: By the pigeonhole principle, $W(2;r) = r+1$.

Inductive step: Fix $k \geq 2$ and suppose that for all $t \geq 2$, the number $W(k;t)$ exists. The following inductive step shows that $W(k+1;r)$ exists. Set $q_0 = 1$ and for each s with $1 \leq s \leq r$, define

$$n_{s-1} = W(k; r^{q_{s-1}}) \quad \text{and} \quad q_s = 2n_{s-1}q_{s-1}.$$

The goal of the proof is to show that $W(k+1;r) \leq q_r$. Fix an r-colouring $\Delta : [1, q_r] \longrightarrow [1, r]$. Using the choices of n_s and q_s, a sequence of arithmetic progressions of blocks are defined recursively in r steps.

Since $q_r = 2n_{r-1}q_{r-1}$, the interval $[1, q_r]$ can be divided into $2n_{r-1}$ blocks each of length q_{r-1} and since $n_{r-1} = W(k, r^{q_{r-1}})$, by Lemma 3.2.3, among the first n_{r-1} blocks in $[1, q_r]$, there is an AP_k of blocks, $\{B(1), B(2), \ldots, B(k)\}$, all with the same colour pattern under Δ. Set $B(k+1) = B(k) + d_1$.

The blocks $B(1), \ldots, B(k)$ are all contained in the first half of the interval $[1, q_r]$ and so $B(k+1) \subseteq [1, q_r]$, but nothing is known about the colouring of $B(k+1)$.

Since $q_{r-1} = 2n_{r-2}q_{r-2}$ the block $B(1)$ can be divided into $2n_{r-2}$ blocks of length q_{r-2}. Since $n_{r-2} = W(k, r^{q_{r-2}})$, in the first half of $B(1)$, there is an AP_k of blocks $\{B(1,1), B(1,2), \ldots, B(1,k)\}$ with difference d_2 which all have the same colour pattern under Δ. Set $B(1, k+1) = B(1,k) + d_2$. Since $B(1,k)$ is contained in the first half of the block $B(1)$, $B(1, k+1) \subseteq B(1)$ but again, nothing is known about the colouring of $B(1, k+1)$.

Translate the AP_{k+1} of blocks $\{B(1,1), \ldots, B(1,k), B(1,k+1)\}$ into the other blocks $B(2), \ldots, B(k+1)$ as follows: For $i = 2, 3, \ldots k+1$ and $j = 1, 2, \ldots, k+1$, define

$$B(i,j) = B(1,j) + (i-1)d_1.$$

Since the AP_k of blocks $\{B(1,1), \ldots, B(1,k)\} \subseteq B(1)$ all have the same colour pattern and the blocks $B(1), B(2), \ldots B(k)$ all have the same colour pattern, for $1 \leq i \leq j \leq k$ all the blocks $B(i,j)$ have the same colour pattern. Also, for $1 \leq i \leq k$ all the blocks $B(i, k+1)$ have the same colour pattern, though not necessarily the same as $B(1,1)$.

In general, for $s < r$ at step s of the recursion, the block $B(\underbrace{1, \ldots, 1}_{s-1})$ will be an interval of length $q_{r-s+1} = 2n_{r-s}q_{r-s}$ and if $B(1, \ldots, 1)$ is partitioned into $2n_{r-s}$ blocks, since $q_{r-s} = W(k, r^{q_{r-s}})$, the first half of $B(1, \ldots, 1)$ contains an AP_k of blocks

$$\{B(\underbrace{1, \ldots, 1}_{s-1}, 1), B(\underbrace{1, \ldots, 1}_{s-1}, 2), \ldots, B(\underbrace{1, \ldots, 1}_{s-1}, k)\}$$

with difference d_s and all with the same colour pattern under Δ. Set

$$B(\underbrace{1, \ldots, 1}_{s-1}, k+1) = B(\underbrace{1, \ldots, 1}_{s-1}, k) + d_s$$

and translate the AP_{k+1} of blocks, $\{B(1, \ldots, 1, 1), \ldots, B(1, \ldots, 1, k+1)\}$ into all the blocks constructed in step $s-1$ of the recursion. Note that if $i_1, \ldots, i_s, j_1, \ldots, j_s \in$

$[1, k]$, then the blocks $B(i_1, i_2, \ldots i_s)$ and $B(j_1, \ldots, j_s)$ have the same colour pattern under Δ.

After step r of the recursion, the blocks are all of size $q_0 = 1$. Since these blocks are all singletons, they will be treated interchangeably as integers or sets.

The following properties of the blocks of integers constructed in this way are worth noting.

If $1 \leq s < r$ and if $1 \leq i_{s+1}, i_{s+2}, \ldots, i_r \leq k+1$, then $B(i_1, \ldots i_s, i_{s+1}, \ldots, i_r)$ appears in the same position in the block $B(i_1, \ldots i_s)$ as $B(j_1, \ldots j_s, i_{s+1}, \ldots, i_r)$ does in the block $B(j_1, \ldots j_s)$. If $1 \leq i_1, \ldots i_s, j_1, \ldots, j_s \leq k$, then since the two blocks have the same colour pattern, the two integers have the same colour under Δ.

For $1 \leq s \leq r$, since $B(i_1, \ldots, i_{s-1}, i_s + 1) = B(i_1, i_2, \ldots, i_{s-1}, i_s) + d_s$ and $B(i_1, \ldots, i_{s-1}, i_s, i_{s+1}, \ldots, i_r)$ and $B(i_1, \ldots, i_{s-1}, i_s + 1, i_{s+1}, \ldots, i_r)$ appear in the same position in their respective blocks,

$$B(i_1, \ldots, i_{s-1}, i_s + 1, i_{s+1}, \ldots, i_r) - B(i_1, \ldots, i_s, \ldots, i_r) = d_s. \tag{3.1}$$

Consider the following $r + 1$ elements. For each $i \in [0, r]$, let

$$b_i = B(\underbrace{1, \ldots, 1}_{i}, \underbrace{k+1, \ldots, k+1}_{r-i}).$$

Since Δ is an r-colouring, by the pigeonhole principle, there must be u and v with $0 \leq u < v \leq r$ so that $\Delta(b_u) = \Delta(b_v)$.

For each $i \in [1, k+1]$, define $a_i = B(\underbrace{1, \dots, 1}_{u}, \underbrace{i, \dots, i}_{v-u}, \underbrace{k+1, \dots, k+1}_{r-v})$. Then,

$$a_1 = b_v \quad \text{and} \quad a_{k+1} = b_u.$$

If $i+1 \leq k$, then by a previous remark, $a_i = B(1, \dots, 1, i, \dots, i, k+1, \dots, k+1)$ and $a_{i+1} = B(1, \dots, 1, i+1, \dots, i+1, k+1, \dots, k+1)$ have the same colour. Since $\Delta(a_1) = \Delta(a_{k+1})$, the set $\{a_1, \dots, a_{k+1}\}$ is monochromatic under Δ. To show that $\{a_1, \dots a_{k+1}\}$ is an AP_{k+1}, fix $i \in [1, k]$ and for each $m \in [0, v-u]$, define

$$a_{i,m} = B(\underbrace{1, \dots, 1}_{u}, \underbrace{i+1, \dots, i+1}_{m}, \underbrace{i, \dots, i}_{v-u-m}, \underbrace{k+1, \dots, k+1}_{r-v})$$

so that $a_{i,0} = a_i$ and $a_{i,v-u} = a_{i+1}$. Now,

$$\begin{aligned} a_{i,m} - a_{i,m-1} &= B(1, \dots, 1, i+1, \dots, i+1, \underbrace{i+1}_{(u+m)-\text{th}}, i, \dots, i, k+1, \dots, k+1) \\ &\qquad - B(1, \dots, 1, i+1, \dots, i+1, \underbrace{i}_{(u+m)-\text{th}}, i, \dots, i, k+1, \dots, k+1) \\ &= d_{u+m} \quad \text{(by eq'n (3.1))}. \end{aligned}$$

The sequence $\{a_{i,m} : 0 \leq m \leq v-u\}$ can be used to write the difference $a_{i+1} - a_i$ as a telescoping series.

$$\begin{aligned} a_{i+1} - a_i &= \sum_{m=1}^{v-u} a_{i,m} - a_{i,m-1} \\ &= \sum_{m=1}^{v-u} d_{u+m} \\ &= d_{u+1} + d_{u+2} + \dots + d_v. \end{aligned}$$

Therefore, the set $\{a_1, a_2, \ldots, a_{k+1}\}$ is a monochromatic AP_{k+1} with difference $d_{u+1} + d_{u+2} + \cdots + d_v$ and so $W(k+1; r) \leq q_r$. Therefore, by induction, for any $k \in \mathbb{Z}^+$, the van der Waerden number $W(k; r)$ exists. □

3.3 Graham-Rothschild proof

In 1974, Graham and Rothschild [50] gave an alternative proof of van der Waerden's theorem. Their proof, presented in this section, is essentially a variant of the argument in the previous section and is included for historical interest.

Definition 3.3.1. For each k, m, let $S(k, m)$ be the statement that for every $r \geq 2$, there is an integer $N(k, m, r)$ such that for any r-colouring of $[1, N(k, m, r)]$, there are $a, d_1, \ldots, d_m \in \mathbb{Z}^+$ such that for each $i \in [1, m]$, the i-fold arithmetic progression of length k with starting point $a + k(d_{i+1} + \cdots + d_m)$,

$$\{a + k(d_{i+1} + \cdots + d_m) + \sum_{j=1}^{i} x_j d_j : x_1, \ldots, x_i \in [0, k-1]\}$$

is monochromatic.

Note that if $N(k, m, r)$ exists, then for every r-colouring of $[1, N(k, m, r)]$ there is a monochromatic m-fold arithmetic progression of length k. In particular, the statement $S(k, 1)$ is precisely van der Waerden's theorem.

Theorem 3.3.2 (Graham and Rothschild [50]). For each $k, m \geq 1$, $S(k, m)$ holds.

Proof. The proof proceeds by double induction on k and m.

Base Case: When $m = 1$ and $k = 1$, the statement $S(1, 1)$ holds since a 1-fold arithmetic progression of length 1 is just a singleton.

Inductive Steps: (i) For any $k, m \geq 1$, $S(k, m) \Rightarrow S(k, m + 1)$.

Note first that $S(k, m) \Rightarrow S(k, 1)$ (simply ignore all but the 1-fold arithmetic progression of length k).

Suppose that for some $k \geq 1$, the statement $S(k, m)$ holds. Fix an $r \in \mathbb{Z}^+$ and set $M = N(k, m, r)$ and $M' = N(k, 1, r^M)$. Let $\Delta : [1, MM'] \to r$ be any r-colouring. For $1 \leq j \leq M'$, define blocks $I_j = [(j-1)M + 1, jM]$. By the choice of M' and Lemma 3.2.3, there is an arithmetic progression of blocks $\{I_{a'}, I_{a'+d'}, \ldots, I_{a'+(k-1)d'}\}$ all with the same colour pattern under Δ.

By the choice of M, there are $a, d_2, \ldots, d_{m+1}$ such that for each $i \in [1, m + 1]$, the set

$$\{a + k(d_{i+1} + \cdots + d_{m+1}) + \sum_{j=2}^{i} x_j d_j : \text{each of } x_1, \ldots, x_i \in [0, k-1]\} \subseteq I_{a'}$$

is monochromatic under Δ.

Set $d_1 = d'M$. For any $i \in [1, m + 1]$ and $x_1, \ldots, x_i \in [0, k - 1]$ the integer $a + k(d_{i+1} + \ldots d_{m+1}) + \sum_{j=2}^{i} x_j d_j$ occupies the same position in $I_{a'}$ as the integer $a + k(d_{i+1} + \ldots d_{m+1}) + \sum_{j=2}^{i} x_j d_j + x_1 d_1$ does in the block $I_{a'+x_1 d'}$. Since these two blocks have the same colour pattern, the set

$$\{a + k(d_{i+1} + \cdots d_{m+1}) + \sum_{j=1}^{i} x_j d_j : x_1, \ldots, x_i \in [0, -1]\}$$

is monochromatic. Therefore, $N(k, m+1, r) \leq MM'$ and the statement $S(k, m+1)$ holds.

(ii) [For all $m \geq 1$, $S(k, m)$] $\Rightarrow S(k+1, 1)$.

Fix $k \geq 1$ and suppose that for all $m \geq 1$, $S(k, m)$ holds. Fix r and let $\Delta : [1, N(k, r, r)] \to [1, r]$ be given. Then, there are $a, d_1, \ldots, d_r$ such that for each $i \in [0, r]$, the set

$$\{a + k(d_{i+1} + \cdots + d_r) + \sum_{j=1}^{i} x_j d_j : x_1, \ldots, x_i \in [0, k-1]\} \tag{3.2}$$

is monochromatic.

For each $s \in [1, r+1]$, consider the sum $a + \sum_{i=s}^{r} kd_i$ where the empty sum is 0. Since there are $r+1$ choices for s and r colours, by the pigeonhole principle, there are $1 \leq u < v \leq r+1$ such that

$$\Delta(a + \sum_{i=u}^{r} kd_i) = \Delta(a + \sum_{i=v}^{r} kd_i).$$

Let $b = a + \sum_{i=v}^{r} kd_i$, $c = \sum_{i=u}^{v-1} d_i$ and consider the sequence $\{b + xc : x \in [0, k]\}$.

The set $\{b + xc : x \in [0, k-1]\} = \{a + \sum_{i=v}^{r} kd_i + \sum_{i=u}^{v-1} xd_i : x \in [0, k-1]\}$ is monochromatic by equation (3.2) and the numbers b and $b + kc$ have the same colour since,

$$\Delta(b + kc) = \Delta\left(\left(a + \sum_{i=v}^{r} kd_i\right) + k\left(\sum_{i=u}^{v-1} d_i\right)\right)$$

$$
\begin{aligned}
&= \Delta\left(a + \sum_{i=u}^{r} kd_i\right) \\
&= \Delta\left(a + \sum_{i=v}^{r} kd_i\right) && \text{(by the choice of u and v)} \\
&= \Delta(b).
\end{aligned}
$$

Therefore the sequence $\{b + xc : x \in [0, k]\}$ is a monochromatic AP_{k+1} and so $S(k+1, 1)$ holds. □

Chapter 4

Density results

4.1 Erdős-Turán function

In a 1936 paper, Erdős and Turán [41] considered the problem of determining the maximum number of elements in an AP_k-free subset of $[1, n]$. The function they introduced has been called the Erdős-Turán function (for example in [80, pp.41-3]).

Definition 4.1.1. For each $n, k \in \mathbb{Z}^+$, let $r_k(n)$ be the maximum size of an AP_k-free subset of $[1, n]$.

While van der Waerden's theorem is what is known as a "partition result", results related to the function r_k are called "density results".

It is worth noting that, by Lemma 3.1.4, the number $r_k(n)$ is also the maximum size of an AP_k-free subset of any interval of length n or of any arithmetic progression of length n.

The notation here has not yet been standardized and $\nu_k(n)$ is sometimes used in place of $r_k(n)$. In a 1969 paper, Szemerédi [107] used τ_k for this function, but subsequently switched to r_k in a 1975 paper [108]. Often, $r(n)$ and $\nu(n)$ are each used to denote $r_3(n)$ and $\nu_3(n)$ respectively, and this function has also been denoted in the past as $A(n)$ by Roth [93] or as $S(n)$ by Graham, Rothschild and Spencer [51].

Throughout this chapter, when colourings are discussed, the letter t will be used for the number of colours instead of the usual r to avoid confusion with the function r_k. The number $r_k(n)$ is related to the van der Waerden numbers in the sense that for some $n, k, t \in \mathbb{Z}^+$, if $r_k(n) < \frac{n}{t}$, then for any t-colouring of $[1, n]$, by the pigeonhole principle, one colour class contains at least $\frac{n}{t}$ elements and hence an AP_k, showing $W(k; t) \leq n$.

Definition 4.1.2. A function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ is *subadditive* if for all $x, y \in \mathbb{Z}^+$, $f(x + y) \leq f(x) + f(y)$.

Lemma 4.1.3 (Behrend [11])**.** For all $k \in \mathbb{Z}^+$, the function r_k is subadditive.

Proof. Fix $k \in \mathbb{Z}^+$ and let $n_1, n_2 \in \mathbb{Z}^+$ be given. If $A \subseteq [1, n_1 + n_2]$ is AP_k-free with $|A| = r_k(n_1 + n_2)$, then neither $A \cap [1, n_1]$ nor $A \cap [n_1 + 1, n_1 + n_2]$ contain any AP_k's and so $r_k(n_1 + n_2) \leq r_k(n_1) + r_k(n_2)$. □

Definition 4.1.4. For each $k \in \mathbb{Z}^+$, set $c_k = \limsup_{n\to\infty} \frac{r_k(n)}{n}$.

Note that for all $n \in \mathbb{Z}^+$, $r_k(n) \leq n$ and so this limit is finite and $c_k \leq 1$.

Lemma 4.1.5 (Behrend [11])**.** For all $k \in \mathbb{Z}^+$, the limit $c_k = \lim_{n\to\infty} \frac{r_k(n)}{n}$ exists and for all $n \in \mathbb{Z}^+$, $\frac{r_k(n)}{n} \geq c_k$.

Proof. Fix $k, n \in \mathbb{Z}^+$ and for each $x \in \mathbb{Z}^+$, let $q_x \in \mathbb{Z}^+$ and $0 \leq s_x < n$ be such that $x = q_x n + s_x$. Then $r_k(x) \leq r_k((q_x+1)n) \leq (q_x+1)r_k(n)$ by the subadditivity of the function r_k. Therefore,

$$\frac{r_k(x)}{x} = \frac{r_k(x)}{q_x n + s_x} \leq \frac{(q_x+1)}{q_x}\frac{r_k(n)}{n}$$

and so

$$\begin{aligned} c_k &= \limsup_{x\to\infty} \frac{r_k(x)}{x} \\ &\leq \limsup_{x\to\infty} \frac{q_x+1}{q_x} \cdot \frac{r_k(n)}{n} \\ &= \frac{r_k(n)}{n} && \text{(since } q_x \to \infty\text{).} \end{aligned}$$

Therefore, $c_k \leq \frac{r_k(n)}{n}$ and since n was arbitrary, $c_k \leq \liminf_{n\to\infty} \frac{r_k(n)}{n}$. Trivially, $c_k \geq \liminf_{n\to\infty} \frac{r_k(n)}{n}$ and hence $\limsup_{n\to\infty} \frac{r_k(n)}{n} = \liminf_{n\to\infty} \frac{r_k(n)}{n}$. Thus $\lim_{n\to\infty} \frac{r_k(n)}{n}$ exists and equals c_k. □

Erdős and Turán conjectured [41] that for all $k \geq 3$, $c_k = 0$. In 1938, Behrend [11] published a proof that either for each $k \geq 3$, $c_k = 0$ or else $\lim_{k\to\infty} c_k = 1$.

By unravelling the definition of the limit, $c_k = 0$ iff for every $\varepsilon > 0$, there is an $N \in \mathbb{Z}^+$ such that for all $n \geq N$, if $A \subseteq [1, n]$ with $|A| \geq \varepsilon n$, then A contains an AP_k. As with van der Waerden's theorem (see Theorem 3.1.1), this problem has an equivalent form regarding infinite sets.

Definition 4.1.6. Given a set $A \subseteq \mathbb{Z}^+$ the *upper density* of A is defined to be

$$\bar{d}(A) = \limsup_{n\to\infty} \frac{|A \cap [1, n]|}{n}.$$

The set A has *positive upper density* iff $\bar{d}(A) > 0$.

The proof of the following theorem is a standard argument, a variant of which can be found in an article by Szemerédi [108].

Lemma 4.1.7. Let $k \in \mathbb{Z}^+$, then $c_k = 0$ iff every set $A \subseteq \mathbb{Z}^+$ with positive upper density contains an AP_k.

Proof. Suppose that $c_k = 0$. Fix $A \subseteq \mathbb{Z}^+$ with positive upper density and set $\bar{d}(A) = \varepsilon_0 > 0$. Let $N_1 \in \mathbb{Z}^+$ be such that for all $n \geq N_1$, if $|B \cap [1, n]| \geq \frac{\varepsilon_0}{3} n$, then B contains an AP_k. Let N_2 be such that if $n \geq N_2$, then $|\sup_{m\geq n} \left\{\frac{|A\cap[1,m]|}{m}\right\} - \varepsilon_0| \leq \frac{\varepsilon_0}{2}$. Set $M = \max\{N_1, N_2\}$. Then since $\sup_{m\geq M} \left\{\frac{|A\cap[1,m]|}{m}\right\} \geq \frac{\varepsilon_0}{2}$, there is some $M_0 \geq M$ with $\frac{|A\cap[1,M_0]|}{M_0} \geq \frac{\varepsilon_0}{3}$ and so $A \cap [1, M_0]$ must contain an AP_k.

For the converse, suppose that $c_k > 0$. Then there is an $\varepsilon_0 > 0$ and an infinite sequence $n_1 < n_2 < \cdots$ such that for each $i \in \mathbb{Z}^+$, there is an AP_k-free set $A_i \subseteq$

$[1, n_i]$ with $|A_i| \geq \varepsilon_0 n_i$. By ignoring some of the indices of the sequence, it can be assumed that for each $m \in \mathbb{Z}^+$,

$$n_m > 2 \sum_{j=1}^{m-1} n_j.$$

For each $m \in \mathbb{Z}^+$, set $B_m = \cup_{i=1}^{m}(A_i + n_i + 2\sum_{j=1}^{i-1} n_j)$. For every $m \geq 1$ the set $B_m \subseteq [1, 2\sum_{j=1}^{m} n_j]$ and since $n_{m+1} \geq 2\sum_{j=1}^{m} n_j$, there are no AP_k's containing elements from both B_m and $(A_{m+1} + n_{m+1} + 2\sum_{j=1}^{m} n_j)$.

Since each of the sets A_m are AP_k-free, so are the sets B_m. Therefore, the set $B = \cup_{m\in\mathbb{Z}^+} B_m$ is an AP_k-free set with upper density at least $\frac{\varepsilon_0}{2} > 0$. □

4.2 Roth's theorem

In 1952, Roth [93], proved that $\lim_{n\to\infty} \frac{r_3(n)}{n} = 0$ (*i.e.* $c_3 = 0$). Refining his own method, in 1953, Roth [94] subsequently proved the following result, which is presented here without proof.

Theorem 4.2.1 (Roth [94])**.** There is a positive constant c such that for all n sufficiently large, if $A \subseteq [1, n]$ with $|A| \geq \frac{cn}{\ln\ln n}$, then A contains an AP_3. That is,

$$r_3(n) < \frac{cn}{\ln\ln n}.$$

The bound given on $r_3(n)$ in the previous theorem implies that $\lim_{n\to\infty} \frac{r_3(n)}{n} = 0$.

In many of the theorems below regarding bounds on r_k, numerous positive constants are required. The symbols c_k will be reserved for the limit value as previously defined, but other symbols such as c, c' or c^* will be freely reused for various positive constants.

Rather than proving Theorem 4.2.1, the weaker result given in 1952 is proved.

Theorem 4.2.2 (Roth [93])**.** For every $\varepsilon > 0$, there is an N sufficiently large so that for all $n \geq N$, if $A \subseteq N$ with $|A| > \varepsilon n$, then A contains an AP_3. That is,

$$c_3 = \lim_{n\to\infty} \frac{r_3(n)}{n} = 0.$$

Proof. The following is an adaptation of the original proof that uses summations instead of integrals and appears in [51, pp. 49–53]. Assume, in hope of a contradiction, that $c_3 > 0$. Set $\varepsilon = \frac{(c_3)^2}{20}$ and let m be large enough so that for each $n \geq 2m+1$,

$$c_3 \leq \frac{r_3(n)}{n} < c_3 + \varepsilon.$$

Let $N \geq 2m+1$ be large (just how large to be determined throughout the proof) and let $A \subseteq [1, 2N]$ with $|A| = r_3(2N) \geq c_3 \cdot 2N$ be AP_3-free.

Let $A = \{u_1, u_2, \ldots, u_r\}$ and denote the even elements of A by $2v_1, \ldots, 2v_s$. By the choice of $|A| = r_3(2N)$ and since $2N \geq 2m+1$,

$$c_3 \cdot 2N \leq r < (c_3 + \varepsilon)2N. \tag{4.1}$$

The density of A on each of the sets of odd and even integers in $[1, 2N]$ must be less than $(c_3 + \varepsilon)$, otherwise A would contain an AP_3 (since $N \geq 2m+1$). Therefore

$s < (c+\varepsilon)N$ and $r - s < (c_3+\varepsilon)N$. Hence

$$s \;>\; r-(c_3+\varepsilon)N \;\geq\; c_3 2N-(c_3+\varepsilon)N \;=\; (c_3-\varepsilon)N,$$

and so,

$$(c_3-\varepsilon)N < s < (c_3+\varepsilon)N. \tag{4.2}$$

Three complex-valued functions will be used throughout the proof. For any real numbers x and α, let

$$\begin{aligned} e(x) &= e^{2\pi i x} && \text{(where } i=\sqrt{-1}\text{)},\\ f(\alpha) &= \sum_{j=1}^{r} e(\alpha u_j) && \text{and}\\ g(\alpha) &= \sum_{k=1}^{s} e(\alpha v_k). \end{aligned}$$

If h is any function of α, let $\sum^* h(\alpha) = \sum_{a=0}^{2N-1} h(\frac{a}{2N})$. In particular, if $u \in \mathbb{Z}$ with $|u| < 2N$, then

$$\sum\nolimits^* e(\alpha u) = \begin{cases} 2N, & \text{if } u = 0,\\[2ex] 0, & \text{if } u \neq 0. \end{cases} \tag{4.3}$$

In the first case, 1 is added $2N$ times and in the second, all the $\left(\frac{2N}{\gcd(u,2N)}\right)$-th roots of unity are added $\gcd(u,2N)$ times. The first part of the proof focuses on approximating the sum,

$$\sum\nolimits^* f(\alpha)g(-\alpha)^2 = \sum\nolimits^* \left(\sum_{j=1}^{r} e(\alpha u_j)\right)\left(\sum_{k=1}^{s} e(-\alpha v_k)\right)^2$$

$$= \sum_{j=1}^{r} \sum_{k=1}^{s} \sum_{\ell=1}^{s} \sum\nolimits^{*} e(\alpha(u_j - v_k - v_\ell)) \qquad \text{(expanding).}$$

Note that $u_j - v_k - v_\ell = 0$ iff $u_j = (2v_k + 2v_\ell)/2$ which is precisely when $\{2v_k, u_j, 2v_\ell\}$ is an AP_3. As A is AP_3-free, this is only possible when $2v_k = u_j = 2v_\ell$. There are exactly s triples (j, k, ℓ) in the above sum with $u_j = 2v_k = 2v_\ell$. For all other triples (j, k, ℓ), $\sum^{*} e(\alpha(u_j - v_k - v_\ell)) = 0$ by equation (4.3). Therefore,

$$\begin{aligned} \sum\nolimits^{*} f(\alpha) g(-\alpha)^2 &= s2N \\ &\le (c_3 + \varepsilon) N \cdot 2N && \text{(by eq'n (4.2))} \\ &= c_3 N^2 2 \left(1 + \frac{c_3}{20}\right) && \text{(since } \varepsilon = c_3^2/20) \\ &\le 3c_3 N^2 && \text{(since } c_3 \le 1). \end{aligned} \tag{4.4}$$

Now $g(-\alpha) = \overline{g(\alpha)}$, the complex conjugate, so that $|g(-\alpha)|^2 = g(\alpha)g(-\alpha)$. Using an argument similar to that in the calculation of inequality (4.4) above gives

$$\begin{aligned} \sum\nolimits^{*} |g(-\alpha)|^2 &= \sum\nolimits^{*} g(\alpha) g(-\alpha) \\ &= \sum_{j=1}^{s} \sum_{k=1}^{s} \sum\nolimits^{*} e(\alpha(v_j - v_k)) \\ &= 2N \cdot s && \text{(since } v_j - v_k = 0 \text{ iff } j = k) \\ &\le 3c_3 N^2 && \text{(as for eq'n (4.4)).} \end{aligned} \tag{4.5}$$

Also,

$$f(0)g(0)^2 = \sum_{j=1}^{r} e(0) \left(\sum_{k=1}^{s} e(0) \right)^2$$

$$\begin{aligned}
&= rs^2 \\
&\geq c_3 2N((c_3-\varepsilon)N)^2 && \text{(by eq'ns (4.1) and (4.2))} \\
&= (c_3)^3 N^3 \cdot 2\left(1-\frac{c_3}{20}\right)^2 && \text{(since } \varepsilon = (c_3)^2/20\text{)} \\
&\geq (c_3)^3 N^3 \cdot 2\left(\frac{19}{20}\right)^2 && \text{(since } c_3 \leq 1\text{)} \\
&\geq (c_3)^3 N^3. && (4.6)
\end{aligned}$$

The idea of the proof is to use the previous bounds together with the inequality

$$\begin{aligned}
|f(0)g(0)^2| &\leq \left|\sum\nolimits^* f(\alpha)g(-\alpha)^2\right| + \left|\sum_{\alpha\neq 0}{}^{*} f(\alpha)g(-\alpha)^2\right| && \text{(triangle inequality)} \\
&\leq \left|\sum\nolimits^* f(\alpha)g(-\alpha)^2\right| + \max_{\alpha\neq 0}|f(\alpha)|\sum\nolimits^* |g(-\alpha)^2| && (4.7)
\end{aligned}$$

to derive an absurd inequality.

The remainder of the proof focuses on finding a bound for $|f(\alpha)|$ when $\alpha \neq 0$. However, this is not be accomplished directly, but rather by calculating bounds of a function related to f. In Graham, Rothschild and Spencer [51, p. 51] this function is called a *smear* of f.

First, note the following bound for the complex-valued function $e(x)$. For any real number x,

$$\begin{aligned}
\left|\frac{e(x)+e(-x)}{2} - 1\right| &= |\cos(2\pi x) - 1| && (e^{i\theta} = \cos\theta + i\sin\theta) \\
&\leq \frac{(2\pi x)^2}{2} && \text{(Taylor series for } \cos x\text{)}
\end{aligned}$$

$$= 2\pi^2 x^2 \tag{4.8}$$

Now, for any real γ,

$$\begin{aligned}
\left|\frac{1}{2m+1}\sum_{|k|\le m} e(k\gamma) - 1\right| &= \left|\frac{1}{2m+1}\left(e(0) - 1 + \sum_{k=1}^{m} e(k\gamma) + e(-k\gamma) - 2\right)\right| \\
&\le \frac{2}{2m+1}\sum_{k=1}^{m}\left|\frac{1}{2}(e(k\gamma) + e(-k\gamma)) - 1\right| \\
&\le \frac{2}{2m+1}\sum_{k=1}^{m} 2\pi^2(k\gamma)^2 \qquad \text{(by eq'n (4.8))} \\
&= \frac{4\pi^2\gamma^2}{2m+1}\cdot\frac{m(m+1)(2m+1)}{6} \\
&= \frac{2\pi^2\gamma^2 m(m+1)}{3} \\
&\le \pi^2(\gamma m)^2 \qquad \text{(for } m \ge 2\text{).}
\end{aligned}$$

Since $|e(\alpha u)| = 1$, multiplying the previous inequality through by $|e(\alpha u)|$ gives,

$$\left|\frac{1}{2m+1}\sum_{|k|\le m} e(\alpha u + k\gamma) - e(\alpha u)\right| \le \pi^2(m\gamma)^2$$

and summing u over A gives,

$$\begin{aligned}
&\left|f(\alpha) - \frac{1}{2m+1}\sum_{u\in A}\sum_{|k|\le m} e(\alpha u + k\gamma)\right| \\
&\qquad = \left|\sum_{u\in A} e(\alpha u) - \frac{1}{2m+1}\sum_{u\in A}\sum_{|k|\le m} e(\alpha u + k\gamma)\right| \\
&\qquad \le \pi^2(m\gamma)^2|A| \\
&\qquad \le \pi^2(m\gamma)^2 2N. \qquad (4.9)
\end{aligned}$$

In order to choose an appropriate value for γ, recall the theorem of Dirichlet (see, for example, [71, p.43]) that states that for any real α and integer M, there exist integers p, q and a real number β with $\alpha = \frac{p}{q} + \beta$, $1 \leq q \leq M$ and $|\beta| \leq \frac{1}{Mq}$.

Fix $\alpha \neq 0$, $M = \lceil N^{1/2} \rceil$ and let p, q, β, from Dirichlet's theorem, be such that $\alpha = \frac{p}{q} + \beta$, $q \in [1, M]$, and $|\beta| \leq \frac{1}{Mq}$. Then,

$$\begin{aligned}
e(\alpha(u+kq)) &= e(\alpha u + k(p+\beta q)) && \text{(since } \alpha q = p + q\beta\text{).} \\
&= e(\alpha u + k\beta q)e(kp) \\
&= e(\alpha u + k\beta q) && \text{(since } kp \in \mathbb{Z}\text{, } e(kp) = e^{2\pi i kp} = 1\text{).}
\end{aligned}$$

Taking $\gamma = \beta q$ in equation (4.9) gives,

$$\begin{aligned}
\left| f(\alpha) - \frac{1}{2m+1} \sum_{u \in A} \sum_{|k| \leq m} e(\alpha(u+kq)) \right| &\leq 2N\pi^2(mq\beta)^2 \\
&\leq 2N\frac{\pi^2 m^2}{M^2} && \text{(since } |\beta| \leq \frac{1}{qM}\text{)} \\
&\leq 2\pi^2 m^2 && \text{(since } M \geq N^{1/2}\text{)} \quad (4.10)
\end{aligned}$$

The focus of the proof now turns to finding a bound for

$$\frac{1}{2m+1} \sum_{u \in A} \sum_{|k| \leq m} e(\alpha(u+kq)).$$

For $x \in [0, 2N-1]$, set $W_x = \{x + kq \pmod{2N} : |k| \leq m\}$. For any fixed $x \in [0, 2N-1]$, the number of pairs (u, k), where $u \in A$ and $k \in \mathbb{Z}$ with $|k| \leq m$, such that $u + kq = x$ is exactly the number of elements of A in $\{x - kq : |k| \leq m\}$.

Therefore, double counting the pairs (u, k),

$$\frac{1}{2m+1}\sum_{u\in A}\sum_{|k|\le m} e(\alpha(u+kq)) = \sum_{x=0}^{2N-1} e(\alpha x)\frac{|W_x \cap A|}{2m+1}. \tag{4.11}$$

Again, as it is difficult to calculate a small enough bound for $\sum_{x=0}^{2N-1}|W_x \cap A|$, a bound on something close will be found. For each $x \in [0, 2N-1]$, set

$$E_x = \frac{|W_x \cap A|}{2m+1} - c_3.$$

If $mq < x < 2N - mq$, then W_x is an $AP_{2m+1} \subseteq [1, 2N]$. For these values of x, by the choice of m and since A contains no AP_3's, $|W_x \cap A| \le (2m+1)(c_3+\varepsilon)$ and hence $E_x \le \varepsilon$. For the other $2mq$ values of x, the bound $E_x \le 1 - c_3 \le 1$ (since $|W_x| \le 2m+1$) will be used.

Each $u \in A$ appears in exactly $2m+1$ different sets W_x and so double counting the elements of $W_x \cap A$,

$$\frac{1}{2m+1}\sum_{x=0}^{2N-1}|W_x \cap A| = \frac{|A|(2m+1)}{2m+1} = |A|.$$

Thus the average value of E_x is

$$\begin{aligned}\frac{1}{2N}\sum_{x=0}^{2N-1} E_x &= \frac{1}{2N}\sum_{x=0}^{2N-1}\left(\frac{|W_x \cap A|}{2m+1} - c_3\right)\\ &= \frac{|A|}{2N} - c_3\\ &\ge 0 && \text{(since } |A| \ge c_3 \cdot 2N\text{)}.\end{aligned}$$

For each $x \in [0, 2N-1]$, let $E_x^+ = \max\{0, E_x\}$. Since the average value of E_x is positive, a bound on $\sum |E_x|$ can be found as follows.

$$\begin{aligned}
\sum_{x=0}^{2N-1} |E_x| &\leq 2\sum_{x=0}^{2N-1} E_x^+ \\
&\leq 2\left(\sum_{x=0}^{mq} E_x^+ + \sum_{x=mq+1}^{2N-mq-1} E_x^+ + \sum_{x=2N-mq}^{2N-1} E_x^+\right) \\
&\leq 2(mq + 2N\varepsilon + mq) \\
&\leq 4N\varepsilon + 4mM \quad \text{(since } q \leq M) \\
&\leq 5N\varepsilon \qquad \text{(taking } N \geq (4m/\varepsilon)^2). \qquad (4.12)
\end{aligned}$$

When $\alpha \in \{\frac{1}{2N}, \frac{2}{2N}, \ldots, \frac{2N-1}{2N}\}$, since $e(\alpha)$ is a root of unity it follows that $\sum_{x=0}^{2N-1} e(\alpha x) = 0$. Therefore, for $\alpha \in \{\frac{1}{2N}, \ldots, \frac{2N-1}{2N}\}$,

$$\begin{aligned}
\left|\sum_{x=0}^{2N-1} e(\alpha x)\frac{|W_x \cap A|}{2m+1}\right| &= \left|\sum_{x=0}^{2N-1} e(\alpha x)(E_x + c_3)\right| \\
&= \left|\sum_{x=0}^{2N-1} e(\alpha x)E_x\right| \qquad \text{(since } \sum_{x=0}^{2N-1} e(\alpha x) = 0) \\
&\leq \sum_{x=0}^{2N-1} |e(\alpha x)E_x| \\
&= \sum_{x=0}^{2N-1} |E_x| \qquad \text{(since } |e(\alpha x)| = 1) \\
&\leq 5N\varepsilon \qquad \text{(by eq'n (4.12))}. \qquad (4.13)
\end{aligned}$$

Recall that by equations (4.10) and (4.11),

$$\left|f(\alpha) - \sum_{x=0}^{2N-1} e(\alpha x)\frac{|W_x \cap A|}{2m+1}\right| \leq 2\pi^2 m^2.$$

Therefore, by the triangle inequality, when $\alpha \in \{1/2N, \dots, (2N-1)/2N\}$,

$$\begin{aligned} |f(\alpha)| &\leq 2\pi^2 m^2 + \left| \sum_{x=0}^{2N-1} e(\alpha x) \frac{|W_x \cap A|}{2m+1} \right| \\ &\leq 2\pi^2 m^2 + 5N\varepsilon && \text{(by eq'n (4.12))} \\ &\leq 6N\varepsilon && \text{(taking } N \geq 2\pi^2 m^2/\varepsilon). \qquad (4.14) \end{aligned}$$

Returning to the inequality (4.7) given in the idea of the proof,

$$\begin{aligned} (c_3)^3 N^3 &\leq f(0)g(0)^2 && \text{(by (4.6))} \\ &\leq \left| \sum\nolimits^* f(\alpha)g(-\alpha)^2 \right| + \left| \sum_{\alpha \neq 0}^* f(\alpha)g(-\alpha)^2 \right| && \text{(triangle inequality)} \\ &\leq \left| \sum\nolimits^* f(\alpha)g(-\alpha)^2 \right| + \max_{\alpha \neq 0} |f(\alpha)| \sum\nolimits^* |g(-\alpha)^2| \\ &\leq 3c_3 N^2 + 6N\varepsilon \cdot 3c_3 N^2 && \text{(by (4.4), (4.14), and (4.5))} \\ &= 3c_3 N^2 + 18 c_3 \varepsilon N^3. \end{aligned}$$

Now,

$$\begin{aligned} (c_3)^3 N^3 \leq 3c_3 N^2 + 18 c_3 \varepsilon N^3 &\Leftrightarrow (c_3)^2 N \leq 3 + 18\varepsilon N \\ &\Leftrightarrow (c_3)^2 N \left(1 - \frac{18}{20}\right) \leq 3 \\ &\Leftrightarrow N \leq \frac{30}{(c_3)^2}. \end{aligned}$$

Therefore, taking N large enough, the assumption that $c_3 > 0$ leads to an absurd inequality. Therefore, $c_3 = 0$ and Roth's theorem holds. □

In 1987, Heath-Brown [61] published a proof of the following improvement on Roth's bound for $r_3(n)$ (Theorem 4.2.1). His result is presented here without proof.

Theorem 4.2.3 (Heath-Brown [61])**.** There are positive constants c, c' so that for n sufficiently large, if $A \subseteq [1, n]$ with $A \geq \frac{cn}{(\ln n)^{c'}}$, then A contains an AP_3. That is,

$$r_3(n) \leq \frac{cn}{(\ln n)^{c'}}.$$

Szemerédi showed that the previous theorem holds with the constant $c' = \frac{1}{20}$ (see [61]).

In 1999, Bourgain [17] gave what is the currently best known bound for $r_3(n)$ using analytic techniques.

Theorem 4.2.4 (Bourgain [17])**.** There is a positive constant c such that for n sufficiently large, if $A \subseteq [1, n]$ with $|A| > cn(\frac{\ln \ln n}{\ln n})^{1/2}$, then A contains an AP_3. That is,

$$r_3(n) < cn\left(\frac{\ln \ln n}{\ln n}\right)^{1/2}.$$

4.3 Szemerédi's theorem

In 1969, Szemerédi [107] was able to extend Roth's theorem to the case $k = 4$, giving a combinatorial proof that $r_4(n) = o(n)$. Roth ([96, 97]) soon after gave an analytic proof of the same result. Then, in 1975 , Szemerédi [108] extended his result to arbitrary k.

Theorem 4.3.1 (Szemerédi [108])**.** Let $A \subseteq \mathbb{Z}^+$ have positive upper density. Then for every $k \in \mathbb{Z}^+$, A contains an AP_k.

By Lemma 4.1.7, Szemerédi's theorem is equivalent to the statement that for all $k \in \mathbb{Z}^+$, $\lim_{n\to\infty} \frac{r_k(n)}{n} = 0$. Later, in 1977, Furstenberg [44] gave an ergodic theory proof of Szemerédi's theorem. More recently, in 2001, Gowers [47] published an analytic proof of Szemerédi's theorem that generalized Roth's technique. Gowers' bound gives some of the best known upper bounds on the van der Waerden numbers $W(k;t)$.

Theorem 4.3.2 (Gowers [47])**.** For each $k \in \mathbb{Z}^+$, set $c(k) = 2^{-2^{k+9}}$. Then for n sufficiently large, if $A \subseteq [1,n]$ with $|A| > \frac{n}{(\log_2 \log_2 n)^{-c(k)}}$ then A contains an AP_k. That is,

$$r_k(n) \leq \frac{n}{(\log_2 \log_2 n)^{-c(k)}}.$$

In Szemerédi's 1969 paper [107], he described how his combinatorial proof of Theorem 4.3.1 in the case $k = 4$ could be adapted to give another proof of Roth's theorem (Theorem 4.2.2) that $r_3(n) = o(n)$.

Theorem 4.2.2 (Roth [93])**.** For every $\varepsilon > 0$, when n is sufficiently large, for any set $A \subseteq [1,n]$ with $|A| \geq \varepsilon n$, then A contains an AP_3. That is $c_3 = \lim_{n\to\infty} \frac{r_3(n)}{n} = 0$.

Szemerédi's proof of Theorem 4.2.2. The details of the following proof appear in full in Graham, Rothschild and Spencer [51, pp. 48–9]. The idea is to look at arithmetic progressions built inside an affine cube. Recall the corollary to Szemerédi's cube lemma, Corollary 2.2.4, which states that for every $\varepsilon > 0$ and $n \in \mathbb{Z}^+$, there is a constant c such that if $m \leq \log_2 \log_2 n + c$, and $A \subseteq [1, n]$ with $|A| \geq \varepsilon n$, then A contains an affine m-cube. In this proof, for simplicity, write $\log_2 n = \log n$.

Again, set $c_3 = \lim_{n\to\infty} \frac{r_3(n)}{n}$ and suppose that $c_3 > 0$. Set $\varepsilon = c_3^2/200 > 0$ and let $m \in \mathbb{Z}^+$ be such that for all $n \geq m$,

$$c_3 \leq \frac{r_3(n)}{n} < c_3 + \varepsilon.$$

Let N be at least large enough so that $0.01c_3^2 \log\log N > m$ (it may be necessary to make N even larger later). Let $A \subseteq [1, N]$ be AP_3-free with $|A| = r_3(N) \geq c_3 N$. Since the function r_3 is subadditive and A is AP_3-free on the intervals $[1, 0.49N] \cup [0.5N, N]$, A contains at most $r_3(0.49N) + r_3(0.5N) \leq 0.99r_3(N) < 0.99N(c_3 + \varepsilon)$ elements.

Since $|A| \geq c_3 N$, within the interval $(0.49N, 0.5N)$, A has at least $c_3 N - 0.99N(c_3 + \varepsilon) = 0.01Nc_3 - 0.99N\varepsilon$ elements and so density at least $c_3 - 99\varepsilon = c_3(1 - \frac{99c_3}{200}) \geq \frac{c_3}{2}$ since $c_3 < 1$.

Divide $(0.49N, 0.5N)$ into smaller intervals, each of length $N^{1/2} + O(1)$. On one of these subintervals, A has density at least $c_3/2$ and so by Szemerédi's cube lemma (Lemma 2.2.3), possibly requiring N to be larger, there is a $k \in \mathbb{Z}^+$ with $k = \log\log N^{1/2} + O(1) = \log\log N + O(1)$ such that there is an affine k-cube

$H = H(a, d_1, \ldots, d_k) \subseteq A \cap (0.49N, 0.5N)$. If N is large enough so that $N^{1/2} \geq O(1)$, then for $i \in [1, k]$, $d_i \leq N^{1/2} + O(1) \leq 2N^{1/2}$.

Set $H_0 = \{a\}$ and for $i \in [1, k]$, define an affine i-cube, $H_i = H(a, d_1, \ldots, d_i)$. For $i \in [0, k]$, set $L_i = \{2h - x : x \in A, x < a, \text{ and } h \in H_i\}$. In particular, L_i is the set of third terms of arithmetic progressions $\{x, h, y\}$ with $x \in A$ and $h \in H_i \subseteq A$.

Since A is AP_3-free, $A \cap L_i = \emptyset$ and since A has density at least $\frac{c_3}{2}$ on $[1, 0.49N]$,

$$\begin{aligned} |L_i| \geq |L_0| = |A \cap [1, a)| &\geq \frac{c_3}{2}(0.49N) \qquad\qquad \text{(since } a < 0.49N\text{)} \\ &= c_3(0.245N). \end{aligned} \tag{4.15}$$

For each $i \in [0, k-1]$, $H_{i+1} = H_i \cup (H_i + d_i)$ and so $L_{i+1} = L_i \cup (L_i + 2d_i)$. Therefore, $L_0 \subseteq L_1 \subseteq \cdots \subseteq L_k \subseteq [1, N]$ and by averaging, there is an $i \in [0, k-1]$ with

$$|L_{i+1} \backslash L_i| < \frac{N}{k}.$$

Partition the set L_i into sets of maximal arithmetic progressions with difference $2d_i$. For each maximal arithmetic progression $\{x, x + 2d_i, \ldots, x + s(2d_i)\}$ of L_i, $x + (s+1)(2d_i) \in L_{i+1} \backslash L_i$. Similarly, each element of $L_{i+1} \backslash L_i$ corresponds to a maximal arithmetic progression in L_i. Therefore, L_i is partitioned into less than N/k classes.

Consider one of the residue classes $\pmod{2d_i}$ of $[1, N]$. If a particular residue class $\pmod{2d_i}$ contains t partition classes of L_i, then $[1, N] \backslash L_i$ can be partitioned into at most $t+1$ maximal arithmetic progressions with difference $2d_i$ contained in that residue class.

Thus, $[1,N]\backslash L_i$ can be partitioned into at most $\frac{N}{k}+2d_i$ maximal arithmetic progressions (at most one more than L_i for each of the $2d_i$ residue classes).

Since $d_i < 2N^{1/2}$ and $k = \log\log N + O(1)$,

$$\begin{aligned}\frac{N}{k}+2d_i &< \frac{N}{\log\log N + O(1)} + 2N^{1/2} \\ &= \frac{N}{\log\log N}\left(\frac{\log\log N}{\log\log N + O(1)} + \frac{2\log\log N}{N^{1/2}}\right) \\ &= \frac{N}{\log\log N}(1+o(1)).\end{aligned}$$

Define a partition class of the set $[1,N]\backslash L_i$ to be *small* if it contains less than $0.01c_3^2\log\log N$ elements and *large* otherwise.

All the small classes together have at most

$$0.01c_3^2\log\log N\left(\frac{N}{\log\log N}(1+o(1))\right) = 0.01c_3^2N + o(N)$$

elements.

Since A is AP_3-free and $0.01c_3^2\log\log N > m$, on every large partition class, A has density less than $c_3+\varepsilon$. Thus on the union of the large classes, A has density less than $c_3+\varepsilon$.

Therefore, since $A\cap L_i = \emptyset$,

$$|A| = |A \cap ([1,N]\backslash L_i)|$$

$$
\begin{aligned}
&< \underbrace{(c_3+\varepsilon)(N-|L_i|)}_{\text{large blocks}} + \underbrace{0.01c_3^2N + o(N)}_{\text{small blocks}} \\
&\leq \left(c_3 + \frac{c_3^2}{200}\right)(N - c_3(0.245N)) + 0.01c_3^2N + o(N) \quad (\text{by (4.15) and } \varepsilon = \frac{c_3^2}{200}) \\
&= c_3N + \frac{c_3^2N}{200} - c_3^2 0.235N - \frac{c_3^3}{200}0.245N + o(N) \\
&< c_3N - 0.23c_3^2N + o(N) \\
&\leq c_3N \qquad\qquad (\text{when } N \text{ is large}).
\end{aligned}
$$

As this contradicts the choice of the set A with $|A| \geq c_3N$, the initial assumption that $c_3 > 0$ is false and so $c_3 = 0$. □

4.4 Varnavides

In 1955, Varnavides [115] used Roth's theorem to show that for any fixed ε, if n is large enough and $|A \cap [1,n]| \geq \varepsilon n$, then A contains not only one arithmetic progression of length 3, but in fact on the order of n^2 arithmetic progressions. Using Szemerédi's theorem, Varnavides' technique can be used to show that the same is true of arithmetic progressions of any fixed length k (see also [116]).

Theorem 4.4.1 (Varnavides [115])**.** Let $k \in \mathbb{Z}^+$ and $\varepsilon > 0$ be given. There is a positive constant $c = c(k, \varepsilon)$ such that for n sufficiently large, if $A \subseteq [1,n]$ satisfies $|A| \geq \varepsilon n$, then A contains at least cn^2 AP_k's.

Proof. By Szemerédi's theorem (Theorem 4.3.1), fix $\ell \in \mathbb{Z}^+$ so that if $X \subseteq [1,\ell]$

with $|X| \geq \frac{\varepsilon\ell}{2}$, then X contains an AP_k. Let $N > \ell$ and let $A \subseteq [1, N]$ be such that $|A| > \varepsilon N$. Call an AP_ℓ's in $[1, N]$ a *good progression* if it contains at least $\frac{\varepsilon\ell}{2}$ elements of A and a *bad progression* otherwise. Note that by the choice of ℓ, every good progression contains at least one AP_k in A.

For each $d \in [1, \frac{N-1}{\ell-1}]$, let $G_d(N)$ be the number of good progressions in $[1, N]$ with difference d. Define $f(a, d) = |A \cap \{a, a+d, \ldots, a+(\ell-1)d\}|$. The idea of the proof is to find both upper and lower bounds for $\sum_{a \geq 1} f(a, d)$ in terms of the number of good progressions in order to get a bound on $G_d(N)$. For a fixed $x \in A$, x can occur in at most ℓ different AP_ℓ's (ℓ possible different positions) and only the numbers x with $1+(\ell-1)d \leq x \leq N-(\ell-1)d$ will occur in exactly ℓ different AP_ℓ's. There are at most $2(\ell-1)d \leq 2\ell d$ elements of A outside the interval $[1+(\ell-1)d, N-(\ell-1)d]$. Therefore, at least $\varepsilon N - 2\ell d$ elements of A appear in exactly ℓ different AP_ℓ's and so

$$\begin{aligned}\sum_{a=1}^{N-(\ell-1)d} f(a,d) &> \ell(\varepsilon N - 2\ell d)\\ &> \ell\left(\varepsilon N - \frac{2\ell\varepsilon N}{8\ell}\right) \qquad \text{(restricting to } d < \frac{\varepsilon N}{8\ell})\\ &= \frac{3}{4}\ell\varepsilon N.\end{aligned}$$

For the upper bound,

$$\sum_{a=1}^{N-(\ell-1)d} f(a,d) < \underbrace{N\frac{\varepsilon\ell}{2}}_{\text{bad AP's}} + \underbrace{G_d(N)\ell}_{\text{good AP's}}.$$

Combining these two inequalities gives $\frac{3}{4}\varepsilon\ell N < \frac{1}{2}\varepsilon\ell N + G_d(N)\ell$, which implies that $G_d(N) > \frac{1}{4}\varepsilon N$.

The total number of good progressions in $[1, N]$ is

$$\begin{aligned}\sum_{d=1}^{\frac{N-1}{\ell-1}} G_d(N) &> \sum_{d=1}^{\frac{\varepsilon N}{8\ell}} G_d(N)\\ &> \frac{1}{4}\varepsilon N \frac{\varepsilon N}{8\ell}\\ &= \frac{\varepsilon^2 N^2}{32\ell}.\end{aligned}$$

By the choice of ℓ, each good progression contains an AP_k in the set A. In order to find a lower bound for the number of AP_k's in A, it suffices to determine how many different good progressions could contain a particular AP_k. Fix an AP_k, $P = \{a, a+d, \ldots, a+(k-1)d\}$ and consider all AP_ℓ's containing P.

Case I: AP_ℓ's with difference d that contain P.

There are exactly $\ell - k + 1$ different positions the first term of P could occupy in an AP_ℓ with difference d and hence $\ell - k + 1$ different AP_ℓ's with difference d that contain P.

Case II: AP_ℓ's with a difference other than d that contain P.

Suppose that $d' \neq d$ and $P \subseteq \{b, b+d', \ldots, b+(\ell-1)d'\}$. Then there are integers i, j with $0 \leq i < j \leq \ell - 1$ so that $a = b + id'$ and $a + d = b + jd'$. Therefore, $d = (a+d) - a = (b+jd') - (b+id') = (j-i)d'$. Since $\ell d' > (\ell-1)d' \geq (k-1)d =$

$(k-1)(j-i)d'$, it follows that $(j-i) < \frac{\ell}{k-1}$. Therefore, there are less than $\frac{\ell}{k-1}$ possible values of d' where P is contained in an AP_ℓ with difference d'. As before, for a fixed value of d', there are at most $\ell - k + 1$ different AP_ℓ's with difference d' that contain P. Thus P is contained in at most $\frac{\ell}{k-1}(\ell - k + 1)$ different AP_ℓ's with a difference other than d.

Therefore, in total, P can appear in at most

$$\begin{aligned}(\ell - k + 1) + \frac{\ell}{k-1}(\ell - k + 1) &= \frac{1}{k-1}(\ell - k + 1)(\ell + k - 1) \\ &= \frac{\ell^2}{k-1} - (k-1) \\ &\leq \frac{\ell^2}{k-1}\end{aligned}$$

different good progressions.

Therefore, there are at least $\frac{\varepsilon^2 N^2}{32\ell} \cdot \frac{k-1}{\ell^2} = \frac{\varepsilon^2(k-1)}{32\ell^3} N^2$ AP_k's in A. Take $c = \frac{\varepsilon^2(k-1)}{32\ell^3}$. Then c is a constant, depending only on ε and k, so that for all N sufficiently large and $A \subseteq [1, N]$ with $|A| \geq \varepsilon N$, the set A contains at least cN^2 different k-term arithmetic progressions. □

4.5 Systems of equations

Both partition and density theorems can be phrased in terms of solutions of matrix equations. Note that distinct numbers $x_1, \ldots, x_k$ form an AP_k iff $[x_1 \; x_2 \; \ldots \; x_k]^T$ is a non-trivial solution to the system

$$\begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.16}$$

The problem of finding solutions to such systems of equations within a particular set can be generalized.

Definition 4.5.1. Let A be an $\ell \times k$ matrix with integral coefficients. Then A is *partition regular in* $\mathbb{Z}^+$ iff for every finite colouring of $\mathbb{Z}^+$ (it is important here that $0 \notin \mathbb{Z}^+$), there is a monochromatic set $\{x_1, \ldots, x_k\}$ such that $A[x_1 \ \cdots \ x_k]^T = [0 \ \cdots \ 0]^T$. Similarly, A is *density regular in* $\mathbb{Z}^+$ iff for every $X \subseteq \mathbb{Z}^+$ with positive upper density, there is a set $\{x_1, \ldots, x_k\} \subseteq X$ with $A[x_1 \ \cdots \ x_k]^T = [0 \ \cdots \ 0]^T$.

Rado [88] characterised the partition regular equations with a set of linear relations between the matrix entries called the *columns condition.* Deuber [29] later proved another characterisation of partition regularity in terms of collections of integers called (m, p, c)*-sets.*

Definition 4.5.2. Given a matrix A with integral coefficients, the system $A\boldsymbol{x} = \mathbf{0}$ is *irredundant* iff there is a solution $\boldsymbol{x} = [x_1 \ \ldots \ x_k]^T$ with not all x_i's equal. A solution $\boldsymbol{x} = [x_1 \ \ldots \ x_k]^T$ is *proper* iff all x_i's are distinct.

The following result (given without proof) extends Varnavides' result to other systems of equations in the integers.

Theorem 4.5.3 (Frankl, Graham and Rödl [43]). Let A be a density regular $\ell \times k$ matrix with rank ℓ so that the system $A\boldsymbol{x} = \mathbf{0}$ is irredundant. Then for every $\varepsilon > 0$, there is a constant $c > 0$ that depends on A and ε, so that for n sufficiently large, if $X \subseteq [1, n]$ with $|X| \geq \varepsilon n$, then X must contain at least $cn^{k-\ell}$ proper solutions to the system $A\boldsymbol{x} = \mathbf{0}$.

Consider the $k \times (k-2)$ matrix in equation (4.16). The corresponding homogeneous system is density regular by Szemerédi's theorem (Theorem 4.3.1), which guarantees non-constant arithmetic progressions and so the system $A\boldsymbol{x} = \mathbf{0}$ is irredundant. Therefore, by Theorem 4.5.3, for every $\varepsilon > 0$, there is a constant c so that for all n sufficiently large, if $X \subseteq [1, n]$ with $|X| \geq \varepsilon n$, then X contains at least $cn^{k-(k-2)} = cn^2$ proper solutions to $A\boldsymbol{x} = \mathbf{0}$. That is, X contains at least cn^2 arithmetic progressions — precisely the statement of Theorem 4.4.1.

4.6 Lower bounds

Behrend adapted techniques from an article by Salem and Spencer [98] to find a bound of the function $r_3(n)$ (recall Definition 4.1.1).

Theorem 4.6.1 (Behrend [12])**.** There is a positive constant c so that for n sufficiently large, there is an AP_3-free set $A \subseteq [1, n]$ with $|A| \geq ne^{-c\sqrt{\ln n}}$. That is, $r_3(n) > ne^{-c\sqrt{\ln n}}$.

Proof. Fix $n \in \mathbb{Z}^+$, let $b \in \mathbb{Z}^+$ and fix $m \in \mathbb{Z}^+$ so that $(2b-1)^{m-1} < n \leq (2b-1)^m$. That is,

$$\frac{\ln n}{\ln(2b-1)} \leq m < \frac{\ln n}{\ln(2b-1)} + 1. \tag{4.17}$$

To each $x \in [1, n]$ associate an m-tuple $(x_0, \ldots, x_{m-1}) \in [0, 2b-2]^m$ corresponding to the base $(2b-1)$ representation of x so that $x = \sum_{i=0}^{m-1} x_i(2b-1)^i$.

For each x, set $M(x) = [\sum_{i=0}^{m-1} x_i^2]^{1/2}$ and for each $s \geq 1$, set

$$A_s = \{x \in [1, n] : \text{for each } i \in [0, m-1],\ 0 \leq x_i \leq b-1 \text{ and } M(x)^2 = s\}.$$

Suppose that for some $s \geq 1$, there were $x, y, z \in A_s$ with $x + y = 2z$. Let $(x_0, \ldots, x_{m-1}), (y_0, \ldots, y_{m-1})$ and $(z_0, \ldots, z_{m-1})$ be the m-tuples corresponding to x, y and z respectively. For each $i \in [0, m-1]$, since $x_i, y_i \leq b-1$, $x_i + y_i \leq 2b-2$ and so there is no carrying in the base $(2b-1)$ addition. Therefore, for each $i \in [0, m-1]$, $x_i + y_i = 2z_i$ and so since $M(x) = M(y) = M(z)$,

$$\begin{aligned} M(x+y) = \left[\sum_{i=0}^{m-1} (x_i + y_i)^2\right]^{1/2} &= \left[\sum_{i=0}^{m-1} (2z_i)^2\right]^{1/2} \\ &= 2M(z) \\ &= M(x) + M(y). \end{aligned}$$

This is only possible when the m-tuples $(x_0,\ldots,x_{m-1})$ and $(y_0,\ldots,y_{m-1})$ are proportional, but since $M(x)=M(y)$, it must be that $(x_0,\ldots,x_{m-1})=(y_0,\ldots,y_{m-1})$ and hence $x=y$. Therefore, for all $s\geq 1$, the set A_s is AP_3-free.

Since

$$\bigcup_{s\geq 1} A_s=\left\{\sum_{i=0}^{m-1} x_i(2b-1)^i : (x_0,\ldots,x_{m-1})\in[0,b-1]^m\backslash\{(0,\ldots,0)\}\right\},$$

it follows that $|\cup_{s\geq 1}A_s|=b^m-1$. When $s>(b-1)^2m$, $A_s=\emptyset$ and so by averaging, there is at least one choice of $s\in[1,(b-1)^2m]$ with

$$|A_s|\geq\frac{b^m-1}{(b-1)^2m}.$$

Since A_s is AP_3-free, $r_3(n)\geq|A_s|$ and so for any $n\in\mathbb{Z}^+$ and any choice of b, with m defined as above,

$$\begin{aligned}
r_3(n)&\geq\frac{b^m-1}{(b-1)^2m}\\
&>\frac{b^{m-2}}{m}\\
&\geq\frac{(n^{1/m}+1)^{m-2}}{2^{m-2}m} && \text{(since } n\leq(2b-1)^m\text{)}\\
&=\frac{n^{\frac{m-2}{m}}}{2^{m-2}m}(1+n^{-1/m})^{m-2}\\
&\geq\frac{n^{1-2/m}}{2^{m-2}m} && \text{(since } 1+n^{-1/m}\geq 1\text{)}\\
&=ne^{-\frac{2}{m}\ln n-(m-2)\ln 2-\ln m}. && (4.18)
\end{aligned}$$

In particular, if $b=\lfloor\frac{1}{2}(e^{\sqrt{\ln n}}+1)\rfloor$, then $\ln(2b-1)\leq\sqrt{\ln n}$ and so by equation (4.17),

$$m \geq \frac{\ln n}{\ln(2b-1)} \geq \ln n\sqrt{\ln n} = \sqrt{\ln n}. \tag{4.19}$$

A small calculation shows that for the choice of b, if $n \geq 4$, then $\ln(2b-1) > \sqrt{\ln n} - 1$ and so

$$\begin{aligned} m < \frac{\ln n}{\ln(2b-1)} + 1 &< \frac{\ln n}{\sqrt{\ln n} - 1} + 1 && \text{(by eq'n (4.17))} \\ &= \sqrt{\ln n} + \frac{1}{\sqrt{\ln n} - 1} + 2 \\ &\leq \sqrt{\ln n} + 3 && \text{(for } \sqrt{\ln n} > 1\text{).} \qquad (4.20) \end{aligned}$$

Thus, by equations (4.18),(4.19) and (4.20),

$$\begin{aligned} r_3(n) &> ne^{-\frac{2}{m}\ln n-(m-2)\ln 2-\ln m} \\ &\geq ne^{-\frac{2}{\sqrt{\ln n}}\ln n-(\sqrt{\ln n}+1)\ln 2-\ln(\sqrt{\ln n}+3)} \\ &= ne^{-2\sqrt{\ln n}-\ln 2\sqrt{\ln n}-\ln 2-\ln(\sqrt{\ln n}+3)} \\ &= ne^{-\sqrt{\ln n}\left(2+\ln 2+\frac{\ln 2}{\sqrt{\ln n}}+\frac{\ln(\sqrt{\ln n}+3)}{\sqrt{\ln n}}\right)} \\ &> ne^{-\sqrt{\ln n}(4+\ln 2)} \end{aligned}$$

when n is large enough. □

Moser [83] applied the same technique to give a lower bound for the van der Waerden numbers. He showed that for some fixed constant $c > 0$,

$$W(k;r) > (k-1)r^{c\ln r}.$$

In 1960, Rankin [92] presented the following extension of Theorem 4.6.1 for $k \geq 3$. Behrend's theorem (Theorem 4.6.1) is the case $k = 3$ in Theorem 4.6.2.

Theorem 4.6.2 (Rankin [92])**.** Let $k \geq 3$, there is a positive constant c so that for n sufficiently large, there is an AP_k-free set $A \subseteq [1, n]$ with $|A| \geq ne^{-c(\ln n)^{1/\lceil \log_2 k \rceil}}$. That is,

$$r_k(n) \geq ne^{-c(\ln n)^{1/\lceil \log_2 k \rceil}}.$$

4.7 Arithmetic progressions of primes

It seems that one of the motivations behind studying the function $r_k(n)$ was the possibility of determining whether the set of primes contains arbitrarily long arithmetic progressions. It is also worth noting that the set of primes contains no infinite arithmetic progression, for if p is any prime and $d \in \mathbb{Z}^+$, the $(p+1)$-th term of the arithmetic progression $\{p + id : i \geq 0\}$, is $p + pd = p(1 + d)$ which cannot be prime since $p \geq 1$ and $d + 1 \geq 1$. Dirichlet's theorem (see, for example, [1, Chapter 7]) states that for any relatively prime integers a and b, there are infinitely many primes in the arithmetic progression $\{a + ib : i \geq 0\}$. However, this does not guarantee that any of those infinitely many primes are themselves in arithmetic progression.

In 1939, van der Corput [112] showed that there are infinitely many AP_3's of primes. A stronger result along the same lines was that of Balog [3] who showed that for any $k \in \mathbb{Z}^+$, there are k primes $p_1, \ldots, p_k$ such that all of the averages $\frac{1}{2}(p_i + p_j)$ are also prime. In other words, for any integers i, j with $1 \leq i < j \leq k$, the set $\{p_i, \frac{1}{2}(p_i + p_j), p_j\}$ is an AP_3 of primes.

One of the approaches towards solving the general conjecture about primes in arithmetic progression has been to consider the function $\pi(n)$, the number of primes in the interval $[1, n]$. The prime number theorem, which was proved by Hadamard [57] and de la Vallée Poussin [28] in 1896, states that $\pi(n) \sim \frac{n}{\ln n}$ (see, for example, [1, Chapter 13]). If it could be shown that for all k, $r_k(n) < \pi(n)$ when n is large enough, this would guarantee that the set of primes contains arbitrarily long arithmetic progressions simply by virtue of their density in the integers. Although a proof of this type has not yet been found, in 2005, Green [53] published an analytic proof that all sets of primes A with $\limsup_{n\to\infty} \frac{|A\cap[1,n]|}{\pi(n)} > 0$ contain infinitely many AP_3's. The same year, Green and Tao [54] settled the conjecture that the primes contain arbitrarily long arithmetic progressions by giving a Szemerédi-type theorem for the primes.

Theorem 4.7.1 (Green and Tao [54])**.** Let A be an infinite set of primes with

$$\limsup_{n\to\infty} \frac{|A \cap [1, n]|}{\pi(n)} > 0,$$

then A contains arbitrarily long arithmetic progressions.

More details about the result and the history of the problem can be found in [78]. In 2006, Tao was awarded a Fields Medal for numerous contributions to various areas of mathematics, including Theorem 4.7.1 (see, for example, [69]). Green and Tao also showed that a careful analysis of their proof of Theorem 4.7.1 gave the following.

Theorem 4.7.2 (Green and Tao [55])**.** For each $k \in \mathbb{Z}^+$, there is an arithmetic progression of primes of length k with all terms less than

$$2^{2^{2^{2^{2^{2^{2^{100k}}}}}}}.$$

Chapter 5

Van der Waerden numbers

5.1 Exact values

The only five non-trivial van der Waerden numbers whose values are known precisely are $W(3;2) = 9$, $W(4;2) = 35$, $W(5;2) = 178$, $W(3;3) = 27$ and $W(3;4) = 76$. Recently, Kouril claims to have found $W(6;2) = 1132$ (see [62]), but the result has yet to be published. More can be said if the van der Waerden numbers are generalized in a fashion similar to that of the Ramsey numbers (Definition 2.1.5).

Definition 5.1.1. For each $r, k_1, k_2, \ldots, k_r \in \mathbb{Z}^+$, let $W(k_1, \ldots, k_r; r)$ be the least integer such that for any $n \geq W(k_1, \ldots, k_r; r)$ and any r-colouring $\Delta : [1, n] \rightarrow [1, r]$, there is an $i \in [1, r]$ such that $\Delta^{-1}(i)$ contains an AP_{k_i}. The numbers $W(k_1, \ldots, k_r; r)$ are called the *mixed van der Waerden numbers*.

That is, instead of looking for an arithmetic progression of a fixed length in any colour class, look only for AP_{k_i}'s in the i-th colour class. The mixed van der Waerden numbers will always exist because if $k_1, \ldots, k_r \in \mathbb{Z}^+$, then $W(k_1, \ldots, k_r; r) \leq W(\max\{k_1, \ldots, k_r\}; r)$. The mixed van der Waerden numbers enjoy the same symmetry properties as the Ramsey numbers. For any integers $k_1, \ldots, k_r$, if σ is a permutation of $[1, r]$, then $W(k_1, \ldots, k_r; r) = W(k_{\sigma(1)}, \ldots, k_{\sigma(r)}; r)$.

Of course, some values of the van der Waerden numbers are trivial. For any $r \in \mathbb{Z}^+$, $W(2; r) = r + 1$ by the pigeonhole principle. It can be shown that for any odd $k \in \mathbb{Z}^+$, $W(k, 2; 2) = 2k$ as follows: Suppose that the two colours are red and blue. The 2-colouring of $[1, 2k - 1]$ where the number k is blue and all other numbers are red has no blue pairs and no red AP_k's. However, for any 2-colouring of $[1, 2k]$, if there is no pair of elements that are blue, then either the first k or the last k numbers in the interval are all red. A similar argument shows that for any even $k \in \mathbb{Z}^+$, $W(k, 2; 2) = 2k - 1$.

This idea was extended by Culver, Landman and Robertson [26] who found exact values for many mixed van der Waerden numbers of the form $W(k, 2, \ldots, 2; r)$. Their results included the following which are presented without proof.

Theorem 5.1.2 (Culver, Landman and Robertson [26]). Fix $k > r \geq 2$. Set $m = \prod\{p : p \leq r,\ p \text{ prime}\}$.

(i) If $\gcd(k, m) = 1$, then $W(k, 2, \ldots, 2; r) = rk$.

(ii) Otherwise, if either (a) $\gcd(k-1,m)=1$ or (b) r is prime and $\gcd(k,m)=r$, then $W(k,2,\ldots,2;r)=rk-r+1$.

According to Landman (personal communication), the results of Theorem 5.1.2 were extended in 2007 by Khodkar and Landman [76].

k_1	k_2	k_3	$W(k_1,\ldots,k_r;r)$	k_1	k_2	k_3	k_4	k_5	$W(k_1,\ldots,k_r;r)$
3	3	-	9	5	2	2	-	-	15 [26]
4	3	-	18 [22]	5	3	2	-	-	32 [19]
4	4	-	35 [22]	5	3	3	-	-	80 [26]
5	3	-	22 [22]	5	4	2	-	-	71 [19]
5	4	-	55 [22]	6	2	2	-	-	16 [26]
5	5	-	178 [105]	6	3	2	-	-	40 [19]
6	3	-	32 [22]	6	4	2	-	-	83 [26]
6	4	-	73 [10]	7	2	2	-	-	21 [26]
7	3	-	46 [22]	7	3	2	-	-	55 [26]
7	4	-	109 [9]	8	2	2	-	-	24 [26]
8	3	-	58 [10]	3	3	2	2	-	17 [19]
9	3	-	77 [10]	3	3	3	2	-	40 [19]
10	3	-	97 [10]	3	3	3	3	-	76 [10]
11	3	-	114 [26]	4	3	2	2	-	25 [19]
12	3	-	135 [26]	4	3	3	2	-	60 [26]
13	3	-	160 [26]	4	4	2	2	-	53 [19]
3	2	2	7 [26]	5	2	2	2	-	20 [26]
3	3	2	14 [19]	5	3	2	2	-	43 [19]
3	3	3	27 [22]	6	2	2	2	-	21 [26]
4	2	2	11 [26]	6	3	2	2	-	48 [26]
4	3	2	21 [19]	7	2	2	2	-	28 [26]
4	3	3	51 [19]	7	3	2	2	-	65 [26]
4	4	2	40 [19]	3	3	2	2	2	20 [26]
4	4	3	89 [26]	3	3	3	2	2	41 [26]

Figure 5.1: Exact van der Waerden numbers

The exact values in Figure 5.1 were given by Chvátal [22], Brown [19], Stevens and Shantaram [105], Beeler and O'Neil [10], Beeler [9], and Culver, Landman and Roberston [26] (see also [87]). Most exact values of mixed van der Waerden numbers were found using computers. A similar table of mixed van der Waerden numbers is given by Culver *et al.* [26]. The Figure 5.1 also includes the values for the numbers $W(k, 2, \ldots, 2; r)$ given by Theorem 5.1.2 that fit within the table.

Calculating exact values of the mixed van der Waerden numbers becomes computationally difficult as the values of k_i and r increase. The remainder of this chapter will explore different techniques that have been used to find either upper or lower bounds for the van der Waerden numbers.

5.2 Probabilistic method

One useful approach to calculating bounds is to assume that the colouring is done at random, so that each integer has an equal chance of being any one colour. Then, for example, if there is a positive probability that some interval $[1, n]$ contains no monochromatic arithmetic progression then there is at least one colouring of $[1, n]$ with no monochromatic arithmetic progressions.

First some definitions are given to formalize this idea. While it is possible to define concepts of probability for infinite sets, the purposes here require only finite probability spaces.

Definition 5.2.1. Let Ω be a finite set and $P : \Omega \to [0,1]$ (where $[0,1] \subseteq \mathbb{R}$) be any function such that $\sum_{x\in\Omega} P(x) = 1$. The pair (Ω, P) is called a *probability space.*

Most often here, the function P used will be the constant function $P(x) = \frac{1}{|\Omega|}$. For this probability function, (Ω, P) is called the *uniform probability space.*

Definition 5.2.2. A subset $A \subseteq \Omega$ is called an *event* and for any event A, define $P(A) = \sum_{x\in A} P(x)$.

Given two events A and B, the event $A \cup B$ (sometimes denoted by $A \vee B$) is the event that either A or B occurs, and $A \cap B$ (also denoted by $A \wedge B$) is the event that both A and B occur. For any event $A \subseteq \Omega$, define $\bar{A} = \Omega \backslash A$, the complement of A.

Definition 5.2.3 (Bayes' formula)**.** For two events A, B, the *conditional probability* of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

It is the probability that the event A occurs assuming that the event B does occur.

Definition 5.2.4. An event A is *independent* of an event B iff $P(A|B) = P(A)$ and A is *mutually independent* from events $B_1, \ldots, B_m$ iff A is independent of any boolean combination of the events $B_1, \ldots, B_m$.

Note that for any $A, B \subseteq \Omega$, since $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, if $\{A_i \subseteq \Omega : i \in I\}$ is any collections of events, then $P(\cup_{i\in I} A_i) \leq \sum_{i\in I} P(A_i)$.

The proof of the following theorem can be easily re-phrased as a counting argument employing the pigeonhole principle, but it serves as a useful demonstration of the probabilistic method.

Theorem 5.2.5 (Erdős-Rado [37])**.** For all $k, r \in \mathbb{Z}^+$, $W(k;r) > \sqrt{2(k-1)}r^{\frac{k-1}{2}}$.

Proof. Fix $k, r \in \mathbb{Z}^+$ and take $n \leq \sqrt{2(k-1)}r^{\frac{k-1}{2}}$. Define a probability space as follows. Let Ω be the set of all r-colourings of $[1,n]$ and let P be the uniform probability for Ω, *i.e.*, for all $x \in \Omega$, $P(x) = \frac{1}{r^n}$. For any AP_k, $S \subseteq [1,n]$, let A_S be the event that S is monochromatic. That is, A_S is the set of all r-colourings of $[1,n]$ where S is monochromatic. Then, $P(A_S) = \frac{r}{r^k} = r^{1-k}$.

By Lemma 1.2.3, the total number of AP_k's contained in $[1,n]$ is less than $\frac{n^2}{2(k-1)}$. Therefore, since $\cup A_S$ is the event that any AP_k in $[1,n]$ is monochromatic and

$$P(\cup A_S) \leq \sum P(A_S) < \frac{n^2}{2(k-1)}r^{1-k} < 1 \quad (\text{since } n \leq \sqrt{2(k-1)}r^{\frac{k-1}{2}}),$$

there is at least one r-colouring of $[1,n]$ with no monochromatic AP_k. Therefore, $W(k;r) \ > n.$ □

In the case of arithmetic progressions, since there is so much overlap, the events associated with two given AP_k's being monochromatic are often not independent. A probabilistic proof that does not take this into account seems likely to be imprecise. Thus a tool is needed that can take into account the interdependencies of a collection of events. The Lovász Local Lemma, proved by Erdős and Lovász in 1975 [36] is such a tool. Before proceeding to the statement of the lemma, a few definitions and identities about dependency are given.

Definition 5.2.6. Let $A_1, A_2, \ldots, A_n$ be events in a probability space. Let $V = [1, n]$ and let $E \subseteq [V]^2$. The graph (V, E) is a *dependency graph* for the events $A_1, \ldots, A_n$ iff for each $i \in [1, n]$, the event A_i is mutually independent from the events $\{A_j : \{i, j\} \notin E\}$. That is, A_i is independent from every boolean combination of the sets $\{A_j : \{i, j\} \notin E\}$.

The following two lemmas are standard applications of Bayes' formula (Definition 5.2.3.

Lemma 5.2.7. For any events A, B, and C in a probability space (Ω, P),

$$P(A|B \cap C) = \frac{P(A \cap B|C)}{P(B|C)}.$$

Proof. Let $A, B, C \subseteq \Omega$, then

$$\begin{aligned} P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= \frac{P(A \cap B \cap C)/P(C)}{P(B \cap C)/P(C)} \\ &= \frac{P(A \cap B|C)}{P(B|C)}. \end{aligned}$$ □

Note that rearranging terms also gives

$$P(A \cap B|C) = P(A|B \cap C) \cdot P(B|C).$$

Lemma 5.2.8. For any events $A_1, \ldots, A_n, B$ in a probability space (Ω, P),

$$P\left(\bigcap_{i=1}^{n} A_i | B\right) = \prod_{i=1}^{n} P\left(A_i | \bigcap_{j=i+1}^{n} A_j \cap B\right)$$

where $\bigcap\limits_{j=n+1}^{n} A_j \cap B = B$.

Proof. The proof is by induction on n.

Base Case: The case $n = 2$ holds by Lemma 5.2.7.

Inductive Step: Suppose that the lemma holds for some integer $n \geq 2$. Let $A_1, \ldots, A_n, A_{n+1}$ and B be any events. Then

$$\begin{aligned} P(\cap_{i=1}^{n+1} A_i | B) &= P(A_1 | \cap_{i=2}^{n+1} A_i \cap B) P(\cap_{i=2}^{n+1} A_i | B) && \text{(by Lemma 5.2.7)} \\ &= P(A_1 | \cap_{i=2}^{n+1} A_i \cap B) \prod_{i=2}^{n+1} P(A_i | \cap_{j=i+1}^{n+1} A_j \cap B) && \text{(by ind. hyp.)} \\ &= \prod_{i=1}^{n+1} P(A_i | \cap_{j=i+1}^{n+1} A_i \cap B). \end{aligned}$$

□

Theorem 5.2.9 (**Lovász Local Lemma**, Erdős and Lovász [36])**.** Let $A_1, \ldots, A_n$ be events with a dependency graph $G = (V, E)$. If there are $x_1, \ldots, x_n \in [0, 1)$ such that for all $i \in [1, n]$,

$$P(A_i) < x_i \prod_{\{i,j\} \in E} (1 - x_j)$$

then

$$P\left(\bigcap_{i=1}^{n} \bar{A}_i\right) \geq \prod_{i=1}^{n} (1 - x_i)$$

and in particular, $P(\cap_{i=1}^{n} \bar{A}_i) > 0$.

Proof. The proof relies on the following claim. For any $S \subsetneq [1, n]$ and any $i \notin S$, $P(A_i | \cap_{j\in S} \bar{A}_j) \le x_i$. The proof of the claim is by strong induction on $|S|$.

Base Step: When $|S| = 0$, $S = \emptyset$ and

$$\begin{aligned} P(A_i | \cap_{j\in\emptyset} \bar{A}_j) = P(A_i) &\le x_i \prod_{\{i,j\}\in E} (1 - x_j) \qquad \text{(by assumption)} \\ &\le x_i\,(1 - x_j) \le x_i. \end{aligned}$$

Inductive Step: Fix some $S \subsetneq [1, n]$ and assume that the claim holds for all smaller subsets of $[1, n]$. Fix some $i \notin S$. Set $S_1 = S \cap \{j : \{i, j\} \in E\}$ and $S_2 = S\backslash S_1$. Relabel the events if necessary so that $S_1 = \{1, 2, \ldots, s\}$ (for some s). Then,

$$\begin{aligned} P(A_i | \cap_{j\in S} \bar{A}_j) &= P(A_i | \cap_{j\in S_1} \bar{A}_j \cap \cap_{j\in S_2} \bar{A}_j) \\ &= \frac{P(A_i \cap \cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j)}{P(\cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j)} \qquad \text{(by Lemma 5.2.7)} \\ &\le \frac{P(A_i | \cap_{j\in S_2} \bar{A}_j)}{P(\cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j)}, \end{aligned}$$

and since A_i is mutually independent of the events $\{A_j : j \in S_2\}$,

$$P(A_i | \cap_{j\in S} \bar{A}_j) \le \frac{P(A_i | \cap_{j\in S_2} \bar{A}_j)}{P(\cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j)} = \frac{P(A_i)}{P(\cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j)}. \tag{5.1}$$

Now, consider the denominator of the last term in equation (5.1):

$$\begin{aligned} P(\cap_{j\in S_1} \bar{A}_j | \cap_{j\in S_2} \bar{A}_j) &= P(\bar{A}_1 \cap \cdots \cap \bar{A}_s | \cap_{j\in S_2} \bar{A}_j) \\ &= \prod_{k=1}^{s} P\Big(\bar{A}_k \Big| \bigcap_{j=k+1}^{s} \bar{A}_j \cap \bigcap_{j\in S_2} \bar{A}_j\Big) \qquad \text{(by Lemma 5.2.8)} \end{aligned}$$

$$
\begin{aligned}
&= \prod_{k=1}^{s} \biggl(1 - P\biggl(A_k \biggm| \bigcap_{j=k+1}^{s} \bar{A}_j \cap \bigcap_{j \in S_2} \bar{A}_j\biggr)\biggr) \\
&\geq \prod_{k=1}^{s} (1 - x_k) \quad \text{(by induction hyp. with } \{k+1, \ldots, s\} \cup S_2) \\
&\geq \prod_{\{i,k\} \in E} (1 - x_k). \qquad (5.2)
\end{aligned}
$$

Therefore, since $P(A_i) \leq x_i \prod_{\{i,k\} \in E} (1 - x_k)$, combining equations (5.1) and (5.2) gives

$$
\begin{aligned}
P(A_i | \cap_{\{i,k\} \in E} \bar{A}_j) &\leq \frac{P(A_i)}{P(\cap_{j \in S_1} \bar{A}_j | \cap_{j \in S_2} \bar{A}_j)} \\
&\leq \frac{x_i \prod_{\{i,k\} \in E} (1 - x_k)}{\prod_{\{i,k\} \in E} (1 - x_k)} \\
&= x_i.
\end{aligned}
$$

The claim can now be used to prove the theorem. Using Lemma 5.2.8 with $B = \Omega$,

$$
\begin{aligned}
P(\cap_{i=1}^{n} \bar{A}_i) &= \prod_{i=1}^{n} P(\bar{A}_i | \cap_{j=i+1}^{n} \bar{A}_j) \\
&= \prod_{i=1}^{n} (1 - P(A_i | \cap_{j=i+1}^{n} \bar{A}_j)) \\
&\geq \prod_{i=1}^{n} (1 - x_i) \quad \text{(by claim, using } S = \{i+1, \ldots, n\}). \qquad \square
\end{aligned}
$$

The Local Lemma is often used in the symmetric case where all the values of x_i are equal.

Theorem 5.2.10 (**Symmetric Lovász Local Lemma**, Erdős and Lovász [36])**.** Let $A_1, \ldots, A_n$ be events in a probability space and let (V, E) be a dependency graph for the events $A_1, \ldots, A_n$. Let $d \in \mathbb{Z}^+$ and $0 \leq p \leq 1$ be such that if $i \in [1, n]$, then $|\{j \in [1, n] : (i, j) \in E\}| \leq d$ and $P(A_i) \leq p$. If $ep(d+1) \leq 1$, then $P(\cap_{i=1}^n \bar{A}_i) > 0$.

Proof. Suppose that $ep(d+1) \leq 1$. Using the inequality, $e^x > 1 + x$ (from the Taylor series), $e \geq (1 + \frac{1}{d})^d$ and so $e^{-1} < (\frac{d}{d+1})^d = (1 - \frac{1}{d+1})^d$. Therefore, for each $i \in [1, n]$,

$$P(A_i) \leq p \leq \frac{1}{d+1} e^{-1} < \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d.$$

Setting $x_1 = x_2 = \cdots = x_n = \frac{1}{(d+1)}$, for each $i \in [1, n]$,

$$P(A_i) < \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d = x_i \prod_{\{i,j\} \in E} (1 - x_j).$$

Therefore, by Theorem 5.2.9, $P(\cap_{i=1}^n \bar{A}_i) > 0$. □

The following standard application of Theorem 5.2.10 to the van der Waerden numbers appears in Graham, Rothschild and Spencer's *Ramsey Theory* [51, p. 97].

Theorem 5.2.11. For all $k, r \in \mathbb{Z}^+$, $W(k; r) > \frac{r^{k-1}}{ke}(1 - o(1))$.

Proof. Fix $k, r \in \mathbb{Z}^+$ and let $n \in \mathbb{Z}^+$. Let Ω be the set of all r-colourings of $[1, n]$ and let P be the uniform probability for Ω. As before, for each AP_k, $S \subseteq [1, n]$, let $A_S = \{\Delta \in \Omega : S \text{ is monochromatic under } \Delta\}$. Let $G = (V, E)$ be the dependency graph defined as follows. Let V be the set of all AP_k's in $[1, n]$ and define edges by $E = \{(S_1, S_2) : S_1 \cap S_2 \neq \emptyset\}$.

Given an AP_k, $S = \{a, a+d, \ldots, a+(k-1)d\}$, consider the number of other arithmetic progressions with which S could have non-empty intersection. For each $1 \leq d' \leq \frac{n-1}{k-1}$ and $0 \leq i \leq k-1$, there are no more than k different AP_k's with difference d' that contain $a + id$. Thus, S intersects no more than $\frac{(n-1)k^2}{k-1}$ other AP_k's in $[1, n]$. Therefore, for each AP_k S, $|\{S_j : (S, S_j) \in E\}| \leq \frac{(n-1)k^2}{k-1}$ and since $P(A_S) = \frac{r}{r^k} = r^{1-k}$, by the symmetric Lovász local lemma, if $er^{1-k}(\frac{n-1}{k-1}k^2 + 1) \leq 1$ then $P(\cap \bar{A}_S) > 0$. Since $\cap \bar{A}_S$ is the event that no AP_k in $[1, n]$ is monochromatic, if $P(\cap \bar{A}_S) > 0$ then there is at least one r-colouring of $[1, n]$ with no monochromatic AP_k and so $W(k; r) > n$.

Let $n \in \mathbb{Z}^+$ be such that $er^{1-k}(\frac{n-1}{k-1}k^2 + 1) \leq 1$, but $er^{1-k}(\frac{n}{k-1}k^2 + 1) > 1$ (that is, n is the greatest integer that satisfies this condition). Then, by the previous paragraph, $W(k; r) > n$. Now,

$$\begin{aligned} er^{1-k}\left(\frac{n}{k-1}k^2 + 1\right) > 1 \Leftrightarrow n &> \left(\frac{r^{k-1}}{e} - 1\right)\frac{k-1}{k^2} \\ &= \frac{r^{k-1}}{ek}\left(\frac{k-1}{k} - \frac{(k-1)e}{kr^{k-1}}\right) \\ &= \frac{r^{k-1}}{ek}(1 - o(1)) \qquad \text{(for a fixed } r\text{).} \end{aligned}$$

Thus for any fixed $r \in \mathbb{Z}^+$, $W(k; r) > \frac{r^{k-1}}{ek}(1 - o(1))$. □

More recently, in 1990, a more careful application of the Local Lemma was used to show the following.

Theorem 5.2.12 (Szabó [106])**.** For every $\varepsilon > 0$, there is a $K = K(\varepsilon)$ so that for all $k \geq K$,

$$W(k;2) \geq \frac{2^k}{k^{\varepsilon}}.$$

5.3 Constructive lower bound

Using techniques of field theory, Berlekamp [15] showed that for any prime p, there is a 2-colouring of a large interval that avoids arithmetic progressions of length $p+1$. The proof of Berlekamp's result requires the following well-known facts (see for example [68, Chapter 5] for proofs) about field extensions and finite fields.

Definition 5.3.1. Given fields F, E with $F \subseteq E$, E is a *field extension* of F if the operations of F are those of E restricted to F. For $a \in E \backslash F$, the set rational functions in a over F, denoted $F(a)$, is a field extension of F. The *degree* of E over F, denoted $[F : E]$, is the dimension of E as a vector space over F.

Lemma 5.3.2. If $D \subseteq E \subseteq F$ are fields, such that E is a field extension of D and F is a field extension of E, then F is a field extension of D and

$$[F : D] = [F : E][E : D].$$

Lemma 5.3.3. For every prime p and for every $n \in \mathbb{Z}^+$, there is a unique field of order p^n called the *Galois field* of order p^n and denoted by $GF(p^n)$. Then $GF(p^n)^* = \langle GF(p^n)\backslash\{0\}, \cdot\rangle$ is a cyclic group of order $p^n - 1$ and $[GF(p^n) : \mathbb{Z}_p] = n$.

Definition 5.3.4. Let $f(x)$ be a polynomial over a commutative ring. Then $f(x)$ is *reducible* over this ring iff there are polynomials $g(x), h(x)$ over this ring of smaller degree with $f(x) = g(x)h(x)$. Otherwise, $f(x)$ is *irreducible*.

Recall that for a field F, $F[x]$ is the ring of polynomials in x over F.

Fact 5.3.5. Let E, F be fields with E a field extension of F. Let $\alpha \in E$ and $f \in F[x]$ be irreducible with $f(\alpha) = 0$. If $g(x) \in F[x]$ with $g(\alpha) = 0$, then $f(x)|g(x)$.

Fact 5.3.6. Let E, F be fields where E is a field extension of F and $[E : F] < \infty$. Then for every $\alpha \in E$, there is an irreducible $f \in F[x]$ such that $f(\alpha) = 0$.

Fact 5.3.7. Let E, F be fields where E is a field extension of F. Let $\alpha \in E$ and $f \in F[x]$ with $f(\alpha) = 0$, then $[F(\alpha) : F] \leq \deg(f)$.

Berlekamp proved a complicated result that gave bounds on the van der Waerden numbers whenever the number of colours is a prime power. The proof of Theorem 5.3.8 is a simplification of Berlekamp's proof in the case when the number of colours is a prime. (The proof of the result for 2 colours can also be found in [51, pp. 96–7].)

Theorem 5.3.8 (Berlekamp [15])**.** If p, q are prime, then

$$W(p+1; q) > p(q^p - 1)/(q-1).$$

Proof. By Lemma 5.3.3, $GF(q^p)^*$ is a cyclic group. Let α be one of its generators. Let $\{v_1, v_2, \ldots, v_p\}$ be a basis for $GF(q^p)$ as a vector space over $\mathbb{Z}_q$ (again by Lemma

5.3.3) and for each $j \in [0, p(q^p - 1)/(q-1)]$, let $r_{1,j}, r_{2,j}, \ldots, r_{p,j} \in \mathbb{Z}_q$ be such that,

$$\alpha^j = \sum_{i=1}^{p} r_{i,j} v_i.$$

Define a partition of $[0, p(q^p - 1)/(q-1) - 1]$ as follows. For each $i \in [0, q-1]$, set

$$S_i = \{j : r_{1,j} = i\}.$$

Suppose one of the sets S_i contains an AP_{p+1}; that is, for some integers a and $d \neq 0$, $\{a, a+d, \ldots, a+pd\} \subseteq S_i$. Note that since $pd \leq a + pd < p(q^p - 1)/(q-1)$,

$$0 < d < \frac{q^p - 1}{q - 1} \leq q^p - 1.$$

Therefore, since the order of α in $GF(q^p)^*$ is $q^p - 1$,

$$\alpha^d \neq 1 \text{ and } \alpha^{d(q-1)} \neq 1. \tag{5.3}$$

Case I: $i = 0$.

Define $T_0 = \text{span}\{v_2, v_3, \ldots, v_p\}$. The set T_0 is a subspace of $GF(q^p)$ of dimension $p-1$ over $\mathbb{Z}_q$ and for all $x \in S_0$, $\alpha^x \in T_0$. Therefore, since $\{a+d, a+2d, \ldots, a+pd\} \subseteq S_0$, if $n \in [1,p]$, then $\alpha^{a+dn} \in T_0$. Then, $\alpha^{a+d}, \ldots, \alpha^{a+pd}$ are p elements in a $(p-1)$-dimensional space and so they are linearly dependent over $\mathbb{Z}_q$. Let $b_1, b_2, \ldots, b_p \in \mathbb{Z}_q$ not all 0 be such that $\sum_{n=1}^{p} b_n \alpha^{a+dn} = 0$. Then $\alpha^{a+d} \sum_{n=1}^{p} b_n \alpha^{d(n-1)}$.

Since $\alpha^{a+d} \neq 0$, α^d is a root of the polynomial $\sum_{n=1}^{p} b_n x^{n-1}$ which has degree at most $p-1$. Therefore, the degree of the extension $\mathbb{Z}_q(\alpha^d)$ over $\mathbb{Z}_q$ is at most $p-1$.

Since $GF(q^p) = \mathbb{Z}_q(\alpha)$, by Lemma 5.3.2,

$$p = [GF(q^p) : \mathbb{Z}_q] = [GF(q^p) : \mathbb{Z}_q(\alpha^d)][\mathbb{Z}_q(\alpha^d) : \mathbb{Z}_q].$$

Thus $[\mathbb{Z}_q(\alpha^d) : \mathbb{Z}_q] \leq p-1$ divides p, a prime. Therefore, $\mathbb{Z}_q(\alpha^d)$ is an extension of degree 1 over $\mathbb{Z}_q$ and so $\alpha^d \in \mathbb{Z}_q$. Since $\langle \mathbb{Z}_q \backslash \{0\}, \cdot \rangle$ is a cyclic group of order $q-1$, $\alpha^{d(q-1)} = 1$ which is impossible by (5.3). Therefore, S_0 cannot contain any AP_{p+1}'s and in fact since only the elements $a+d, \ldots, a+pd$ are used in this part of the proof, S_0 can only contain AP_p's with difference $d \geq (q^p - 1)/(q-1)$.

Case II: $i \in [1, q-1]$.

Suppose that for some $i \neq 0$, $\{a, a+d, \ldots, a+pd\} \subseteq S_i$, then for each $j \in [0, p]$, $\alpha^{a+jd} - \alpha^a = \alpha^a(\alpha^{jd} - 1) \in T_0$. As in Case I, there are $b_1, b_2, \ldots, b_p \in \mathbb{Z}_q$ such that $\sum_{i=1}^p b_i \alpha^a (\alpha^{di} - 1) = 0$. That is

$$\begin{aligned} 0 = \sum_{i=1}^{p} b_i \alpha^a (\alpha^{di} - 1) &= \alpha^a(\alpha^d - 1)(b_1 + \sum_{i=2}^{p} b_i(\alpha^{d(i-1)} + \alpha^{d(i-2)} + \cdots + \alpha^d + 1)) \\ &= \alpha^a(\alpha^d - 1) \sum_{i=0}^{p-1} (\alpha^d)^i \left(\sum_{j=i+1}^{p} b_j \right). \end{aligned}$$

Since $\alpha^d \neq 1$ and $\alpha^a \neq 0$, α^d is a zero of a polynomial in $\mathbb{Z}_q$ of degree $p-1$, contradicting $\alpha^{d(q-1)} \neq 1$ as in Case I. □

In [80], it is mentioned that this proof can be adapted to show that for all p, q prime with $p \geq 5$, $W(p+1; q) > p(q^p - 1)$, but I have been unable to confirm that this is true.

The following gives the details of Berlekamp's strengthening of the bound from Theorem 5.3.8 in the case $q = 2$ and yields the best known lower bounds for these van der Waerden numbers. The idea of the proof is to show that given a careful choice of basis elements for $GF(2^p)$, the partition given in the previous proof can be extended without introducing any arithmetic progressions.

Corollary 5.3.9 (Berlekamp [15])**.** If p is prime, $W(p+1; 2) > p\,2^p$.

Proof. The result holds for $p = 2$ since $W(3; 2) = 9 > 2 \cdot 2^2$. For a prime number $p > 2$, as in the proof of Theorem 5.3.8, let α be a generator for $GF(2^p)^*$. Define the basis elements for $GF(2^p)$ over $\mathbb{Z}_2$ as follows. Set

$$v_1 = 1,\ v_2 = 1+\alpha,\ v_3 = 1+\alpha^2,\ \ldots,\ v_{\frac{p+1}{2}} = 1+\alpha^{\frac{p-1}{2}},$$

$$v_{\frac{p+1}{2}+1} = 1+\alpha^{-1},\ v_{\frac{p+1}{2}+2} = 1+\alpha^{-2},\ \ldots,\ v_p = \alpha^{-(\frac{p-1}{2})}.$$

Claim. The set $\{v_1, v_2, \ldots, v_p\}$ is linearly independent.

Proof of Claim. Let $x_1, \ldots, x_p \in \mathbb{Z}_2$ be such that $\sum_{i=1}^{p} x_i v_i = 0$. Then,

$$\begin{aligned} 0 &= \alpha^{(p-1)/2}\sum_{i=1}^{p} x_i v_i = \sum_{i=1}^{p} x_i\,\alpha^{(p-1)/2} v_i \\ &= x_p + x_{p-1}\alpha + \cdots x_{\frac{p+1}{2}+1}\alpha^{\frac{p-1}{2}-1} + (x_1 + x_2 + \cdots + x_p)\alpha^{\frac{p-1}{2}} \\ &\quad + x_2\alpha^{\frac{p-1}{2}+1} + x_3\alpha^{\frac{p-1}{2}+2} + \cdots x^{\frac{p+1}{2}}\alpha^{p-1}. \end{aligned} \tag{5.4}$$

If any of the constants $x_1, x_2, \ldots, x_p$ is non-zero, equation (5.4) gives a polynomial of degree no more than $p-1$ for which α is a zero. However, since $\mathbb{Z}_2(\alpha) = GF(2^p)$, and $[GF(2^p) : \mathbb{Z}_2] = p > p-1$, it must be that $x_1 = x_2 = \cdots = x_p = 0$. Therefore, the set $\{v_1, v_2, \ldots, v_p\}$ is linearly independent □

Define the sets S_0 and S_1 as in the proof of Theorem 5.3.8 in the case $q = 2$. Since $v_1 = 1 = \alpha^0$, $0 \in S_1$. Furthermore, $[1, \frac{p-1}{2}] \subseteq S_1$ since for each $i \in [1, (p-1)/2]$,

$$\alpha^i = 1 + (1 + \alpha^i) = v_1 + v_{i+1}.$$

Similarly, $[p(2^p-1)-(p-1)/2, p(2^p-1)-1] \subseteq S_1$ since for each $j \in [1, (p-1)/2]$,

$$\alpha^{p(2^p-1)-j} = \alpha^{-j} = 1 + (1 + \alpha^{-j}) = v_1 + v_{\frac{p+1}{2}+j}.$$

Thus,

$$\left[0, \frac{p-1}{2}\right] \subseteq S_1 \quad \text{and} \quad \left[p(2^p-1) - \frac{p-1}{2}, p(2^p-1)-1\right] \subseteq S_1. \tag{5.5}$$

Set

$$\begin{aligned} S_0' &= \left[-\frac{p-1}{2}, -1\right], \\ S_0'' &= \left[p(2^p-1), p(2^p-1) + \frac{p-1}{2}\right], \qquad\qquad \text{and} \\ S_0^+ &= S_0 \cup S_0' \cup S_0''. \end{aligned}$$

As before, S_1 contains no AP_{p+1}'s, so assume that there are integers a and d with $d > 0$ such that $P = \{a, a+d, \ldots, a+pd\} \subseteq S_0^+$. Since S_0 is also AP_{p+1}-free, there are three possible forms the arithmetic progression could take.

Case I: The set P includes one element from S_0' (or S_0'') and an AP_p in S_0. By a remark at the end of Case I of the proof of Theorem 5.3.8, this can only occur when $d \geq 2^p - 1$. But this is impossible since then P spans $p(2^p - 1)$ elements and so must contain an element of S_1, by (5.5).

Case II: The set P includes one element from S_0' and one from S_0''. It must be that $a \in S_0'$ and $a + pd \in S_0''$. Then for some $1 \leq i \leq (p-1)/2$, $a = -i$ and for some $0 \leq j \leq (p-1)/2$, $a + pd = p(2^p - 1) + j$. Now, $pd = a + pd - a$ is divisible by p, but $p(2^p - 1) + j - (-i) = p(2^p - 1) + j + i$ is not since $1 \leq i + j \leq p - 1$ and p is prime.

Case III: The set P includes two elements from S_0' (or S_0''). But then $d \leq (p-1)/2$ and so it must also include an element from S_1 by (5.5).

Therefore, S_1 and S_0^+ partition the integers $[-(p-1)/2, p\ (2^p - 1) + (p-1)/2]$ (an interval of length $p\ 2^p$) and neither contains an AP_{p+1}. Therefore, by Lemma 3.1.4, $W(p+1; 2) > p\ 2^p$. □

5.4 Hypergraph techniques

Definition 5.4.1. Let S be any set and $\mathcal{E} \subseteq \mathcal{P}(S) \setminus \{\emptyset\}$. Then $\mathcal{H} = (S, \mathcal{E})$ is called a *hypergraph*. The elements of the set S are called the *vertices of* $\mathcal{H}$ and the sets in $\mathcal{E}$ are called the *hyperedges*.

The *independence number* of $\mathcal{H}$, $\alpha(\mathcal{H})$, is the maximal size of a subset of S that contains no member of $\mathcal{E}$. The *chromatic number* of $\mathcal{H}$, $\chi(\mathcal{H})$, is the least integer r so that there is an r-colouring of S with no member of $\mathcal{E}$ monochromatic. Note that if the $\mathcal{E}$ contains any singletons, the chromatic number is undefined. For the purposes here, there are no hypergraphs with singletons for hyperedges.

A hypergraph $\mathcal{H}$ is called *uniform* iff there is an integer k so that $\mathcal{E} \subseteq [S]^k$.

Consider any hypergraph $\mathcal{H} = (S, \mathcal{E})$ and any $r \in \mathbb{Z}^+$, with $\chi(\mathcal{H}) > r$. Then for any r-colouring of S, there is a monochromatic member of $\mathcal{E}$. On the other hand, if $\chi(\mathcal{H}) \leq r$, then there is an r-colouring of S with no monochromatic member of $\mathcal{E}$. Thus, bounds on the chromatic number can be used to find bounds on the numbers associated with Ramsey-type problems.

Definition 5.4.2. Let $\mathcal{H} = (S, \mathcal{E})$ be a hypergraph. A permutation σ of S is an *automorphism of* $\mathcal{H}$ iff for every $E \in \mathcal{E}$, $\sigma(E) \in \mathcal{E}$. The group $G = \text{aut}(\mathcal{H})$ of automorphisms of $\mathcal{H}$ is *transitive* iff or each $s_1, s_2 \in S$, there is a $\sigma \in G$ such that $\sigma(s_1) = s_2$.

Definition 5.4.3. The hypergraph $\mathcal{H}$ is *symmetric* iff the group $\text{aut}(\mathcal{H})$ is transitive.

The Symmetric Hypergraph Theorem can be found in Graham, Rothschild and Spencer [51, pp. 98–103], however I have been unable to determine its original source. The Symmetric Hypergraph Theorem, gives a tool for finding upper bounds on the chromatic number of certain types of hypergraphs. The proof is omitted.

Theorem 5.4.4 (Symmetric Hypergraph Theorem [51]). Let $\mathcal{H} = (S, \mathcal{E})$ be a symmetric hypergraph with $|S| = m$ and $\alpha = \alpha(\mathcal{H})$. Then,

$$\chi(\mathcal{H}) \leq 1 + \frac{\ln m}{-\ln(1 - \alpha/m)}.$$

Whenever $\alpha < m$, the Taylor series for $\ln(1 + x)$ can be used to show that $\ln(1 - \frac{\alpha}{m}) > -\frac{\alpha}{m}$ and hence

$$\begin{aligned}\chi(\mathcal{H}) &\leq 1 + \frac{\ln m}{-\ln(1 - \alpha/m)} \\ &< 1 + \frac{m}{\alpha} \ln m \\ &= \frac{m}{\alpha}(\ln m)(1 + o(1)).\end{aligned}$$

In order to apply Theorem 5.4.4 to the van der Waerden numbers $W(3; r)$, it is necessary to define an appropriate hypergraph. For each $n \in \mathbb{Z}^+$, let $S_n = [1, n]$, let $\mathcal{E}_n$ be the set of all AP_3's in $[1, n]$ and let $\mathcal{H}_n = (S_n, \mathcal{E}_n)$ be the corresponding hypergraph.

Since $\alpha(\mathcal{H}_n) = r_3(n)$ (recall Definition 4.1.1), any lower bound on $r_3(n)$ can be used together with Theorem 5.4.4 to find an upper bound on $\chi(\mathcal{H}_n)$ and from that a lower bound on the van der Waerden number $W(3; r)$.

Since $(S_n, \mathcal{E}_n)$ is not a symmetric hypergraph, in order to apply the Symmetric Hypergraph Theorem, it is necessary to define a new hypergraph. Set $Y_n = \mathbb{Z}_{2n-1}$ and $\mathcal{E}'_n = \{\{a, a+d, a+2d\} : a \in \mathbb{Z}_{2n-1} \text{ and } d \leq n/2\}$ with addition modulo $2n-1$. Define the hypergraph $\mathcal{H}'_n = (Y_n, \mathcal{E}'_n)$. That is, $\mathcal{E}'_n$ is the set of all AP_3's in $\mathbb{Z}_{2n-1}$

(with addition modulo $2n-1$) that are contained in a block of n consecutive integers. Therefore, $[S_n]^3 \cap \mathcal{E}'_n = \mathcal{E}_n$ and so $(S_n, \mathcal{E}_n)$ is a subhypergraph of $(Y_n, \mathcal{E}'_n)$. Further, since all of the permutations $\sigma_i : x \mapsto x + i \pmod{2n-1}$ are automorphisms of $\mathcal{H}'_n$, for any $a, b \in \mathbb{Z}^+$, there is a $\sigma_{b-a} \in \text{aut}(\mathcal{H}'_n)$ with $\sigma_{b-a}(a) = b$. Therefore, $(S'_n, \mathcal{E}'_n)$ is a symmetric hypergraph.

Since any independent set in $(S_n, \mathcal{E}_n)$ will also be independent in $(Y_n, \mathcal{E}'_n)$, it follows that $\alpha(\mathcal{H}'_n) \geq \alpha(\mathcal{H}_n) = r_3(n)$. For any $r \in \mathbb{Z}^+$ and any r-colouring of Y_n with no monochromatic members of $\mathcal{E}'_n$ will induce a r-colouring of S_n with no monochromatic members of $\mathcal{E}_n$. Thus, $\chi(\mathcal{H}_n) \leq \chi(\mathcal{H}'_n)$.

Theorem 5.4.5 (Graham, Rothschild and Spencer [51])**.** There is a fixed constant c such that, for r sufficiently large,

$$W(3; r) > r^{c \ln r}.$$

Proof. The following proof is due to Spencer [104]. Let c' be such that for n large enough, $r_3(n) \geq ne^{-c'\sqrt{\ln n}}$ (Theorem 4.6.1). Fix $r \in \mathbb{Z}^+$ and set $n = r^{c \ln r}$ where c is a constant small enough so that $\sqrt{c}c' = 0.9$. Then, $\ln n = \ln r^{c \ln r} = c(\ln r)^2$ and so

$$r_3(n) \geq ne^{-c'\sqrt{\ln n}} = ne^{-c'\sqrt{c} \ln r} = nr^{-0.9}.$$

By Theorem 5.4.4,

$$\chi(\mathcal{H}_n) \leq \chi(\mathcal{H}'_n) < \frac{(2n-1)}{\alpha(\mathcal{H}'_n)} \ln{(2n-1)}(1 + o(1))$$

$$\begin{aligned}
&< \frac{2n\ln n}{r_3(n)}(1+o(1)) \\
&\leq \frac{2n\ln n}{nr^{-0.9}}(1+o(1)) \\
&< 2r^{0.9}c(\ln r)^2(1+o(1)) \\
&= r\left(\frac{2c(\ln r)^2}{r^{0.1}}\right)(1+o(1)).
\end{aligned}$$

Take r large enough so that $\frac{2c(\ln r)^2}{r^{0.1}}(1+o(1)) < 1$. In that case, when $n = r^{c'\ln r}$, $\chi(\mathcal{H}_n) < r$. That is, there is an r-colouring of S_n with no monochromatic elements of $\mathcal{E}_n$ and hence $W(3;r) > n = r^{c'\ln r}$. □

Repeating the proof of Theorem 5.4.5 with 3 replaced by arbitrary k (and using the bound $r_k(n) > ne^{-c(\ln n)^{1/\lceil \log_2 k\rceil}}$ given by Theorem 4.6.2 instead of Theorem 4.6.1) shows that for any $k \in \mathbb{Z}^+$, there is a constant $c' = c'(k)$ such that for all $r \in \mathbb{Z}^+$ sufficiently large,

$$W(k;r) > e^{c'(\ln r)^{\lceil \log_2 k\rceil}}. \tag{5.6}$$

Although the proof of equation 5.6 is nearly identical to the proof of Theorem 5.4.5, this result does not seem to appear in the literature.

Another approach that can be used to find bounds is to use results on the possible number of hyperedges in uniform hypergraphs with a chromatic number larger than two.

Lemma 5.4.6 (Schmidt [100]). Let $n \in \mathbb{Z}^+$, $\mathcal{E} \subseteq [n]^k$ and $\mathcal{H} = ([n], \mathcal{E})$ be such that for every 2-colouring of $[1,n]$ there is a monochromatic member of $\mathcal{E}$, that is

$\chi(\mathcal{H}) > 2$. Then,

$$|\mathcal{E}| \geq \frac{2^k}{1 + 2k^{-1}}.$$

A probabilistic proof of Schmidt's lemma can be found in [39]. The following bound is achieved using Lemma 5.4.6.

Theorem 5.4.7 (Erdős and Spencer [39])**.** For all $k \in \mathbb{Z}^+$,

$$W(k;2) \geq 2^{\frac{k+1}{2}}\sqrt{k-1}(1 - o(1)).$$

Proof. Let $n \geq W(k;2)$ and let $\mathcal{E}$ be the set of all AP_k's in $[1,n]$. By Lemma 1.2.3, $|\mathcal{E}| \leq \frac{n^2}{2(k-1)}$.

Since $n \geq W(k;2)$, for every 2-colouring of $[1,n]$ there is a monochromatic member of $\mathcal{E}$. Thus, by Schmidt's lemma (Lemma 5.4.6),

$$\frac{2^k}{1 + 2k^{-1}} \leq |\mathcal{E}| \leq \frac{n^2}{2(k-1)}.$$

Therefore,

$$n \geq \sqrt{\frac{2^{k+1}(k-1)}{1 + 2k^{-1}}} = 2^{\frac{k+1}{2}}\sqrt{k-1}(1 - o(1)).$$

□

Another result by Schmidt [99] showed that there is a constant $c > 0$ so that for any integers k and r, $W(k;r) \geq r^{k - c(k \log k)^{1/2}}$. However, this lower bound for the numbers $W(k;r)$ is smaller than that given by Theorem 5.2.11.

5.5 Upper bounds

While van der Waerden was able to prove the existence of the numbers $W(k;r)$, the upper bounds given by his proof grew incredibly quickly. To describe the size of the bounds, a fast-growing function called an "Ackermann function" is needed.

Definition 5.5.1. For any function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ and any $n \in \mathbb{Z}$, let $f^{(n)}$ denote the composition of f with itself n times. That is, for any x in the domain of f,

$$f^{(n)}(x) = \underbrace{f(f(\ldots(f}_{n \text{ times}}(x))\ldots)).$$

Define a series of functions on the integers $\{f_i : i \in \mathbb{Z}^+\}$ recursively as follows. Let $f_1(k) = 2k$ and for each $i \geq 1$, having defined the function f_i, let

$$f_{i+1}(k) = f_i^{(k)}(1).$$

For example, $f_2(k) = 2^k$, $f_3(k) = \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}_{k \text{ twos}}$ and for each $k \geq 2$, $f_4(k) = \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}_{f_4(k-1) \text{ twos}}$. Finally, define a function $f_\omega : \mathbb{Z}^+ \to \mathbb{Z}^+$ as follows. For each $n \in \mathbb{Z}^+$, set $f_\omega(n) = f_n(n)$. The function f_ω grows incredibly quickly and the original proof of van der Waerden's theorem was only able to guarantee that $W(k;2) \leq f_\omega(k)$. In 1988, Shelah [102] provided a proof that showed that for some constant c, $W(k;2) \leq f_4(ck)$ (see [51, pp.60–6] for a detailed discussion). According to Shelah, some mathematicians (for example Solovay, see [102]) attempted to show that in fact $W(k;r)$ was of the same order of magnitude as the function f_ω. This was supported by the fact that a careful analysis of Furstenberg's ergodic proof [44] by Girard (see Shelah [102]) yielded large bounds similar to those from the original inductive proof.

At present, the best known upper bounds come from the density proofs and bounds on the function $r_k(n)$. The following upper bound for the van der Waerden numbers $W(3;r)$ uses the bound for $r_3(n)$ from Theorem 4.2.4 and comes closest to the best known lower bound for these van der Waerden numbers. Recall that Bourgain (Theorem 4.2.4) proved that there is a constant c so that for n sufficiently large, $r_3(n) < cn(\frac{\ln\ln n}{\ln n})^{1/2}$.

Corollary 5.5.2 (Bourgain [17])**.** There is a constant c' so that for r sufficiently large,

$$W(3;r) \leq e^{c'r^2 \ln r}.$$

Proof. Let c be the constant from Theorem 4.2.4, set $c' = 4c^2$, let r be large and set $n = e^{c'r^2 \ln r}$.

Then $\ln n = c'r^2 \ln r$ and when r is large enough, $\ln\ln n = \ln c' + 2\ln r + \ln\ln r \leq 4\ln r$.

Then,

$$r_3(n) < cn\left(\frac{4\ln r}{c'r^2 \ln r}\right)^{1/2} = \frac{2c}{\sqrt{c'}} \cdot \frac{n}{r} = \frac{n}{r}.$$

and hence $W(3;r) \leq n = e^{c'r^2 \ln r}$. □

In fact, Green [52] showed that the constant c' in Corollary 5.5.2 can be taken to be 2^{56}. Theorem 5.4.5 together with Corollary 5.5.2 show that for r sufficiently large and constants c and c', $r^{c' \ln r} < W(3;r) < r^{c'r^2}$.

For $k > 3$, Gowers' upper bound for $r_k(n)$ (recall Theorem 4.3.2) gives the best upper bounds in general for the van der Waerden numbers $W(k;r)$. Recall that Gowers showed that for every k, if $c(k) = 2^{-2^{k+9}}$, then $r_k(n) < n(\log_2 \log_2 n)^{-c(k)}$.

Corollary 5.5.3 (Gowers [47])**.** For every $r, k \in \mathbb{Z}^+$,

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

Proof. Set $n = 2^{2^{r^{2^{2^{k+9}}}}}$. Then $\log_2 \log_2 n = r^{2^{2^{k+9}}}$ and so $(\log_2 \log_2 n)^{2^{-2^{k+9}}} = r^{2^{2^{k+9}-2^{k+9}}} = r$. Thus, for any r-colouring of $[1,n]$, one colour class will contain at least

$$\frac{n}{r} = \frac{n}{(\log_2 \log_2 n)^{c(k)}} > r_k(n)$$

elements and hence that colour class contains an AP_k. □

The bounds given in Corollaries 5.5.2 and 5.5.3 are based on density results. Thus, for example, Corollary 5.5.2 guarantees that for some constant c, when $n \geq e^{c'r^2 \ln r}$, for any r-colouring of $[1,n]$, the largest colour class will contain an AP_3. However, since the proofs of both lemmas are based on analytic techniques, they are not constructive. Although weaker than the bound given in Corollary 5.5.2, the next bound to come (Theorem 5.5.5), due to Haung and Yang [60], is the best known constructive upper bound for $W(3;r)$.

The following terminology is adapted from Tao [109], but may not be optimal.

Definition 5.5.4. A set $F \subseteq \mathbb{Z}^+$ is called a *fan of t different AP_k's* iff there are $a_1, \ldots, a_t, d_1, \ldots, d_t \in \mathbb{Z}^+$ (all d_i's distinct) such that

$$F = \bigcup_{i=1}^{t} \{a_i + j \cdot d_i : 0 \leq j \leq k-1\}$$

and either $a_1 = \cdots = a_t$ or else $a_1 + (k-1)d_1 = \cdots = a_t + (k-1)d_t$. That is, F is a collection of t different AP_k's either all sharing the same beginning point or all sharing the same endpoint.

For each $i \in [1, t]$, the AP_{k-1} in F formed by removing the common element from the set $\{a_i + j \cdot d_i : 0 \leq j \leq k-1\}$ is called a *spoke*. Given a colouring Δ of F, the fan F is called *weakly polychromatic* (with respect to Δ) if all the spokes are monochromatic but all different colours from each other and *strongly polychromatic* if in addition, the common element of all the AP_k's in F is a different colour from all the spokes.

The following proof closely follows that of Haung and Yang with some slight alterations. It states that for $r > 4$, $W(3; r) < (\frac{r}{4})^{3^r}$ although for some small values of r, this does not seem to follow from their proof. When r is large, the bound given here is smaller than that from the original paper.

Theorem 5.5.5 (Haung and Yang [60])**.** For each $r \geq 2$,

$$W(3; r) \leq \frac{5}{2} \left(\frac{r}{2}\right)^{\frac{3^r - 1}{2}}.$$

Proof. Fix $r \in \mathbb{Z}^+$ (the number of colours). Recursively define a sequence $\{R_t\}_{t=0}^r$ by $R_0 = r$ and for each $t \in [1, r]$, having defined R_{t-1}, let

$$R_t = \binom{R_{t-1}+1}{2} R_{t-1}(r-t). \tag{5.7}$$

For each $t \in [0, r]$, define $S_t = 1 + 2\sum_{i=0}^{t-1} R_i$ (where the empty sum is 0). Note that for each $t \in [1, r]$,

$$S_t = S_{t-1} + 2R_{t-1}. \tag{5.8}$$

The goal of the proof is to first show that $W(3; r) \leq S_r$ and then to find a bound for S_r.

Consider fans of AP_3's of the following particular form. For each $t \in [1, r]$ let $a_1, \ldots, a_t, d_1, \ldots, d_t$ be any integers such that $1 \leq a_1 < a_1 + d_1 \leq R_0 + 1$ and for each $i \in [2, t]$, $0 \leq a_i < a_i + d_i \leq R_{i-1}$. Define

$$F(a_1, d_1; \ldots; a_t, d_t) = \bigcup_{\ell=1}^{t} \left\{ \underbrace{\sum_{i=1}^{\ell-1}(a_i + 2d_i) + \left(\sum_{i=\ell}^{t} a_i\right)}_{\text{starting point}} + j \cdot \underbrace{\sum_{i=\ell}^{t} d_i}_{\text{difference}} \ : j \in \{0, 1, 2\} \right\}.$$

For each $\ell \in [1, t]$, the endpoint of the ℓ -th AP_3 in the above union is

$$\sum_{i=1}^{\ell-1}(a_i + 2d_i) + \sum_{i=\ell}^{t} a_i + 2\sum_{i=\ell}^{t} d_i = \sum_{i=1}^{t}(a_i + 2d_i).$$

Thus $F(a_1, d_1; \ldots; a_t, d_t)$ is a fan of t AP_3's where every AP_3 has the same endpoint $\sum_{i=1}^{t}(a_i + 2d_i)$.

Claim 1. Let $t \in [1, r]$ and let $a_1, \ldots, a_t, d_1, \ldots, d_t \in \mathbb{Z}^+$ be such that

$1 \leq a_1 < a_1 + d_1 \leq R_0 + 1$ and for each $i \in [2, t]$, $0 \leq a_i < a_i + d_i \leq R_{i-1}$. Then $F(a_1, d_1; \ldots; a_t, d_t) \subseteq [1, S_t]$.

Proof of Claim 1. Since $1 \leq a_1 < a_1 + d_1 \leq R_0 + 1$, $a_1 + 2d_1 \leq 2R_0 + 1$ and for each $i \in [2, t]$, since $0 \leq a_i < a_i + d_i \leq R_{i-1}$ it must be that $a_i + 2d_i \leq 2R_{i-1}$. Thus,

$$\sum_{i=1}^{t} (a_i + 2d_i) \leq 2R_0 + 1 + \sum_{i=2}^{t} (2R_{i-1}) = 1 + \sum_{i=1}^{t} 2R_{i-1} = S_t$$

and since $\sum_{i=1}^{t} (a_i + 2d_i)$ is the endpoint of all the t AP_3's in $F(a_1, d_1; \ldots; a_t, d_t)$, $F(a_1, d_1; \ldots; a_t, d_t) \subseteq [1, S_t]$. □

Claim 2. For each $t \in [1, r]$,

(i) for any r-colouring of $[1, S_t]$, either there is a monochromatic AP_3 in $[1, S_t]$ or else there are $a_1, \ldots, a_t, d_1, \ldots, d_t$ so that $F(a_1, d_1; \ldots; a_t, d_t)$ is a strongly polychromatic fan of t AP_3's, and

(ii) there are at most R_t pairs (F, Δ) where for some $a_1, \ldots, a_t, d_1, \ldots, d_t$, $F = F(a_1, d_1; \ldots; a_t, d_t)$ and Δ is an r-colouring of F under which F is strongly polychromatic.

Before proving Claim 2, observe that it implies that $W(3; r) \leq S_r$ since for any r-colouring of $[1, S_r]$, since there can be no strongly polychromatic fan of any kind with r spokes (this would require $r + 1$ different colours), by the part (i) of the claim, there must be a monochromatic AP_3 and so $W(3; r) \leq r$.

Proof of Claim 2. The proof given here proceeds by induction on t, showing that for each $t \in [1, r]$, both (i) and (ii) hold.

Base Case: For $t = 1$, fix $\Delta_0 : [1, S_1] \to [1, r]$. Since $S_1 = 2r + 1$, by the pigeonhole principle, there are a_1, d_1 with $1 \leq a_1 < a_1 + d_1 \leq r + 1$ and $\Delta_0(a_1) = \Delta_0(a_1 + d_1)$. Then $\{a_1, a_1 + d_1, a_1 + 2d_1\} = F(a_1, d_1)$ is either a monochromatic AP_3 or else a strongly polychromatic fan with one spoke. There are $\binom{r+1}{2} = \binom{R_0+1}{2}$ choices for the pair $\{a_1, a_1 + d_1\}$, $r = R_0$ possible choices for the colour of a_1 and $a_1 + d_1$, and $r - 1$ choices for the colour of $a_1 + 2d_1$, so that $F(a_1, d_1)$ is strongly polychromatic. By equation (5.7) there are at most $\binom{R_0+1}{2} R_0(r-1) = R_1$ different pairs $(F(a_1, d_1), \Delta)$ where Δ is an r-colouring of the fan $F(a_1, d_1)$ which makes the fan strongly polychromatic.

Inductive step: Fix some $t \in [1, r - 1]$ and suppose that both (i) and (ii) hold for this t. Fix an r-colouring $\Delta_0 : [1, S_{t+1}] \to [1, r]$ and assume that there are no monochromatic AP_3's in hopes of showing that there must be a strongly polychromatic fan with $t + 1$ spokes.

By equation (5.8) $S_{t+1} = 2R_t + S_t$, and so for each $\ell \in [0, 2R_t]$, $\Delta_0|_{[1+\ell, S_t+\ell]}$ is an r-colouring of an interval of length S_t with no monochromatic AP_3's and so by the induction hypothesis, each such colouring produces a strongly polychromatic fan with t spokes. Since there are only R_t pairs (F, Δ) of a fan F with t spokes and an r-colouring Δ that makes F strongly polychromatic, there must be integers

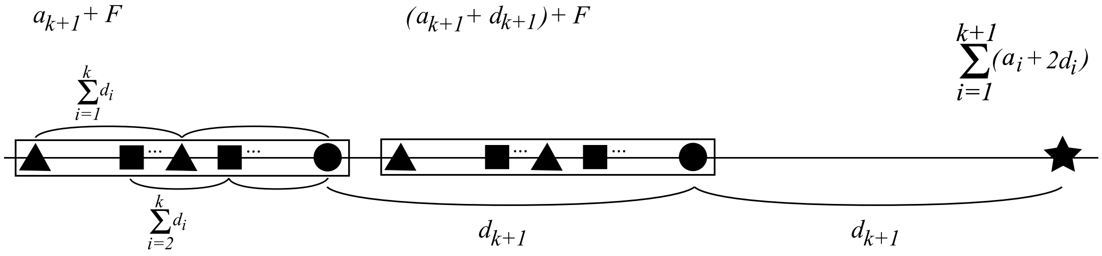

Figure 5.2: Fans with the same colour pattern

a_{t+1}, d_{t+1} with $0 \le a_{t+1} < a_{t+1} + d_{t+1} \le R_t$ and a fan $F_0 = F(a_1, d_1; \ldots; a_t, d_t)$ so that both fans $a_{t+1} + F_0$ and $(a_{t+1} + d_{t+1}) + F_0$ have the same colour pattern under Δ_0 and are both strongly polychromatic.

For each $\ell \in [1, t]$,

$$\Delta_0\left(\sum_{i=1}^{\ell-1}(a_i + 2d_i) + \sum_{i=\ell}^{t+1} a_i\right) = \Delta_0\left(\sum_{i=1}^{\ell-1}(a_i + 2d_i) + \sum_{i=\ell}^{t}(a_i + d_i) + a_{t+1}\right)$$

(since $a_{t+1} + F_0$ is polychromatic)

$$= \Delta_0\left(\sum_{i=1}^{\ell-1}(a_i + 2d_i) + \sum_{i=\ell}^{t}(a_i + d_i) + (a_{t+1} + d_{t+1})\right)$$

(since both fans have the same colour pattern)

and

$$\Delta_0\left(\sum_{i=1}^{t}(a_i + 2d_i) + a_{t+1}\right) = \Delta_0\left(\sum_{i=1}^{t}(a_i + 2d_i) + (a_{t+1} + d_{t+1})\right)$$

(since the two fans have the same colour pattern).

Therefore, since F_0 is strongly polychromatic, $F(a_1, d_1; \ldots; a_t, d_t; a_{t+1}, d_{t+1}) = F'$ is weakly polychromatic. If the endpoint of F' is the same colour as any of its spokes, there would be a monochromatic AP_3. Since it was assumed that this does not happen, $F(a_1, d_1; \dots; a_t, d_t; a_{t+1}, d_{t+1})$ must be strongly polychromatic.

As in the base case, for any strongly polychromatic fan $F(a_1, d_1; \dots; a_t, d_t)$, there are at most $\binom{R_t+1}{2}$ possible choices for the pair $\{a_{t+1}, a_{t+1} + d_{t+1}\}$ so that $0 \leq a_{t+1} < a_{t+1} + d_{t+1} \leq R_t$ and $(r - t)$ choices for the colour of the endpoint of $F(a_1, d_1; \dots; a_t, d_t; a_{t+1}, d_{t+1})$ so that the new fan is also strongly polychromatic. Therefore, there are at most $\binom{R_t+1}{2} R_t(r - t) = R_{t+1}$ (by equation (5.7)) different pairs $(F(a_1, d_1; \dots; a_{t+1}, d_{t+1}), \Delta)$ where the fan is strongly polychromatic under Δ. □

Thus the claim holds and by the remark immediately following the claim, for any $r \in \mathbb{Z}^+$, $W(3; r) \leq S_r$.

The following recursive bound on the sequence $\{R_i\}_{i=1}^{r}$ is useful for finding a bound on the number S_r in terms of r. For any $i \in [1, r]$,

$$\begin{aligned} 2R_i &= 2\binom{R_{i-1}+1}{2} R_{i-1}(r-i) \\ &= R_{i-1}^2 r \left(1 + \frac{1}{R_{i-1}}\right)\left(1 - \frac{i}{r}\right) \\ &\leq R_{i-1}^3 r \left(1 + \frac{1}{r}\right)\left(1 - \frac{1}{r}\right) \quad \text{(since } R_{i-1} \geq r \text{ and } i \leq r\text{)} \\ &= R_{i-1}^3 r \left(\frac{r^2-1}{r^2}\right) \\ &\leq R_{i-1}^3 r. \end{aligned} \tag{5.9}$$

Therefore,

$$
\begin{aligned}
S_r &= 1 + 2\sum_{i=0}^{r-1} R_i \\
&= 1 + 2r + \sum_{i=1}^{r-1} 2R_i && \text{(since } R_0 = r\text{)} \\
&< 1 + 2r + \sum_{i=1}^{r-1} R_{i-1}^3 r && \text{(by eq'n (5.9))} \\
&= 1 + 2r + r^3 \cdot r + \sum_{i=1}^{r-2} R_i^3 r \\
&< 1 + 2r + r^{3+1} + \sum_{i=1}^{r-2} \left(\frac{R_{i-1}^3 r}{2}\right)^3 r \\
&= 1 + 2r + r^{3+1} + \frac{r^{3^2+3+1}}{2^3} + \sum_{i=1}^{r-3} R_i^{3^2} \frac{r^{3+1}}{2^3} \\
&< 1 + 2r + r^{3+1} + \frac{r^{3^2+3+1}}{2^3} + \sum_{i=1}^{r-3} \left(\frac{R_{i-1}^3 r}{2}\right)^{3^2} \frac{r^{3+1}}{2^3} \\
&= 1 + 2r + r^{3+1} + \frac{r^{3^2+3+1}}{2^3} + \frac{r^{3^3+3^2+3+1}}{2^{3^2+3}} + \sum_{i=1}^{r-4} R_{i-1}^{3^3} \frac{r^{3^2+3+1}}{2^{3^2+3}} \\
&\vdots \\
&< 1 + 2r + r^{3+1} + \frac{r^{3^2+3+1}}{2^3} + \frac{r^{3^3+3^2+3+1}}{2^{3^2+3}} + \cdots + \frac{r^{3^{r-1}+3^{r-2}+\cdots+3+1}}{2^{3^{r-2}+\cdots 3}} \\
&= 1 + \frac{r^{(3^1-1)/2}}{2^{(3^0-3)/2}} + \frac{r^{(3^2-1)/2}}{2^{(3^1-3)/2}} + \frac{r^{(3^3-1)/2}}{2^{(3^2-3)/2}} + \cdots + \frac{r^{(3^r-1)/2}}{2^{(3^{r-1}-3)/2}} \\
&< \frac{5}{4}\frac{r^{(3^r-1)/2}}{2^{(3^{r-1}-3)/2}} && \text{(for } r \geq 2\text{)} \\
&= \frac{5}{2}\left(\frac{r}{2}\right)^{\frac{3^r-1}{2}}.
\end{aligned}
$$

□

Chapter 6

Discrepancy theory

6.1 Preliminaries

Van der Waerden's theorem (Theorem 3.1.1) shows that for any $k \in \mathbb{Z}^+$, there is an $n \in \mathbb{Z}^+$ so that every 2-colouring of $[1, n]$ produces a monochromatic AP_k. However, considering the known bounds on the van der Waerden numbers (Chapter 5), it is possible that n might be very large compared to k. Fixing an integer n and looking at the collection of all arithmetic progressions in $[1, n]$, is it possible to guarantee that for every 2-colouring of $[1, n]$, there is at least one arithmetic progression with many more elements of one colour than the other? Conversely, is it possible to find a 2-colouring of $[1, n]$ such that, for every arithmetic progression, the difference between the number of elements of each colour is relatively small?

Discrepancy theory is the study of this type of problem for arbitrary set systems. It will be useful to develop some general theory before turning specifically to the problem with regard to arithmetic progressions. Let S be a finite set, $\mathcal{H} \subseteq \mathcal{P}(S)$ and $H = (S, \mathcal{H})$ (often simply referred to by $\mathcal{H}$). Is it possible to find a partition of S that splits each member of $\mathcal{H}$ as equally as possible? Discrepancy theory seeks to measure how equally the "best" partition splits the members of $\mathcal{H}$.

For the purposes of discrepancy theory, bi-partitions (2-colourings) of S will be denoted by colouring functions $\Delta : S \to \{-1, 1\}$. The choice of image $\{-1, 1\}$ has the main advantage that if $A \in \mathcal{H}$, then $|\sum_{x \in A} \Delta(x)|$ is exactly the difference between the number of elements of A in each partition set $\Delta^{-1}(1)$ and $\Delta^{-1}(-1)$.

Definition 6.1.1. Let S be a finite set and $\mathcal{H} \subseteq \mathcal{P}(S)$. The *discrepancy* of $\mathcal{H}$ is defined to be

$$\mathcal{D}(\mathcal{H}) = \min_{\Delta : S \to \{-1,1\}} \left\{ \max_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right| \right\}.$$

Though it may be difficult to exactly calculate the discrepancy of a given set system, it is often possible to find upper and lower bounds. Unraveling the definition of discrepancy gives the following meanings to upper and lower bounds. Given a set S and $\mathcal{H} \subseteq \mathcal{P}(S)$, if $U \in \mathbb{R}$ is such that $\mathcal{D}(\mathcal{H}) \leq U$, then there is a 2-colouring of S where each $A \in \mathcal{H}$ is split with no more than U elements more in one colour class than in the other. On the other hand, if $L \in \mathbb{R}$ is such that $\mathcal{D}(\mathcal{H}) \geq L$, then for every 2-colouring of S, there is at least one $A \in \mathcal{H}$ where there are at least L

elements of A more in one colour class than in the other.

A different measure of discrepancy defined in a similar fashion can also be useful.

Definition 6.1.2. Let S be a finite set and $\mathcal{H} \subseteq \mathcal{P}(S)$. The ℓ_2-*discrepancy* of $\mathcal{H}$ is

$$\mathcal{D}_2(\mathcal{H}) = \min_{\Delta : S \to \{-1,1\}} \left\{ \left(\sum_{A \in \mathcal{H}} \left(\sum_{x \in A} \Delta(x) \right)^2 \right)^{1/2} \right\}.$$

Other forms of discrepancy such as hereditary discrepancy, linear discrepancy and weighted discrepancy can be found for example in [8, 21]. The following standard inequality appears in the same survey article.

Lemma 6.1.3. Let $(S, \mathcal{H})$ be a finite hypergraph. Then

$$\frac{\mathcal{D}_2(\mathcal{H})}{\sqrt{|\mathcal{H}|}} \leq \mathcal{D}(\mathcal{H}) \leq \mathcal{D}_2(\mathcal{H}).$$

Proof. For the lower bound, let $\Delta : S \to \{-1, 1\}$ be any 2-colouring. Then

$$\begin{aligned} \sum_{A \in \mathcal{H}} \left(\sum_{x \in A} \Delta(x) \right)^2 &= \sum_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right|^2 \\ &\leq \sum_{A \in \mathcal{H}} \left(\max_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right|^2 \right) \\ &\leq |\mathcal{H}| \left(\max_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right|^2 \right) \\ &= \sum_{A \in \mathcal{H}} \left(\max_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right| \right)^2. \end{aligned}$$

Therefore, $\left(\sum_{A \in \mathcal{H}} \left(\sum_{x \in A} \Delta(x) \right)^2 \right)^{1/2} \leq \sqrt{|\mathcal{H}|} \max_{A \in \mathcal{H}} \left| \sum_{x \in A} \Delta(x) \right|$ and taking the minimum over all 2-colourings gives $\mathcal{D}_2(\mathcal{H}) \leq \sqrt{|\mathcal{H}|} \mathcal{D}(\mathcal{H})$.

For the upper bound, again let $\Delta : S \to \{-1,1\}$ be any 2-colouring. Then,

$$\begin{aligned}\left(\max_{A\in\mathcal{H}}\left|\sum_{x\in A}\Delta(x)\right|\right)^2 &= \max_{A\in\mathcal{H}}\left[\sum_{x\in A}\Delta(x)\right]^2\\ &\leq \sum_{A\in\mathcal{H}}\left[\sum_{x\in A}\Delta(x)\right]^2.\end{aligned}$$

Therefore, $\max_{A\in\mathcal{H}}\left|\sum_{x\in A}\Delta(x)\right| \leq \left(\sum_{A\in\mathcal{H}}\left[\sum_{x\in A}\Delta(x)\right]^2\right)^{1/2}$ and thus, taking the minimum over all 2-colourings, $\mathcal{D}(\mathcal{H}) \leq \mathcal{D}_2(\mathcal{H})$. □

6.2 Lower bounds

One way to obtain bounds on the discrepancy of a hypergraph is to consider the incidence matrix of the hypergraph and turn to matrix theory to answer questions on discrepancy.

Recall that if $S = [1,n]$ and $\mathcal{H} = \{A_1,\ldots,A_m\} \subseteq \mathcal{P}(S)$, then the $m\times n$ incidence matrix, $M = (m_{ij})_{m\times n}$, for $(S,\mathcal{H})$ is defined by

$$m_{ij} = \begin{cases} 1, & \text{if } j \in A_i;\\ 0, & \text{otherwise.}\end{cases}$$

For any real-valued matrix B with real eigenvalues, let $\lambda_{\min}(B)$ be the minimum eigenvalue of B. The following theorem, stated without proof, provides a connection between discrepancy and eigenvalues. (The result is attributed to Lovász and Sós in [8].)

Theorem 6.2.1 (Lovász and Sós)**.** Let $(S, \mathcal{H})$ be a hypergraph with $|\mathcal{H}| = m$ and $|S| = n$ and let M be the incidence matrix for the hypergraph $(S, \mathcal{H})$. Then $\mathcal{D}_2(\mathcal{H}) \geq (n\lambda_{\min}(M^T M))^{1/2}$.

For each $n \in \mathbb{Z}^+$, set $S_n = [1, n]$ and let $\mathcal{H}_n$ be the collection of all arithmetic progressions in $[1, n]$. In 1964, Roth [95] proved that there is a positive constant c so that $\mathcal{D}(\mathcal{H}_n) > cn^{1/4}$. According to Prömel and Voigt [85], the constant c can be taken to be $1/60$. Roth's proof used analytic techniques and later, another proof was given that made use of matrix theory. (The following version of the proof appears in a survey article by Beck and Sós [8] and I have been unable to ascertain its origin though it may be due to Lovász.)

Theorem 6.2.2 (Roth [95])**.** For n sufficiently large, $\mathcal{D}(\mathcal{H}_n) > \frac{1}{20}n^{1/4}$.

Proof. Fix $n \in \mathbb{Z}^+$ and set $k = \lfloor \sqrt{n/6} \rfloor$. Let $\mathcal{H}$ be the hypergraph consisting of the collection of all k-subsets of $[1, n]$ of the form

$$A(a, d) = \{a + td \pmod{n} : a \in [1, n],\ 0 \leq t \leq k - 1 \text{ and } d \leq 6k\}.$$

Then, $|\mathcal{H}| \leq n \cdot 6k$. For any $a \in [1, n]$ and $d \leq 6k$, the arithmetic progression $A(a, d)$ spans at most $6k^2 \leq n$ elements. Therefore, $A(a, d)$ is the union of at most

2 proper arithmetic progressions (with the usual addition). If A_1 and A_2 are two disjoint arithmetic progressions with $A(a,d) = A_1 \cup A_2$, then

$$\max\left\{\left|\sum_{x\in A_1}\Delta(x)\right|, \left|\sum_{x\in A_2}\Delta(x)\right|\right\} \geq \frac{1}{2}\left|\sum_{x\in A(a,d)}\Delta(x)\right| \tag{6.1}$$

and thus $\mathcal{D}(\mathcal{H}_n) \geq \frac{1}{2}\mathcal{D}(\mathcal{H})$.

Let $M = (m_{st})_{|\mathcal{H}|\times n}$ be the incidence matrix for $(S_n, \mathcal{H})$. Recall that any matrix $B = (b_{s,t})_{n\times n}$ is *circulant* iff for each $s, t \in [1, n]$, $b_{s,t} = b_{s+1,t+1}$, where the addition of indices is taken modulo n. If $M = [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \cdots\ \boldsymbol{c}_n]$ (where the $\boldsymbol{c}_j$'s are all column vectors), then the (s,t)-th entry of the matrix M^TM is $\boldsymbol{c}_s^T \cdot \boldsymbol{c}_t$. The column vector $\boldsymbol{c}_{s+1}$ can be obtained from $\boldsymbol{c}_s$ by permuting the rows of the matrix M, moving the row corresponding to the arithmetic progression $A(a,d)$ to the row corresponding to $A(a+1,d)$. Thus, if $s,t \in [1,n]$ then $\boldsymbol{c}_s^T \cdot \boldsymbol{c}_t = \boldsymbol{c}_{s+1}^T \cdot \boldsymbol{c}_{t+1}$ and so the (s,t)-th entry of M^TM is equal to its $(s+1,t+1)$-th entry. Thus M^TM is circulant and there are $b_1, \ldots, b_n$ such that

$$M^TM = \begin{bmatrix} b_1 & b_2 & \ldots & b_n \\ b_n & b_1 & \ldots & b_{n-1} \\ \vdots & & & \vdots \\ b_2 & b_3 & \ldots & b_1 \end{bmatrix}.$$

When ω is any n-th root of unity,

$$M^TM\begin{bmatrix}\omega\\ \omega^2\\ \vdots\\ \omega^n=1\end{bmatrix} = \begin{bmatrix}b_1 & b_2 & \dots & b_n\\ b_n & b_1 & \dots & b_{n-1}\\ \vdots & & & \\ b_2 & b_3 & \dots & b_1\end{bmatrix}\begin{bmatrix}\omega\\ \omega^2\\ \vdots\\ \omega^n\end{bmatrix}$$

$$= \begin{bmatrix}\sum_{j=1}^n b_j\omega^j\\ \sum_{j=1}^n b_j\omega^{j+1}\\ \vdots\\ \sum_{j=1}^n b_j\omega^{j+n-1}\end{bmatrix}$$

$$= \sum_{j=1}^n b_j\omega^{j-1}\begin{bmatrix}\omega\\ \omega^2\\ \vdots\\ \omega^n\end{bmatrix}.$$

Thus $\lambda_\omega = \sum_{j=1}^n b_j\omega^{j-1}$ is an eigenvalue for M^TM and $\boldsymbol{\omega} = [\omega\ \omega^2\ \cdots\ \omega^n]^T$ is a corresponding eigenvector. Since there are exactly n such n-th roots of unity, all eigenvalues of M^TM must be of the form $\lambda_\omega = \sum_{j=1}^n b_j\omega^{j-1}$. A lower bound for $\lambda_{\min}(M^TM)$ can be obtained by finding a lower bound for an arbitrary eigenvalue λ_ω. For any complex valued matrix $B = (b_{i,j})$ denote by $B^* = (\overline{b_{i,j}})^T$ the conjugate transpose of B.

Since M is a real-valued matrix, $M^T = M^*$. Thus $\lambda_\omega\boldsymbol{\omega} = M^TM\boldsymbol{\omega} = M^*M\boldsymbol{\omega}$

and so

$$\begin{aligned}
\lambda_\omega &= \lambda_\omega \frac{\boldsymbol{\omega}^* \cdot \boldsymbol{\omega}}{\|\boldsymbol{\omega}\|^2} \\
&= \frac{1}{\|\boldsymbol{\omega}\|^2} \boldsymbol{\omega}^* \lambda_\omega \boldsymbol{\omega} \\
&= \frac{1}{\|\boldsymbol{\omega}\|^2} \boldsymbol{\omega}^* M^* M \boldsymbol{\omega} \\
&= \frac{1}{|\omega|^2 + |\omega^2|^2 + \cdots |\omega^n|^2} (M\boldsymbol{\omega})^*(M\boldsymbol{\omega}) \\
&= \frac{1}{n} \sum_{s=1}^{n \cdot 6k} \left| \sum_{t=1}^{n} m_{st} \omega^t \right|^2 \\
&= \frac{1}{n} \sum_{A \in \mathcal{H}} \left| \sum_{t \in A} \omega^t \right|^2 && \text{(since } m_{st} = 1 \text{ iff } t \in A_j\text{)} \\
&= \frac{1}{n} \sum_{a=1}^{n} \left(\sum_{d=1}^{6k} \left| \sum_{j \in A(a,d)} \omega^j \right|^2 \right). && (6.2)
\end{aligned}$$

Let $A = A(a,d) \in \mathcal{H}$ and $x \in [1,n]$. If $A' = A(a+x,d)$, then

$$\left| \sum_{j \in A'} \omega^j \right| = \left| \sum_{j \in A} \omega^x \omega^j \right| = |\omega^x| \left| \sum_{j \in A} \omega^j \right| = \left| \sum_{j \in A} \omega^j \right|.$$

Thus, for any d, if $a \in [1,n]$, then $|\sum_{j \in A(a,d)} \omega^j| = |\sum_{j \in A(n,d)} \omega^j|$ and continuing from equation (6.2) above,

$$\begin{aligned}
\lambda_\omega &= \frac{n}{n} \sum_{d=1}^{6k} \left| \sum_{j \in A(n,d)} \omega^j \right|^2 \\
&= \sum_{d=1}^{6k} \left| \sum_{j=0}^{k-1} \omega^{n+jd} \right|^2 \\
&= \sum_{d=1}^{6k} \left| \sum_{j=1}^{k-1} \omega^{jd} \right|^2 && \text{(since } \omega^n = 1\text{).} && (6.3)
\end{aligned}$$

Consider the complex numbers $\omega^0, \omega^1, \ldots, \omega^{6k}$. For any complex number z, $\mathrm{Arg}(z) \in [0, 2\pi)$, and so by the pigeonhole principle, there are d_1 and d_2 with $0 \leq d_1 < d_2 \leq 6k$ such that $|\mathrm{Arg}(\omega^{d_1}) - \mathrm{Arg}(\omega^{d_2})| \leq \frac{2\pi}{6k} = \frac{\pi}{3k}$. Set $d_0 = d_1 - d_2$. Then, $1 \leq d_0 \leq 6k$ and $|\mathrm{Arg}(\omega^{d_0})| = |\mathrm{Arg}(\omega^{d_1}) - \mathrm{Arg}(\omega^{d_2})| \leq \frac{\pi}{3k}$. Therefore, if $j \in [0, k-1]$, then $\frac{\pi}{3k} j \leq \frac{\pi}{3}$ and hence the real part of $\omega^{d_0 j}$ satisfies $\Re(\omega^{j d_0}) \geq \frac{1}{2}$. Combining this with equation (6.3) gives,

$$\lambda_\omega \geq \left| \sum_{j=0}^{k-1} \omega^{j d_0} \right|^2 \geq \left[\Re\left(\sum_{j=0}^{k-1} \omega^{j d_0} \right) \right]^2 \geq \left[\sum_{j=0}^{k-1} \frac{1}{2} \right]^2 = \frac{k^2}{4}. \tag{6.4}$$

Since ω was arbitrary, $\lambda_{\min}(M^T M) \geq \frac{k^2}{4}$. This bound together with Theorem 6.2.1 can be used to find a lower bound for $\mathcal{D}(\mathcal{H})$.

$$\begin{aligned}
\mathcal{D}(\mathcal{H}) &\geq \frac{1}{\sqrt{|\mathcal{H}|}} \mathcal{D}_2(\mathcal{H}) && \text{(by Lemma 6.1.3)} \\
&\geq \frac{1}{\sqrt{|\mathcal{H}|}} [n \lambda_{\min}(M^T M)]^{1/2} && \text{(by Theorem 6.2.1)} \\
&= \left[\frac{n}{n \cdot 6k} \lambda_{\min}(M^T M) \right]^{1/2} && \\
&\geq \left(\frac{1}{6k} \cdot \frac{k^2}{4} \right)^{1/2} && \text{(by eq'n (6.4))} \\
&= \left(\frac{k}{24} \right)^{1/2} && \\
&\geq \left(\frac{\sqrt{\frac{n}{6}} - 1}{24} \right)^{1/2} && \text{(since } k = \lfloor \sqrt{n/6} \rfloor\text{)} \\
&= n^{1/4} \left(\frac{1 - \sqrt{\frac{6}{n}}}{24\sqrt{6}} \right)^{1/2} && \\
&> n^{1/4} \frac{1}{10} && \text{(for } n \geq 36\text{).}
\end{aligned}$$

Finally, since $\mathcal{D}(\mathcal{H}_n) > \frac{1}{2}\mathcal{D}(\mathcal{H})$, by eq'n (6.1),

$$\mathcal{D}(\mathcal{H}_n) > \frac{1}{2}\mathcal{D}(\mathcal{H}) > \frac{1}{20}n^{1/4}. \qquad \square$$

6.3 Upper bounds

Erdős [33] (in Hungarian) showed that $\mathcal{D}(\mathcal{H}_n) < 10n^{1/2}$. (The proof appears in English in [39, §8].) Using an extension of Erdős's method, Spencer [103] improved this bound, showing that $\mathcal{D}(\mathcal{H}_n) < 100\sqrt{\frac{n \ln\ln n}{\ln n}}$. It was conjectured (see [39, p. 39]) that if $\varepsilon > 0$, then for n sufficiently large, $\mathcal{D}(\mathcal{H}_n) > n^{\frac{1}{2}-\varepsilon}$. This was shown to be false by Sárközy and Montgomery (see [38, p. 39] and [38, Problem 10]) who showed that there is a constant c so that $\mathcal{D}(\mathcal{H}_n) < cn^{1/3}(\ln n)^{2/3}$. The focus of this section is to present Beck's proof [6] that there is a constant c' so that for n sufficiently large, $\mathcal{D}(\mathcal{H}_n) < c'n^{1/4}(\ln n)^{5/2}$.

Definition 6.3.1. For any hypergraph $(S, \mathcal{H})$ the *maximum degree* of $\mathcal{H}$ is defined to be $\text{maxdeg}(\mathcal{H}) = \max_{x\in S} |\{A \in \mathcal{H} : x \in A\}|$.

The following theorem, stated without proof, provides a connection between discrepancy and the maximum degree of a hypergraph $\mathcal{H}$.

Theorem 6.3.2 (Beck-Fiala [7])**.** Let $\mathcal{H}$ be a finite hypergraph. Then

$$\mathcal{D}(\mathcal{H}) < \text{maxdeg}(\mathcal{H}).$$

The following theorem can be used to find an upper bound for the discrepancy of arithmetic progressions and provides a more complicated link between discrepancy and degree. Essentially, it says that if the hypergraph contains large hyperedges that are relatively spread out, then the discrepancy will be small.

Theorem 6.3.3 (Beck [6])**.** Let $(S, \mathcal{H})$ be a finite hypergraph and $t \in \mathbb{R}$ be such that $\text{maxdeg}(\{A \in \mathcal{H} : |A| \geq t\}) \leq t$. Then there is a positive constant c so that

$$\mathcal{D}(\mathcal{H}) \leq ct^{1/2}(\ln|\mathcal{H}|)^{1/2}\ln|S|.$$

Beck and Sós [8] provide a more general version of this result that gives bounds on $\mathcal{D}(\mathcal{H})$ if $\text{maxdeg}(\{A \in \mathcal{H} : |A| \geq m\}) \leq t$. The proof of the next theorem, which uses only the symmetric case, shows that Roth's bound is nearly sharp. Throughout (as before), for each $n \in \mathbb{Z}^+$, set $S_n = [1, n]$ and let $\mathcal{H}_n$ be the collection of all arithmetic progressions in $[1, n]$.

Theorem 6.3.4 (Beck [6])**.** There is a constant $c > 0$ so that for each $n \in \mathbb{Z}^+$ sufficiently large, $\mathcal{D}(\mathcal{H}_n) < cn^{1/4}(\ln n)^{5/2}$.

Proof. For $a, d \in \mathbb{Z}^+$ and $i \leq j$ set $AP(a, d, i, j) = \{a + xd : i \leq x \leq j\}$, the arithmetic progression with difference d, beginning at $a+id$ and ending with $a+jd$.

Call an arithmetic progression A *elementary* if there are integers $d \geq 1$, $1 \leq b \leq d$, $i \geq 0$, and $s \geq 0$ with

$$A = AP(b, d, i2^s, (i+1)2^s - 1).$$

Let $\mathcal{H}_n^*$ be the family of all elementary arithmetic progressions in $[1,n]$. For any arithmetic progression $A = \{a+d, \ldots, a+kd\} \subseteq [1,n]$, let $p_1 \geq 0$ and $1 \leq b < d$ be such that $a = b + p_1 d$ and set $p_2 = k + p_1$. Set $A_1 = AP(b,d,0,p_1)$ and $A_2 = AP(b,d,0,p_2)$. Then $A = A_2 \backslash A_1$ and for any 2-colouring $\Delta : S_n \to \{-1,1\}$,

$$\begin{aligned}\left|\sum_{x\in A}\Delta(x)\right| &= \left|\sum_{x\in A_2}\Delta(x) - \sum_{x\in A_1}\Delta(x)\right| \\ &\leq \left|\sum_{x\in A_2}\Delta(x)\right| + \left|\sum_{x\in A_1}\Delta(x)\right| \\ &\leq 2\max_{i\in\{1,2\}}\left|\sum_{x\in A_i}\Delta(x)\right|. \qquad (6.5)\end{aligned}$$

For each $i \in \{1,2\}$, if $s(i,1) > s(i,2) > \cdots > s(i,\ell_i)$ are such that $p_i + 1 = \sum_{j=1}^{\ell_i} 2^{s(i,j)}$, then $A_i = \bigcup_{t=1}^{\ell_i} AP(b,d,\sum\limits_{j=1}^{t-1} 2^{s(i,j)}, \sum\limits_{j=1}^{t} 2^{s(i,j)} - 1)$, a union of elementary arithmetic progressions. Since $p_i + 1 \geq 2^{\ell_i} - 1$, the arithmetic progression A_i is the disjoint union of no more than $\ell_i \leq \log_2(p_i+2) \leq \log_2 n + 1 \leq 2\log_2 n$ elementary arithmetic progressions. If $\{B_{i,j}\}_{j=1}^{2\log_2 n} \subseteq \mathcal{H}_n^*$ are such that $A_i = \bigcup_{j=1}^{2\log_2 n} B_{i,j}$, then for any 2-colouring $\Delta : S_n \to \{-1,1\}$,

$$\begin{aligned}\left|\sum_{x\in A_i}\Delta(x)\right| &\leq \sum_{j=1}^{\log_2 n}\left|\sum_{x\in B_{i,j}}\Delta(x)\right| \\ &\leq 2\log_2 n \max_{1\leq j\leq \log_2 n}\left|\sum_{x\in B_{i,j}}\Delta(x)\right|. \qquad (6.6)\end{aligned}$$

Thus, combining equations (6.5) and (6.6), $\mathcal{D}(\mathcal{H}) \leq 4\log_2 n\ \mathcal{D}(\mathcal{H}_n^*)$. Also, $\mathcal{H}_n^*$ is contained in the family of all arithmetic progressions in $[1,n]$ whose lengths are a power of 2. By Lemma 1.2.3, for any $n, k > 1$, the number of AP_k's in $[1,n]$ is no more than $\frac{n^2}{2(k-1)}$. Thus, for some constant c,

$$|\mathcal{H}_n^*| \leq n + \sum_{k=2}^{\log_2 n} \frac{n^2}{2(k-1)} \leq cn^2. \tag{6.7}$$

In order to use Theorem 6.3.3 to find an upper bound for $\mathcal{D}(\mathcal{H}_n^*)$ it is necessary to find a bound for the maximum degree of certain sub-collections of $\mathcal{H}_n$: For any $m \in \mathbb{Z}^+$,

$$\begin{aligned}
&\operatorname{maxdeg}(\{A \in \mathcal{H}_n^* : |A| \geq m\}) \\
&\qquad = \max_{x\in[1,n]} |\{A \in \mathcal{H}_n^* : |A| \geq m \text{ and } x \in A\}| \\
&\qquad \leq \max_{x\in[1,n]} \sum_{1\leq d\leq \frac{n-1}{m-1}} \sum_{\substack{1\leq b\leq d \\ b\equiv x \bmod d}} |\{s : 2^s \geq m,\ (b+(2^s-1)d) \leq n\}| \\
&\qquad \leq \max_{x\in[1,n]} \sum_{1\leq d\leq \frac{n-1}{m-1}} \sum_{\substack{1\leq b\leq d \\ b\equiv x \bmod d}} \left(\log_2\left(\frac{n-b+d}{d}\right) - \log_2 m + 1\right) \\
&\qquad \leq \max_{x\in[1,n]} \sum_{1\leq d\leq \frac{n-1}{m-1}} \sum_{\substack{1\leq b\leq d \\ b\equiv x \bmod d}} c' \log_2\left(\frac{n}{md}\right) \qquad \text{(for some constant } c'\text{)} \\
&\qquad = \sum_{1\leq d\leq \frac{n-1}{m-1}} c' \log_2 \frac{n}{md} \\
&\qquad = c' \log_2\left(\prod_{1\leq d\leq \frac{n-1}{m-1}} \frac{n}{md}\right) \\
&\qquad = c' \log_2\left(\left(\frac{n}{m}\right)^{(\frac{n-1}{m-1})} \frac{1}{(\frac{n-1}{m-1})!}\right).
\end{aligned} \tag{6.8}$$

Now, by Stirling's formula that $n! \sim \left(\frac{n}{e}\right)\sqrt{2\pi n}$ (see [25, pp.361–4]),

$$\left(\frac{n-1}{m-1}\right)! \sim \left(\frac{(n-1)/(m-1)}{e}\right)^{\left(\frac{n-1}{m-1}\right)} \sqrt{2\pi\left(\frac{n-1}{m-1}\right)}$$

$$> \left(\frac{n-1}{m-1}\right)^{(\frac{n-1}{m-1})} e^{-(\frac{n-1}{m-1})}.$$

Thus,

$$\begin{aligned} c' \log_2 \left(\left(\frac{n}{m}\right)^{(\frac{n-1}{m-1})} \frac{1}{(\frac{n-1}{m-1})!} \right) &< c' \log_2 \left(\left(\frac{n/m}{(n-1)/(m-1)} \right)^{(\frac{n-1}{m-1})} e^{(\frac{n-1}{m-1})} \right) \\ &= c' \left(\frac{n-1}{m-1}\right) \log_2 \left(\frac{n/m}{(n-1)/(m-1)} e \right) \\ &< c^* \frac{n}{m} \qquad \text{(for some constant } c^*\text{).} \qquad (6.9) \end{aligned}$$

Taking $m = \sqrt{c^* n}$ in the above, and combining equations (6.8) and (6.9)

$$\operatorname{maxdeg}(\{A \in \mathcal{H}^* : |A| \geq \sqrt{c^* n}\}) \leq \frac{c^* n}{\sqrt{c^* n}} = \sqrt{c^* n}.$$

Therefore, by applying Theorem 6.3.3,

$$\begin{aligned} \mathcal{D}(\mathcal{H}_n^*) &\leq c(c^* n)^{1/4} (\ln \mathcal{H}_n^*)^{1/2} \ln n \\ &\leq c(c^* n)^{1/4} (\ln c' n^2)^{1/2} \ln n \qquad \text{(by eq'n (6.7))} \\ &\leq c'' n^{1/4} (\ln n)^{3/2} \qquad \text{(for some constant } c'\text{).} \end{aligned}$$

By the remark at the beginning of the proof,

$$\begin{aligned} \mathcal{D}(\mathcal{H}_n) &\leq 4 \log_2 n \; \mathcal{D}(\mathcal{H}_n^*) \\ &\leq c n^{1/4} (\ln n)^{5/2} \qquad \text{(for some constant } c\text{).} \qquad \square \end{aligned}$$

It has since been shown by Matoušek and Spencer [82] that in fact Roth's original bound is sharp up to the constant.

Theorem 6.3.5 (Matoušek and Spencer [82]). There is a constant c so that for n sufficiently large $\mathcal{D}(\mathcal{H}_n) \leq cn^{1/4}$.

The results of Roth and Beck on discrepancy of arithmetic progressions have been extended by Valkó [111] to generalizations of arithmetic progressions in $(\mathbb{Z}^+)^d$.

Definition 6.3.6. For each $n, d \in \mathbb{Z}^+$, set

$$\mathcal{H}_{n,d} = \{\{\boldsymbol{a} + i\boldsymbol{b} : 0 \leq i \leq k-1\} : \boldsymbol{a}, \boldsymbol{b} \in [1,n]^d \text{ and } k \geq 1\}.$$

In the case $d = 1$, $\mathcal{H}_{n,1} = \mathcal{H}_n$, the hypergraph given by the collection of all arithmetic progressions in the interval $[1, n]$.

Theorem 6.3.7 (Valkó [111]). For each $d \in \mathbb{Z}^+$, there are constants c_1, c_2 (depending on d) so that for sufficiently large $n \in \mathbb{Z}^+$,

$$c_1 n^{\frac{d}{2d+2}} \leq \mathcal{D}(\mathcal{H}_{n,d}) \leq c_2 n^{\frac{d}{2d+2}} (\ln n)^{5/2}.$$

The proof of this theorem, which is omitted here, relies on extending the techniques of Roth's proof [95] of Theorem 6.2.2 for the lower bound and the techniques of Beck's proof [6] of Theorem 6.3.4 for the upper bound. When $d = 1$, Theorem 6.3.7 gives (up to the constant) both Theorem 6.2.2 and Theorem 6.3.4.

Chapter 7

Hales-Jewett theorem

7.1 Definitions

The Hales-Jewett theorem is a purely combinatorial generalization of van der Waerden's theorem. The Hales-Jewett theorem examines problems related to a generalized version of the game Tic-Tac-Toe and guarantees that for any integer k, there is an n so that for every 2-colouring of the set $[1, k]^n$, there will be a monochromatic set of k elements all in a line. Associating n-tuples with integers either by adding all the coordinates or by treating coordinates as a base decomposition, these lines of k n-tuples can be associated with arithmetic progressions of length k. If arithmetic structures are not of concern, there is no need for the underlying set to be $[1, k]^n$ and the problem can be stated purely combinatorially for any finite set A, colouring

A^n. Before this can be done, some notation and definitions are necessary.

Throughout, let A be a finite set, called an *alphabet*. It will be convenient to let A^n be interchangeably both the set of n-tuples over A and the set of all functions $\{f : [1, n] \to A\}$.

Definition 7.1.1. Let $m \leq n \in \mathbb{Z}^+$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ be distinct symbols not in A called *parameters*. A function $f : [1, n] \to (A \cup \{\lambda_1, \ldots, \lambda_m\})$ is called an *m-parameter word of length n over A* iff for each $1 \leq i < j \leq m$, $f^{-1}(\lambda_i) \neq \emptyset$ and $\min f^{-1}(\lambda_i) < \min f^{-1}(\lambda_j)$.

The second condition is to ensure that the first occurrences of each parameter appear in increasing order, but it is for the most part unnecessary.

For any $m, n \in \mathbb{Z}^+$, the set of all m-parameter words of length n over A is denoted by $[A]\binom{n}{m}$. Note that the elements of A^n can be considered 0-parameter words and so $A^n = [A]\binom{n}{0}$.

Definition 7.1.2. Given $f \in [A]\binom{n}{m}$ and $g \in [A]\binom{m}{k}$, the composition of f and g, $f \circ g \in [A]\binom{n}{k}$, is defined as follows. For each $i \in [1, n]$,

$$f \circ g(i) \begin{cases} f(i), & \text{if } f(i) \in A; \\ \\ g(j), & \text{if for some } 1 \leq j \leq m,\ f(i) = \lambda_j. \end{cases}$$

Definition 7.1.3. A subset $M \subseteq A^n$ is a *combinatorial m-space* iff there is an $f \in [A]\binom{n}{m}$ such that $M = f \circ A^m = \{f \circ (a_1, \ldots, a_m) : (a_1, \ldots, a_m) \in A^m\}$. A combinatorial 1-space is called a *combinatorial line.*

For example, if $A = \{0, 1\}$, consider the 2-parameter word $f = (\lambda_1, \lambda_2, 0, \lambda_1) \in [A]\binom{4}{2}$. The corresponding combinatorial 2-space is

$$f \circ A^2 = \{(0, 0, 0, 0), (1, 0, 0, 1), (0, 1, 0, 0), (1, 1, 0, 1)\}.$$

The notion of a combinatorial space is not the same concept as an affine space. For example, in $\{0, 1, 2\}^2$, while $\{(0, 0), (1, 1), (2, 2)\}$ is a combinatorial line corresponding to the 1-parameter word (λ_1, λ_1), the set $\{(0, 2), (1, 1), (2, 0)\}$ is not, even though its elements satisfy the linear equation $y = 2 - x$ over $\mathbb{Z}_2$.

It is now possible to state the Hales-Jewett theorem precisely in terms of combinatorial m-spaces. The proof is deferred until Section 7.3.

Theorem 7.1.4 (Hales, Jewett [58]). Let A be an alphabet and let m, r be positive integers. There exists a positive integer $n = HJ(|A|, m; r)$ such that for every r-colouring, $\Delta : A^n \to [1, r]$, there is a monochromatic combinatorial m-space in A^n.

Definition 7.1.5. Given $f \in [A]\binom{n}{m}$ and $g \in [A]\binom{\ell}{k}$, the *concatenation* of f and g, $f^\frown g \in [A]\binom{n+\ell}{m+k}$ is defined as follows:

$$f^\frown g(i) = \begin{cases} f(i), & \text{if } 1 \leq i \leq n; \\ g(i-n), & \text{if } i > n \text{ and } g(i-n) \in A; \\ \lambda_{j+m} & \text{if } i > n \text{ and } g(i-n) = \lambda_j. \end{cases}$$

In other words, the n-tuple associated with f and the ℓ-tuple associated with g are concatenated as usual, but the parameters in g are renamed to ensure that they are different from the parameters in f.

7.2 Shelah Cube Lemma

The Shelah Cube Lemma is a tool that can be used to show that the functions $HJ(t, m; r)$, associated with the Hales-Jewett theorem, and $W(k; r)$, the van der Waerden numbers, are primitive recursive (as described in Chapter 5). Recall that $[n]^2$ denotes the set of all 2-element subsets of $[1, n] = \{1, \dots n\}$. (See also [84].)

Theorem 7.2.1 (Shelah [102])**.** For all integers m and r, there is a least integer $n = Sh(m; r)$ such that, for any sequence of m r-colourings $(1 \leq i \leq m)$:

$$\Delta_i : \underbrace{[n]^2 \times \cdots \times [n]^2}_{i-1} \times n \times \underbrace{[n]^2 \times \cdots \times [n]^2}_{m-i} \to [1, r]$$

there are m pairs $a_1 < b_1, a_2 < b_2, \dots, a_m < b_m$ where for each $i \in [1, m]$, $\{a_i, b_i\} \in [n]^2$ and

$$\Delta_i(\{a_1, b_1\}, \dots, \{a_{i-1}, b_{i-1}\}, a_i, \{a_{i+1}, b_{i+1}\}, \dots, \{a_m, b_m\})$$

$$= \Delta_i(\{a_1, b_1\}, \dots, \{a_{i-1}, b_{i-1}\}, b_i, \{a_{i+1}, b_{i+1}\}, \dots, \{a_m, b_m\}).$$

Proof. The proof proceeds by induction on m.

Base Case: If $m = 1$, then $Sh(1; r) = r + 1$, for if $\Delta_1 : [1, r + 1] \to [1, r]$, then by the pigeonhole principle, there is a pair $a_1 < b_1 \leq n$ such that $\Delta_1(a_1) = \Delta_1(b_1)$.

Inductive Step: Suppose that for some $k \geq 1$, $Sh(k;r)$ exists.

Set $n = Sh(k;r)$ and $N = 1 + r^{\binom{n}{2}^k}$, then it can be shown that $Sh(k+1;r) \leq N$.

For $i \in [1, k+1]$, let

$$\Delta_i : \underbrace{[N]^2 \times \cdots \times [N]^2}_{i-1} \times N \times \underbrace{[N]^2 \times \cdots \times [N]^2}_{k+1-i} \rightarrow [1, r]$$

be any r-colourings.

Consider the induced colouring, $\Delta^{'}_{k+1} : N \rightarrow r^{\binom{n}{2}^k}$ defined by

$$\Delta'_{k+1}(x) = \prod_{x_1<y_1\leq n, \ldots, x_k<y_k\leq n} \{\Delta_{k+1}(\{x_1, y_1\}, \ldots, \{x_k, y_k\}, x)\}.$$

That is, each colour is an $\binom{n}{2}^k$-tuple with entries from $[1, r]$. By the pigeonhole principle, since $N = 1 + r^{\binom{n}{2}^k}$, there are two elements $a_{k+1} < b_{k+1} \leq N$ such that $\Delta^{'}_{k+1}(a_{k+1}) = \Delta'_{k+1}(b_{k+1})$. That is, for all $x_1 < y_1 \leq n, \ldots, x_k < y_k \leq n$, $\Delta_{k+1}(\{x_1, y_1\}, \ldots, \{x_k, y_k\}, a_{k+1}) = \Delta_{k+1}(\{x_1, y_1\}, \ldots, \{x_k, y_k\}, b_{k+1})$.

For each $i \in [1, k]$, define an induced r-colouring,

$$\Delta'_i : \underbrace{[n]^2 \times \cdots \times [n]^2}_{i-1} \times n \times \underbrace{[n]^2 \times \cdots \times [n]^2}_{k-i} \rightarrow [1, r]$$

as follows

$$\Delta'_i(\{x_1, y_1\}, \ldots, x, \ldots, \{x_k, y_k\})$$

$$= \Delta_i(\{x_1, y_1\}, \ldots, x, \ldots, \{x_k, y_k\}, \{a_{k+1}, b_{k+1}\}).$$

By the induction hypothesis, since $n = Sh(k;r)$, there are k pairs of integers $a_1 < b_1 \leq n, \ldots, a_k < b_k \leq n$ such that for each $i \in [1,k]$,

$$\Delta'_i(\{a_1,b_1\},\ldots,a_i,\ldots\{a_k,b_k\}) = \Delta'_i(\{a_1,b_1\},\ldots,b_i,\ldots,\{a_k,b_k\})$$

and hence

$$\Delta_i(\{a_1,b_1\},\ldots a_i \ldots\{a_{k+1},b_{k+1}\}) = \Delta_i(\{a_1,b_1\},\ldots,b_i,\ldots,\{a_{k+1},b_{k+1}\}).$$

Finally, by the choice of a_k and b_k,

$$\Delta_{k+1}(\{a_1,b_1\},\ldots,\{a_k,b_k\},a_{k+1}) = \Delta_{k+1}(\{a_1,b_1\},\ldots,\{a_k,b_k\},b_{k+1}).$$

Therefore, $Sh(k+1;r) \leq N$ and so by induction, for all $m, r \in \mathbb{Z}^+$, the number $Sh(m;r)$ exists. □

7.3 Proof of the Hales-Jewett theorem

The original proof of the Hales-Jewett theorem used a double induction. In this section, a proof of the Hales-Jewett theorem is given that uses only a single induction and the Shelah cube lemma (see [84] for a simple write-up of the proof).

Theorem 7.1.4 (Hales, Jewett [58])**.** Let A be an alphabet and let m, r be positive integers. There exists a positive integer $n = HJ(|A|, m; r)$ such that for every r-colouring, $\Delta : A^n \to [1,r]$, there is a monochromatic combinatorial m-space in A^n.

Proof. Fix $r \in \mathbb{Z}^+$. First it is shown, by induction on t, that for all $t \geq 1$, the number $HJ(t,1;r)$ exists.

Base Case: For any integer r, $HJ(1,1;r) = 1$.

Inductive Step: Suppose that for some $t \geq 1$, $HJ(t,1,r)$ exists. Set $m = HJ(t,1;r)$ and $n = Sh(m; r^{(t+1)^m})$. Then it can be shown that $HJ(t+1,1;r) \leq mn$.

Let A be an alphabet with $|A| = t+1$ and let $\Delta : A^{mn} \to [1,r]$ be any r-colouring. Fix two elements $c, d \in A$, let $B = A \backslash \{d\}$. For each $i \in [0, n-1]$, define $h_i \in A^n$ by

$$h_i = (\underbrace{c, \dots, c}_{i \text{ times}}, \underbrace{d, \dots, d}_{(n-i) \text{ times}}),$$

and for each $0 \leq i < j \leq n-1$, define a 1-parameter word,

$$g_{i,j} = (\underbrace{c, \dots, c}_{i \text{ times}}, \underbrace{\lambda_1, \dots \lambda_1}_{j-i \text{ times}}, \underbrace{d, \dots, d}_{n-j \text{ times}}).$$

Note that $g_{i,j} \circ c = h_j$ and $g_{i,j} \circ d = h_i$. For each $i \in [1,m]$, define an r-colouring

$$\Delta_i : \underbrace{[n]^2 \times \cdots \times [n]^2}_{i \text{ times}} \times n \times \underbrace{[n]^2 \times \cdots \times [n]^2}_{m-1-i \text{ times}} \to [1, r^{|A|^m}]$$

as follows. For each sequence of pairs $x_1 < y_1, \dots, x_m < y_m \leq n$, and $x \leq n$, set

$$\Delta_i(\{x_1, y_1\}, \dots, \{x_{i-1}, y_{i-1}\}, x, \{x_{i+1}, y_{i+1}\}, \dots, \{x_m, y_m\})$$

$$= \prod_{(\alpha_1, \dots, \alpha_m) \in A^m} \{\Delta(g_{x_1,y_1} \circ (\alpha_1)^\frown \cdots {}^\frown g_{x_{i-1},y_{i-1}} \circ (\alpha_{i-1})$$

$$\widehat{\ } h_x \widehat{\ } g_{x_{i+1},y_{i+1}} \circ (\alpha_{i+1}) \widehat{\ } \cdots \widehat{\ } g_{x_m,y_m} \circ (\alpha_m))\}.$$

By Theorem 7.2.1 and the choice of $n = Sh(m; r^{(t+1)^m})$, there are m pairs $a_1 < b_1, \ldots, a_m < b_m \leq n$ such that for each $i \in [1, m]$,

$$\Delta_i(\{a_1, b_1\}, \ldots, a_i, \ldots, \{a_m, b_m\}) = \Delta_i(\{a_1, b_1\}, \ldots, b_i, \ldots, \{a_m, b_m\}).$$

That is, for each $i \in [1, m]$ and all $(\alpha_1, \ldots, \alpha_m) \in A^m$,

$$\begin{aligned} &\Delta(g_{a_1,b_1} \circ (\alpha_1) \widehat{\ } \cdots \widehat{\ } h_{a_i} \widehat{\ } \cdots \widehat{\ } g_{a_m,b_m} \circ (\alpha_m)) \\ &\quad = \Delta(g_{a_1,b_1} \circ (\alpha_1) \widehat{\ } \cdots \widehat{\ } h_{b_i} \widehat{\ } \cdots \widehat{\ } g_{a_m,b_m} \circ (\alpha_m)). \end{aligned} \tag{7.1}$$

Consider the m-parameter word $g = g_{a_1,b_1} \widehat{\ } \cdots \widehat{\ } g_{a_m,b_m}$. Recall that $B = A \backslash \{d\}$ and consider the r-colouring $\Delta^* : B^m \to [1, r]$ defined by,

$$\Delta^*(\alpha_1, \ldots, \alpha_m) = \Delta(g \circ (\alpha_1, \ldots, \alpha_m)).$$

By the choice of $m = HJ(t, 1; r)$ and since $|B| = t$, there is a combinatorial line $h \in [B]\binom{m}{1}$ in B^m that is monochromatic with respect to Δ^*. That is, $g \circ h$ is a combinatorial line in A^{mn} for which $g \circ h \circ [B]$ is monochromatic with respect to Δ.

To see that $\Delta(g \circ h \circ (d)) = \Delta(g \circ h \circ (c))$, note that if for some $i \in [1, m]$, $\alpha_i = d$, then

$$\Delta(g \circ (\alpha_1, \ldots, d, \ldots, \alpha_m))$$

$$= \Delta(g_{a_1,b_1} \circ (\alpha_1)^\frown \cdots {}^\frown g_{a_i,b_i} \circ (d)^\frown \cdots {}^\frown g_{a_m,b_m} \circ (\alpha_m))$$

$$= \Delta(g_{a_1,b_1} \circ (\alpha_1)^\frown \cdots {}^\frown h_{a_i}{}^\frown \cdots {}^\frown g_{a_m,b_m} \circ (\alpha_m))$$

$$= \Delta(g_{a_1,b_1} \circ (\alpha_1)^\frown \cdots {}^\frown h_{b_i}{}^\frown \cdots {}^\frown g_{a_m,b_m} \circ (\alpha_m)) \qquad \text{(by eq'n (7.1))}$$

$$= \Delta(g_{a_1,b_1} \circ (\alpha_1)^\frown \cdots {}^\frown g_{a_i,b_i} \circ (c)^\frown \cdots {}^\frown g_{a_m,b_m} \circ (\alpha_m))$$

$$= \Delta(g \circ (\alpha_1, \ldots, c, \ldots, \alpha_m))$$

Thus, $g \circ h$ is monochromatic in A^{mn}. Therefore, by induction, for all $t, r \in \mathbb{Z}^+$, the number $HJ(t, 1; r)$ exists.

Finally, it can be shown that for all t, m, r, $HJ(t, m; r) \leq m \cdot HJ(t^m, 1; r)$ and hence $HJ(t, m; r)$ exists. To see this, let A be a set with $|A| = t$ and set $n = HJ(t^m, 1, r)$. The elements of $(A^m)^n$ correspond to $A^{m \cdot n}$ by

$$(a_1, \ldots, a_{mn}) = ((a_1, \ldots, a_m), \ldots, (a_{m(n-1)+1}, \ldots, a_{mn}))$$

and a combinatorial line in $(A^m)^n$ corresponds to a combinatorial m-space in $A^{m \cdot n}$.

□

The following are two generalizations of the Hales-Jewett theorem, stated without proof. The first, known as the Graham-Rothschild theorem gives a result comparable to the Hales-Jewett theorem where the objects coloured are the combinatorial k-spaces of A^n rather than the points of the set A^n.

Theorem 7.3.1 (Graham and Rothschild [49]). Given positive integers $m, k, r \in \mathbb{Z}^+$ and a finite set A, there is a least integer $n = GR(|A|, k, m; r)$ such that for every

r-colouring $\Delta : [A]\binom{n}{k} \to [1, r]$, there is an $f \in [A]\binom{n}{m}$ such that the set $f \circ [A]\binom{m}{k}$ is monochromatic.

The next theorem brings this result from the realm of combinatorial spaces to that of vector spaces. The following theorem shows that the statement of the Graham-Rothschild theorem still holds if combinatorial spaces are replaced with vector spaces and subspaces.

Theorem 7.3.2 (Graham, Leeb and Rothschild [48])**.** Let F be any finite field with $|F| = q$. For all $k, \ell, r \in \mathbb{Z}^+$, there is an N such that for all $n \geq N$, the following holds. For any n-dimensional vector space V over F, if the k-dimensional subspaces of V are r-coloured, then there exists an ℓ-dimensional subspace of V all of whose k-dimensional subspaces have the same colour.

7.4 Homothetic copies

By relating numbers to n-tuples, other Ramsey-type theorems can be seen as consequences of the Hales-Jewett theorem. The first is van der Waerden's theorem.

Corollary 7.4.1 (Hales and Jewett [58])**.** For any $k, r \in \mathbb{Z}^+$, the van der Waerden number $W(k; r)$ satisfies $W(k; r) \leq (k - 1)HJ(k, 1; r) + 1$.

Proof. Fix $k, r \in \mathbb{Z}^+$. Set $A = [0, k - 1]$, $n = HJ(k, 1; r)$ and let $\Delta : [0, n(k - 1)] \to [1, r]$ be any r-colouring.

Define an r-colouring of A^n by $\Delta^*(a_1, \ldots, a_n) = \Delta(\sum_{i=1}^n a_i)$. By the choice of n, there is a combinatorial line $f \in [A]\binom{n}{1}$ so that $f \circ A$ is monochromatic under Δ^*. Let $M = \{i \in [1, n] : f(i) = \lambda_1\}$, then for each $x \in [0, k-1]$,

$$\begin{aligned}\Delta^*(f \circ x) &= \Delta\bigg(\sum_{i \in [1,n] \setminus M} f(i) + \sum_{i \in M} x\bigg) \\ &= \Delta\bigg(\sum_{i \in [1,n] \setminus M} f(i) + |M|x\bigg).\end{aligned}$$

Therefore, the AP_k starting at $\sum_{i \in [1,n] \setminus M} f(i)$ with difference $|M|$ is monochromatic under Δ. □

Using the Hales-Jewett numbers $HJ(k, m; r)$ a similar proof also shows that for every $k, m, r \in \mathbb{Z}^+$, there is an $n \in \mathbb{Z}^+$ so that for every r-colouring of $[1, n]$, there is a monochromatic m-fold arithmetic progression of length k.

Van der Waerden's theorem can be phrased in terms of what are called "homothetic copies" of a set of integers: simply a scaled and translated copy of the set.

Definition 7.4.2. Given a set $V \subseteq \mathbb{R}^n$, for any $\boldsymbol{a} \in \mathbb{R}^n$ and $d \in \mathbb{R} \setminus \{0\}$, the set $\boldsymbol{a} + dV = \{\boldsymbol{a} + d\boldsymbol{v} : \boldsymbol{v} \in V\}$ is a *homothetic copy* of V.

Corollary 7.4.3. For every finite set S, $r \in \mathbb{Z}^+$ and any r-colouring of $\mathbb{Z}^+$, there exist $a, d \in \mathbb{Z}^+$ so that the set $a + dS$ is monochromatic.

Proof. Let $\Delta : \mathbb{Z}^+ \to [1, r]$ be any r-colouring and set $k = \max S$. By van der

Waerden's theorem, there are $a, d \in \mathbb{Z}^+$ so that

$$\{a, a+d, \ldots, a+kd\} = a + d \cdot [1,k] \supseteq a + d \cdot S$$

is monochromatic. □

The following theorem due to Gallai (see Rado [89]) and Witt [119] provides an extension of this result to any dimension. The following proof is not the original, but uses the Hales-Jewett theorem and can be found in [58].

Theorem 7.4.4 (Gallai-Witt [89], [119])**.** Let $V \subseteq \mathbb{R}^m$ be a finite set and $r \in \mathbb{Z}^+$. For every r-colouring of $\mathbb{R}^m$ there is a monochromatic homothetic copy of V.

Proof. Fix $r \in \mathbb{Z}^+$, set $|V| = t$ and $N = HJ(t, 1; r)$. Let $\Delta : \mathbb{R}^m \to [1, r]$ be any r-colouring.

Define an r-colouring of V^N by $\Delta^*(\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}) = \Delta(\sum_{i=1}^n \boldsymbol{a_i})$. By the choice of N, there is an $f \in [V]\binom{N}{1}$ so that $f \circ V$ is monochromatic under Δ^*. Let $M = \{i \in [1,n] : f(i) = \lambda_1\}$, then for each $\boldsymbol{v} \in V$,

$$\begin{aligned}\Delta^*(f \circ \boldsymbol{v}) &= \Delta\bigg(\sum_{i \in [1,n]\setminus M} f(i) + \sum_{i \in M} \boldsymbol{v}\bigg)\\ &= \Delta\bigg(\sum_{i \in [1,n]\setminus M} f(i) + |M|\boldsymbol{v}\bigg).\end{aligned}$$

Therefore, the set $\sum_{i \in [1,n]\setminus M} f(i) + |M|V$ is a monochromatic homothetic copy of V. □

Taking $V = [0, k-1]$ in the Gallai-Witt theorem gives exactly the statement of van der Waerden's theorem.

Chapter 8

Ultrafilters

8.1 Ramsey ultrafilters

The topic of ultrafilters may seem far afield from Ramsey theory, but there are numerous connections. It is possible to show that every Ramsey-type theorem corresponds to the existence of a particular ultrafilter (Theorem 8.1.6 to come) and there are proofs using ultrafilters of a number of Ramsey-type theorems. In particular, there is a proof of van der Waerden's theorem that uses ultrafilters, the presentation of which is the goal of this chapter. The lemmas presented in this section are standard results on filters and ultrafilters that can be found in any reference on the subject (for example [70]). Other references on ultrafilters and Ramsey theory can be found for example in [5].

Definition 8.1.1. Given a set X, a non-empty collection $p \subseteq \mathcal{P}(X)$ is called a *filter on* X iff it satisfies the following three conditions:

(i) $\emptyset \notin p$;

(ii) $A, B \in p \Rightarrow A \cap B \in p$ (p is closed under intersection) and

(iii) $A \in p$ and $A \subseteq B \subseteq X \Rightarrow B \in p$ (p is upward closed).

A filter p on X is defined to be an *ultrafilter* if in addition, for every $A \subseteq X$, either $A \in p$ or $X \backslash A \in p$.

Definition 8.1.2. Let X be any set and $a \in X$. The collection $p_a = \{A \subseteq X : a \in A\}$ is an ultrafilter and is called the *principal ultrafilter* at a. Similarly, for any non-empty set $B \subseteq X$, the collection $\{A \subseteq X : B \subseteq A\}$ is a filter.

Lemma 8.1.3. A filter p is an ultrafilter iff p is a maximal filter.

Proof. Let p be an ultrafilter on X and let $q \subseteq \mathcal{P}(X)$ with $p \subsetneq q$. Let $A \in q \backslash p$, then since p is an ultrafilter, $X \backslash A \in p \subseteq q$. Therefore, $\emptyset = A \cap (X \backslash A) \in q$ and so q is not a filter. Thus, p is a maximal filter.

Let p be a filter and suppose that it is maximal. Fix $A \notin p$ and consider the collection

$$q = p \cup \{B \subseteq X : \text{for some } P \in p,\ P \cap (X \backslash A) \subseteq B \text{ or } (X \backslash A) \subseteq B\}.$$

The collection q is upward closed and closed under intersections. If $\emptyset \in q$ then for some $P \in p$, $P \cap (X \backslash A) = \emptyset$. If so, then $P \subseteq A$ and so $A \in p$ contrary to the initial assumption. Therefore, $\emptyset \notin q$ and q is a filter and since p is maximal, $p = q$ and in particular, $X\backslash A \in p$ and so p is an ultrafilter. □

In some sources (for example [42]), the definition of an ultrafilter is given as a maximal filter and the condition that for all $A \subseteq X$, either $A \in p$ or $(X\backslash A) \in p$ is given as a consequence of the definition.

For any filter p, the properties (i) and (ii) show that if $A, B \subseteq X$, then $A \cap B \in p$ iff $A \in p$ and $B \in p$. If in addition p is an ultrafilter, something comparable is also true for $A \cup B$.

Lemma 8.1.4. Let u be an ultrafilter on X. Let $A, B \subseteq X$ be such that $A \cup B \in u$, then either $A \in u$ or $B \in u$.

Proof. Suppose that $A \notin u$. Then since u is an ultrafilter, $X\backslash A \in u$ and hence $(A \cup B) \cap (X\backslash A) \in u$. Since u is upward closed and $B \supseteq (A \cup B) \cap (X\backslash A)$, then also $B \in u$. Thus either $A \in u$ or $B \in u$. □

Applying this lemma inductively shows that if u is an ultrafilter, given any set $A_1, A_2, \ldots, A_r \subseteq X$ with $\bigcup_{i=1}^{r} A_i \in u$, then there is an $i_0 \in [1, r]$ such that $A_{i_0} \in u$. In particular, since every ultrafilter contains X, for any ultrafilter u and any r-colouring of X, u contains one of the colour classes.

Lemma 8.1.5. Assuming the Axiom of Choice, every filter is contained in an ultrafilter.

Proof. The proof uses Zorn's Lemma. Let p be a filter and consider the collection $\mathcal{F} = \{q \subseteq \mathcal{P}(X) : q \text{ is a filter and } p \subseteq q\}$, partially ordered by inclusion. Then, $\mathcal{F} \neq \emptyset$ since $p \in \mathcal{F}$. Let $\mathcal{C}$ be any chain in $\mathcal{F}$ and let $q_0 = \cup\mathcal{C}$. For each $q \in \mathcal{C}$, $p \subseteq q$ and so $p \subseteq q_0$. In order to show that $q_0 \in \mathcal{F}$, it remains to show that q_0 is a filter.

(i) For each $q \in \mathcal{C}$, $\emptyset \notin q$ and hence $\emptyset \notin q_0$.

(ii) If $A, B \in q_0$ then there are filters $q_A, q_B \in \mathcal{C}$ with $A \in q_A$ and $B \in q_B$. Since $\mathcal{C}$ is linearly ordered, without loss of generality $q_A \subseteq q_B$. Thus $A, B \in q_B$ and since q_B is a filter, $A \cap B \in q_B \subseteq q_0$.

(iii) If $A \in q_0$ and $A \subseteq B$, then for some $q_A \in \mathcal{C}, A \in q_A$ and since q_A is a filter, $B \in q_A \subseteq q_0$.

Therefore, q_0 is a filter and it is an upper bound in $\mathcal{F}$ for the chain $\mathcal{C}$.

Thus, by Zorn's Lemma, $\mathcal{F}$ has a maximal element u. The filter u is a maximal filter containing p and hence an ultrafilter containing p. □

This lemma guarantees the existence of non-principal ultrafilters on any infinite set. Let X be an infinite set and consider the filter $p = \{B \subseteq X : |X \backslash B| < \infty\}$. This filter is called the *Fréchet filter* or the *co-finite filter*. By Lemma 8.1.5, there is an ultrafilter u that contains p, but u cannot be principal because if $x \in X$, then $X \backslash \{x\} \in u$ and so $\{x\} \notin u$.

The following theorem due to Hindman shows that many Ramsey-type results correspond to the existence of a particular ultrafilter.

Theorem 8.1.6 (Hindman [65])**.** Let X be a set and $\mathcal{F} \subseteq \mathcal{P}(X)\backslash\{\emptyset\}$. The following are equivalent:

(i) For every finite colouring of X, there is an $F \in \mathcal{F}$ that is monochromatic.

(ii) There is an ultrafilter u on X such that for every $A \in u$, there is an $F \in \mathcal{F}$ with $F \subseteq A$.

Proof. [(i) $\Rightarrow$ (ii)] Suppose that for every finite colouring of X, there is an $F \in \mathcal{F}$ that is monochromatic. Set $\mathcal{B} = \{A \subseteq X : \text{for all } F \in \mathcal{F}, A \cap F \neq \emptyset\}$. Fix any $A_1, \ldots, A_k \in \mathcal{B}$. If X is partitioned into 2^k parts by the Venn diagram for the sets $A_1, \ldots, A_k$, then, by assumption, there is an $F \in \mathcal{F}$ contained in one part. Since for each $i \in [1, k]$, $A_i \cap F \neq \emptyset$, the only possibility is that $F \subseteq A_1 \cap \cdots \cap A_k$. As $F \neq \emptyset$, also $A_1 \cap \cdots A_k \neq \emptyset$. That is, the family $\mathcal{B}$ has the finite intersection property. Thus, since $\mathcal{B}$ is upward closed, the set $\mathcal{B}^* = \{B_1 \cap \cdots \cap B_k : B_1, \ldots, B_k \in \mathcal{B}\}$ is a filter. By Lemma 8.1.5, there is an ultrafilter u with $\mathcal{B}^* \subseteq u$.

For each $A \in u$, since u is an ultrafilter, $X\backslash A \notin u$ and hence $X\backslash A \notin \mathcal{B}$. Therefore, there is an $F \in \mathcal{F}$ such that $F \cap (X\backslash A) = \emptyset$ and hence $F \subseteq A$.

[(ii) $\Rightarrow$ (i)] Now suppose there is an ultrafilter u on X such that for every $A \in u$, there is an $F \in \mathcal{F}$ with $F \subseteq A$. Fix $r \in \mathbb{Z}^+$ and let $\Delta : X \rightarrow [1, r]$ be any r-colouring of X. By Lemma 8.1.4, there is an $i \in [1, r]$ such that $\Delta^{-1}(i) \in u$. Take $F \in \mathcal{F}$ with $F \subseteq \Delta^{-1}(i)$; F is monochromatic. □

Now, in terms of arithmetic progressions, take $X = \mathbb{Z}^+$ and for some $k \in \mathbb{Z}^+$ let

$\mathcal{F}$ be the set of all AP_k's in $\mathbb{Z}^+$. Theorem 8.1.6 together with van der Waerden's theorem shows that there is an ultrafilter u on $\mathbb{Z}^+$ with the property that for every $A \in u$, there are integers a and $d > 0$ such that $\{a, a+d, \ldots, a+(k-1)d\} \subseteq A$. Conversely, if it can be shown without use of van der Waerden's theorem that such an ultrafilter exists, this provides another proof of van der Waerden's theorem.

In order to find such an ultrafilter guaranteed by Theorem 8.1.6, some other properties are used. A semigroup operation that extends the addition of integers can be defined on the ultrafilters on $\mathbb{Z}^+$ which can also be given a topology. Before proceeding to ultrafilters on $\mathbb{Z}^+$, some general results about semigroups with topologies are given which are needed for the proof of van der Waerden's theorem.

8.2 Semigroups

8.2.1 Topology review

This section begins with a few standard definitions and results from topology that are given without proof but can be found in any standard reference (for example [118]). Throughout, let X be a topological space with topology τ unless stated otherwise.

Definition 8.2.1. A topological space X is *Hausdorff* iff for every $x, y \in X$ with $x \neq y$, there are disjoint open sets U_x and U_y with $x \in U_x$ and $y \in U_y$.

Definition 8.2.2. A topological space X is *compact* iff for every collection of open sets $\{U_i : i \in I\}$ with $X = \bigcup_{i \in I} U_i$, there is a finite set $F \subseteq I$ so that $X = \bigcup_{f \in F} U_f$

Definition 8.2.3. Let (X, τ) be a topological space. A collection $\{B_i : i \in I\} \subseteq \mathcal{P}(X)$ is called an *open basis* for τ iff $\tau = \{\bigcup_{j \in J} B_j : J \subseteq I\}$. The basis $\{B_i : i \in I\}$ is said to *generate* τ and the sets $\{B_i : i \in I\}$ are called *basic open sets.* The sets $\{X \backslash B_i : i \in I\}$ are called *basic closed sets.*

Fact 8.2.4. Let X be a set and $\mathcal{B} \subseteq \mathcal{P}(X)$. Then $\{\cup \mathcal{C} : \mathcal{C} \subseteq \mathcal{B}\}$ forms a basis for a topology on X iff

(i) $\bigcup_{B \in \mathcal{B}} B = X$ and

(ii) For every $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there is a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Definition 8.2.5. A collection of sets $\{A_i : i \in I\}$ has the *finite intersection property* (abbreviated FIP) iff for every finite subset $F \subseteq I$, $\cap_{i \in F} A_i \neq \emptyset$.

Fact 8.2.6. A topological space X is compact iff for every collection of closed sets $\{C_i : i \in I\}$ with the finite intersection property, $\cap_{i \in I} C_i \neq \emptyset$.

Fact 8.2.7. Let X be compact and Hausdorff. Then a set $C \subseteq X$ is closed iff C is compact.

Fact 8.2.8. Let $f : X \to Y$ be a continuous function. If $K \subseteq X$ is compact, then $f(K) \subseteq Y$ is compact.

Recall that if (X, τ) is a topological space and $n \in \mathbb{Z}^+$ the product topology on X^n is given by the open basis $\{U_1 \times \ldots \times U_n : \text{for each } i \in [1, n],\ U_i \in \tau\}$.

Fact 8.2.9. If X is compact, then X^n with the product topology is compact.

Fact 8.2.10. If X is Hausdorff, then X^n with the product topology is Hausdorff.

Fact 8.2.11. For $i \in [1, n]$, let $f_i : X \to Y$ be continuous functions. Then the function $f : X^n \to Y^n$ given by $f(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$ is continuous in the product topologies on X^n and Y^n.

8.2.2 Semigroups with topologies

In this section, some properties of semigroups with topologies are given that can later be applied to a semigroup operation and topology on the ultrafilters on $\mathbb{Z}^+$. The following approaches to proving Ramsey theorems using ultrafilters can be found in an article by Bergelson, Furstenberg, Hindman and Katznelson [13], an article by Furstenberg and Katznelson [46] and notes by Hart [59].

Throughout, let $(X, *)$ be a semigroup with a topology τ on X.

Definition 8.2.12. For every $x \in X$, define $\lambda_x : X \to X$ by $\lambda_x(y) = x * y$. The semigroup operation $*$ is *left-continuous* iff for every $x \in X$, λ_x is continuous.

Definition 8.2.13. The semigroup $(X, *)$ with the topology τ, is *semi-topological* iff (X, τ) is compact and Hausdorff and $*$ is left-continuous.

Definition 8.2.14. An element $x \in X$ is called an *idempotent* iff $x * x = x$.

The following result on the existence of idempotents in semi-topological semigroups has become known as the Idempotent Lemma and plays a key role in many of the results that follow.

Lemma 8.2.15 (**Idempotent Lemma**, Ellis [32])**.** Let $(X, *)$ be a semi-topological semigroup, then X contains an idempotent.

Proof. Let $\mathcal{A} = \{A \subseteq X : A \neq \emptyset,\ A \text{ is closed and } A * A \subseteq A\}$ be ordered by inclusion. Any chain in $\mathcal{A}$ will have the finite intersection property and hence nonempty intersection (by Fact 8.2.6) since the sets $A \in \mathcal{A}$ are closed and X is compact. If $\mathcal{C}$ is any chain in $\mathcal{A}$, then $C = \cap\mathcal{C}$ is also closed. If $c_1, c_2 \in C$, then for all $A \in \mathcal{C}$, $c_1, c_2 \in A$ and so $c_1 * c_2 \in A$. Thus $c_1 * c_2 \in C$ and so $C * C \subseteq C$ and $C \in \mathcal{A}$. Therefore, C is a lower bound for $\mathcal{C}$ in $\mathcal{A}$. Since $X \in \mathcal{A}$, $\mathcal{A} \neq \emptyset$ and thus Zorn's lemma can be used to find a minimal element A_0 of $\mathcal{A}$.

Fix $x \in A_0$. Since A_0 is closed and so compact, $x * A_0 = \lambda_x(A_0)$ is compact and hence closed. Since $x \in A_0$ and $A_0 * A_0 \subseteq A_0$, $x * A_0 \subseteq A_0$ and so,

$$\begin{aligned}(x * A_0) * (x * A_0) &\subseteq x * (A_0 * A_0) \\ &\subseteq x * A_0.\end{aligned}$$

Therefore, $x * A_0 \in \mathcal{A}$ and since A_0 is minimal, $x * A_0 = A_0$. In particular, there is at least one $y \in A_0$ with $x * y = x$.

Consider the set $Y = \{y \in A_0 : x{*}y = x\} = A_0 \cap \lambda_x^{-1}(\{x\})$. By the above, $Y \neq \emptyset$ and Y is closed since A_0 is closed and λ_x is continuous. To show that $Y * Y \subseteq Y$, let $y, z \in Y$, then

$$\begin{aligned} x * (y * z) &= (x * y) * z \\ &= x * z && \text{(since } y \in Y\text{)} \\ &= x && \text{(since } z \in Y\text{).} \end{aligned}$$

Therefore $y * z \in Y$ and so $Y \in \mathcal{A}$. Again, since A_0 is minimal, $Y = A_0$ and in particular, since $x \in A_0$, $x * x = x$. □

Definition 8.2.16. A set $R \subseteq X$ is a *right ideal* iff for every $x \in X$, $R * x \subseteq R$. Similarly, L is a *left ideal* iff whenever $x \in X$, then $x * L \subseteq L$. A set $I \subseteq X$ is a *two-sided ideal* iff it is both a left and a right ideal.

Lemma 8.2.17. Every right ideal in X contains a minimal right ideal.

Proof. Let R be a right ideal. Then for every $r \in R$, $\lambda_r(X) = r * X \subseteq R$ is a closed right ideal contained in R.

Let $\mathcal{F} = \{C \subseteq R : C \neq \emptyset \text{ and } C \text{ is a closed right ideal}\}$ be ordered by inclusion. Then $\mathcal{F} \neq \emptyset$. Let $\mathcal{C}$ be any chain in $\mathcal{F}$ and let $C_0 = \cap\mathcal{C}$. Then $C_0 \neq \emptyset$ since X is compact and $\mathcal{C}$ has the finite intersection property. The set C_0 is closed since all sets in $\mathcal{C}$ are closed and C_0 is a right ideal since it is an intersection of right ideals. Thus $C_0 \in \mathcal{F}$. Therefore, by Zorn's lemma, $\mathcal{F}$ contains a minimal element. This minimal element will be a minimal right ideal contained in R. In addition, this shows that any minimal right ideal is closed. $\square$

It will be useful to note the following property of ideals. Let R be any closed right ideal. The ideal R is closed under $*$ since it is an ideal, R is compact since it is closed, R is Hausdorff since X is Hausdorff, and $*$ is left-continuous on R since $*$ is left-continuous on X. Therefore, any closed right ideal is itself a semi-topological semigroup and hence contains an idempotent by Lemma 8.2.15. Further, since every right ideal contains a closed right ideal, every right ideal contains an idempotent.

8.2.3 Sticky diagonal theorem

This section and the next are independent of each other and provide two different approaches to a proof of van der Waerden's theorem using ultrafilters. Each of these two proofs are presented in Section 8.4. The results in this section are due to Bergelson, Furstenberg, Hindman and Katznelson [13].

Lemma 8.2.18. Let R be a minimal right ideal in X and $q \in R$ be an idempotent. Then for every $r \in R$, $q * r = r$.

Proof. Since $q \in R$ and R is a right ideal, $q * R \subseteq R$. Since $q * R$ is a right ideal and R is minimal, $q * R = R$. Thus, for each $r \in R$ there is an $a_r \in R$ with $q * a_r = r$.

Then,

$$\begin{aligned} q * r &= q * (q * a_r) \\ &= q * a_r \qquad \text{(since } q \text{ is an idempotent)} \\ &= r. \end{aligned}$$

□

Fix $k \in \mathbb{Z}^+$ and consider the semigroup on X^k with the semigroup operation defined component-wise by $*$ and with the product topology on X^k. The space X^k is compact and Hausdorff since X is compact and Hausdorff and $*$ is left-continuous on X^k since $*$ is left-continuous on X. Therefore $(X^k, *)$ is also a semi-topological semigroup. The elements of X^k are denoted by $\boldsymbol{x} = (x_1, \ldots, x_k) \in X^k$.

The following theorem is called the "Sticky Diagonal Theorem" in [120]. This result on the diagonal elements of a semigroup X^k seems to be related to the "Central Sets Theorem" which appears in [67] where it is attributed to Furstenberg.

Theorem 8.2.19 (**Sticky diagonal theorem**, Bergelson *et al.* [13])**.** Fix $k \in \mathbb{Z}^+$ and let $E \subseteq X^k$ be a semi-topological semigroup with $\{(x, \ldots, x) \in X^k : x \in X\} \subseteq E$ and I a two-sided ideal in E. Then there exists a $p \in X$ with $(p, \ldots, p) \in I$.

Proof. Let R be a minimal right ideal in X and let $p \in R$. Set $\boldsymbol{p} = (p, \ldots, p) \in E$. Since I is a left ideal in E and $\boldsymbol{p} \in E$, $\boldsymbol{p} * I \subseteq I$. Since I is a right ideal, $\boldsymbol{p} * I$ is also a right ideal and so $\boldsymbol{p} * I$ contains an idempotent $\boldsymbol{q} = (q_1, \ldots, q_k)$. Since $\boldsymbol{q} \in \boldsymbol{p} * I$, there is $(x_1, \ldots, x_k) \in I$ with $\boldsymbol{q} = \boldsymbol{p} * (x_1, \ldots, x_k)$. Then, for each $i \in [1, k]$, $q_i = p * x_i \in R$

since $p \in R$ and R is a right ideal. Since $*$ is defined component-wise, for each $i \in [1, k]$, q_i is an idempotent in X. Therefore, by Lemma 8.2.18, since R is a minimal right ideal, whenever $i \in [1, k]$, then $q_i * p = p$ and so $\boldsymbol{q} * \boldsymbol{p} = \boldsymbol{p}$. Therefore, $(p, \ldots, p) = \boldsymbol{q} * \boldsymbol{p} \in I$ since $\boldsymbol{q} \in I$ and I is a right ideal. □

8.2.4 Minimal idempotents

This section presents another set of results on semigroups that can be used to give a proof of van der Waerden's theorem. The focus is to show that there is a partial order on the idempotents of X so that any two-sided ideal in X contains all of the minimal idempotents of X. The results in this section are due to Furstenberg and Katznelson [46] and can also be found in notes by Hart [59].

A partial ordering is defined on the idempotents of X. For idempotents x and y, $x \preccurlyeq y$ iff $x = x * y = y * x$. For any idempotent x, since $x * x = x$, $x \preccurlyeq x$. If $x \preccurlyeq y$ and $y \preccurlyeq x$, then $x = x * y = y$. If $x \preccurlyeq y$ and $y \preccurlyeq z$ then,

$$\begin{aligned} x * z &= x * y * z && \text{(since } x \preccurlyeq y) \\ &= x * y && \text{(since } y \preccurlyeq z) \\ &= x. \end{aligned}$$

Thus, the relation $\preccurlyeq$ is indeed a partial order.

The next lemma demonstrates some of the connection between the right ideals of X and the ordering $\preccurlyeq$.

Lemma 8.2.20. Let x be an idempotent and $R \subseteq x * X$ be a minimal right ideal. Then there is an idempotent $y \in R$ with $y \preccurlyeq x$.

Proof. As noted previously, since R is a closed right ideal, R contains an idempotent r. Since $r \in R \subseteq x * X$, there is an $a \in X$ with $r = x * a$. Set $y = r * x$. Since $r \in R$ and R is a right ideal, $y \in R$ and since $r = x * a$, $y = x * a * x$. Now, y is an idempotent since

$$\begin{aligned} y * y &= (r * x) * (x * a * x) && \text{(since } y = r * x = x * a * x\text{)} \\ &= r * x * a * x && \text{(since } x * x = x\text{)} \\ &= r * r * x && \text{(since } r = x * a\text{)} \\ &= r * x && \text{(since } r * r = r\text{)} \\ &= y. \end{aligned}$$

To see that $y \preccurlyeq x$, note that

$$y * x = r * x * x = r * x = y,$$

and

$$x * y = x * x * a * x = x * a * x = y. \qquad \square$$

Lemma 8.2.21 (Furstenberg and Katznelson [46])**.** Let x be an idempotent. Then x is minimal iff x is contained in a minimal right ideal.

Proof. Suppose x is minimal. Since $x * X$ is a right ideal, by Lemma 8.2.17, there is a minimal right ideal $R \subseteq x * X$. By Lemma 8.2.20, there is an idempotent $y \in R$ with $y \preccurlyeq x$. Since x is minimal, $x = y$ and so $x \in R$, a minimal right ideal.

For the converse, suppose R is a minimal right ideal with $x \in R$. Suppose there is an idempotent y with $y \preccurlyeq x$. Since $y = x * y$, $y \in R$ and as R is minimal, $y * X = R$. Then there is an $a \in X$ with $y * a = x$. Then,

$$x = y * a = y * y * a = y * x = y.$$

Therefore, x is minimal. □

Until this point, it has not yet been shown that there are any minimal idempotents. The previous results can be used to show that there is at least one.

Corollary 8.2.22. Let x be an idempotent. Then there exists a minimal idempotent y with $y \preccurlyeq x$.

Proof. Let R be a minimal right ideal with $R \subseteq x * X$ and by Lemma 8.2.20, let $y \in R$ be an idempotent with $y \preccurlyeq x$. By Lemma 8.2.21, y is minimal. □

The only special property of the minimal idempotents needed for what follows is that they can always be found in 2-sided ideals.

Lemma 8.2.23. Let R be a minimal right ideal and I be a 2-sided ideal. Then $R \subseteq I$.

Proof. For any $x \in R$ and $y \in I$, $x*y \in R \cap I$ and so $R \cap I \neq \emptyset$. Therefore $R \cap I \subseteq R$ is a right ideal and since R is minimal, $R = R \cap I$ and so $R \subseteq I$. □

Note that since every minimal idempotent belongs to a minimal right ideal, if I is a 2-sided ideal then I contains all the minimal idempotents of X.

8.3 Ultrafilters on $\mathbb{Z}^+$

The ultrafilters on $\mathbb{Z}^+$ can be given a topology so that the space is isomorphic to what is called the "Stone-Čech compactification" of $\mathbb{Z}^+$. A space X is Tychonoff iff for every closed set $C \subseteq X$ and $x \notin C$, there is a continuous function $f : X \to [0, 1]$ (interval of real numbers) such that $f(C) = 0$ and $f(x) = 1$. For a Tychonoff space X, the *Stone-Čech compactification* of X, denoted by βX is a compact Hausdorff space for which $X \subseteq \beta X$ is a dense subspace and is characterised up to isomorphism by the following property. For any compact Hausdorff space K and continuous function $f : X \to K$, there is a continuous function $F : \beta X \to K$ so that $F_{|X} = f$. Any space with the discrete topology (where all sets are open), is Tychonoff since all functions are continuous.

Let $\beta\mathbb{Z}^+$ be the Stone-Čech compactification of $\mathbb{Z}^+$ with the discrete topology. Though it will not be proved here, $\beta\mathbb{Z}^+$ is isomorphic to the set of all ultrafilters on $\mathbb{Z}^+$ with a particular topology. All the results needed to prove van der Waerden's theorem using ultrafilters can be proved by defining the topology directly on the space of ultrafilters on $\mathbb{Z}^+$ and without using the full strength of the Stone-Čech compactification (further information on the Stone-Čech compactification of a semigroup can be found in [67]). For each $A \subseteq \mathbb{Z}^+$, set $U_A = \{u \in \beta\mathbb{Z}^+ : A \in u\}$.

Lemma 8.3.1. The sets $\{U_A : A \subseteq \mathbb{Z}^+\}$ form a basis for a topology on $\beta\mathbb{Z}^+$ and for every $A \subseteq \mathbb{Z}^+$, the set U_A is also closed under this topology.

Proof. The proof uses Fact 8.2.4. For any $u \in \beta\mathbb{Z}^+$, if $A \in u$ then $u \in U_A$. Thus $\bigcup_{A \subseteq \mathbb{Z}^+} U_A = \beta\mathbb{Z}^+$.

Let $A_1, A_2 \subseteq \mathbb{Z}^+$ and $u \in U_{A_1} \cap U_{A_2}$, then $A_1 \in u$ and $A_2 \in u$. Since u is a filter, $A_1 \cap A_2 \in u$ and hence $u \in U_{A_1 \cap A_2}$. For any $v \in U_{A_1 \cap A_2}$, since v is upward closed and both $A_1 \cap A_2 \subseteq A_1$ and $A_1 \cap A_2 \subseteq A_2$, both $A_1 \in v$ and $A_2 \in v$. Therefore $v \in U_{A_1} \cap U_{A_2}$ and since v was an arbitrary ultrafilter, $u \in U_{A_1 \cap A_2} = U_{A_1} \cap U_{A_2}$.

Therefore the collection $\{U_A : A \subseteq \mathbb{Z}^+\}$ generates a topology on $\beta\mathbb{Z}^+$. To show that every basic open set in this topology is also closed, let $A \subseteq \mathbb{Z}^+$. Then

$$\begin{aligned}\beta\mathbb{Z}^+ \backslash U_A &= \{u \in \beta\mathbb{Z}^+ : A \notin u\} \\ &= \{u \in \beta\mathbb{Z}^+ : \mathbb{Z}^+ \backslash A \in u\} \qquad \text{(since } u \text{ is an ultrafilter)} \\ &= U_{\mathbb{Z}^+ \backslash A}\end{aligned}$$

is an open set and so U_A is a closed set and the sets $\{U_A : A \subseteq \mathbb{Z}^+\}$ are also basic closed sets. □

For simplicity, for every $n \in \mathbb{Z}^+$ associate the integer n with the principal ultrafilter p_n and in this way, assume that $\mathbb{Z}^+ \subseteq \beta\mathbb{Z}^+$. The symbols n and p_n are used interchangeably for the principal ultrafilter at n, but when n is used both as an integer and as an ultrafilter, the difference will be noted.

Lemma 8.3.2. The space $\beta\mathbb{Z}^+$ is Hausdorff, compact and $\mathbb{Z}^+$ is dense in $\beta\mathbb{Z}^+$.

Proof. To prove that $\beta\mathbb{Z}^+$ is Hausdorff, fix $u, v \in \beta\mathbb{Z}^+$ with $u \neq v$. For any $A \in u \backslash v$, since $A \notin v$, $\mathbb{Z}^+ \backslash A \in v$. Thus $u \in U_A$ while $v \in U_{\mathbb{Z}^+ \backslash A}$ and $U_A \cap U_{\mathbb{Z}^+ \backslash A} = U_{A \cap \mathbb{Z}^+ \backslash A} = U_\emptyset = \emptyset$, since no ultrafilters contain the empty set. Therefore $\beta\mathbb{Z}^+$ is Hausdorff.

In order to show that $\beta\mathbb{Z}^+$ is compact, it suffices to show that every collection of basic closed sets with the finite intersection property has non-empty intersection by Fact 8.2.6. Suppose that $\{U_{A_i}\}_{i \in I}$ has the finite intersection property. For any finite subcollection, $U_{A_1} \cap U_{A_2} \cap \cdots \cap U_{A_n} = U_{A_1 \cap A_2 \cap \cdots \cap A_n} \neq \emptyset$. Hence, since only $U_\emptyset = \emptyset$, $A_1 \cap \cdots \cap A_n \neq \emptyset$. Therefore, the collection $\{A_i\}_{i \in I}$ has the finite intersection property and hence $\mathcal{A} = \{A \subseteq \mathbb{Z}^+ : A \supseteq A_{i_1} \cap \cdots \cap A_{i_n} \text{ for some } i_1, \ldots i_n \in I\}$ is a filter. By Lemma 8.1.5, $\mathcal{A}$ is contained in an ultrafilter u. For each $i \in I$, $A_i \in u$ and so $u \in U_{A_i}$. Thus $u \in \cap_{i \in I} U_{A_i}$, so $\cap_{i \in I} U_{A_i} \neq \emptyset$ and thus $\beta\mathbb{Z}^+$ is compact.

Finally, to see that $\mathbb{Z}^+$ is dense in $\beta\mathbb{Z}^+$, let U_A be any basic open set and fix $n \in A$. Then $A \in p_n$ and so $p_n \in U_A$. Thus $\mathbb{Z}^+ \cap U_A \neq \emptyset$ and so $\mathbb{Z}^+$ is dense in $\beta\mathbb{Z}^+$. □

For each $n \in \mathbb{Z}^+$ and $A \subseteq \mathbb{Z}^+$, set $A - n = \{x \in \mathbb{Z}^+ : x + n \in A\}$. The addition of integers $+$ is extended to $\beta\mathbb{Z}^+$ as follows. For $u, v \in \beta\mathbb{Z}^+$, define

$$u + v = \{A \subseteq \mathbb{Z}^+ : \{x \in \mathbb{Z}^+ : A - x \in u\} \in v\}.$$

Lemma 8.3.3. For each $u, v \in \beta\mathbb{Z}^+$, $u + v \in \beta\mathbb{Z}^+$.

Proof. Let u, v be ultrafilters. For any $x \in \mathbb{Z}^+$, $\emptyset - x = \emptyset$, therefore

$$\begin{aligned} \emptyset \in u + v &\Leftrightarrow \{x \in \mathbb{Z}^+ : \emptyset - x \in u\} \in v \\ &\Leftrightarrow \emptyset \in v && \text{(since } \emptyset \notin u\text{).} \end{aligned}$$

Since v is an ultrafilter, $\emptyset \notin v$ and hence $\emptyset \notin u + v$.

To show that $u + v$ is closed under intersections, let $A, B \in u + v$. Then $\{x : A - x \in u\} \in v$ and $\{x : B - x \in u\} \in v$ and so $\{x : A - x \in u\} \in v \cap \{x : B - x \in u\} \in v$. Since

$$\begin{aligned} \{x : A \cap B - x \in u\} &= \{x : (A - x) \cap (B - x) \in u\} \\ &= \{x : A - x \in u\} \cap \{x : B - x \in u\}, \end{aligned}$$

$\{x : A \cap B - x \in u\} \in v$ and hence $A \cap B \in u + v$.

To show that $u + v$ is upward closed, let $A \in u + v$ and $B \supseteq A$. Then $\{x : A - x \in u\} \in v$ and since u is upward closed, $\{x : B - x \in u\} \supseteq \{x : A - x \in u\} \in v$ and since v is upward closed, $\{x : B - x \in u\} \in v$ and hence $B \in u + v$.

Finally, fix $A \subseteq \mathbb{Z}^+$ with $A \notin u+v$, then $\{x : A - x \in u\} \notin v$ and since v is an ultrafilter, $\mathbb{Z}^+ \backslash \{x : A - x \in u\} \in v$. Consider the set $\mathbb{Z}^+ \backslash A$.

$$\begin{aligned} \{x : (\mathbb{Z}^+ \backslash A) - x \in u\} &= \{x : \mathbb{Z}^+ \backslash (A - x) \in u\} \\ &= \mathbb{Z}^+ \backslash \{x : A - x \in u\} \in v. \end{aligned}$$

Thus $\mathbb{Z}^+ \backslash A \in u + v$ and so $u + v$ is an ultrafilter. □

Lemma 8.3.4. The operation $+$ is associative on $\beta\mathbb{Z}^+$.

Proof. Given ultrafilters $u, v, w \in \beta\mathbb{Z}^+$ and $A \subseteq \mathbb{Z}^+$:

$$\begin{aligned} A \in (u+v)+w &\Leftrightarrow \{x : A - x \in u+v\} \in w \\ &\Leftrightarrow \{x : \{y : (A-x) - y \in u\} \in v\} \in w \\ &\Leftrightarrow \{x : \{y : A - y \in u\} - x \in v\} \in w \\ &\Leftrightarrow \{y : A - y \in u\} \in (v+w) \\ &\Leftrightarrow A \in u + (v+w). \end{aligned}$$

Therefore, $+$ is associative. □

Lemma 8.3.5. The operation $+$ is left-continuous on $\beta\mathbb{Z}^+$.

Proof. Recall (Definition 8.2.12) that $+$ is left-continuous iff for every $u \in \beta\mathbb{Z}^+$, the function $\lambda_u(v) = u + v$ is continuous. For each $u \in \beta\mathbb{Z}^+$, and basic open set U_A,

$$\lambda^{-1}(U_A) = \{v \in \beta\mathbb{Z}^+ : u + v \in U_A\}$$

$$\begin{aligned} &= \{v \in \beta\mathbb{Z}^+ : A \in u + v\} \\ &= \{v \in \beta\mathbb{Z}^+ : \{x : A - x \in u\} \in v\} \\ &= U_{\{x:A-x\in u\}} \end{aligned}$$

is an open set. Therefore λ_u is continuous. □

Lemma 8.3.6. For $n, m \in \mathbb{Z}^+$, the principal ultrafilters p_n, p_m and p_{n+m} satisfy $p_n + p_m = p_{n+m}$.

Proof. For any $A \subseteq \mathbb{Z}^+$,

$$\begin{aligned} A \in p_n + p_m &\Leftrightarrow \{x : A - x \in p_n\} \in p_m \\ &\Leftrightarrow A - m \in p_n \\ &\Leftrightarrow n \in A - m \\ &\Leftrightarrow n + m \in A \\ &\Leftrightarrow A \in p_{n+m} \end{aligned}$$ □

Definition 8.3.7. For each $u \in \beta\mathbb{Z}^+$, let $\rho_u : \beta\mathbb{Z}^+ \to \beta\mathbb{Z}^+$ be defined by $\rho_u(v) = v + u$.

While all of the functions λ_u are continuous, it can be shown that the right addition function, ρ_u, is only continuous when u is a principal ultrafilter.

Lemma 8.3.8. For any $n \in \mathbb{Z}^+$, ρ_n is a continuous function.

Proof. Fix $n \in \mathbb{Z}^+$ and let U_A be any basic open set. For any $u \in \beta\mathbb{Z}^+$,

$$\begin{aligned} A \in u + n &\Leftrightarrow \{x : A - x \in u\} \in n \\ &\Leftrightarrow A - n \in u. \end{aligned}$$

Thus,

$$\begin{aligned} \rho_n^{-1}(U_A) &= \{v \in \beta\mathbb{Z}^+ : v + n \in U_A\} \\ &= \{v \in \beta\mathbb{Z}^+ : A \in v + n\} \\ &= \{v \in \beta\mathbb{Z}^+ : A - n \in v\} \\ &= U_{A-n} \end{aligned}$$

is an open set and hence ρ_n is continuous. □

8.4 Proof of van der Waerden's theorem

The last section showed that the space $\beta\mathbb{Z}^+$ is a semi-topological semigroup and so the results of sections 8.2.3 and 8.2.4 for these can be applied to give another proof of van der Waerden's theorem.

Theorem 8.4.1 (Bergelson *et al.* [13], Furstenberg and Katznelson [46])**.** For every $k \in \mathbb{Z}^+$, there is an ultrafilter $u \in \beta\mathbb{Z}^+$ such that for every $A \in u$, A contains an AP_k.

Combining Theorem 8.4.1 with Theorem 8.1.6 gives van der Waerden's theorem. Another version of the following proof that uses "piecewise syndetic sets" can be found in [67]. The idea of the proof of Theorem 8.4.1 is as follows. Fix $k \in \mathbb{Z}^+$ and consider the semi-topological semigroup $(\beta\mathbb{Z}^+)^k$. It will be first shown that the proof of Theorem 8.4.1, relies on properties of the following sets:

$$\begin{aligned} I^* &= \{(a, a+d, \ldots, a+(k-1)d) : a, d \in \mathbb{Z}^+\} \\ E^* &= \{(a, a+d, \ldots, a+(k-1)d) : a \in \mathbb{Z}^+, d \in \mathbb{Z}^+ \cup \{0\}\} \\ I &= cl(I^*) \\ E &= cl(E^*) \end{aligned}$$

where the closure is taken in $(\beta\mathbb{Z}^+)^k$. Recall that $0 \notin \mathbb{Z}^+$ so that the set I^* contains sequences of AP_k's with non-zero difference, while E^* is the set of all sequences of AP_k's with possibly 0 difference.

If it can be shown that there is a $u \in \beta\mathbb{Z}^+$ with $(u, \ldots, u) \in I$, then for any $A \in u$, $(U_A)^k$ is an open set in $(\beta\mathbb{Z}^+)^k$ with $(u, \ldots, u) \in (U_A)^k$. Since $(u, \ldots, u) \in cl(I^*)$, $(U_A)^k \cap I^* \neq \emptyset$ and so there exist integers a and $d > 0$ with $(a, a+d, \ldots, a+(k-1)d) \in (U_A)^k$. That is, for $0 \leq i \leq k-1$,

$$\begin{aligned} a + id \in U_A \Rightarrow A &\in a + id && (a+id, \text{ the principal ultrafilter}) \\ \Rightarrow a + id &\in A && (a+id, \text{ the integer}). \end{aligned}$$

Thus, $\{a, a+d, \ldots, a+(k-1)d\} \subseteq A$.

In order to show that such an ultrafilter u can be found more information about E and I is needed. Recall that the elements of $(\beta\mathbb{Z}^+)^k$ are denoted by $\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}$.

Lemma 8.4.2. The set E is a semi-topological semigroup and I is a 2-sided ideal in E.

Proof. Since E is closed in $(\beta\mathbb{Z}^+)^k$, E is compact. Since E is also Hausdorff and $+$ is left-continuous on E, in order to show that E is a semi-topological semigroup, it remains only to show that E is closed under $+$.

Let $\boldsymbol{p}, \boldsymbol{q} \in E$. It will be shown that $\boldsymbol{p}+\boldsymbol{q} \in E$ (and so E is closed under $+$) and that if either $\boldsymbol{p} \in I$ or $\boldsymbol{q} \in I$, then $\boldsymbol{p}+\boldsymbol{q} \in I$ (and so I is a 2-sided ideal in E).

Let U be open in $(\beta\mathbb{Z}^+)^k$ with $\boldsymbol{p}+\boldsymbol{q} \in U$ and set $V_1 = \lambda_{\boldsymbol{p}}^{-1}(U)$. Then $\boldsymbol{q} \in V_1$ and since $\lambda_{\boldsymbol{p}}$ is continuous, V_1 is open. Since $\boldsymbol{q} \in cl(E^*)$, there exists $a \in \mathbb{Z}^+$ and $d \geq 0$ with $(a, a+d, \ldots, a+(k-1)d) \in E^* \cap V_1$. (If $\boldsymbol{q} \in I = cl(I^*)$ then $d > 0$.)

Then $\boldsymbol{p}+(a, \ldots, a+(k-1)d) \in U$ and since $\rho_{(a,\ldots,a+(k-1)d)}$ is continuous, the set $V_2 = \rho^{-1}_{(a,\ldots,a+(k-1)d)}(U)$ is open and $\boldsymbol{p} \in V_2$. As with $\boldsymbol{q}$, since $\boldsymbol{p} \in cl(E^*)$, there exists $b \in \mathbb{Z}^+$ and $c \geq 0$ with $(b, b+c, \ldots, b+(k-1)c) \in E^* \cap V_2$. (If $\boldsymbol{p} \in I$ then $c > 0$.)

Then

$$\begin{aligned}&(b, \ldots, b+(k-1)c) + (a, \ldots, a+(k-1)d)\\ &\qquad\qquad = (b+a, \ldots b+a+(k-1)(c+d)) \in U \cap E^*.\end{aligned}$$

Thus, if U is an open set with $\boldsymbol{p}+\boldsymbol{q} \in U$ then $U \cap E^* \neq \emptyset$ and so $\boldsymbol{p}+\boldsymbol{q} \in$

$cl(E^*) = E$. Note that if either $\boldsymbol{p} \in I$ or $\boldsymbol{q} \in I$, then $c + d > 0$ which implies that $(b + a, \ldots b + a + (k - 1)(c + d)) \in U \cap I^*$ and therefore $\boldsymbol{p} + \boldsymbol{q} \in I$. □

Two proofs of Theorem 8.4.1. The following are two proofs, one using minimal idempotents and the other using Theorem 8.2.19.

(i) Minimal idempotent route:

Let u be a minimal idempotent in $\beta\mathbb{Z}^+$. Since $+$ is performed component-wise, $(u, \ldots, u)$ is a minimal idempotent in $(\beta\mathbb{Z}^+)^k$ and $(u, \ldots, u) \in E$ since $(u, \ldots, u) \in cl(\{(n, \ldots, n) : n \in \mathbb{Z}^+\}) \subseteq E$. By Lemma 8.2.23, $(u, \ldots, u) \in I$ since I is a 2-sided ideal in E.

(ii) Sticky diagonal theorem route:

Since I is a two-sided ideal in E, by Theorem 8.2.19, there is a $u \in \beta\mathbb{Z}^+$, such that $(u, \ldots, u) \in I$. □

8.5 Sum sets

There are a number of Ramsey-type theorems with ultrafilter proofs, among them is Hindman's theorem on finite sum sets. Using ultrafilters it has also been shown that van der Waerden's theorem can be extended. For any finite colouring of $\mathbb{Z}^+$, one colour class contains not only arithmetic progressions, but also other sets with a nice pattern.

Definition 8.5.1. For a set $A \subseteq \mathbb{Z}$, the *finite sum set* of A is defined to be

$$FS(A) = \left\{ \sum_{x \in F} x : \emptyset \neq F \subseteq A,\ |F| < \infty \right\}.$$

Similarly, the *finite product set* is defined to be

$$FP(A) = \left\{ \prod_{x \in F} x : \emptyset \neq F \subseteq A,\ |F| < \infty \right\}.$$

The following theorem was originally proved using purely combinatorial techniques by Hindman [64] and subsequently by Baumgartner [4], but can also be proved by the methods of ultrafilters. The ultrafilter proof of the following theorem is due to Glazer (see [51, pp. 168–9]). In this proof, it is important that $0 \notin \mathbb{Z}^+$ because $p_0 + p_0 = p_0$ is an idempotent and the proof requires that no principal ultrafilter be an idempotent.

Theorem 8.5.2 (Hindman [64])**.** For every $r \in \mathbb{Z}^+$ and every r-colouring of $\mathbb{Z}^+$, there is an infinite set A such that $FS(A)$ is monochromatic.

Proof. Fix an idempotent $p \in \beta\mathbb{Z}^+$. Let $\Delta : \mathbb{Z}^+ \to [1, r]$ be any r-colouring, by Lemma 8.1.4, there is an $i \in [1, r]$ such that $\Delta^{-1}(i) \in p$. Set $A_0 = \Delta^{-1}(i)$.

The following notation will be used in this proof. Given $B \subseteq \mathbb{Z}^+$, let $B^* = \{n \in \mathbb{Z}^+ : B - n \in p\}$. Note that:

$$B \in p = p + p \Rightarrow \{n \in \mathbb{Z}^+ : B - n \in p\} \in p$$

$$\Rightarrow B^* \in p$$

Further, note that for any principal ultrafilter p_i, $p_i + p_i = p_{2i} \neq p_i$. Therefore, the idempotent ultrafilter is not principal and contains no singleton sets.

Since $A_0 \in p$, $A_0^* \in p$ and hence $A_0 \cap A_0^* \in p$. Since $\emptyset \notin p$, $A_0 \cap A_0^* \neq \emptyset$ so pick $a_1 \in A_0 \cap A_0^*$ and set $A_1 = A_0 \cap (A_0 - a_1) \backslash \{a_1\}$.

Since $a_1 \in A_0^*$, $A_0 - a_1 \in p$ and since p is not principal, $\mathbb{Z}^+ \backslash \{a_1\} \in p$. Thus $A_1 \in p$. Further, $A_1 \subseteq A_0$ and $a_1 + A_1 \subseteq A_0$.

In general, having chosen $A_n \in p$, select $a_{n+1} \in A_n \cap A_n^*$ and set $A_{n+1} = A_n \cap (A_n - a_{n+1}) \backslash \{a_{n+1}\}$. Then $A_{n+1} \subseteq A_n$, $a_{n+1} + A_{n+1} \subseteq A_n$ and $A_{n+1} \in p$.

This gives a nested sequence of sets $A_0 \supset A_1 \supset \cdots$ and elements $a_n \in A_{n-1}$ such that for all $n \geq 1$, $a_n + A_n \subseteq A_{n-1}$. Thus given any finite subset of these elements, $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$, where $i_1 < i_2 < \ldots i_n$, $a_{i_1} + a_{i_2} + \ldots + a_{i_n} \in A_{i_1 - 1} \subseteq A_0$.

Let $A = \{a_n\}_{n \geq 0}$. Then $FS(A) \subseteq A_0 = \Delta^{-1}(i)$ is monochromatic. □

Hindman's theorem can also be used to show the existence of an infinite set whose finite product set is monochromatic. Given any r-colouring $\Delta : \mathbb{Z}^+ \rightarrow [1, r]$, define a new r-colouring $\Delta^* : \mathbb{Z}^+ \rightarrow [1, r]$ by $\Delta^*(n) = \Delta(2^n)$. Applying Theorem 8.5.2 to the colouring Δ^* shows that there is an infinite set A and $i \in [1, r]$ so that $FS(A) \subseteq \Delta^{*-1}(i)$. Let $B = \{2^n : n \in A\}$. Given any distinct $2^{n_1}, 2^{n_2}, \ldots, 2^{n_k} \in B$, since $n_1, n_2, \ldots, n_k \in A$, $i = \Delta^*(n_1 + n_2 + \cdots + n_k) = \Delta(2^{n_1 + n_2 + \cdots + n_k}) = \Delta(2^{n_1} \cdot 2^{n_2} \cdots 2^{n_k})$.

Thus $FP(B) \subseteq \Delta^{-1}(i)$. The following theorem shows a connection between these results and arithmetic progressions.

Theorem 8.5.3 (Hindman [66])**.** For any $r \in \mathbb{Z}^+$ and any r-colouring $\Delta : \mathbb{Z}^+ \to [1, r]$, there exist an $i \in [1, r]$ and infinite sets $A, B \subseteq \mathbb{Z}^+$ so that $FS(A) \cup FP(B) \subseteq \Delta^{-1}(i)$ and so that $\Delta^{-1}(i)$ contains arbitrarily long arithmetic progressions.

Chapter 9

Generalizations and applications

9.1 Polynomial van der Waerden

Van der Waerden's theorem guarantees that for any finite colouring of $\mathbb{Z}^+$, there are arbitrarily long monochromatic arithmetic progressions, but says nothing about the possible differences of these arithmetic progressions. Is it possible to guarantee a monochromatic arithmetic progression whose difference is a perfect square or a perfect cube? In 1996, this question was answered in the affirmative by Bergelson and Leibman who used ergodic theory techniques to prove generalizations of both van der Waerden's theorem and Szemerédi's theorem that involve polynomials with integer coefficients. The theorems in this section are stated without proof.

Theorem 9.1.1 (**Polynomial van der Waerden**, Bergelson and Leibman [14]).

Let $p_1, p_2, \ldots, p_k$ be polynomials with integer coefficients such that for each $i \in [1, k]$, $p_i(0) = 0$. Then, for every $r \in \mathbb{Z}^+$ and any r-colouring of $\mathbb{Z}^+$, there are integers $a, d \in \mathbb{Z}^+$ such that the set

$$\{a\} \cup \{a + p_i(d) : 1 \leq i \leq k\}$$

is monochromatic.

Fix $k \in \mathbb{Z}^+$ and for each $i \in [1, k-1]$, let p_i be the polynomial $p_i(x) = ix$. Then Theorem 9.1.1 is exactly van der Waerden's theorem. Alternatively, applying the theorem to the polynomials $x^2, 2x^2, \ldots, (k-1)x^2$ shows that any finite colouring of $\mathbb{Z}^+$ yields a monochromatic AP_k whose difference is a perfect square.

In 2000, Walters [117] gave purely combinatorial proofs for both Theorem 9.1.1 and a polynomial version of the Hales-Jewett theorem (recall Theorem 7.1.4). He also showed how these results can be used to deduce a polynomial version of the Gallai-Witt theorem (recall Theorem 7.4.4).

The polynomial version of Szemerédi's theorem given by Bergelson and Leibman uses a slightly different notion of density than the upper density used in Szemerédi's theorem.

Definition 9.1.2. For a set $A \subseteq \mathbb{Z}$, A has *positive upper Banach density* iff for some $\varepsilon > 0$ there are sequences $\{a_n\}_{n\geq 0}$, $\{b_n\}_{n\geq 0} \subseteq \mathbb{Z}$ with $\lim_{n\to\infty} (b_n - a_n) = \infty$

and such that for every $n \geq 0$,

$$\frac{|A \cap [a_n, b_n]|}{|[a_n, b_n]|} > \varepsilon.$$

Positive upper Banach density is a weaker notion than upper density (recall Definition 4.1.6) since any set with positive upper density will also have positive upper Banach density.

In 1978, Furstenberg and Katznelson [45] proved a multidimensional version of Szemerédi's theorem which was extended, in 1996, by Bergelson and Leibman to a multidimensional "Polynomial Szemerédi theorem". In order to simplify the notation used, only the 1-dimensional version of the Polynomial Szemerédi Theorem is presented here, without proof.

Theorem 9.1.3 (**Polynomial Szemerédi**, Bergelson and Leibman [14])**.** Let $A \subseteq \mathbb{Z}$ have positive upper Banach density. Let $p_1, \ldots, p_k$ be polynomials with rational coefficients that take integer values on the integers and such that for each $i \in [1, k]$, $p_i(0) = 0$. For any $v_1, \ldots, v_k \in \mathbb{Z}$, there exist $a, d \in \mathbb{Z}$ so that $\{a + p_i(d)v_i : 1 \leq i \leq k\} \subseteq A$.

In particular, using the polynomials $x, 2x, \ldots, (k-1)x$ and taking $v_1 = v_2 = \cdots = v_{k-1} = 1$, this provides a version of the usual Szemerédi's theorem using upper Banach density.

The Polynomial Szemerédi Theorem has also been extended to the set of prime numbers.

Theorem 9.1.4 (Tao and Ziegler [110])**.** Let A be an infinite set of primes with

$$\limsup_{n\to\infty} \frac{|A \cap [1,n]|}{\pi(n)} > 0,$$

and let $p_1, \ldots, p_k$ be polynomials with integer coefficients and such that for each $i \in [1,k]$, $p_i(0) = 0$. Then there are $a, d \in \mathbb{Z}^+$ such that

$$\{a + p_1(d), a + p_2(d), \ldots, a + p_k(d)\} \subseteq A.$$

9.2 Rainbow results

Rainbow Ramsey theory is sometimes also called "Anti-Ramsey theory". While Ramsey theory concerns itself with finding monochromatic sets of a certain form, rainbow Ramsey theory seeks sets that are either monochromatic or have one element of each colour. Most results here are recorded without proof but can be found, for example in [2, 23, 72, 73, 74].

Definition 9.2.1. Given a set X, $r \in \mathbb{Z}^+$ and an r-colouring $\Delta : X \to [1,r]$, a set $A \subseteq X$ is *rainbow* iff for every $i \in [1,r]$, $|A \cap \Delta^{-1}(i)| \leq 1$. That is, A is rainbow if no two elements of A are the same colour.

The following theorem is often called a "canonical" version of van der Waerden's theorem, shows a connection between finding monochromatic arithmetic progressions and rainbow arithmetic progressions.

Theorem 9.2.2 (Erdős and Graham [35])**.** *For every $k \geq 3$, there is an integer $n(k)$ such that for every $n \geq n(k)$ and every colouring of $[1, n]$, there is either a monochromatic AP_k or else a rainbow AP_k.*

One of the rainbow Ramsey problems that has been investigated is that of determining what conditions on a colouring of an interval $[1, n]$, or $\mathbb{Z}^+$, will guarantee the existence of a rainbow AP_k. The following example due to Jungić *et al.* [72] shows that taking n sufficiently large is not enough.

For any $n \in \mathbb{Z}^+$, let $\Delta : [1, n] \to [1, \lfloor \log_3 n \rfloor + 1]$ be defined by $\Delta(i) = \max\{t \geq 0 : 3^t | i\}$. For any $x \in [1, n]$, $\Delta(x) = \Delta(2x)$ and if $x, y \in [1, n]$ are such that $\Delta(x) \neq \Delta(y)$, then $\Delta(x + y) = \min\{\Delta(x), \Delta(y)\}$. Therefore, for any $\{a, a + d, a + 2d\} \subseteq [1, n]$, if $\Delta(a) \neq \Delta(a + 2d)$, then

$$\Delta(a + d) = \Delta(2(a + d)) = \Delta(a + a + 2d) = \min\{\Delta(a), \Delta(a + 2d)\}.$$

Thus $[1, n]$ cannot contain any rainbow AP_3's and so also no longer arithmetic progressions can be rainbow.

In the above example, at least two thirds of the integers in $[1, n]$ are in the colour class $\Delta^{-1}(0)$. Many of the problems related to rainbow sets consider only colourings where the sizes of the colour classes are all approximately equal.

Definition 9.2.3. Given $n, r \in \mathbb{Z}^+$, an r-colouring $\Delta : [1, n] \to [1, r]$ is *equinumerous* iff for each $i \in [1, r]$, $\Delta^{-1}(i) \in \{\lfloor n/k \rfloor, \lceil n/k \rceil\}$. That is, the sizes of the colour classes are as equal as possible.

When dealing with problems related to rainbow AP_3's, the colours will be denoted $\{R, B, G\}$ (red, blue, and green) and for any 3-colouring $\Delta : [1, n] \to \{R, B, G\}$, denote by $\Delta^{-1}(R) = \mathcal{R}_\Delta$, $\Delta^{-1}(B) = \mathcal{B}_\Delta$ and $\Delta^{-1}(G) = \mathcal{G}_\Delta$. Given a 3-colouring $\Delta : \mathbb{Z}^+ \to \{R, B, G\}$, denote $\Delta^{-1}(R) \cap [1, n] = \mathcal{R}_\Delta(n)$, $\Delta^{-1}(B) \cap [1, n] = \mathcal{B}_\Delta(n)$ and $\Delta^{-1}(G) \cap [1, n] = \mathcal{G}_\Delta(n)$.

In 2003, Jungić *et. al.* [72] showed that in any 3-colouring of $\mathbb{Z}^+$ where each colour class has density greater than $\frac{1}{6}$, there will be a rainbow AP_3.

Theorem 9.2.4 (Jungić, Licht, Maholian, Nešetřil and Radoičić [72])**.** For any 3-colouring $\Delta : \mathbb{Z}^+ \to \{R, B, G\}$ with

$$\limsup_{n\to\infty} \{\min\{\mathcal{R}_\Delta(n), \mathcal{B}_\Delta(n), \mathcal{G}_\Delta(n)\} - \frac{n}{6}\} = \infty,$$

$\mathbb{Z}^+$ contains a rainbow AP_3.

It was conjectured that something similar held in the finite case. In their 2003 paper, Jungić *et al.* [72] gave the following two examples of colourings of finite intervals that showed that if such a conjecture were true, the result would be best possible.

For $n \neq 2 \pmod 6$ let $\Delta : [1, n] \to \{R, B, G\}$ be defined by

$$\Delta(i) = \begin{cases} R & \text{if } i \equiv 1 \pmod 6, \\ B & \text{if } i \equiv 4 \pmod 6, \\ G & \text{otherwise.} \end{cases}$$

The smallest colour class $\Delta^{-1}(B)$ has $\lfloor\frac{n+2}{6}\rfloor$ elements. Let $\{a, a+d, a+2d\} \subseteq [1,n]$ be any AP_3. It cannot be that one of $\{a, a+2d\}$ is red and the other blue since then $2d = (a+2d) - a \equiv 3 \pmod 6$ is an odd number. If either the pair $\{a, a+d\}$ or $\{a+d, a+2d\}$ contains one red element and one blue element, then the difference d must be equal to $3 \pmod 6$. In either case, the remaining term of the AP_3 must either be red or blue. Thus, under this colouring, $[1,n]$ contains no rainbow AP_3's.

For $n \equiv 2 \pmod 6$, let $q \in\geq 0$ be such that $n = 6q+2$. The colouring $\Delta : [1,n] \to \{R, B, G\}$, defined by

$$\Delta(i) = \begin{cases} R & \text{if } i \leq 2q+1 \text{ and } i \text{ is odd,} \\ B & \text{if } i \geq 4q+2 \text{ and } i \text{ is even,} \\ G & \text{otherwise.} \end{cases}$$

The smallest colour classes are $\Delta^{-1}(R)$ and $\Delta^{-1}(B)$ with $|\Delta^{-1}(R)| = |\Delta^{-1}(B)| = q+1 = \frac{n+4}{6}$ elements. Again, red and blue elements cannot be the first and last terms of an AP_3 since their difference would be and odd number. If red and blue elements are consecutive terms in an arithmetic progression, the difference of the progression must be at least $2q+1$ which is impossible.

In 2004, Axenovich and Fon-Der-Flass showed that indeed, any 3-colouring where the smallest colour class is larger than those in the two previous examples will contain a rainbow AP_3.

Theorem 9.2.5 (Axenovich and Fon-Der-Flaas [2]). For every $n \geq 3$ and every

3-colouring $\Delta : [1,n] \rightarrow \{R, B, G\}$ such that

$$\min\left(|\mathcal{R}_\Delta|, |\mathcal{B}_\Delta|, |\mathcal{G}_\Delta|\right) > \begin{cases} \lfloor (n+2)/6 \rfloor & \text{if } n \neq 2 \pmod 6, \\ (n+4)/6 & \text{if } n = 2 \pmod 6) \end{cases}$$

$[1,n]$ contains a rainbow AP_3.

This theorem shows that every equinumerous 3-colouring of a finite interval yields a rainbow AP_3. However, it was later shown that no such result holds for longer arithmetic progressions.

Theorem 9.2.6 (Conlon, Jungić and Radoičic [23])**.** For every $n \in \mathbb{Z}^+$ with $n \equiv 0 \pmod 8$, there is an equinumerous 4-colouring of $[1,n]$ with no rainbow AP_4.

Theorem 9.2.7 (Axenovich and Fon-Der-Flaas [2])**.** For every $k \geq 5$ and $n \in \mathbb{Z}^+$, there is an equinumerous k-colouring of $[1,n]$ with no rainbow AP_k.

Thus, an equinumerous k-colouring of an interval $[1,n]$ guarantees a rainbow AP_k only when $k = 3$.

The problem of how to guarantee rainbow AP_k's has also been studied in the context of modular arithmetic. Theorem 9.2.4 can be used to show that for any $n \in \mathbb{Z}^+$ and any 3-colouring of $\mathbb{Z}_n$ where the smallest colour class has density greater than $\frac{1}{6}$ will contain a rainbow AP_3. The following theorem provides an improvement on this result for AP_3's in $\mathbb{Z}_n$.

Theorem 9.2.8 (Jungić, Licht, Maholian, Nešetřil and Radoičić [72]). Let n be odd and q be the smallest prime factor of n. Then for every 3-colouring of $\mathbb{Z}_n$ where $\min\{|\Delta^{-1}(R)|, |\Delta^{-1}(B)|, |\Delta^{-1}(G)|\} > \frac{n}{q}$, there is a rainbow AP_3.

Another variation of the problem of finding rainbow arithmetic progressions is to let the number of colours be larger than the length of the arithmetic progressions. For any fixed $k \in \mathbb{Z}^+$, what is the minimal $t \in \mathbb{Z}^+$ such that for every $n \in \mathbb{Z}^+$ and every equinumerous t-colouring of $[1, tn]$, there is a rainbow AP_k? Denote this minimal t, T_k. In their 2003 paper, Jungić *et al.* [72] showed that for each $k \geq 3$, $\lfloor \frac{k^2}{4} \rfloor < T_k < \frac{k(k-1)^2}{2}$.

9.3 Applications

Historically, the first application of van der Waerden's theorem may be due to Brauer who proved a conjecture of Schur about quadratic residues. Brauer used a generalization of van der Waerden's theorem, which he attributed to Schur. The following theorem is a further generalization of Brauer's result. The proof is now folklore and I have been unable to locate the original source. The details appear, for example, in [51, p. 70].

Theorem 9.3.1 ([51]). For every $k, r \in \mathbb{Z}^+$, there is a least integer $n = SB(k, s; r)$

such that for every r-colouring of $[1, n]$, there exist $a, d \in \mathbb{Z}^+$ so that the set

$$\{sd\} \cup \{a + id : 0 \leq i \leq k - 1\}$$

contained in $[1, n]$ is monochromatic.

Proof. Fix $k, s \in \mathbb{Z}^+$. The proof is by induction on r and uses van der Waerden's theorem.

Base Case: For $r = 1$, take $n = max\{s, k\}$. Then for any 1-colouring of $[1, n]$, the set $\{s\} \cup [1, k]$ is monochromatic.

Inductive Step: Suppose that for some $r \geq 1$, $SB(k, s; r)$ exists. Set $n = s \cdot W(k \cdot SB(k, s; r) + 1; r + 1)$ and let $\Delta : [1, n] \to [1, r + 1]$ be any $(r + 1)$-colouring. By the choice of n (and van der Waerden's theorem), there are $a, d \in \mathbb{Z}^+$ such that the arithmetic progression

$$\{a + id : 0 \leq i \leq k \cdot SB(k, s; r)\} \ \subseteq [1, W(k \cdot SB(k, s; r) + 1; r + 1)]$$

is monochromatic.

If, for some $j \in [1, SB(k, s; r)]$, $\Delta(a) = \Delta(s \cdot jd)$, then the set

$$\{s \cdot jd\} \cup \{a + i(jd) : 0 \leq i \leq k - 1\}$$

is monochromatic.

Otherwise, the colouring Δ restricted to the set $\{s \cdot jd : 1 \leq j \leq SB(k, s; r)\}$ is an r-colouring and so by the definition of $SB(k, s; r)$, there are $a', d' \in \mathbb{Z}^+$ such that the set

$$\{sd(a' + id') : 0 \leq i \leq k-1\} \cup \{sd(sd')\}$$

is monochromatic. That is,

$$\{sda' + i(sdd') : 0 \leq i \leq k-1\} \cup \{s(sdd')\}$$

is monochromatic and is contained in $[1, n]$. □

Definition 9.3.2. Given $a, n \in \mathbb{Z}^+$, a is a *quadratic residue* $\pmod{n}$ iff there is a solution to the equation $x^2 \equiv a \pmod{n}$. Otherwise, a is a *quadratic non-residue* $\pmod{n}$.

An important fact about quadratic residues is that if $a, b \in \mathbb{Z}_n$ are both quadratic residues or both quadratic non-residues $\pmod{n}$ then ab is a quadratic residue modulo n. If a is a quadratic residue and b is a quadratic non-residue, then ab is a quadratic non-residue $\pmod{n}$.

Van der Waerden's theorem can be applied directly to show, colouring integers according to whether they are a quadratic residue or a quadratic non-residue, that for every k and $n \geq W(k; 2)$, $[1, n]$ either contains an AP_k of quadratic residues $\pmod{n}$ or else an AP_k of quadratic non-residues $\pmod{n}$. Using Theorem 9.3.1 and the properties of quadratic residues, Brauer was able to show that more is in fact true.

Theorem 9.3.3 (Brauer [18])**.** For every $k \in \mathbb{Z}^+$ there is a least integer $n = B(k)$ such that for every prime $p \geq n$, there exist k consecutive quadratic residues $\pmod{p}$ as well as k consecutive quadratic non-residues $\pmod{p}$.

Proof. Fix $k \in \mathbb{Z}^+$ and let $p > SB(k, 1; 2)$ be prime. Define a 2-colouring of $[1, p-1]$ by

$$\Delta(x) = \begin{cases} 1 & \text{if } x \text{ is a quadratic residue} \pmod{p}, \\ -1 & \text{otherwise.} \end{cases}$$

By the choice of n, there are $a, d \in \mathbb{Z}^+$ such that

$$\{d\} \cup \{a + id : 0 \leq i \leq k-1\}$$

is monochromatic under Δ. Since p is prime and $d \in [1, p-1]$, d is invertible in $\mathbb{Z}_p$ and so d^{-1} is a quadratic residue $\pmod{p}$ iff d is a quadratic residue. Therefore, since the product of two quadratic residues or two quadratic non-residues is a quadratic residue, the set

$$\{d^{-1}(a + id) : 0 \leq i \leq k-1\} = \{ad^{-1} + i : 0 \leq i \leq k-1\}$$

is a sequence of k consecutive quadratic residues $\pmod{p}$, thus completing the first part of the proof.

Now, set $\ell = k! \cdot k + 1$ and let $p > SB(\ell; 2)$ be a prime. By the first part of the proof, there is a $b \in \mathbb{Z}^+$ such that the set $\{b, b+1, \ldots, b+(\ell-1)\}$ is a sequence of ℓ

consecutive quadratic residues $\pmod p$. Let $q \in [1, p-1]$ be the smallest quadratic non-residue $\pmod p$. Two different cases are considered.

Case I: Suppose that $q \leq k!$. Then there is a c with $1 \leq c \leq q \leq k!$ such that $q|(b+c)$. For each $j \in [1, k]$,

$$c + jq \leq k! + (k-1)k! = k! \cdot k = \ell - 1$$

and so $b + (c + jq)$ is a quadratic residue and since q^{-1} is a quadratic non-residue, so is $q^{-1}(b + c + jq)$. Thus

$$\{q^{-1}(b + c + jq) : 0 \leq j \leq k-1\} = \{q^{-1}(b+c) + j : 0 \leq j \leq k-1\}$$

is a sequence of k-consecutive quadratic non-residues modulo p.

Case II: Suppose that $q > k!$. Then there exists a c with $0 \leq c \leq k! - 1$ so that $k!|(q-c)$. For every $j \in [1, k]$, $j|(c-q)$ and since $c < k! < q$, $\frac{c-q}{j} < 0$. Therefore, $\frac{c-q}{j} + q < q$ and since q is the least quadratic non-residue, if $j \in [1, k]$, then $\frac{c-q}{j} + q$ is a quadratic residue. Since $k < q$ each $j \in [1, k]$ is also a quadratic residue and hence for each $j \in [1, k]$, the integer

$$j\left(\frac{c-q}{j} + q\right) = (c-q) + jq$$

is also a quadratic residue. Finally, since q is a quadratic non-residue, q^{-1} is also a quadratic non-residue and the set

$$\{q^{-1}(c - q + jq) : 1 \leq j \leq k\} = \{q^{-1}c - 1 + j : 1 \leq j \leq k\}$$

is a sequence of k consecutive quadratic non-residues. □

A generalization of Theorem 9.3.3 that extends the result to the Gaussian integers ($\mathbb{Z}[i]$) was proved by Rabung [86] in 1975.

Another application of results related to van der Waerden's theorem is due to Coppersmith and Winograd [24]. They were able to use the large AP_3-free sets given by Salem and Spencer (see Theorem 4.6.1) to produce an algorithm for fast matrix multiplication.

9.4 Open problems

There are a number of variations of van der Waerden's theorem and related problems that remain unsolved. A few of these are presented here.

While Szemerédi's theorem states that all sets of integers with positive upper density contain arbitrarily long arithmetic progressions, there are sets with zero upper density that contain arbitrarily long arithmetic progressions. For example, let $A = \bigcup_{k=0}^{\infty} [2^k, 2^k + k]$. The set A contains arbitrarily long strings of consecutive integers (and hence arbitrarily long arithmetic progressions), but has zero density. Erdős (see for example, [34]) conjectured that if X is any set satisfying $\sum_{x \in X} \frac{1}{x} = \infty$, then X contains arbitrarily long arithmetic progressions. It can be shown that $\sum_{p \text{ prime}} \frac{1}{p} = \infty$. Thus, a proof of Erdős's conjecture would provide another proof of

Green and Tao's theorem (Theorem 4.7.1) that the primes contain arbitrarily long arithmetic progressions. Note that for the above set A, $\sum\limits_{a\in A}\frac{1}{a}$ is finite and hence the condition that $\sum\limits_{x\in X}\frac{1}{x}=\infty$ is not a necessary condition for a set of integers to contain arbitrarily long arithmetic progressions.

Only a small number of exact values of van der Waerden numbers or mixed van der Waerden numbers are known (see Figure 5.1). In most cases, the known upper and lower bounds for van der Waerden numbers are of vastly different orders of magnitude (see Chapter 5). Erdős has conjectured (see for example [34]) that $\lim_{k\to\infty}\frac{W(k;2)}{2^k}=\infty$ and $\lim_{k\to\infty}(W(k;2))^{1/k}=\infty$. Landman [79] has conjectured that the mixed van der Waerden numbers $W(3,k;2)$ are quadratic in k.

In a problem related to discrepancy (see Chapter 6) Erdős suggested examining the following problem (see [34]). For any $k\in\mathbb{Z}^+$ and $\frac{k}{2}<\ell\leq k$, define $f(\ell,k)$ to be the smallest integer such that if $[1,f(\ell,k)]$ is 2-coloured, there exists an AP_k with ℓ elements monochromatic. According to Gunderson, Erdős conjectured that for every $\varepsilon>0$ there is a constant c so that for all ℓ, $f((\frac{1}{2}+\varepsilon)\ell,\ell)\leq c2^\ell$. That is, for every $\varepsilon>0$ there is a constant c so that for all ℓ, and $n\geq c2^\ell$, for every 2-colouring of $[1,n]$, there is an AP_ℓ so that $(\frac{1}{2}+\varepsilon)\ell$ elements are monochromatic.

There are many other problems related to van der Waerden's theorem that remain unsolved. Some of these can be found, for example, in [34, 51, 80].

Bibliography

[Note:] Numbers following each bibliography item indicate where item is cited.

[1] T. M. Apostol, *Introduction to analytic number theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1976. 72, 73

[2] M. Axenovich and D. Fon-Der-Flaas, On rainbow arithmetic progressions, *Electron. J. Combin.* **11** (2004), no. 1, Research Paper 1, 7 pages (electronic). 168, 171, 172

[3] A. Balog, Linear equations in primes, *Mathematika* **39** (1992), 367–378. 72

[4] J. E. Baumgartner, A short proof of Hindman's theorem, *J. Combinatorial Theory Ser. A* **17** (1974), 384–386. 162

[5] J. E. Baumgartner and A. D. Taylor, Partition theorems and ultrafilters, *Trans. Amer. Math. Soc.* **241** (1978), 283–309. 137

[6] J. Beck, Roth's estimate of the discrepancy of integer sequences is nearly sharp, *Combinatorica* **1** (1981), 319–325. 118, 119, 123

[7] J. Beck and T. Fiala, "Integer-making" theorems, *Discrete Appl. Math.* **3** (1981), 1–8. 118

[8] J. Beck and V. Sós, Discrepancy Theory, Chapter 26 in *Handbook of Combinatorics*, (eds. R. L. Graham, M. Grötschel and L. Lovász), North-Holland (1995), 1405–1446. 111, 113, 119

[9] M. Beeler, A new van der Waerden number, *Discrete Appl. Math.* **6** (1983), 207. 77, 78

[10] M. Beeler and P. O'Neil, Some new van der Waerden numbers, *Discrete Math.* **28** (1979), 135–146. 77, 78

[11] F. A. Behrend, On sequences of integers containing no arithmetic progression, *Časopis Mat. Fys. Praha* **67** (1938), 235–239. 45, 46

[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. 69

[13] V. Bergelson, H. Furstenberg, N. Hindman, and Y. Katznelson, An algebraic proof of van der Waerden's theorem, *Enseign. Math. (2)* **35** (1989), 209–215. 144, 147, 148, 158

[14] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. Amer. Math. Soc.* **9** (1996), no. 3, 725–753. 165, 167

[15] E. R. Berlekamp, A construction for paritions which avoid long arithmetic progressions, *Canad. Math. Bull.* **11** (1968), 409–414. 87, 88, 91

[16] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics **184**, Springer-Verlag, New York, 1998. 12

[17] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.* **9** (1999), 968–984. 58, 100

[18] A. Brauer, Über Sequenzen von Potenzresten, *Sitzungsber. Preuß. Akad. Wiss. Math. Phy. Kl.* (1928), 9–16. 176

[19] T. C. Brown, Some new van der Waerden numbers (preliminary report), *Notices Amer. Math. Soc.* **21** (1974), A-432. 77

[20] T. C. Brown, Variations on van der Waerden's and Ramsey's Theorems, *Amer. Math. Monthly* **82** (1975), 993–995. 31

[21] B. Chazelle, *Discrepancy Method*, Cambridge University Press, 2001. 111

[22] V. Chvátal, Some unknown van der Waerden numbers, in *Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969)*, Gorden and Breach, New York (1970), 31–33. 77

[23] D. Conlon, V. Jungić and R. Radoičić, On the existence of rainbow 4-term arithmetic progressions, preprint available from `http://www.math.rutgers.edu/~rados/#publications`. 168, 172

[24] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progressions, *J. Symbolic Logic* **9** (1990), 251–280. 178

[25] R. Courant, *Differential and Integral Calculus Vol. I*, Trans. E. J. McShane, 2nd ed., Blackie & Son Ltd., London and Glasgow, 1937. 121

[26] C. Culver, B. Landman and A. Roberston, Some new exact van der Waerden numbers, in *Topics in Combinatorial Number Theory: Proceedings of the INTEGERS Conference 2003*, Prague (2005), 123–134. 76, 77, 78

[27] N. G. de Bruijn and P. Erdős, A colour problem for infinite graphs and a problem in the theory of relations, *Indag. Math.* **13** (1951), 371–373. 29

[28] Ch. de la Vallé Poussin, Recherches analytiques sur la théorie des nombres premiers, *Ann. Soc. Sci. Bruxelles* **20** (1896), 183–256 and 281–297. 73

[29] W. Deuber, Partitionen und lineare Gleichungssysteme, *Math. Z.* **133** (1973), 109–123. 67

[30] 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 Angew. Math.* **135** (1909), 181–188. 23, 24

[31] L. E. Dickson, Fermat's Last Theorem and the Origin and Nature of the Theory of Algebraic Numbers, *Ann. of Math.* **18** (1917), no. 4, 161–187. 23

[32] R. Ellis, *Lectures on Topological Dynamics*, Benjamin, New York, 1969. 145

[33] P. Erdős, Szamelmeleti megjegyzesek V. Extremalis problemak a szamelmeletben, II, *Mat. Lapok* (1966), 135–155. 118

[34] P. Erdős, A survey of problems in combinatorial number theory. *Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978). Ann. Discrete Math.* **6** (1980), 89–115. 178, 179

[35] P. Erdős, R. L. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathématique **28**, Université de Genève, L'Enseignement Mathématique, Geneva, 1980. 169

[36] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, *Infinite and finite sets (Colloq., Keszthely, 1973), Vol. II* 609–627. 80, 82, 84

[37] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. London Math. Soc. (3)* **2** (1952), 417–439. 80

[38] P. Erdős and A. Sárközy, Some solved and unsolved problems in combinatorial number theory, *Math. Slovaca* **28** (1978), 407–421. 118

[39] P. Erdős and J. Spencer, *Probabilistic methods in combinatorics*, Probability and Mathematical Statistics, Vol. 17. Academic Press , New York-London, 1974. 98, 118

[40] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* (1935) **2**, 463–470. 13, 14

[41] P. Erdös and P. Turán, On some sequences of integers, *J. London Math. Soc.* **11** (1936), 261–264. 44, 46

[42] R. Fraïssé, *Theory of Relations*, translated from the French, Studies in Logic and the Foundations of Mathematics, **118**, North-Holland Publishing Co., Amsterdam, 1986. 139

[43] P. Frankl, R. L. Graham and V. Rödl, Quantitative Theorems for Regular Systems of Equations, *J. Combin. Theory Ser. A* **47** (1988), 246–261. 68

[44] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* **31** (1977), 204–256. 59, 100

[45] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. Analyse Math.* **34** (1978), 275–291. 167

[46] H. Furstenberg and Y. Katznelson, Idempotents in compact semigroups and Ramsey theory, *Israel J. Math.* **68** (1989), 257–270. 144, 149, 150, 158

[47] W. T. Gowers, A new proof of Szemerédi's Theorem, *Geom. Funct. Anal.* **11** (2001), 465–588. 59, 101

[48] R. L. Graham, K. Leeb, and B. L. Rothschild, Ramsey's theorem for a class of categories, *Adv. in Math.* **8** (1972), 417–433. 133

[49] R. L. Graham and B. L. Rothschild, Ramsey's theorem for n-parameter sets, *Trans. Amer. Math. Soc.* **159** (1971), 257–292. 132

[50] 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. 40

[51] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory*, Second Edition, John Wiley & Sons, New York, 1990. 14, 22, 27, 45, 49, 52, 59, 85, 88, 94, 95, 96, 99, 162, 173, 179

[52] B. Green, On triples in arithmetic progession by Jean Bourgain, (1999) available at `http://www.maths.bris.ac.uk/~mabjg/papers/bourgain.pdf`. 100

[53] B. Green, Roth's theorem in the primes, *Ann. of Math. (2)* **161** (2005), 1609–1636. 73

[54] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, to appear in *Ann. of Math.*, available at `http://www.math.ucla.edu/~tao/preprints/acnt.html`, (2005). 73

[55] B. Green and T.Tao, A bound for progressions of length k in the primes, Preprint, available at `http://www-math.mit.edu/~green/preprints.html`. 74

[56] D. S. Gunderson and V. Rödl, Extremal Problems for affine cubes of integers, *Combin. Probab. Comput.* **7** (1998), 65–79. 21

[57] J. Hadamard, Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, *Bull. Soc. Math. France* **24** (1896), 199–220. 73

[58] A. W. Hales and R. I. Jewett, Regularity and positional games. *Trans. Amer. Math. Soc.* **106** (1963), 222-229. 126, 129, 133, 135

[59] K. P. Hart, Applications in combinatorics, Chapter 6 in *The Čech-Stone compactification*, Lecture notes (2002-2003), available at `http://dutiaw37.twi.tudelft.nl/~kp/onderwijs/wi4041/2002-2003/chapter6.pdf`. 144, 149

[60] Y. R. Haung and J. S. Yang, New upper bounds for van der Waerden numbers $W(3,n)$ (Chinese), *Chinese Annals of Math. Series A* **21** (2000), 631–634. 101, 102

[61] D. R. Heath-Brown, Integer sets containing no arithmetic progressions, *J. London Math. Soc. (2)* **35** (1987), 385–394. 58

[62] P. R. Herwig, M.J.H. 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. 75

[63] D. Hilbert, Über die Irreduzibilität ganzer reationaler Funktionen mit ganzzahligen Keffizienten, *J. Reine Angew. Math.* **110** (1892), 104–129. 16

[64] N. Hindman, Finite sums from sequences within cells of a partition of N, *J. Combin. Theory Ser. A* **17** (1974), 1–11. 162

[65] N. Hindman, Ultrafilters and Combinatorial Number Theory, *Number Theory Carbondale 1979*, Lecture Notes in Math. **751**, Springer, Berlin, 1979. 141

[66] N. Hindman, Ramsey's theorem for sums, products, and arithmetic progressions, *J. Combin. Theory Ser. A* **38** (1985), 82–83. 164

[67] N. Hindman and D. Strauss, *Algebra in the Stone-Čech compactification: Theory and applications*, Walter de Gruyter, Berlin and New York, 1998. 148, 153, 159

[68] T. W. Hungerford, *Algebra*, Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1974. 23, 24, 87

[69] International Mathematical Union, *Fields Medal 2006 for Terence Tao*, press release, 2006, accessed 03.30.07, available at `http://www.mathunion.org/General/Prizes/2006/TaoENG.pdf` 74

[70] T. Jech, *Set theory*, Pure and Applied Mathematics **79**. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978. 137

[71] S. Jukna, *Extremal Combinatorics*, Springer-Verlag, Berlin, 2001. 54

[72] V. Jungić, J. Licht, M. Maholian, J. Nešetřil, and R. Radoičić, Rainbow arithmetic progressions and anti-Ramsey results, *Combin. Probab. Comput.* **12** (2003), 599–620. 168, 169, 170, 173

[73] V. Jungić, J. Nešetřil and R. Radoičić, Rainbow Ramsey Theory, *Integers* **5** (2005), A9, 13 pages (electronic). 168

[74] V. Jungić and R. Radoičić, Rainbow 3-term Arithmetic Progressions, *Integers* **3** (2003), A18, 8 pages (electronic). 168

[75] A. Y. Khinchin, *Three pearls of number theory*, Graylock Press, Rochester, N. Y., 1952. 34

[76] A. Khodkar and B. Landman, Recent progress in Ramsey theory on the integers, in *Combinatorial Number Theory: Dedicated to Ron Graham on his 70th Birthday*, Ed. Landman, Nathanson, Nešetřil, Nowakowski, Pomerance, de Gruyter, Berlin, 2007. 77

[77] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its Applications in Graph Theory, *Combinatorics, Paul Erdös is Eighty (Vol. 2) Keszthely (Hungary) 1993*, Bolyai Soc. Math. Stud., **2**, János Bolyai Math. Soc., Budapest 1996, 295–352. 7

[78] B. Kra, The Green-Tao Theorem on arithmetic progressions in the primes: an ergodic point of view, *Bull. Amer. Math. Soc. (N.S.)* **43** (2006), 3–23. 73

[79] B. Landman, personal communication, 2006. 179

[80] B. Landman and A. Robertson, *Ramsey theory on the integers.*, Student Mathematical Library, **24**. American Mathematical Society, Providence, RI, 2004. 27, 44, 90, 179

[81] L. Lovász, *Combinatorial problems and exercises*, Second edition, North-Holland Publishing Co., Amsterdam, 1993. 17, 18

[82] J. Matoušek and J. Spencer, Discrepancy in arithmetic progressions, *J. Amer. Math. Soc.* **9** (1996), 195–204. 122, 123

[83] L. Moser, Notes on number theory II: On a theorem of van der Waerden, *Can. Math. Bull.* **3** (1960), 23–25. 71

[84] A. Nilli, Shelah's proof of the Hales-Jewett theorem, in *Mathematics of Ramsey theory*, Algorithms Combin., **5**, Springer, Berlin, 1990, 150–151. 127, 129

[85] H. J. Prömel and B. Voigt, Aspects of Ramsey-Theory II: Arithmetic Progressions, preprint, 1989. 21, 26, 113

[86] J. R. Rabung, On applications of van der Waerden's theorem, *Math. Mag.* **48** (1975), 142–148. 31, 32, 178

[87] J. R. Rabung, Some progression-free partitions constructed using Folkman's method, *Canad. Math. Bull.* **22** (1979), 87–91. 78

[88] R. Rado, Studien zur Kombinatorik, *Math. Z.* **36** (1933), 425–480. 67

[89] R. Rado, Note on combinatorial analysis, *Proc. London Math. Soc.* **48** (1943), 122–160. 135

[90] R. Rado, Axiomatic treatment of rank in infinite sets, *Canad. J. Math.* **1** (1949), 337–343. 29

[91] F. P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286. 11, 12

[92] R. A. Rankin, Sets of integers containing not more than a given number of terms in arithmetical progression, *Proc. Roy. Soc. Edinburgh Sect. A* **65** (1960), 332–344. 72

[93] K. Roth, Sur quelques ensembles d'entiers, *C. R. Acad. Sci. Paris* **234** (1952), 388–390. 45, 48, 49, 59

[94] K. Roth, On Certain Sets of Integers, *J. London Math. Soc.* **28** (1953), 104–109. 48

[95] K. Roth, Remark concerning integer sequences, *Acta Arith.* **9** (1964), 257–260. 113, 123

[96] K. Roth, Irregularities of sequences relative to arithmetic progressions III, *J. Number Theory* **2** (1970), 125–142. 58

[97] K. Roth, Irregularities of sequences relative to arithmetic progressions IV, *Period. Math. Hungar.* **2** (1972), 301–326. 58

[98] R. Salem and D. C. Spencer, On sets of integers which contain no three terms in arithmetical progression, *Proc. Nat. Acad. Sci. U. S. A.* **28** (1942), 561–563. 68

[99] W. M. Schmidt, Two combinatorial theorems on arithmetic progressions, *Duke Math. J.* **29** (1962), 129–140. 98

[100] W. M. Schmidt, Ein kombinatorisches Problem von P. Erdős und A. Hajnal, *Math. Acad. Sci. Hungar* **15** (1964), 373–374. 97

[101] I. Schur, Über die Kongruenz $x^m + y^m \equiv z^m \pmod p$, *Jahresber. Deutsch. Math.-Verein* **25** (1916), 114–116. 21, 24

[102] S. Shelah, Primitive recursive bounds for van der Waerden numbers, *J. Amer. Math. Soc.* **1** (1988), 683–697. 99, 100, 127

[103] J. H. Spencer, A remark on coloring integers, *Canad. Math. Bull.* **15** (1972), 43–44. 118

[104] J. H. Spencer, *personal communication*, (2006). 96

[105] R. Stevens and R. Shantaram, Computer-generated van der Waerden partitions, *Math. Comp.* **32** (1978), 635–636. 77, 78

[106] Z. Szabó, An application of Lovász' Local Lemma — A new lower bound for the van der Waerden number, *Random Structures Algorithms* **1** (1990), 343–360. 86

[107] E. Szemerédi, On sets of integers containing no 4 elements in arithmetic progression, *Acta Math. Acad. Sci. Hungar.* **20** (1969), 89–104. 18, 45, 58, 59

[108] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245. 45, 47, 58, 59

[109] T. Tao, Arithmetic Ramsey Theory, note available online `http://www.math.ucla.edu/~tao/preprints/Expository/ramsey.dvi`, accessed 12.14.06. 102

[110] T. Tao and T. Ziegler, The primes contain arbitrarily long polynomial progressions, preprint available online `http://lanl.arxiv.org/PS_cache/math/pdf/0610/0610050.pdf`, accessed 03.19.07. 168

[111] B. Valkó, Discrepancy of arithmetic progressions in higher dimmensions, *J. Number Theory* **92** (2002), 117–130. 123

[112] J. G. van der Corput, Über Summen von Primzahlen und Primzahlquadraten, *Math. Ann.* **116** (1939), 1–50. 72

[113] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw. Arch. Wiskd.*, **15** (1927), 212–216. 26, 34

[114] B. L. van der Waerden, How the proof of Baudet's Conjecture was found, in *Studies in Pure Mathematics*, ed. L. Mirsky, Academic Press, London, 1971, 251–260. 26, 27, 30

[115] P. Varnavides, Note on a theorem of Roth, *J. London Math. Soc.* **30** (1955), 325–326. 63

[116] P. Varnavides, On certain sets of positive density, *J. London Math. Soc.* **34** (1959), 358–360. 63

[117] 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. 166

[118] S. Willard, *General topology*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970. 142

[119] E. Witt, Ein Kombinatorischer Satz der Elementargeometrie, *Math. Nachr.* **6** (1951), 261–262. 135

[120] P. Zeitz, Semigroups, dynamics and combinatorics, notes based on lectures by H. Furstenberg, 1990. 148

Index