Ramsey Theory

Gregory E. W. Taylor
326639

March 2006

I warrant that the content of this dissertation is the direct result of my own work and that any use made in it of published or unpublished material is fully and correctly referenced.

Signed ....

Date ....

Contents

1 Introduction

In 1928 the English mathematician Frank Plumpton Ramsey published his paper *On a problem of formal logic* [13] in which he proved what would become known as Ramsey's Theorem. The paper has led to a large area of combinatorics now known as Ramsey Theory. We shall explore some major results in Ramsey Theory which all, broadly speaking, find some degree of order within a large disordered set.

The field of Ramsey Theory has only relatively recently come together to be viewed as one body and many seemingly basic results are still not known, and there is no prospect of them being known in the near future.

In 1916 Issai Schur proved that in any finite colouring of the natural numbers there must exist three monochromatic elements, x, y and z such that $x + y = z$. This basic result was generalised by Richard Rado in 1933 to give a characterisation of the homogeneous systems to which a monochromatic solution can be found in any finite colouring of the natural numbers. We examine these results in Chapter 4.

Between Schur proving this theorem in 1916 and Rado publishing his theorem in 1933, Ramsey and Van der Waerden published theorems now considered central to Ramsey Theory.

We shall begin by examining Ramsey's Theorem, initially for graphs, and then, more generally, for sets. For example Ramsey's theorem for graphs states that in any large enough finitely coloured complete graph there must exist some large monochromatic substructure. Little is known about the actual orders that these complete graphs must have to ensure that they contain some particular monochromatic substructure.

Van der Waerden's Theorem was proved in 1927, a year earlier than Ramsey's. Van der Waerden proved that in any finite colouring of the natural numbers there must exist some monochromatic arithmetic progression with k terms. We shall give an elegant and short proof based on colour focussing.

Finally we shall turn to Hindman's Theorem, the most recent theorem which we shall examine, it was proved in 1974. Hindman's Theorem states that, for every finite colouring of the natural numbers there exists some infinite subset $S \subseteq \mathbb{N}$ such that all the finite sums of the elements of S are monochromatic. Although it is not hard to understand this theorem, it requires some powerful mathematical tools in its proof. In Chapter 5 we shall build up these ideas before finally proving the result.

2 Ramsey's Theorem

A result relating to many problems in Ramsey Theory is the Pigeonhole principle, we introduce it here.

2.1 The Pigeonhole principle

The pigeonhole principle, also known as the Dirichlet pigeonhole principle, simply states that if there exists n pigeonholes containing $n+1$ pigeons, one of the pigeonholes must contain at least two pigeons. This can be generalised to say that if there are a finite number of pigeonholes containing an infinite number of pigeons at least one of the pigeonholes must contain an infinite number of pigeons.

2.2 Small Ramsey numbers

To understand Ramsey numbers and Ramsey's Theorem we must first understand what is meant by a *coloured graph*.

Definition 2.2.1. *A 2-coloured graph is a graph whose edges have been coloured with 2 different colours.*

Example. Three ways in which [1]K_4 could be 2-coloured are given in Figures 1, 2 and 3.

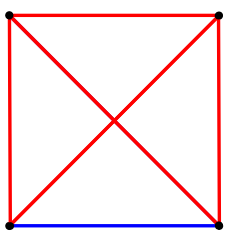


Fig. 1

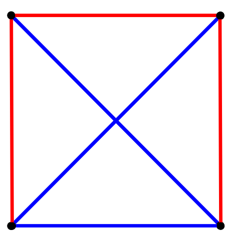


Fig. 2

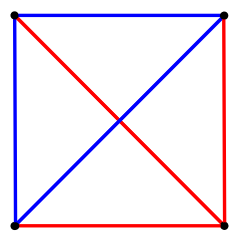


Fig. 3

Ramsey's Theorem assets that there exists a number $R(s)$ such that that any complete 2-coloured graph of order $n \geq R(s)$ must contain a complete monochromatic subgraph of order s. That is, in any 2-colouring of K_n with the colours red and blue there must exist either a red or a blue K_s. Equivalently, every graph of order $n \geq R(s)$ has either a complete or empty subgraph of order s. These two statements are equivalent because, any graph, G, of order n gives rise to a 2-colouring of K_n, since we may colour G with one colour and its complement the other colour. Colouring a graph is simply a convenient way of splitting it's edges into separate subgraphs.

Definition 2.2.2. *The Ramsey number, $R(s,t)$, is the order of the smallest complete graph which, when 2-coloured, must contain a red K_s or a blue K_t.*

[1]K_x denotes the complete graph of order x.

$R(s,t) = R(t,s)$ since the colour of each edge can be swapped. Two simple results are $R(s,1) = 1$ and $R(s,2) = s$. $R(s,1) = 1$ is trivial since K_1 has no edges and so no edges to colour, thus any colouring of K_1 will always contain a blue K_1. $R(s,2) = s$ is also a simple result; if all the edges of K_s are coloured red, it will contain a red K_s, however if one edge is coloured blue it will contain a blue K_2. The edges of any graph of order less that s could all be coloured red in which case the graph would contain neither a red K_s or a blue K_2.

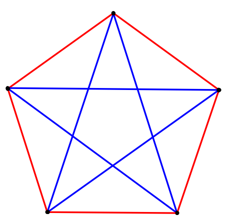
The values of these Ramsey numbers are, perhaps surprisingly, very difficult to determine and only a small number of them are known, for example $R(5,5)$ is still unknown. The known non-trivial Ramsey numbers for two colours are listed in the table below.

R(s,t)	3	4	5	6	7	8	9
3	6	9	14	18	23	28	36
4	9	18	25				
5	14	25					
6	18						
7	23						
8	28						
9	36						

Given below are two examples which illustrate the methods by which Ramsey numbers may be found.

Example. $R(3,3) = 6$.

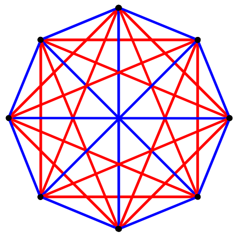
We see first that $R(3,3) > 5$ from the colouring of K_5 below. This colouring shows K_5 may be 2-coloured such that it does not contain a red or blue K_3 as a subgraph.



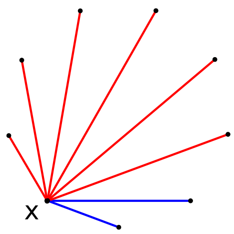
It is then simple to see that $R(3,3) \leq 6$ and so $R(3,3) = 6$. Indeed, in any colouring of K_6 each vertex must be incident to at least three red or three blue edges by the pigeonhole principle. We take a vertex, say x, which is incident to at least three red edges. These edges are clearly incident to three other vertices. If every edge between these three vertices is blue then we have a blue K_3 and so we assume that at least one of these edges is red. This red edge, together with the two edges incident to x will form a red K_3. If there does not exist a vertex x which is incident to three red edges then every vertex must be incident to at least three blue edges, causing a monochromatic K_3 to arise in a similar way.

Example. $R(4,3) = 9$.

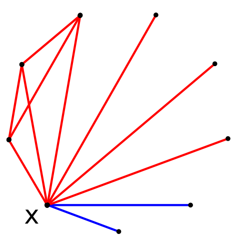
We see first that $R(4,3) > 8$ from the colouring of K_8 below. This colouring shows K_8 may be 2-coloured such that it does not contain a red K_4 or a blue K_3 as a subgraph.



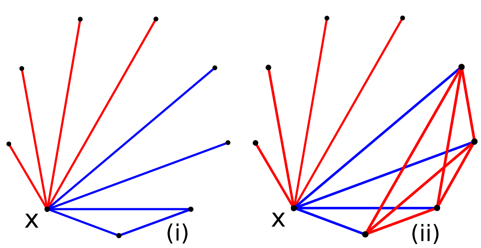
To show that $R(4,3) \leq 9$, we consider any 2-colouring of K_9. In any graph the number of vertices with odd degree must be even. For this reason there cannot exist a red 5-regular subgraph of K_9 or a blue 3-regular subgraph of K_9. This implies that in a complete 2-coloured graph of order nine there must be at least one vertex which is incident to at least six red or at least four blue edges.



We already have that $R(3,3) = 6$ so taking a vertex, say x, which is incident to six red edges, the six vertices connected to these red edges must induce a red or blue K_3. We are done if we have a blue K_3 so we assume that we have a red K_3, this red K_3 together with the edges connecting K_3 to x must induce a red K_4, and we are done.



We now turn to the case where x is incident to four blue edges. Between two of the vertices connected to x by blue edges there must exist either a blue edge or all the edges must be red. If there exists a blue edge we have, together with the edges incident to x, a blue K_3, we call this case (i). Otherwise all four vertices are connected by red edges and we have a red K_4, we call this case (ii).



In either case there must exist a red K_4 or a blue K_3 and so in any 2-coloured complete K_9 there must exist either a red K_4 or a blue K_3 as a subgraph.

2.3 Ramsey's Theorem for coloured graphs

Theorem 2.3.1. *For any two natural numbers, s and t, there exists a natural number, $R(s,t) = n$, such that any 2-coloured complete graph of order at least n, coloured red and blue, must contain a monochromatic red K_s or blue K_t.*

Proof. We prove that $R(s,t)$ exists by proving it is bounded. We shall use proof by induction first assuming that $R(s-1,t)$ and $R(s,t-1)$ exist. As was shown earlier $R(s,2) = R(2,s) = s$ and $R(s,1) = R(1,s) = 1$ are trivial results.
Claim. $R(s,t) \leq R(s-1,t) + R(s,t-1)$.

We first take a 2-colouring of a complete graph with $n = R(s-1,t)+R(s,t-1)$ vertices. We now pick one of the vertices in K_n, say x. We then produce two sets, R_x and B_x, R_x is the set of vertices adjacent to x such that every edge connecting a vertex in R_x to x is red. Similarly B_x is the set of vertices adjacent to x such that every edge connecting a vertex in B_x to x is blue.

Since K_n is a complete graph $B_x = [n] \backslash (R_x \cup \{x\})$ and so $|R_x|+|B_x| = n-1$. If $|R_x| < R(s-1,t)$ and $|B_x| < R(s-1,t)$ then since $n = R(s-1,t)+R(s,t-1)$ we must have $|R_x| + |B_x| \leq n-2$, a contradiction. So $|B_x| \geq R(s,t-1)$ or $|R_x| \geq R(s-1,t)$.

If $|B_x| \geq R(s,t-1)$ and B_x induces a red K_s we are done. If B_x induces a blue K_{t-1} then K_n must contain a blue K_t since $B_x \cup \{x\}$ must induce a blue K_t. Indeed, each edge xt is blue for all $t \in B_x$, from the definition of B_x. So $B_x \cup \{x\}$ must induce a blue K_t if B_x contains a blue K_{t-1}. The case for R_x is completely symmetric, that is, if R_x induces a blue K_t we are done and if R_x induces a red K_{s-1} then K_n must contain a red K_s since $R_x \cup \{x\}$ must induce a red K_s.

We have shown that a 2-coloured complete graph of order $R(s-1,t) + R(s,t-1)$ must contain a red K_s or a blue K_t, proving that $R(s,t) \leq R(s-1,t) + R(s,t-1)$. This completes our induction. □

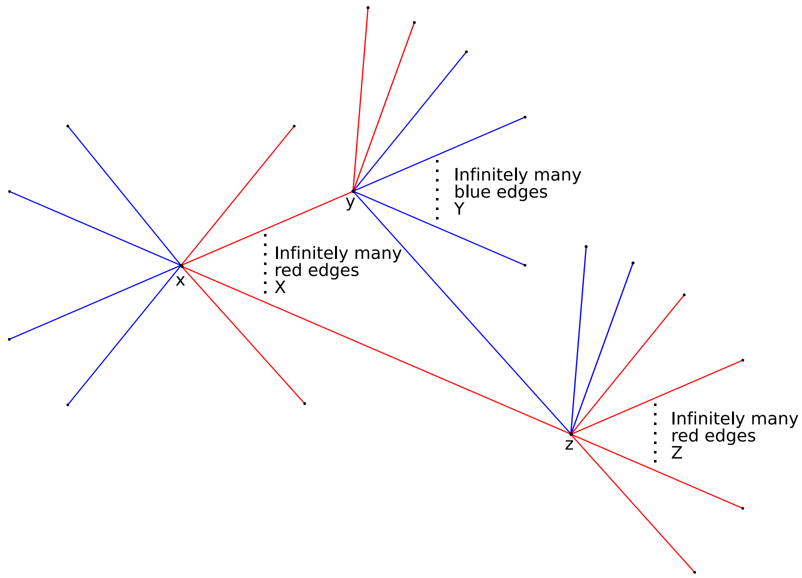
We now prove the infinite case of Ramsey's Theorem for two colours.

Definition 2.3.2. *$K_{\mathbb{N}}$ is the complete graph whose vertex set is countably infinite.*

Theorem 2.3.3. *Every 2-coloured $K_{\mathbb{N}}$ must contain a countably infinite monochromatic complete graph.*

Proof. Fix a 2-colouring of the edges of the complete graph, $K_{\mathbb{N}}$. We label each vertex with an element from $\mathbb{N}$ and take the vertex, x, which we have labeled 1, we now consider all the edges incident with x. Since the graph is infinite, using the pigeonhole principle, there must be an infinite set of red (or blue) edges incident with x. Define X to be the infinite set of vertices connected to x by a red (or blue) edge. Now consider a vertex within X, say $y > 1$. Again because the set X is infinite there must be an infinite set of blue (or red) edges incident with y and some vertex in X. Define $Y \subset X$ to be the infinite set of vertices which are connected to y by a blue (or red) edge. Now consider a vertex within Y say z, where $z > y$. Again because the set Y is infinite there must be infinite number of red (or blue) edges connecting z to vertices in Y.

Define $Z \subset Y$ to be the infinite set of vertices which are connected to z by a red (or blue) edge.



We can continue picking successive vertices indefinitely since our graph is infinite, this will result in a set of vertices $V = \{x, y, z, \dots\} \subseteq K_{\mathbb{N}}$. We define E to be the set of edges connecting the vertices in V, so E is $\{xy, xz, \dots, yz, \dots\}$. From this definition of the set E it is clear that the colour of any edge in E is determined by the smaller of its end vertices. That is, if we assume that the colour of each edge in the set of edges $\{x\bar{x} \,|\, \bar{x} \in X\}$, is red, each edge in the set of edges $\{y\bar{y} \,|\, \bar{y} \in Y\}$, is blue and each edge in the set of edges $\{z\bar{z} \,|\, \bar{z} \in Z\}$, is red. Then any edge xv for $v \in V$ must be red, any edge yv for $v \in V \setminus \{x\}$ must be blue and any edge zv for $v \in V \setminus \{x, y\}$ must be red. We can now produce a 2-colouring of V, we colour any vertex in V, say p, red if every edge in $\{p\bar{p} \,|\, \bar{p} \in P\}$, is coloured red, where P is defined in the same way that X, Y and Z were. Similarly, we colour any vertex in V, say q, blue if every edge in $\{q\bar{q} \,|\, \bar{q} \in Q\}$, is coloured blue, where Q is defined in the same way that X, Y and Z were. In our case the colouring of V is $\{x, y, z, \dots\}$. Since V consists of infinitely many vertices, coloured with only two colours, the pigeonhole principle allows us to conclude there must be an infinite monochromatic set within V, we call this set M. This infinite set of monochromatic vertices induces an infinite monochromatic complete subgraph of $K_{\mathbb{N}}$. Each vertex in M is adjacent to every other vertex in M. Every vertex in M is the same colour, so every edge in the graph induced by M must have the same colour. Thus the graph induced by the vertex set M is a countably infinite monochromatic complete graph. □

Definition 2.3.4. *A graph is r-coloured if we colour each edge of the graph with one of r colours.*

Definition 2.3.5. *The Ramsey Number, $R_r(s)$, is the order of the smallest complete graph which, when r-coloured, must contain a monochromatic K_s.*

$R_r(s)$ can also be written $R(\overbrace{s, s, \ldots, s}^{\text{r times}})$. Generally, as above, we write $R(s)$ for $R_2(s)$ but we could also be write $R(s, s)$. Any complete r-coloured graph of order $n \geq R_r(s)$ must contain a complete monochromatic subgraph of order s. We call Ramsey numbers of the form $R(s, s, \ldots, s)$ diagonal Ramsey numbers. Other Ramsey numbers are of the form $R(s, t, \ldots, n)$.

We may now deduce Ramsey's Theorem for a finite number of colours directly from Theorem 2.3.3.

Theorem 2.3.6. *Every r-coloured $K_{\mathbb{N}}$ must contain a countably infinite monochromatic complete graph, where $1 \leq r < \infty$.*

Proof. Suppose that $K_{\mathbb{N}}$ is coloured with r colours, say $k_1, k_2, \ldots, k_r$. We may produce a graph $K^1_{\mathbb{N}}$ by changing the colouring of $K_{\mathbb{N}}$. Each edge of $K^1_{\mathbb{N}}$ that was coloured with k_1 in $K_{\mathbb{N}}$ is coloured with l_1 in $K^1_{\mathbb{N}}$. Each edge that was coloured with one of $k_2, k_3, \ldots, k_{r-1}$ or k_r in $K_{\mathbb{N}}$ is coloured with l_2 in $K^1_{\mathbb{N}}$.

From Theorem 2.3.3, since $K^1_{\mathbb{N}}$ is a complete countably infinite graph coloured using only two colours, l_1 and l_2, $K^1_{\mathbb{N}}$ must contain a countably infinite monochromatic complete graph coloured with either l_1 or l_2. If $K^1_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph coloured with l_1 we are done since $K_{\mathbb{N}}$ must then also contain this countably infinite monochromatic complete graph. If $K^1_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph coloured with l_2, then we produce the graph $K^2_{\mathbb{N}}$.

We define $K^2_{\mathbb{N}}$ to be the countably infinite monochromatic complete graph coloured with l_2 in $K^1_{\mathbb{N}}$. Each edge of $K^2_{\mathbb{N}}$ that was coloured with k_2 in $K_{\mathbb{N}}$ is coloured with m_1 in $K^2_{\mathbb{N}}$. Each edge of $K^2_{\mathbb{N}}$ that was coloured with any of $k_3, k_4, \ldots, k_{r-1}$ or k_r in $K_{\mathbb{N}}$ is coloured with m_2 in $K^2_{\mathbb{N}}$. From Theorem 2.3.3, since $K^2_{\mathbb{N}}$ is a countably infinite complete graph coloured using only two colours, m_1 and m_2, $K^2_{\mathbb{N}}$ must contain a countably infinite monochromatic complete graph coloured with either m_1 or m_2. If $K^2_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph coloured with m_1 we are done since $K_{\mathbb{N}}$ must then also contain this countably infinite monochromatic complete graph. If $K^2_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph coloured with m_2, then we produce the graph $K^3_{\mathbb{N}}$. We may define $K^3_{\mathbb{N}}$ in exactly the same way we defined $K^2_{\mathbb{N}}$ using $K^1_{\mathbb{N}}$.

We may continue in this way, however, since $K_{\mathbb{N}}$ is only coloured with a finite number of colours at some step we must find either a countably infinite monochromatic complete graph coloured with one of $k_1, k_2, k_3, \ldots k_{r-2}$ or we shall define $K^{r-1}_{\mathbb{N}}$. $K^{r-1}_{\mathbb{N}}$ must be a complete infinite graph, coloured using only k_{r-1} and k_r. Again from Theorem 2.3.3 $K^{r-1}_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph coloured with either k_{r-1} or k_r. Since both the sets of edges coloured with k_{r-1} and k_r in $K^{r-1}_{\mathbb{N}}$ are also in $K_{\mathbb{N}}$ we must have that $K_{\mathbb{N}}$ contains some countably infinite monochromatic complete graph.

□

The second of the results, proved as a direct consequence of Theorem 2.3.3 is Theorem 2.3.1, we give a second proof here. We must first give a definition.

Definition 2.3.7. *If we label the vertices of K_n with the natural numbers, $1, 2, 3, \ldots, n$, then we may restrict a colouring of K_n to a colouring of K_m where $m \leq n$. We restrict the colouring by only colouring the complete graph on the first m vertices in K_n. In this restricted colouring K_m is coloured in exactly the same way it was coloured in K_n.*

Second proof of Theorem 2.3.1. This is a proof by contradiction. We first assume we can find a 2-colouring of K_n which does not contain a red K_s or a blue K_t for every $n \in \mathbb{N}$. Let C_n be such a 2-colouring of K_n.

We first note that for every $n \in \mathbb{N}$ there are a finite number, $2^{\binom{n}{2}}$, ways of 2-colouring K_n. We take a subsequence of the colourings $C_2, C_3, C_4, \ldots$, consisting only of the colourings which when restricted to K_2 colour it's edge in exactly the same way. K_2 can only be coloured in two different ways with two colours so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i for $i \in \mathbb{N}$ and $i \geq 2$, which colour K_2 in the same way. We call this subsequence of colourings $\mathcal{C}_2$ and define I_2 such that $C_i \in \mathcal{C}_2$ only if $i \in I_2$. We can then find a subset $\mathcal{C}_3 \subseteq \mathcal{C}_2$, consisting only of the colourings which when restricted to K_3 colour all edges in exactly the same way. K_3 can only be coloured in eight different ways with two colours so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i for $i \in I_2$, which colour K_3 in the same way. We call this subset of colourings $\mathcal{C}_3$ and define I_3 such that $C_i \in \mathcal{C}_3$ only if $i \in I_3$. We note that $I_3 \subseteq I_2$. We may continue in this way indefinitely.

We now produce a 2-colouring, C, of $K_{\mathbb{N}}$. We define C, as follows. For every $R \in \mathbb{N}$ with $2 \leq R \leq x$ we set $C(K_R) = C_i(K_R)$ where $C_i \in \mathcal{C}_x$. This colouring is well defined, if $R \leq x \leq y$ then defining the colouring $C(K_R)$ using $\mathcal{C}_x$ or $\mathcal{C}_y$ will give equal results since $C(K_R) = C_i(K_R) = C_j(K_R)$ for $C_i \in \mathcal{C}_x$ and $C_j \in \mathcal{C}_y$. This is clear since the subset of colourings $\mathcal{C}_y$ consists of the colourings C_i such that $i \in I_y \subseteq I_x \subseteq I_R$, so they all colour K_R in the same way.

We now assume that there exists some natural number, n, such that under the colouring C the graph K_n contains some red K_s or blue K_t. We can see from the definition of C that $C(K_n) = C_i(K_n)$ for some $C_i \in \mathcal{C}_\nu$ where $n \leq \nu$. However, by definition any colouring in $\mathcal{C}_\nu$ colours any complete graph of order n so that is does not contain a red K_s or a blue K_t. So we have a contradiction. Thus under the 2-colouring C there does not exist a natural number, n, such that $C(K_n)$ contains some red K_s or blue K_t. This is also a contradiction, from Theorem 2.3.3, every 2-colouring of $K_{\mathbb{N}}$ contains a countably infinite monochromatic complete graph. C is a 2-colouring of $K_{\mathbb{N}}$ which does not contain a red K_s or blue K_t so certainly does not contain a countably infinite monochromatic complete graph. So there must exist some $n \in \mathbb{N}$ such that in any 2-colouring of K_n there exists a red K_s or a blue K_t. $\square$

2.4 Ramsey's Theorem for sets

Definition 2.4.1. *For some set, A, and natural number, k, the subsets of A of size k are called k-tuples. The set of all k-tuples in A is $A^{(k)}$.*

Example. The 4-tuples of the set $A = \{1, 2, 3, 4, 5\}$ are the sets $\{1, 2, 3, 4\}$,

$\{1,2,3,5\}$, $\{1,2,4,5\}$, $\{1,3,4,5\}$ and $\{2,3,4,5\}$. The set $A^{(4)}$ is $\{\{1,2,3,4\}, \{1,2,3,5\}, \{1,2,4,5\}, \{1,3,4,5\}, \{2,3,4,5\}\}$. A 3-colouring of $A^{(4)}$ is $\{\{1,2,3,4\}, \{1,2,3,5\}, \{1,2,4,5\}, \{1,3,4,5\}, \{2,3,4,5\}\}$.

Theorem 2.4.2. *Let χ be an r-colouring of $A^{(k)}$ where $1 \leq k < \infty$ and A is a countably infinite set. Then A contains a monochromatic infinite set, M, that is, $M^{(k)}$ is monochromatic.*

Proof. This proof is one of induction on k. The result for $k = 1$ is trivial, since the set is split into sets of size one, that is, we colour each element with one of the r colours. Since these elements are coloured with a finite number of colours and the set is infinite, by the pigeonhole principle, there must be a monochromatic infinite set of these elements.

We assume that the theorem is true for subsets of size less than k. In particular, for every r-colouring of $A^{(q)}$ where A is a countably infinite set and q is a natural number such that $q < k$, A contains a monochromatic infinite set. We first fix an r-colouring, χ, of $A^{(k)}$. We now set $A_0 = A$ and choose an element $a_0 \in A_0$. Let $B_1 = A_0 \setminus \{a_0\}$. We then define an r-colouring, χ_1, of the $(k-1)$-tuples of B_1. The colouring of each $(k-1)$-tuple, τ, in B_1 is defined by $\chi_1(\tau) = \chi(\tau \cup \{a_0\})$. From our induction hypothesis, B_1 must contain an infinite set, A_1, all of whose $(k-1)$-tuples are monochromatic. We now take $a_1 \in A_1$ and define $B_2 = A_1 \setminus \{a_1\}$ and an r-colouring, χ_2, of the $(k-1)$-tuples of B_2. The colouring of each $(k-1)$-tuple, τ, in B_2 is defined by $\chi_2(\tau) = \chi(\tau \cup \{a_1\})$. Again by the induction hypothesis, B_2 must contain an infinite set A_2 all of whose $(k-1)$-tuples are monochromatic. This argument can be continued indefinitely to obtain an infinite sequence of r-coloured elements $\{a_0, a_1, a_2, \ldots\}$. Each element, a_i, of this sequence is given the same colour as the infinite set of monochromatic $(k-1)$-tuples, A_{i+1}. An infinite sequence of nested sets $A_0 \supset A_1 \supset A_2 \supset \cdots$ with $a_n \in A_n$ is also produced since A_i is an infinite set in $A_{i-1} \setminus \{a_{i-1}\}$ and therefore contained in A_{i-1}. All the k-tuples in A whose only element outside A_i is a_{i-1} must have the same colour because of the way χ_i coloured the $(k-1)$-tuples. Any k-tuple in $\{a_0, a_1, a_2, \ldots\}$ which contains a_i and $k-1$ elements from $\{a_{i+1}, a_{i+2}, \ldots\}$ must be red if a_i is red. From the pigeonhole principle there must be infinitely many monochromatic elements in $\{a_0, a_1, a_2, \ldots\}$. Each k-tuple of this monochromatic infinite set must be monochromatic. That is, if in $\{a_0, a_1, a_2, \ldots\}$ there are an infinite number of red elements then our infinite monochromatic set, M, would consist of every red element a_i. $\square$

From Theorem 2.4.2, Ramsey's Theorem for countably infinite sets sets, we may deduce Theorem 2.4.3, Ramsey's Theorem for finite sets. We will only prove this theorem for two colours, however the proof can be simply generalised for any finite colouring.

Theorem 2.4.3. *Given natural numbers, s and k, there exists some natural number n such that for any 2-colouring of $[n]^{(k)}$ there is a monochromatic set, S, of size s, such that $S \subseteq [n]$.*

Proof. This is a proof by contradiction. We first assume we can find a 2-colouring of $[n]^{(k)}$ which does not contain a monochromatic set of size s for every $n \in \mathbb{N}$. Let C_n be such a 2-colouring.

We first note that for every $n \in \mathbb{N}$ there are a finite number, $2^{\binom{n}{k}}$, ways of 2-colouring $[n]^{(k)}$. We take a subsequence of the colourings, $C_k, C_{k+1}, C_{k+2}, \ldots$, consisting only of the colourings which when restricted to $[k]^{(k)}$ colour it in exactly the same way. There are only $2^{\binom{k}{k}} = 2$ possible colourings of $[k]^{(k)}$ so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i, for $i \in \mathbb{N}$ and $i \geq k$, which colour $[k]^{(k)}$ in the same way. We call this subset of colourings $\mathcal{C}_k$ and define I_k such that $C_i \in \mathcal{C}_k$ only if $i \in I_k$. We can then find a subset $\mathcal{C}_{k+1} \subseteq \mathcal{C}_k$, consisting only of the colourings which when restricted to $[k+1]^{(k)}$ colour it in exactly the same way. $[k+1]^{(k)}$ can only be coloured in a finite number of ways with two colours so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i for $i \in I_k$, which colour $[k+1]^{(k)}$ in the same way. We call this subset of colourings $\mathcal{C}_{k+1}$ and define I_{k+1} such that $C_i \in \mathcal{C}_{k+1}$ only if $i \in I_{k+1}$. We note that $I_{k+1} \subseteq I_k$. We can continue to produce subsets of the colourings in this way. A subset of colourings, $\mathcal{C}_x$, each of which colour $[x]^{(k)}$ in exactly the same way can be restricted to a subset of colourings, $\mathcal{C}_y$, each of which colour $[y]^{(k)}$ in exactly the same way, whenever $y \geq x$.

We now produce a 2-colouring, C, of $\mathbb{N}^{(k)}$. We define the 2-colouring, C, as follows. For every $R \in \mathbb{N}$ with $k \leq R \leq x$ we set $C([R]^{(k)}) = C_i([R]^{(k)})$ where $C_i \in \mathcal{C}_x$. This colouring is well defined, since if $[R] \subseteq [x] \subseteq [y]$ then defining the colouring $C([R]^{(k)})$ using $\mathcal{C}_x$ or $\mathcal{C}_y$ will give equal results since $C([R]^{(k)}) = C_i([R]^{(k)}) = C_j([R]^{(k)})$ for $C_i \in \mathcal{C}_x$ and $C_j \in \mathcal{C}_y$. This is clear since the subset of colourings $\mathcal{C}_y$ is just the set of colourings C_i such that $i \in I_y \subseteq I_x \subseteq I_R$, so they all colour $[R]^{(k)}$ in the same way.

We now assume that there exists some natural number, n, such that $C([n]^{(k)})$ contains a monochromatic set of size s. We can see from the definition of C that $C([n]^{(k)}) = C_i([n]^{(k)})$ for some $C_i \in \mathcal{C}_\nu$, where $\nu \geq n$. However, by definition any colouring in $\mathcal{C}_\nu$ colours $[n]^{(k)}$ so that it does not contain a monochromatic set of size s. So we have a contradiction. Thus under the colouring C there does not exist a natural number, n, such that $C([n]^{(k)})$ contains a monochromatic set of size s. This is is also a contradiction, from Theorem 2.4.2, every 2-colouring of $\mathbb{N}^{(k)}$ contains an infinite monochromatic set. C is a 2-colouring of $\mathbb{N}^{(k)}$ which does not contain a monochromatic set of size s, so certainly does not contain an infinite monochromatic set. Therefore, we may conclude there must exist some natural number, n, such that for any 2-colouring of $[n]^{(k)}$ there is a monochromatic set, S, of size s, such that $S \subseteq [n]$. $\square$

3 Van der Waerden's Theorem

In 1927, a year earlier than Ramsey published his theorem, the Dutch mathematician Bartel Leendert van der Waerden published his paper *Beweis einer Baudetschen Vermutung* [14] in which he proved what would become known as Van der Waerden's Theorem.

Van der Waerden's Theorem states that *for all positive integers, k and r, there exists a natural number $W(k,r)$ such that, if the set of natural numbers $\{1, 2, \ldots, W(k,r)\}$ is r-coloured, then it must contain at least one monochromatic k-term arithmetic progression.*

3.1 Small Van der Waerden numbers

Van der Waerden numbers, similarly to Ramsey numbers, are not extensively known. The known non-trivial Van der Waerden numbers are are listed in the table below.

$W(k,r)$	r=2	3	4
k=3	9	27	76
4	35		
5	178		

Trivially $W(1,r) = 1$, $W(2,r) = r+1$ and $W(k,1) = k$.

Example. $W(3,2) = 9$.

To find $W(3,2)$ we find the set of natural numbers $\{1, 2, \ldots, W(3,2)\}$ which when 2-coloured must contain some monochromatic 3-term arithmetic progression. We can see $W(3,2) > 8$ from the 2-coloured set $\{1, 2, 3, 4, 5, 6, 7, 8\}$ which does not contain a monochromatic 3-term arithmetic progression.

The proof that $W(3,2) \leq 9$ is simple since the 2-colourings of $\{1, \ldots, 9\}$ can be reduced to a small number of cases. We may 2-colour $\{1, 2, 3, 4\}$ in sixteen ways, half of which colour 1 red and half of which colour 1 blue. We may simply consider the eight cases where 1 is coloured blue, the other eight cases are completely symmetric. We may also ignore the colourings of $\{1, 2, 3, 4\}$ which contain a monochromatic 3-term arithmetic progression, so we are left with only five cases. We take the first case, $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, since 3 and 4 are red, colouring 5 red would form a monochromatic 3-term arithmetic progression, therefore we colour 5 blue. We must then colour 8 red since otherwise a blue 3-term arithmetic progression would be formed by 2,5 and 8. To avoid the red 3-term arithmetic progression in the terms 4,6 and 8 we colour 6 blue, however, this then forces 7 to be coloured red to avoid producing a blue 3-term arithmetic progression in the terms 5,6 and 7. For these reasons this colouring of $\{1, 2, 3, 4\}$ forces the colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ where 9 cannot be coloured without producing a monochromatic 3-term arithmetic progression. Similarly
$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ forces the colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$,
$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ forces the colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$,
$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ forces the colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and
$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ forces the colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. These colourings show that it is impossible to 2-colour $\{1, 2, \ldots, 9\}$ without producing a mo-

nochromatic 3-term arithmetic progression. We have shown $8 < W(3,2) \leq 9$, so $W(3,2) = 9$.

For larger Van der Waerden numbers the numerical method used above becomes increasingly protracted. As with Ramsey numbers, while upper and lower bounds for many Van der Waerden numbers have been found, their exact values are still unknown and without a huge amount of computing or a mathematical breakthrough will remain so. An example of how an upper bound for one Van der Waerden number is found is given below. In Section 3.2 we will generalise the method used in Proposition 3.1.1 to prove Van der Waerden's Theorem.

Proposition 3.1.1. $W(3,2) \leq 325$

Proof. The set of natural numbers, $\{1, 2, \ldots, 325\}$, are first divided into 65 blocks, each of order five. That is, $\{1, 2, \ldots, 325\} = \{1, \ldots, 5\} \cup \{6, \ldots, 10\} \cup \ldots \cup \{321, \ldots, 325\}$. We then label these block $B_1, B_2, \ldots, B_{65}$ where $B_1 = \{1, \ldots, 5\}, B_2 = \{6, \ldots, 10\}, \ldots, B_{65} = \{321, \ldots, 325\}$. Each of these blocks could be coloured in any one of $2^5 = 32$ ways. We now label each element of $\{1, 2, \ldots, 325\}$, $b_{x,y}$, where x denotes the index of the block which the element is in and y denotes its position within B_x, for example 6 is labeled $b_{2,1}$. Since there are only 32 ways in which any of the blocks may be 2-coloured, by the pigeonhole principle, at least two of the first 33 blocks must be coloured in the same way. We call these blocks B_a and B_{a+d}. Since each block is coloured with only two colours, again by the pigeonhole principle, we can say that at least two of the first three elements in each block must be monochromatic. We call the two monochromatic elements in the first three elements of B_a, $b_{a,\alpha}$ and $b_{a,\alpha+\delta}$. Since both these elements are in the first three terms of B_a the difference between $b_{a,\alpha}$ and $b_{a,\alpha+\delta}$ can only be one or two, so $\delta = 1$ or $\delta = 2$. There are five elements in each block so the third term in the arithmetic progression, $b_{a,\alpha}$, $b_{a,\alpha+\delta}$, must be an element of B_a, namely $b_{a,\alpha+2\delta}$. If $b_{a,\alpha+2\delta}$ has the same colour as $b_{a,\alpha}$ and $b_{a,\alpha+\delta}$ then we have a monochromatic 3-term arithmetic progression. Therefore we assume $b_{a,\alpha+2\delta}$ is not of the same colour as $b_{a,\alpha}$, and turn to the block B_{a+2d}. $b_{a,\alpha}$ and $b_{a+d,\alpha+\delta}$ are monochromatic, since B_a and B_{a+d} are equally coloured. Therefore if $b_{a+2d,\alpha+2\delta}$ is of the same colour as $b_{a,\alpha}$ and $b_{a+d,\alpha+\delta}$ we have a monochromatic 3-term arithmetic progression and we are done. Therefore, we assume $b_{a+2d,\alpha+2\delta}$ is a different colour to $b_{a,\alpha}$ and $b_{a+d,\alpha+\delta}$. Since there are only two colours which any of the numbers can be coloured with, $b_{a+2d,\alpha+2\delta}$ must be the same colour as $b_{a,\alpha+2\delta}$. Since B_a and B_{a+d} are equally coloured $b_{a,\alpha+2\delta}$, $b_{a+d,\alpha+2\delta}$ and $b_{a+2d,\alpha+2\delta}$ must be monochromatic, so we have a monochromatic 3-term arithmetic progression. Thus we have shown that no matter how $\{1, 2, \ldots, 325\}$ is 2-coloured there must always exist some monochromatic 3-term arithmetic progression, so $W(3,2) \leq 325$. □

Example. We take the example when the two monochromatic blocks in $\{1, 2, \ldots, 165\}$ are B_{11} and B_{31}. A colouring of B_{11} might be $B_{11} = \{51, 52, 53, 54, 55\}$ and since the colourings of B_{11} and B_{31} are equal

$B_{31} = \{151, 152, 153, 154, 155\}$. We can see there is no monochromatic 3-term arithmetic progression in these two sets. In this case we use B_{51} to find a monochromatic 3-term arithmetic progression. We arrive at the two possible 2-colourings of B_{51}. They are, $B_{51} = \{251, 252, 253, 254, 255\}$ and $B_{51} = \{251, 252, 253, 254, 255\}$ where the uncoloured numbers may be red or blue. In the first case we have the monochromatic 3-term arithmetic progression $\{51, 152, 253\}$, and in the second we have $\{53, 153, 253\}$.

3.2 Van der Waerden's Theorem

To prove Van der Waerden's Theorem for the general case we must first understand *Colour focusing*.

Definition 3.2.1. *In some r-colouring of the natural numbers, t different k-term arithmetic progressions are colour focused if*

- *each k-term arithmetic progression is monochromatic,*
- *none of the t arithmetic progressions have the same colour,*
- *the $(k+1)^{th}$ terms of each of the t arithmetic progressions are equal.*

The $(k+1)^{th}$ terms of the t arithmetic progressions is called the colour focus of those arithmetic progressions.

In a t-colouring the colour focus of the t monochromatic k-term arithmetic progressions must have the same colour as one of the t progressions and so a monochromatic $(k+1)$-term arithmetic progression must be formed.

Example. A colour focus was used in the previous example. The colour focus for the two monochromatic 2-term arithmetic progressions in the 2-colouring of $\{1, 2, \ldots, 325\}$ was 253.

Example. In the 2-colouring $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$, where the uncoloured numbers may be red or blue, 13 is a colour focus. Indeed, 13 is the fourth term of both the monochromatic arithmetic progressions $1, 5, 9$ and $4, 7, 10$. Also, no matter whether 13 is coloured red or blue there must be a monochromatic 4-term arithmetic progression in the set.

We now prove Van der Waerden's Theorem.

Theorem 3.2.2. *For all positive integers, k and r, there exists a natural number $W(k, r)$ such that, if the set of natural numbers $\{1, 2, \ldots, W(k, r)\}$ is r-coloured, then this set must contain at least one monochromatic k-term arithmetic progression.*

Proof. We shall prove that $W(k, r)$ exists by showing it is bounded. We use a proof by induction, on k. We already have that we can find a natural number $W(1, r)$. We now assume that, for any $q \leq k$ and any l we can find $W(q, l)$. We now show that $W(k+1,r)$ exists for every r.

Claim. For any t, such that $t \leq r$, there exists a natural number $W(t,k,r)$ such that whenever the set, $\{1,2,\ldots,W(t,k,r)\}$ is r-coloured, it must contain either a monochromatic $(k+1)$-term arithmetic progression or t colour focused monochromatic k-term arithmetic progressions together with their colour focus.

We prove this claim by induction, on t. We have previously assumed that we can find a natural number $W(k,r)$. Since there exists one monochromatic k-term arithmetic progression in $\{1,2,\ldots,W(k,r)\}$ it must be colour focused and its focus must be its $(k+1)^{th}$-term. The arithmetic progression's $(k+1)^{th}$-term must be less than or equal to $2W(k,r)$. Therefore $\{1,2,\ldots,2W(k,r)\}$ must contain a colour focused monochromatic k-term arithmetic progression together with its colour focus. So $W(1,k,r) = 2W(k,r)$.

We now assume that $W(t,k,r)$ exists and must prove the existence of $W(t+1,k,r)$.

We begin by taking the natural number, $X = 2W(t,k,r)W(k,r^{W(t,k,r)})$. We may then split the interval, $[1,X]$, into blocks, each of order $W(t,k,r)$. We label each block B_i where i denotes the blocks position in $[1,X]$. So we have

$$\begin{aligned}[1,X] =& \{1,2,\ldots,W(t,k,r)\} \cup \{W(t,k,r)+1, W(t,k,r)+2,\ldots,2W(t,k,r)\} \cup \\ & \ldots \cup \{X-(W(t,k,r)-1), X-(W(t,k,r)-2),\ldots,X\}, \\ =& B_1 \cup B_2 \cup \cdots \cup B_{2W(k,r^{W(t,k,r)})-1} \cup B_{2W(k,r^{W(t,k,r)})}.\end{aligned}$$

We now consider an r-colouring of $\{1,2,\ldots,X\}$. There are $r^{W(t,k,r)}$ ways in which a set of order $W(t,k,r)$ can be r-coloured, so each block, B_i, must be coloured in one of these $r^{W(t,k,r)}$ ways.

If, when the interval $[1,X]$ is r-coloured, any of the blocks of order $W(t,k,r)$ contain a monochromatic $(k+1)$-term arithmetic progression we are done. So we assume that each block contains t colour focused monochromatic k-term arithmetic progressions.

From the definition of $W(k,r^{W(t,k,r)})$, the set of natural numbers $\{1,2,\ldots, W(k,r^{W(t,k,r)})\}$ must contain a monochromatic k-term arithmetic progression when $r^{W(t,k,r)}$-coloured. Our r-colouring of $\{1,2,\ldots,W(t,k,r)W(k,r^{W(t,k,r)})\}$ induces an $r^{W(t,k,r)}$-colouring of the set of blocks, $\{B_1,B_2,\ldots,B_{W(k,r^{W(t,k,r)})}\}$, since each block has size $W(t,k,r)$ and thus is r-coloured in one of $r^{W(t,k,r)}$ ways. Therefore, the first $W(k,r^{W(t,k,r)})$ blocks must contain a monochromatic k-block arithmetic progression. That is, there must exist k identically coloured blocks, $B_a, B_{a+d}, B_{a+2d},\ldots,B_{a+(k-1)d}$, whose indices form an arithmetic progression. Since each block is of order $W(t,k,r)$ we may assume that they all contain t colour focused monochromatic k-term arithmetic progressions together with their colour focus, since otherwise one of the blocks must contain a monochromatic $(k+1)$-term arithmetic progression and we would be done.

We now label each element in $\{1,2,\ldots,X\}$, $b_{x,y}$, where x denotes the index of the block the element is in and y denotes that elements position in B_x. We denote the t colour focused monochromatic k-term arithmetic progressions in

B_a as

$$\begin{aligned}
P_{a,1} &= b_{a,\alpha}, b_{a,\alpha+\delta}, b_{a,\alpha+2\delta}, \dots, b_{a,\alpha+(k-1)\delta},\\
P_{a,2} &= b_{a,\mu}, b_{a,\mu+\nu}, b_{a,\mu+2\nu}, \dots, b_{a,\mu+(k-1)\nu},\\
&\vdots\\
P_{a,t} &= b_{a,\phi}, b_{a,\phi+\psi}, b_{a,\phi+2\psi}, \dots, b_{a,\phi+(k-1)\psi}.
\end{aligned}$$

These progressions each have their colour focus at $b_{a,f}$. That is, $b_{a,\alpha+k\delta} = b_{a,\mu+k\nu} = \cdots = b_{a,\phi+k\psi} = b_{a,f}$. Since all of the k blocks, $B_a, B_{a+d}, B_{a+2d}, \dots, B_{a+(k-1)d}$, are identically coloured there must exist monochromatic k-term arithmetic progressions,

$$\begin{aligned}
P_{a,1} &= b_{a,\alpha}, b_{a,\alpha+\delta}, b_{a,\alpha+2\delta}, \dots, b_{a,\alpha+(k-1)\delta},\\
P_{a+d,1} &= b_{a+d,\alpha}, b_{a+d,\alpha+\delta}, b_{a+d,\alpha+2\delta}, \dots, b_{a+d,\alpha+(k-1)\delta},\\
&\vdots\\
P_{a+(k-1)d,1} &= b_{a+(k-1)d,\alpha}, b_{a+(k-1)d,\alpha+\delta}, b_{a+(k-1)d,\alpha+2\delta}, \dots, b_{a+(k-1)d,\alpha+(k-1)\delta},\\
P_{a,2} &= b_{a,\mu}, b_{a,\mu+\nu}, b_{a,\mu+2\nu}, \dots, b_{a,\mu+(k-1)\nu},\\
&\vdots\\
P_{a+(k-1)d,2} &= b_{a+(k-1)d,\mu}, b_{a+(k-1)d,\mu+\nu}, b_{a+(k-1)d,\mu+2\nu}, \dots, b_{a+(k-1)d,\mu+(k-1)\nu},\\
&\vdots\\
P_{a+(k-1)d,t} &= b_{a+(k-1)d,\phi}, b_{a+(k-1)d,\phi+\psi}, b_{a+(k-1)d,\phi+2\psi}, \dots, b_{a+(k-1)d,\phi+(k-1)\psi},
\end{aligned}$$

such that

$$\begin{aligned}
\chi(P_{a,1}) &= \chi(P_{a+d,1}) = \cdots = \chi(P_{a+(k-1)d,1}),\\
\chi(P_{a,2}) &= \chi(P_{a+d,2}) = \cdots = \chi(P_{a+(k-1)d,2}),\\
&\vdots\\
\chi(P_{a,t}) &= \chi(P_{a+d,t}) = \cdots = \chi(P_{a+(k-1)d,t}),
\end{aligned}$$

where $\chi(P_{i,j})$ denotes the colour of the elements of the progression $P_{i,j}$. Together, each of the t progressions in each of the k blocks produce $t+1$ colour focused monochromatic k-term arithmetic progression. Indeed, consider the following k-term arithmetic progressions,

$$\begin{aligned}
F_1 &= b_{a,\alpha}, b_{a+d,\alpha+\delta}, b_{a+2d,\alpha+2\delta}, \dots, b_{a+(k-1)d,\alpha+(k-1)\delta},\\
F_2 &= b_{a,\mu}, b_{a+d,\mu+\nu}, b_{a+2d,\mu+2\nu}, \dots, b_{a+(k-1)d,\mu+(k-1)\nu},\\
&\vdots\\
F_t &= b_{a,\phi}, b_{a+d,\phi+\psi}, b_{a+2d,\phi+2\psi}, \dots, b_{a+(k-1)d,\phi+(k-1)\psi}.
\end{aligned}$$

Since each of the terms in F_i were taken from $P_{j,i}$, where $j \in \{a, a+d, \dots, a+(k-1)d\}$, each F_i must be monochromatic. Clearly the $(k+1)^{th}$ term of each of these progressions is equal, that is $b_{a+kd,\alpha+k\delta} = b_{a+kd,\mu+k\nu} = \cdots = b_{a+kd,\phi+k\psi}$.

This element is in X since $X = 2W(t,k,r)W(k,r^{W(t,k,r)})$ and each element we have used so far we have taken from the first $W(t,k,r)W(k,r^{W(t,k,r)})$ elements. Thus each of the t monochromatic k-term arithmetic progressions we have produced, $F_1, F_2, \ldots, F_t$, have their colour focus at $b_{a+kd,\alpha+k\delta} = b_{a+kd,\mu+k\nu} = \cdots = b_{a+kd,\phi+k\psi} = b_{a+kd,f}$. Clearly $b_{i,f}$ must be the same colour in every block in the monochromatic k-block arithmetic progression. Therefore the colour focuses of the blocks, $B_a, B_{a+d}, \ldots, B_{a+(k-1)d}$, also form a monochromatic k-term arithmetic progression. These terms, along with the other t monochromatic k-term arithmetic progressions, have their colour focus at $b_{a+kd,f}$. The k-term arithmetic progression, $b_{a,f}, b_{a+d,f}, \ldots, b_{a+(k-1)d,f}$, must have a different colour to each of $F_1, F_2, \ldots, F_t$, since otherwise a monochromatic $k+1$ term arithmetic progression would have been formed in one of $B_a, B_{a+d}, \ldots, B_{a+(k-1)d}$, from the definition of a colour focus. Thus $b_{a+kd,f}$ is the colour focus for $t+1$ monochromatic k-term arithmetic progressions. Therefore $X = W(t+1,k,r)$ and our claim is proved.

Since we have that $W(t,k,r)$ must exist for all $t \leq r$ we have that $W(r,k,r)$ must exist. That is, we can always find r colour focused k-term arithmetic progressions or a monochromatic $(k+1)$-term arithmetic progression in the r-coloured set of natural numbers $\{1, 2, \ldots, W(r,k,r)\}$. If there exists a monochromatic $(k+1)$-term arithmetic progression in this set we are done, so we assume one does not exist. Since we have only used r colours to colour this set of natural numbers, the colour focus of all the r arithmetic progressions must be coloured with one of the r colours. Therefore the colour focus must have the same colour as one of the r k-term arithmetic progression. Together with the colour focus this arithmetic progression then forms a monochromatic $(k+1)$-term arithmetic progression . Therefore, by induction for all positive integers, k and r, there exists a natural number $W(k,r)$ so that, if the set of natural numbers $\{1, 2, \ldots, W(k,r)\}$ is r-coloured, there is at least one monochromatic k-term arithmetic progression. $\square$

Van der Waerden's Theorem cannot be extended to the infinite case, that is, there exists an r-colouring of $\mathbb{N}$, for $r > 1$, under which there is no monochromatic infinite arithmetic progression. We give an example of one case below.

Example. We can 2-colour the natural numbers and avoid an infinite monochromatic arithmetic progression. We colour 1 red, 2,3 blue, 4,5,6 red, and so on, ad infinitum. The colouring produced in this way is
$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, \ldots\}$. We then label these monochromatic blocks B_k, where k represents the block's order and position. For example, $B_4 = \{7, 8, 9, 10\}$ is the fourth monochromatic block and $|B_4| = 4$. We can now label each element of $\mathbb{N}$, $b_{x,y}$, where x represents the index of the block in which the number is and y represents the number's position in the block B_x. We assume that we can find some monochromatic infinite arithmetic progression, say P. We take two successive elements of P, say $b_{k,m}$ and $b_{l,n}$. The difference between $b_{k,m}$ and $b_{l,n}$ is $(\frac{l(l-1)}{2}+n)-(\frac{k(k-1)}{2}+m)$. Therefore the difference between each term in P must be $(\frac{l(l-1)}{2}+n)-(\frac{k(k-1)}{2}+m)$.

Since the size of successive blocks, B_i, increase at a constant rate, there must exist some block, B_h, such that $h > (\frac{l(l-1)}{2} + n) - (\frac{k(k-1)}{2} + m)$, where B_h is coloured a different colour to P and $h > k$. There cannot exist a number in B_h which is also in P, but then there must exist two successive terms in P whose difference is greater than $(\frac{l(l-1)}{2} + n) - (\frac{k(k-1)}{2} + m)$, a contradiction. Since each block is of finite size it is impossible that an infinite monochromatic arithmetic progression can be contained inside one block. We have thus shown an infinite monochromatic arithmetic progression cannot exist in this 2-colouring of $\mathbb{N}$.

4 Rado's Theorem

In Richard Rado's 1933 thesis, *Studien zur Kombinatorik* [15], he generalised a basic result, proved by his supervisor Issai Schur to give Rado's Theorem. Schur's Theorem, proved in 1916, in *Über die Kongruenz* $x^m + y^m \equiv z^m \mod p$ [16], is the earliest result we will look at, and one of the earliest in Ramsey Theory. Rado's Theorem gives the properties which a system of linear homogeneous equations must have in order for them to have a monochromatic solution in any r-colouring of the natural numbers, for $r \in \mathbb{N}$.

We shall first examine what is meant by a system being regular and satisfying the Columns condition, while doing this we will consider Schur's Theorem. We will then move on to generalise Schur's Theorem to all single linear homogeneous equations. We shall then give the proof of Rado's Theorem which generalises this result even further.

Theorem 4.0.1. *For a matrix C, the system $C\mathbf{x} = \mathbf{0}$ is regular if and only if C satisfies the Columns condition.*

Theorem 4.0.1 is Rado's Theorem, however, to understand it we must first understand what it mean for a system to be *regular* and to satisfy the *Columns Condition*.

4.1 Regular systems

In this section we will define what is meant by a regular system and give two examples, one of a system which is regular and another of a system which is not.

Definition 4.1.1. *Let $S = S(x_1, \ldots, x_n)$ denote a system of linear homogeneous equations with variables $x_1, \ldots, x_n$. S is r-regular over A, the set on which S is defined, if given any r-colouring of A there exists a monochromatic set $\{x_1, \ldots, x_n\} \in A$ so that $S(x_1, \ldots, x_n)$ holds. S is regular over A if it is r-regular for all positive integers r. Equivalently, a matrix C is said to be regular over A if the system $C\mathbf{x} = \mathbf{0}$ has a monochromatic solution, $\mathbf{x}$, in every finite colouring of A. Generally we will be looking at systems on $\mathbb{N}$, for this reason if a system is said to be regular we will mean that the system is regular over $\mathbb{N}$.*

4.1.1 Two examples

An example of a simple regular system is given by Schur's Theorem.

Theorem 4.1.2. *For every $r \in \mathbb{N}$ there exists some natural number, n, such that in any r-colouring of $[n]$ there must exist a monochromatic set $x, y, z \in [n]$ such that $x + y = z$.*

Proof. Take n such that [2] $R_r(3) = n + 1$, that is, if we r-colour a complete graph of order $n + 1$, it must contain a monochromatic K_3. We now fix an r-colouring, say χ, of $[n]$, and define an r-colouring, say χ^*, of K_{n+1} by $\chi^*(i, j) =$

[2] $R_r(3)$ is defined as in Ramsey's Theorem.

$\chi(|i-j|)$. Since K_{n+1} is a complete graph on $n+1$ vertices, from the definition of $R_r(3)$ it must contain a monochromatic K_3, so there must exist a set of vertices $\{i,j,k\}$ where $k < j < i$, such that $\chi^*(i,j) = \chi^*(j,k) = \chi^*(k,i)$. Setting $x = i-j$, $y = j-k$ and $z = i-k$ gives $\chi(x) = \chi(y) = \chi(z)$ so we have our monochromatic set $\{x,y,z\}$ where $x+y=z$. We must specify that $k < j < i$ so that $x,y,z \in [n]$. $\square$

Example. The system $\begin{pmatrix} 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is not 2-regular. We can show this by constructing a red-blue colouring of $\mathbb{N}$ in which there is no monochromatic vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ such that

$$\begin{pmatrix} 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We define a colouring of $\mathbb{N}$ such that for every $x \in \mathbb{N}$, the numbers x and $2x$ have different colours. We colour the odd natural numbers blue and the even natural numbers the opposite colour to half their value. Thus $\mathbb{N}$ will be coloured $\{1,2,3,4,5,6,7,8,9,\ldots\}$. This colouring does not contain a monochromatic solution, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, to $\begin{pmatrix} 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$. Indeed, we would need a monochromatic vector, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, satisfying $x_1 = 2x_2$, so x_1 must be even. We defined our colouring so that any even term was coloured the opposite colour to half it's value. $x_2 = \frac{1}{2}x_1$ and so x_1 and x_2 can never be monochromatic.

4.1.2 Basic properties of regular systems

We now develop our understanding of what it means to be regular over countably infinite sets, namely $\mathbb{N}$, $\mathbb{Z}\setminus\{0\}$ and $\mathbb{Q}\setminus\{0\}$. In Lemma 4.1.3 we will prove that any system which is regular over $\mathbb{N}$, $\mathbb{Z}\setminus\{0\}$ or $\mathbb{Q}\setminus\{0\}$ is also regular over some finite subset of $\mathbb{N}$, $\mathbb{Z}\setminus\{0\}$ or $\mathbb{Q}\setminus\{0\}$ respectively.

Lemma 4.1.3. *Let $S(x_1,\ldots,x_n) = 0$ be a regular system of linear homogeneous equations.*

1. *For every $r \in \mathbb{N}$ there exists an $R \in \mathbb{N}$ such that any r-colouring of $[R]$ contains a monochromatic solution to the system, S.*
2. *For every $r \in \mathbb{N}$ there exists a finite set $R \subset \mathbb{Z}\setminus\{0\}$ such that any r-colouring of R contains a monochromatic solution to the system, S.*
3. *For every $r \in \mathbb{N}$ there exists a finite set $R \subset \mathbb{Q}\setminus\{0\}$ such that any r-colouring of R contains a monochromatic solution to the system, S.*

Proof. We shall first prove *1.* The proofs of *2.* and *3.* follow.

This is a proof by contradiction. We first assume we can find an r-colouring of $[R]$ which does not contain a monochromatic solution to the system S for every $R \in \mathbb{N}$. Let C_R be such an r-colouring of $[R]$.

We first note that for every $R \in \mathbb{N}$ there are a finite number, r^R, ways of r-colouring $[R]$. We take a subsequence of the colourings $C_1, C_2, C_3, \ldots$, consisting only of the colourings which colour $[1]$ in exactly the same way. $[1]$

can only be coloured in r ways with r colours so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i for $i \in \mathbb{N}$, which colour [1] in the same way. We call this subset of colourings $\mathcal{C}_1$ and define I_1 such that $C_i \in \mathcal{C}_1$ only if $i \in I_1$. We can then find a subset $\mathcal{C}_2 \subseteq \mathcal{C}_1$, consisting only of the colourings which when restricted to [2] colour the interval in exactly the same way. The set [2] can only be coloured in r^2 different ways with r colours so, by the pigeonhole principle, there must be an infinite number of the colourings, C_i for $i \in I_1$, which colour [2] in the same way. We call this subset of colourings $\mathcal{C}_2$ and define I_2 such that $C_i \in \mathcal{C}_2$ only if $i \in I_2$. We note that $I_2 \subseteq I_1$. We may continue in this way indefinitely.

We now produce an r-colouring, C, of $\mathbb{N}$. We define C as follows. For every $R \in \mathbb{N}$ with $R \leq x$ we set $C([R]) = C_i([R])$ where $C_i \in \mathcal{C}_x$. This colouring is well defined, if $R \leq x \leq y$ then defining the colouring $C([R])$ using $\mathcal{C}_x$ or $\mathcal{C}_y$ will give equal results since $C([R]) = C_i([R]) = C_j([R])$ for all $C_i \in \mathcal{C}_x$ and all $C_j \in \mathcal{C}_y$. This is clear since the subset of colourings $\mathcal{C}_y$ consists of the colourings C_i such that $i \in I_y \subseteq I_x \subseteq I_R$, so they all colour $[R]$ in the same way.

We now assume that there exists some natural number, n, such that under the colouring C the set $[n]$ contains a monochromatic solution to the system S. We can see from the definition of C that $C([n]) = C_i([n])$ for some $C_i \in \mathcal{C}_\nu$ where $n \leq \nu$. However, by definition any colouring in $\mathcal{C}_\nu$ colours $[n]$ such that it does not contain a monochromatic solution to the system S. So we have a contradiction. Thus under the r-colouring, C, there does not exist a natural number, n, such that $[n]$ contains a monochromatic solution to the system S. This is a contradiction, since $S(x_1, \ldots, x_n) = 0$ is a regular system it must have a solution in every finite colouring of $\mathbb{N}$ from the definition of regularity. C is an r-colouring of $\mathbb{N}$ which does not contain a monochromatic solution to the system S. Therefore we can conclude that for any finite number of colours, r, we can find a natural number, R, such that any r-colouring of $[R]$ contains a monochromatic solution to the system S.

Since $\mathbb{Z} \setminus \{0\}$ and $\mathbb{Q} \setminus \{0\}$ are countably infinite the proof given above for part *1.* can be simply altered to prove parts *2.* and *3.* $\square$

We are now able to prove that regularity over $\mathbb{N}$, $\mathbb{Z} \setminus \{0\}$ and $\mathbb{Q} \setminus \{0\}$ are equivalent.

Lemma 4.1.4. *The system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ if and only if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$.*

Proof. Clearly if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ then $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z}\setminus\{0\}$. Indeed, we may take any finite colouring of $\mathbb{Z} \setminus \{0\}$ and use it to define a finite colouring of $\mathbb{N}$. We simply define a finite colouring, χ, of $\mathbb{N}$ by colouring each natural number with the same colour it is coloured with in the finite colouring of $\mathbb{Z} \setminus \{0\}$. If $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ then under the χ colouring there must exist some monochromatic solution, $\mathbf{x} \in \mathbb{N}^n$, to the system $C\mathbf{x} = \mathbf{0}$. Since in the χ colouring each element of $\mathbb{N}$ was coloured with the same colour as in the finite colouring of $\mathbb{Z} \setminus \{0\}$, the monochromatic solution found in $\mathbb{N}^n$ must also be a monochromatic solution in $(\mathbb{Z} \setminus \{0\})^n$. Therefore $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$.

We now prove that if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$ then $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ by contradiction. We assume that there exists a finite colouring, χ, of $\mathbb{N}$ in which there does not exist a monochromatic vector, $\mathbf{x}$, such that $C\mathbf{x} = \mathbf{0}$. We label the colours used to colour $\mathbb{N}$, $r_1, \dots, r_n$, we then find a set of different colours $r'_1, \dots, r'_n$. We can then produce a colouring $\bar{\chi}$ of $\mathbb{Z} \setminus \{0\}$ in which $-\mathbb{N}$ is coloured in the same way as $\mathbb{N}$ but with the new set of colours, the colouring of $\mathbb{N}$ is the same as under the χ colouring. That is, if x is coloured using r_1 under the colouring χ then under the colouring $\bar{\chi}$ we use r'_1 to colour $-x$. We can see that under the colouring $\bar{\chi}$, of $\mathbb{Z} \setminus \{0\}$, there is no monochromatic vector, $\mathbf{x}$, such that $C\mathbf{x} = \mathbf{0}$. This can be seen by considering any vector

$$\mathbf{x}' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \in (\mathbb{Z} \setminus \{0\})^n$$

such that

$$\begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ c_{1n} & \cdots & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

From our definition of the colouring $\bar{\chi}$, we can see that every vector, $\mathbf{x}'$, with only positive entries cannot be a monochromatic vector since $\bar{\chi}$ colours $\mathbb{N}$ in the same way as χ. However, there are also no monochromatic vectors, $\mathbf{x}'$, with only negative entries satisfying $C\mathbf{x}' = \mathbf{0}$ under the $\bar{\chi}$ colouring. Indeed, if there existed a monochromatic vector, $\mathbf{x}'$, with only negative entries satisfying $C\mathbf{x}' = \mathbf{0}$ under the $\bar{\chi}$ colouring of $\mathbb{Z} \setminus \{0\}$ then $-\mathbf{x}'$ would be a monochromatic vector with entries only from $\mathbb{N}$. This is a contradiction since under the $\bar{\chi}$ colouring there are no monochromatic solutions to $C\mathbf{x} = \mathbf{0}$ in $\mathbb{N}$. Therefore $\mathbf{x}'$ cannot be a vector with only negative or only positive entries.

Under the $\bar{\chi}$ colouring we have coloured $\mathbb{Z}^-$ and $\mathbb{Z}^+$ with two different set of colours, namely $r_1, \dots, r_n$ and $r'_1, \dots, r'_n$, so we cannot produce a monochromatic vector which contain both negative and positive entries. We have thus shown that $C\mathbf{x} = \mathbf{0}$ is not regular over $\mathbb{Z} \setminus \{0\}$ and we have a contradiction. Therefore if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$ then $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$. $\square$

Lemma 4.1.5. *The system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$ if and only if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$.*

Proof. Clearly if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$ then $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$. Indeed, we may take any finite colouring of $\mathbb{Q} \setminus \{0\}$ and use it to define a finite colouring of $\mathbb{Z} \setminus \{0\}$. We simply define the finite colouring, χ, of $\mathbb{Z} \setminus \{0\}$ by colouring each element with the same colour it is coloured with in the finite colouring of $\mathbb{Q} \setminus \{0\}$. If $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$ then under the χ colouring there must exist some monochromatic solution, $\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^n$, to the system $C\mathbf{x} = \mathbf{0}$. Since in the χ colouring each element of $\mathbb{Z} \setminus \{0\}$ was coloured with the same colour as in the finite colouring of $\mathbb{Q} \setminus \{0\}$, the monochromatic solution found in $\mathbb{Z} \setminus \{0\}$ must also be a monochromatic solution in $\mathbb{Q} \setminus \{0\}$. Therefore $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$.

It remains to prove that if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$ then $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$.

Lemma 4.1.3 implies that since $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$, for every $r \in \mathbb{N}$ there must exist some finite set, $Q \subset \mathbb{Q} \setminus \{0\}$, such that in every r-colouring of Q there exists a monochromatic vector,

$$\mathbf{x}' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \in Q^n,$$

such that

$$\begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ c_{1n} & \cdots & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We now take some $k \in \mathbb{N}$ such that $kQ \in \mathbb{Z} \setminus \{0\}$, this is possible since Q is finite. For a given r-colouring of $\mathbb{Z} \setminus \{0\}$ we may now define an r-colouring of Q. We colour each element of Q, say q, with the colour given to kq under the r-colouring of $\mathbb{Z} \setminus \{0\}$. There must exist some monochromatic solution, say $\mathbf{q}$, to $C\mathbf{x} = \mathbf{0}$ in Q^n. Therefore, since kQ is coloured in exactly the same way as Q, there must also exist a monochromatic solution, $k\mathbf{q} \in (kQ)^n \subset (\mathbb{Z} \setminus \{0\})^n$. Therefore, if $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Q} \setminus \{0\}$ then in any colouring of $\mathbb{Z} \setminus \{0\}$ there must exist a monochromatic solution to $C\mathbf{x} = \mathbf{0}$, that is, $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{Z} \setminus \{0\}$. $\square$

4.2 The Columns condition

In this section we will define what is meant by a system satisfying the Columns condition and consider the examples given in Section 4.1.1.

Definition 4.2.1. *An $m \times n$ matrix $C = (c_{ij})$ is said to satisfy the Columns condition if its columns can be partitioned as $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_k$, where each $\mathcal{C}_i$ is a set of column vectors from C, such that the following conditions hold. The column vectors of the first partition set sum to zero, that is $\sum_{\mathbf{c}_i \in \mathcal{C}_1} \mathbf{c}_i = 0$, and for all $j > 1$ the sum $\sum_{\mathbf{c}_i \in \mathcal{C}_j} \mathbf{c}_i$ can be written as a linear combination of vectors from the set $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_{j-1}$.*

We now look back to two examples given in Section 4.1.1, one of a matrix which is regular and one of a matrix which is not. We will show that the regular matrix satisfies the Columns condition and the matrix which is not regular does not satisfy the Columns condition. This will help us gain a better understanding of what it means for a matrix to satisfy the Columns condition.

Example. The system $\begin{pmatrix} 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ does not satisfy the Columns condition. Indeed, there is no partition under which the vectors in the first partition set sum to give the zero vector. The equation in Schur's Theorem corresponds to the system $\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$. This system does satisfy the Columns condition. We can label the columns of the vector as $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ and produce the partition $\mathcal{C}_1 = \{\mathbf{c}_1, \mathbf{c}_3\}$ and $\mathcal{C}_2 = \{\mathbf{c}_2\}$. Thus, $\mathbf{c}_1 + \mathbf{c}_3 = \mathbf{0}$ and $\mathbf{c}_2 = \mathbf{c}_1$ so the Columns condition is satisfied.

4.3 Preliminary results

Although the results in this section do not appear to be immediately related to Rado's Theorem it will become clear that they are essential to it's proof. They are proved here in advance of the proof of Rado's Theorem to aid the clarity of the main proof. The following theorem is a strengthening of Van der Waerden's Theorem given in the last chapter. Along with a monochromatic arithmetic progression we also prove that we can find some multiple of the difference of the terms in the arithmetic progression which has the same colour.

Theorem 4.3.1. *For all $k, r, s \in \mathbb{N}$ there exists a natural number $n = n(k, r, s)$ such that, if $[n]$ is r-coloured, there exist $a, d \in \mathbb{N}$ so that*

$$\{a, a+d, a+2d, \ldots, a+(k-1)d\} \cup \{sd\} \tag{4.3.1}$$

is a monochromatic subset of $[n]$.

Proof. We shall prove this result using a proof by induction on r, the number of colours. First, suppose that $r = 1$. Trivially $n(k, 1, s) = \max\{k, s\}$. Indeed, we only colour $[n(k, 1, s)]$ with one colour and so any series in $[n(k, 1, s)]$ must be monochromatic. For $n(k, 1, s)$ we take $a = 1$ and $d = 1$. If $s \leq a+(k-1)d = 1+(k-1) = k$ we only need $n = k$. If $s > a+(k-1)d = k$ we must take $n = s$ to ensure $sd = s \in [n(k, 1, s)]$.

We now assume that $r > 1$ and $n(k, r-1, s)$ exists. We must prove that $n(k, r, s)$ exists.

Claim. We may take $n(k, r, s) = sW(kn(k, r-1, s), r)$.

Take some r-colouring of $[n(k, r, s)]$. From the definition of [3]$W(k, r)$ we can now find a monochromatic, say red, set $\{a + id' \mid 0 \leq i \leq kn(k, r-1, s) - 1\}$ in the first $W(kn(k, r-1, s), r)$ natural numbers.

If there exists some $1 \leq j \leq n(k, r-1, s)$ such that $sd'j$ is red then setting $d = jd'$ we find $\{a, a+d, a+2d, \ldots, a+(k-1)d\} \cup \{sd\}$ is red. Clearly we can find a red set $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$ since $\{a+id' \mid 0 \leq i \leq kn(k, r-1, s)-1\}$ is red. Indeed, $sd \in [sW(kn(k, r-1, s), r)]$ since $d = d'j \in [W(kn(k, r-1, s), r)]$.

If there does not exist some $1 \leq j \leq n(k, r-1, s)$ such that $sd'j$ is red then $\{sd'j \mid 1 \leq j \leq n(k, r-1, s)\}$ is at most $(r-1)$-coloured. From the definition

[3]$W(k, r)$ is defined as in Van der Waerden's Theorem.

of $n(k, r-1, s)$, which we already assumed to exist, we have the desired result. Indeed, we may define an $(r-1)$-colouring of $[n(k, r-1, s)]$ by giving every number x in this set the colour of $sd'x$. There must exist a monochromatic set, $\{a, a+d, \ldots, a+(k-1)d\} \cup \{sd\}$ in $[n(k, r-1, s)]$ from the definition of $n(k, r-1, s)$. From the definition of the colouring of $[n(k, r-1, s)]$, the set $\{sd'a, sd'a+sd'd, \ldots, sd'a+(k-1)sd'd\} \cup \{s^2d'd\}$ must be monochromatic since $\{a, a+d, a+2d, \ldots, a+(k-1)d\} \cup \{sd\}$ is monochromatic in $[n(k, r-1, s)]$.

□

We now turn to an example where we can describe the set $\{a, a+d, a+2d, \ldots, a+(k-1)d\} \cup \{sd\}$ as the solution of a homogeneous system of equations, thus showing that the corresponding matrix must be regular. The matrix also satisfies the Columns condition.

Example. We take the matrix

$$C = \begin{pmatrix} -1 & 2 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 & 0 \\ -s & s & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 \end{pmatrix}.$$

The system, $C\mathbf{x} = \mathbf{0}$, corresponding to the matrix C is

$$\begin{aligned} -x_1 + 2x_2 - x_3 \qquad\qquad\qquad\qquad &= 0, \\ -x_2 + 2x_3 - x_4 \qquad\qquad\qquad &= 0, \\ -x_3 + 2x_4 - x_5 \qquad\qquad &= 0, \\ -x_4 + 2x_5 - x_6 \qquad &= 0, \\ \ddots \qquad &\ \vdots \\ -x_{n-3} + 2x_{n-2} - x_{n-1} &= 0, \\ -sx_1 + sx_2 \qquad\qquad\qquad\qquad - x_n &= 0, \end{aligned}$$

which could equally be written as

$$\begin{aligned} x_3 - x_2 &= x_2 - x_1, \\ x_4 - x_3 &= x_3 - x_2, \\ &\vdots \\ x_{n-1} - x_{n-2} &= x_{n-2} - x_{n-3}, \\ x_n &= s(x_2 - x_1). \end{aligned}$$

Thus any solution to the the system $C\mathbf{x} = \mathbf{0}$ has the form $x_1 = a, x_2 = a+d, x_3 = a+2d, \ldots, x_n = sd$. We have now shown that this system is equivalent to the set in Theorem 4.3.1. In Theorem 4.3.1 it was shown that in any finite colouring of $\mathbb{N}$ we can find a monochromatic set of this type, therefore the system, $C\mathbf{x} = \mathbf{0}$, is regular. We can also see that the matrix C satisfies the Column condition. Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ be the columns of C, we can take the partitions to be $\mathcal{C}_1 = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \ldots, \mathbf{c}_{n-1}\}$ and $\mathcal{C}_2 = \{\mathbf{c}_n\}$ then $\sum_{i=1}^{n-1} \mathbf{c}_i = 0$ and $\mathbf{c}_n = \frac{1}{s}\mathbf{c}_1 - \sum_{i=3}^{n-1} \frac{i-2}{s}\mathbf{c}_i$. So C satisfies the Columns condition.

Corollary 4.3.2. *For all $k, r, s \in \mathbb{N}$ there exists a natural number $n = n'(k, r, s)$ such that, if $[n]$ is r-coloured, there exists $a, d \in \mathbb{N}$ such that*

$$\{a + \lambda d \,|\, |\lambda| \leq (k-1)\} \cup \{sd\}$$

is a monochromatic subset of $[n]$.

Proof. Using Theorem 4.3.1 and replacing $(k-1)$ by $2(k-1)$ we can find $a', d' \in \mathbb{N}$ such that $\{a', a'+d', a'+2d', \ldots, a'+2(k-1)d'\} \cup \{sd'\}$ is monochromatic. This is clearly possible since k can be any natural number. We can then define $d = d'$ and $a = a' + (k-1)d'$ so that the elements in $\{a-(k-1)d, a-(k-2)d, \ldots, a-d, a, a+d, \ldots, a+(k-2)d, a+(k-1)d\} \cup \{sd\}$ correspond to the elements in $\{a', a'+d', a'+2d', \ldots, a'+2(k-1)d'\} \cup \{sd'\}$ which we found to be monochromatic. Therefore, since $n(k, r, s)$ was proved to exist in Theorem 4.3.1, for all $k, r, s \in \mathbb{N}$ there must exists a natural number $n'(k, r, s)$. □

4.4 Rado's Theorem for a single linear homogeneous constraint

We first define the $\mathcal{N}_p$ colouring of the natural numbers.

Definition 4.4.1. *For any $n \in \mathbb{N}$ we may write out the p-expansion of n, that is, we may write $n = n_0 + n_1 p + n_2 p^2 + \cdots + n_k p^k$ where $n_i \in \{0, 1, \ldots, p-1\}$ for all i and $p \geq 2$. We define the $\mathcal{N}_p$ colouring of $\mathbb{N}$ by setting $\mathcal{N}_p(z) = n_i$ where i is the smallest natural number such that $n_i \neq 0$ for all $z \in \mathbb{N}$. The rank of z under the $\mathcal{N}_p$ colouring is i, where i is the smallest natural number such that $n_i \neq 0$.*

Lemma 4.4.2. *If $\mathcal{N}_p(y) = \mathcal{N}_p(z)$ then $\mathcal{N}_p(p^{-m_1} y) = \mathcal{N}_p(p^{-m_1} z)$ where m_1 is the minimum of the rank of y and that of z.*

Proof. Since $\mathcal{N}_p(y) = \mathcal{N}_p(z)$ we must have that $y = ap^{m_i} + n_{m_{i+1}} p^{m_{i+1}} + \cdots + n_{m_j} p^{m_j}$ and $z = ap^{m_k} + n_{m_{k+1}} p^{m_{k+1}} + \cdots + n_{m_l} p^{m_l}$ for some $i, j, k, l \in \mathbb{N}$. Without loss of generality we may assume that $m_i \leq m_k$. Now,

$$\begin{aligned} p^{-m_i} y &= p^{-m_i} a p^{m_i} + p^{-m_i} n_{m_{i+1}} p^{m_{i+1}} + \cdots + p^{-m_i} n_{m_j} p^{m_j}, \\ &= a + n_{m_{i+1}} p^{m_1} + \cdots + n_{m_j} p^{m_{j-i}}, \end{aligned}$$

and

$$\begin{aligned} p^{-m_i} z &= p^{-m_i} a p^{m_k} + p^{-m_i} n_{m_{k+1}} p^{m_{k+1}} + \cdots + p^{-m_i} n_{m_l} p^{m_l}, \\ &= a p^{m_{k-i}} + n_{m_{k+1}} p^{m_{k+1-i}} + \cdots + n_{m_l} p^{m_{l-i}}. \end{aligned}$$

Therefore $\mathcal{N}_p(p^{-m_i} y) = a$ and $\mathcal{N}_p(p^{-m_i} z) = a$, so $\mathcal{N}_p(p^{-m_i} y) = \mathcal{N}_p(p^{-m_i} z)$. □

We now use the $\mathcal{N}_p$ colouring to prove Theorem 4.4.3, Rado's Theorem for a single linear homogeneous constraint.

Theorem 4.4.3. *Let $S(x_1, \ldots, x_n)$ be the system given by the single linear homogeneous constraint*

$$c_1x_1 + \cdots + c_nx_n = 0$$

for $c_i \in \mathbb{Z} \setminus \{0\}$ and $x_i \in \mathbb{N}$. S is regular if and only if some non-empty subset of the c_i sums to zero.

Proof. We fix a finite colouring of $\mathbb{N}$ and reorder the coefficients if necessary so that we can write $c_1 + \cdots + c_k = 0$ for some $k \leq n$. If $k = n$ we can set $x_1 = x_i$, for all $1 \leq i \leq k$, this set must be monochromatic since it is made up of only one element. Since the coefficients sum to zero, we have

$$\begin{aligned} c_1x_1 + c_2x_2 + \cdots + c_nx_n &= c_1x_1 + c_2x_1 + \cdots + c_nx_1, \\ &= x_1(c_1 + c_2 + \cdots + c_n), \\ &= 0. \end{aligned}$$

We now assume that $k < n$. We define $A = hcf(c_1, c_2, \ldots, c_k)$ and $B = c_{k+1} + \cdots + c_n$. We set $s = \frac{A}{hcf(A,B)}$. Note that $s \in \mathbb{N}$. If $B = 0$ then, as before, we can easily find a monochromatic solution. If $B \neq 0$, we can find some $t \in \mathbb{Z}$ so that $At + Bs = 0$, indeed, $t = \frac{-B}{hcf(A,B)}$. By the fundamental theorem of arithmetic we can then then find $\lambda_1, ..., \lambda_k \in \mathbb{Z}$ such that $c_1\lambda_1 + ... + c_k\lambda_k = At$, since $A = hcf(c_1, c_2, \ldots, c_k)$. We can now produce a parametric solution to $c_1x_1 + \cdots + c_nx_n = 0$. Set

$$x_i = \begin{cases} a + \lambda_i d & \text{if} \quad 1 \leq i \leq k, \\ sd & \text{if} \quad k < i \leq n. \end{cases}$$

Note that

$$\sum_{i=1}^{k} c_i(a + \lambda_i d) = \sum_{i=1}^{k} c_i a + \sum_{i=1}^{k} c_i \lambda_i d = 0 + Atd = Atd,$$

and

$$\sum_{i=k+1}^{n} c_i sd = \left(\sum_{i=k+1}^{n} c_i \right) sd = Bsd.$$

Therefore

$$\begin{aligned} c_1x_1 + \cdots + c_nx_n &= \sum_{i=1}^{k} c_i(a + \lambda_i d) + \sum_{i=k+1}^{n} c_i sd, \\ &= Atd + Bsd, \\ &= (At + Bs)d, \\ &= 0d, \\ &= 0. \end{aligned}$$

Because of the form of the parametric solution, by Corollary 4.3.2, we must be able to find $a, d \in \mathbb{N}$ such that $\{x_i \,|\, 1 \leq i \leq n\}$ is monochromatic. Our parametric solution lies in $\mathbb{N}$, proving that S is regular over $\mathbb{N}$.

This completes the proof that *if some non-empty subset of the c_i sums to zero then S is regular.* To complete the proof of Theorem 4.4.3 we will prove that *if S is regular then some non-empty subset of the c_i sums to zero.*

We shall prove this using a proof by contradiction. We suppose that we can find some prime, p, such that the sum of any non-empty subsets of $\{c_i \,|\, 1 \leq i \leq n\}$ is indivisible by p. Therefore no subset of $\{c_i \,|\, 1 \leq i \leq n\}$ will sum to zero.
Claim. Under this condition $c_1x_1 + \cdots + c_nx_n = 0$ has no monochromatic solution in $\mathbb{N}$ when coloured using $\mathcal{N}_p$.

If our claim is true S is not regular over $\mathbb{N}$. We will prove this claim by contradiction, first we assume that we have a set $x_1, \ldots, x_n$ which, under the $\mathcal{N}_p$ colouring, forms a monochromatic solution to $c_1x_1+\cdots+c_nx_n = 0$. We may assume that $p \nmid x_i$ for some i since Lemma 4.4.2 implies that $p^{-k}x_1, \ldots, p^{-k}x_n$ also forms a monochromatic solution to $c_1x_1+\cdots+c_nx_n = 0$ in which $p \nmid p^{-k}x_i$ for some i, where k is the minimum of the ranks of all the x_i. We can now reorder the equation such that $p \nmid x_i$ for $1 \leq i \leq k$ and $p \mid x_i$ for $k < i \leq n$. Now,

$$c_1x_1 + \cdots + c_nx_n \equiv 0 \mod p,$$

that is

$$\sum_{i=1}^{n} \bar{c}_i\bar{x}_i = \bar{0}.$$

Clearly for $k < i \leq n$, $\bar{x}_i = \bar{0}$. All the x_i were defined to be monochromatic integers, so from the definition of $\mathcal{N}_p(x_i)$ and since $p \nmid x_i$ we see that for all $1 \leq i, j \leq k$, $\bar{x}_1 = \bar{x}_i = \bar{x}_j$. Therefore

$$\sum_{i=1}^{n} \bar{c}_i\bar{x}_i = \sum_{i=1}^{k} \bar{c}_i\bar{x}_i = (\sum_{i=1}^{k} \bar{c}_i)\bar{x}_1 = \bar{0}.$$

Since $\bar{x}_1 \neq \bar{0}$ we have $\sum_{i=1}^{k} \bar{c}_i = \bar{0}$ which implies that p divides the sum of a non-empty subset of $\{c_i \,|\, 1 \leq i \leq n\}$, contrary to the assumption. We have proved the claim, $c_1x_1 + \cdots + c_nx_n = 0$ has no monochromatic solutions when coloured using the $\mathcal{N}_p$ colouring. This proves that for $c_i \in \mathbb{Z}$, *if S is regular over $\mathbb{N}$ then some non-empty subset of the c_i sums to zero.* □

Example. To illustrate the proof of Theorem 4.4.3, consider the linear homogeneous equation

$$x_1 + 9x_2 + 5x_3 - x_4 - 2x_5 + 3x_6 - 7x_7 = 0. \tag{4.4.1}$$

We have, $c_1 = 1, c_2 = 9, c_3 = 5, c_4 = -1, c_5 = -2, c_6 = 3, c_7 = -7$. Since

$$c_1 + c_2 + c_4 + c_5 + c_7 = 0,$$

we can relabel the sum as

$$x_1 + 9x_2 - x_3 - 2x_4 - 7x_5 + 5x_6 + 3x_7 = 0.$$

Now, $k = 5$, $A = hcf(1, 9, -1, -2, -7) = 1$, $B = 5 + 3 = 8$, $s = \frac{1}{hcf(1,8)} = \frac{1}{1} = 1$ and $t = \frac{-8}{hcf(1,8)} = \frac{-8}{1} = -8$. We then take $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 1, \lambda_4 = 0, \lambda_5 = 1$ which satisfy $c_1\lambda_1 + c_2\lambda_2 + c_3\lambda_3 + c_4\lambda_4 + c_5\lambda_5 = At$, that is $\lambda_1 + 9\lambda_2 - \lambda_3 - 2\lambda_4 - 7\lambda_5 = -1 - 7 = -8$. Our monochromatic solution is then

$$\begin{aligned} x_1 &= a, \\ x_2 &= a, \\ x_3 &= a + d, \\ x_4 &= a, \\ x_5 &= a + d, \\ x_6 &= d, \\ x_7 &= d, \end{aligned}$$

and equation 4.4.1 can now be written

$$a + 9(a) - (a + d) - 2(a) - 7(a + d) + 5d + 3d = 0.$$

4.5 Rado's Theorem for regular homogeneous systems

We now extend Rado's Theorem from one linear homogeneous equation to a system of linear homogeneous equations. We begin by proving a result seemingly unrelated to Rado's Theorem.

Lemma 4.5.1. *Let $\mathbf{a}, \mathbf{c}_1, \ldots, \mathbf{c}_k \in \mathbb{Z}^n$. Suppose that $\mathbf{a}$ is not in the vector space over $\mathbb{Q}$ generated by $\mathbf{c}_i$. Then for all but a finite number of primes p, $\mathbf{a}$ cannot be expressed as a linear combination of the $\mathbf{c}_i \mod p$. Moreover, for all but a finite number of primes p, $\mathbf{a}p^m$ cannot be expressed as a linear combination of the $\mathbf{c}_i \mod p^{m+1}$ for any $m \geq 0$.*

Proof. We shall only prove that for all but a finite number of primes p, $\mathbf{a}p^m$ cannot be expressed as a linear combination of the $\mathbf{c}_i \mod p^{m+1}$ for any $m \geq 0$. Clearly, setting $m = 0$ we will then have also proved that for all but a finite number of primes p, $\mathbf{a}$ cannot be expressed as a linear combination of the $\mathbf{c}_i \mod p$.

We begin by noting that since $\mathbf{a}$ is not in the vector space generated by $\mathbf{c}_i$ we can find a vector, $\mathbf{u} \in \mathbb{Q}^n$, such that $\mathbf{u} \cdot \mathbf{c}_i = 0$, for $1 \leq i \leq k$, but $\mathbf{u} \cdot \mathbf{a} \neq 0$. We can then easily multiply $\mathbf{u}$ by a suitable constant so that $\mathbf{u} \in \mathbb{Z}^n$. We set $\mathbf{u} \cdot \mathbf{a} = s$. Since $\mathbf{a} \in \mathbb{Z}^n$ and $\mathbf{u} \in \mathbb{Z}^n$ we know $s \in \mathbb{Z} \setminus \{0\}$. Suppose that $\mathbf{a}p^m$ can be expressed as a linear combination of the $\mathbf{c}_i \mod p^{m+1}$, that is,

$$\mathbf{a}p^m \equiv \mathbf{c}_1 x_1 + \cdots + \mathbf{c}_k x_k \mod p^{m+1}.$$

Then multiplying by $\mathbf{u}$ gives,

$$\mathbf{u} \cdot \mathbf{a}p^m \equiv \sum_{i=1}^{k} \mathbf{u} \cdot \mathbf{c}_i x_i \mod p^{m+1}.$$

However, we have already defined $\mathbf{u} \cdot \mathbf{a} = s$ and $\mathbf{u} \cdot \mathbf{c}_i = 0$ so,

$$sp^m \equiv 0 \mod p^{m+1}.$$

This implies that $p^{m+1} | sp^m$ but then $p|s$, that is $p|\mathbf{a} \cdot \mathbf{u}$. Clearly this is only true for a finite number of primes. $\square$

For our next result we must first define some terminology.

Definition 4.5.2. *Given a matrix C with column vectors $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k$ such that $\mathbf{c}_1 + \mathbf{c}_2 + \cdots + \mathbf{c}_k \neq 0$, we define $E(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k)$ to be the set of all primes, p, for which $\mathbf{c}_1 + \mathbf{c}_2 + \cdots + \mathbf{c}_k \equiv 0 \mod p$.*

Definition 4.5.3. *Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$ and $\mathbf{a}$ be vectors such that $\mathbf{a}$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$. We define $E(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k; \mathbf{a})$ to be the set of all primes, p, for which $\mathbf{a}p^m \equiv \mathbf{c}_1 x_1 + \mathbf{c}_2 x_2 + \cdots + \mathbf{c}_k x_k \mod p^{m+1}$ for some $m \in \mathbb{N}$.*

Definition 4.5.4. *E is the union of $E(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k)$ and $E(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k; \mathbf{a})$. The vectors used to find E will be clear from the context.*

We may now begin to prove Rado's Theorem for regular homogeneous systems.

Lemma 4.5.5. *Take a matrix C. Let p be a prime such that $p \notin E$. If $C\mathbf{x} = \mathbf{0}$ has a monochromatic solution under the $\mathcal{N}_p$ colouring then C satisfies the Columns condition.*

Proof. Let $x_1, \ldots, x_n$ be a monochromatic solution to $C\mathbf{x} = \mathbf{0}$ under the $\mathcal{N}_p$ colouring. We first order $x_1, \ldots, x_n$ by their rank, as given in Definition 4.4.1, so that

$$rank(x_i) = \begin{cases} m_1 & \text{if} \quad 1 \leq i \leq k_1, \\ m_2 & \text{if} \quad k_1 < i \leq k_2, \\ \vdots & \qquad \vdots \\ m_s & \text{if} \quad k_{s-1} < i \leq n. \end{cases}$$

We can assume that $m_1 = 0$ since Lemma 4.4.2 implies we can replace each x_i with $x_i p^{-m_1}$ and they will remain monochromatic. We take r to be the colour by which all the x_i are coloured, that is $\mathcal{N}_p(x_i) = r$. From Definition 4.4.1 we have that for all x_i such that $i \leq k_1$ there must exist some $z_i \in \mathbb{N}$ such that $x_i = r + z_i p$. Clearly $p|x_i$ for all $i \geq k_1$ so there must exist some $z_i \in \mathbb{N}$ such that $x_i = z_i p$. So the system of equations $C\mathbf{x} = \mathbf{0}$ can be written as

$$\begin{aligned} \mathbf{0} &= C\mathbf{x}, \\ &= \mathbf{c}_1 x_1 + \cdots + \mathbf{c}_n x_n, \\ &= \mathbf{c}_1(r + z_1 p) + \cdots + \mathbf{c}_{k_1}(r + z_{k_1} p) + \mathbf{c}_{k_1+1} z_{k_1+1} p + \cdots + \mathbf{c}_n z_n p, \\ &\equiv \mathbf{c}_1 r + \cdots + \mathbf{c}_{k_1} r \mod p, \\ &\equiv (\mathbf{c}_1 + \cdots + \mathbf{c}_{k_1}) r \mod p. \end{aligned}$$

From the definition of the $\mathcal{N}_p$ colouring we can see that $r \not\equiv 0 \mod p$ since $0 < r < p$. Therefore $\mathbf{c}_1 + \cdots + \mathbf{c}_{k_1} \equiv \mathbf{0} \mod p$ but since $p \notin E$ and therefore $p \notin E(\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{k_1})$ we must have $\mathbf{c}_1 + \mathbf{c}_2 + \cdots + \mathbf{c}_{k_1} = 0$.

Let $1 < j \leq s$. Then

$$\begin{aligned}\mathbf{0} &\equiv \mathbf{c}_1 x_1 + \cdots + \mathbf{c}_n x_n \mod p^{m_j+1}, \\ &\equiv \sum_{i=1}^{k_{j-1}} \mathbf{c}_i x_i + \sum_{i=k_{j-1}+1}^{k_j} \mathbf{c}_i p^{m_j} r \mod p^{m_j+1}.\end{aligned}$$

We can then multiply by r^{-1}, which must exist since $\mathbb{Z}_{p^{m_j+1}}$ is a field, giving

$$\sum_{i=k_{j-1}+1}^{k_j} \mathbf{c}_i p^{m_j} \equiv -\sum_{i=1}^{k_{j-1}} \mathbf{c}_i r^{-1} x_i \mod p^{m_j+1}.$$

This can be rewritten as

$$\mathbf{a} p^{m_j} \equiv -\sum_{i=1}^{k_{j-1}} \mathbf{c}_i r^{-1} x_i \mod p^{m_j+1},$$

where $\mathbf{a} = \sum_{i=k_{j-1}+1}^{k_j} \mathbf{c}_i$. Again we note $p \notin E$, from the above equation we can see that $\mathbf{a} p^{m_j} \equiv -\sum_{i=1}^{k_{j-1}} \mathbf{c}_i r^{-1} x_i \mod p^{m_j+1}$. Thus from Definition 4.5.3 we can see that $\mathbf{a}$ is a linear combination of the vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k_{j-1}}$. Since j was an arbitrary natural number such that $1 < j \leq s$ we see that the matrix C must satisfy the Columns condition. $\square$

Lemma 4.5.5 proves that *if the system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ then C satisfies the Columns condition*. We now go on to finish the proof of Rado's Theorem by proving that *if C satisfies the Columns condition then the system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$*.

We must first introduce families and their relation to the proof of Rado's Theorem.

Definition 4.5.6. *A family, $\mathfrak{F}$, of finite subsets of $\mathbb{N}$ is called homogeneous if for every set $A \in \mathfrak{F}$ and $a \in \mathbb{N}$ we have that $aA \in \mathfrak{F}$.*

Example. The set of solutions to a homogeneous system is a homogeneous family. That is, the sets of entries of the vectors, $\mathbf{x}$, satisfying the system $C\mathbf{x} = \mathbf{0}$ together form a family of solutions to $C\mathbf{x} = \mathbf{0}$. These solutions form a homogeneous family since if we multiply a vector $\mathbf{x}$ by some natural number, it remains a solution to the system.

Definition 4.5.7. *A family, $\mathfrak{F}$, of finite subsets of $\mathbb{N}$ is called regular if in every finite colouring of $\mathbb{N}$ there exists a monochromatic set $A \in \mathfrak{F}$.*

Example. The set of solutions to a regular homogeneous system is a regular homogenous family. For a regular homogeneous system we can find a monochromatic vector, $\mathbf{x} \in \mathbb{N}^n$, satisfying the system $C\mathbf{x} = \mathbf{0}$ in any finite colouring of the natural numbers. Therefore we may find a monochromatic set in the family of solutions to the system $C\mathbf{x} = \mathbf{0}$ in any finite colouring of $\mathbb{N}$.

We now prove two results involving families so that we may use them in the proof of Rado's Theorem.

Theorem 4.5.8. *Let $\mathfrak{F}$, a family of finite subsets of $\mathbb{N}$, be homogeneous and regular. Let M be a positive integer. If $\mathbb{N}$ is finitely coloured there exists $A \in \mathfrak{F}$ and some positive integer d, such that*

$$\{a + \lambda d \,|\, a \in A,\, |\lambda| \leq M\}$$

is a monochromatic subset of $\mathbb{N}$.

Proof. We first take an arbitrary number of colours, r say. Clearly the proof of Lemma 4.1.3 may be simply modified to show that there exists some $R \in \mathbb{N}$ such that any r-colouring of $[R]$ must contain a monochromatic set $A' \in \mathfrak{F}$.

Given an r-colouring, χ, of $\mathbb{N}$, we can define an r^R-colouring, χ^* of $\mathbb{N}$ by $\chi^*(\alpha) = \chi^*(\beta)$ if and only if $\chi(\alpha i) = \chi(\beta i)$ for $1 \leq i \leq R$. That is two natural numbers, α and β, are monochromatic under χ^* if and only if $\chi(\alpha) = \chi(\beta), \chi(2\alpha) = \chi(2\beta), \ldots, \chi(R\alpha) = \chi(R\beta)$. We can see that χ^* is an r^R-colouring since there are r^R possible colourings of $\{\alpha, 2\alpha, 3\alpha, \ldots, R\alpha\}$ for $\alpha \in \mathbb{N}$ under the χ colouring.

We now set $T = MR^{n-1}$. By Van der Waerden's Theorem we can find a monochromatic arithmetic progression of length $2T + 1$ in the χ^* colouring of $\mathbb{N}$. Clearly Van der Waerden's Theorem also implies that there exists $g, e \in \mathbb{N}$ such that the set $\{g + \mu e \,|\, |\mu| \leq T\}$ is monochromatic and lies in $\mathbb{N}$.

We may define a new r-colouring of $[R]$ by colouring each element $m \in [R]$ the same colour as $gm \in g[R]$, where $g[R]$ is coloured using the χ colouring. Since in the new r-colouring there must exist some monochromatic set $A' \in \mathfrak{F}$ such that $A' \in [R]$, from the way in which $[R]$ was coloured there must also exist a monochromatic set $gA' \in g[R]$. Since $A' \in \mathfrak{F}$ and $\mathfrak{F}$ is a homogeneous family we must have that $gA' \in \mathfrak{F}$. We label the elements of the finite set $A' = \{y_1, y_2, \ldots, y_n\}$ and therefore $gA' = \{gy_1, gy_2, \ldots, gy_n\}$, we also set $gA' = A$.

We now set $a_i = gy_i$ and $d = ey$ where $y = lcm(y_1, \ldots, y_n)$. We will show that A and d are as required in the statement of the Theorem, however, we first note that

$$a_i + \lambda d = gy_i + \lambda ey = y_i \left[g + \lambda e \frac{y}{y_i} \right]$$

for all $|\lambda| \leq M$. We showed earlier that the set $\{g + \mu e \,|\, |\mu| \leq T\}$ is monochromatic under the χ^* colouring. We now note that $|\lambda(\frac{y}{y_i})| \leq MR^{n-1} = T$ since $|\lambda| \leq M$, $y_i \in [R]$ and $y = lcm(y_1, \ldots, y_n)$ so $|\frac{y}{y_i}| \leq R^{n-1}$. We can therefore, say the set $\{g + \lambda(\frac{y}{y_i})e \,|\, |\lambda| \leq M, i = 1, 2, \ldots, n\}$ is monochromatic under the χ^* colouring. That is,

$$\chi^*(g) = \chi^*\left(g + \lambda e\left(\frac{y}{y_1}\right)\right) = \chi^*\left(g + \lambda e\left(\frac{y}{y_2}\right)\right) = \cdots = \chi^*\left(g + \lambda e\left(\frac{y}{y_n}\right)\right)$$

for $|\lambda| \leq M$. From the definition of the colouring χ^* we can see

$$\begin{aligned}
\chi(y_1 g) &= \chi\left(y_1\left[g + \lambda e\left(\frac{y}{y_1}\right)\right]\right),\\
\chi(y_2 g) &= \chi\left(y_2\left[g + \lambda e\left(\frac{y}{y_2}\right)\right]\right),\\
&\vdots\\
\chi(y_n g) &= \chi\left(y_n\left[g + \lambda e\left(\frac{y}{y_n}\right)\right]\right).
\end{aligned}$$

We have already shown that $a_i + \lambda d = y_i[g + \lambda e(\frac{y}{y_i})]$ and so we can now write

$$\begin{aligned}
\chi(y_1 g) &= \chi(a_1 + \lambda d),\\
\chi(y_2 g) &= \chi(a_2 + \lambda d),\\
&\vdots\\
\chi(y_n g) &= \chi(a_n + \lambda d).
\end{aligned}$$

We previously stated that the set, $\{gy_1, gy_2 \dots, gy_n\} \in \mathfrak{F}$ is monochromatic under the χ^* colouring. So $\chi^*(gy_1) = \chi^*(gy_2) = \cdots = \chi^*(gy_n)$. This implies that $\chi(gy_1) = \chi(gy_2) = \cdots = \chi(gy_n)$ since if $\chi^*(\alpha) = \chi^*(\beta)$ then $\chi(\alpha) = \chi(\beta)$. Now we may say that for $|\lambda| \leq M$, $\chi(a_1 + \lambda d) = \chi(a_2 + \lambda d) = \cdots = \chi(a_n + \lambda d)$ as desired. □

Corollary 4.5.9. *Let $\mathfrak{F}$, a family of subsets of $\mathbb{N}$, be homogeneous and regular. Let M and c be positive integers. In any finite colouring of $\mathbb{N}$ there must exist some $A \in \mathfrak{F}$ and some positive integer, d, such that*

$$\{a + \lambda d \,|\, a \in A,\ |\lambda| \leq M\} \cup \{cd\}$$

is a monochromatic subset of $\mathbb{N}$.

Proof. We first define $T = T(M, r, c)$ to be a natural number such that for any r-colouring of $[T]$ there must exist $d \in \mathbb{N}$ and $A \in \mathfrak{F}$ such that the set $\{a + \lambda d \,|\, a \in A,\ |\lambda| \leq M\} \cup \{cd\}$ is monochromatic. We shall prove the result by proving that T exists for all $r \in \mathbb{N}$, using a proof by induction on the number of colours, r. First, suppose that $r = 1$. Trivially $T(M, 1, c) = \max\{m_{\mathfrak{F}} + M, c\}$ where $m_{\mathfrak{F}} = \min\{\max\{a \in A\} \,|\, A \in \mathfrak{F}\}$. Indeed, we only colour $[T(M, 1, c)]$ with one colour so any set in $[T(M, 1, c)]$ must be monochromatic. Therefore we can take $d = 1$ and $A \in \mathfrak{F}$ such that $m_{\mathfrak{F}} = \max\{a \in A\}$. If $m_{\mathfrak{F}} + M \leq c$ then we must take $T(M, 1, c) = c$ to ensure $cd = c \in [T(M, 1, c)]$. If $c \leq m_{\mathfrak{F}} + M$ then we must take $T(M, 1, c) = m_{\mathfrak{F}} + M$ to ensure $a + \lambda d = a + \lambda \leq m_{\mathfrak{F}} + M \in [T(M, 1, c)]$ for all $a \in A$.

We now assume that $r > 1$ and $T = T(M, r-1, c)$ exists. We must prove that $T' = T(M, r, c)$ exists.

By Theorem 4.5.8 in any r-colouring of $\mathbb{N}$ there exists $A \in \mathfrak{F}$ and $d' > 0$ such that

$$\{a + \lambda d' \,|\, a \in A,\ |\lambda| \leq TM\}$$

is monochromatic. If we can find some $\mu \leq T$ such that $\{a+\lambda d' \mid a \in A,\ |\lambda| \leq TM\} \cup \{\mu c d'\}$ is monochromatic then, by setting $d = \mu d'$, we have a monochromatic set

$$\{a + \lambda d \mid a \in A,\ |\lambda| \leq M\} \cup \{cd\}$$

of the desired form. However, if there is no such μ then $\{\mu c d' \mid \mu \leq T\}$ is at most $(r-1)$-coloured. From the definition of $T(M, r-1, s)$, which we already assumed to exist, we have the desired result. Indeed, we may define a new $(r-1)$-colouring of $[T(M, r-1, c)]$ by colouring every number x in this set the same colour as xcd'. There must exist $d \in \mathbb{N}$ and $A \in \mathfrak{F}$ such that the set, $\{a+\lambda d \mid a \in A,\ |\lambda| \leq M\} \cup \{cd\} \subseteq [T(M, r-1, c)]$ is monochromatic from the definition of $T(M, r-1, c)$. Therefore, from the definition of the new colouring of $[T(M, r-1, c)]$, the set $\{a + \lambda c d' d \mid a \in cd'A,\ |\lambda| \leq M\} \cup \{c^2 d d'\}$ must be monochromatic in the old colouring. Since $\mathfrak{F}$ is homogeneous we must have that $cd'A \in \mathfrak{F}$. Therefore we have the desired monochromatic set, where d is taken to be $cd'd$ and A is $cd'A$ in the statement of Corollary 4.5.9. □

To complete the proof of Rado's Theorem we must introduce (m, p, c)-sets.

Definition 4.5.10.

$$\mathbb{Z}_{m,p,c} = \{(\lambda_1, \ldots, \lambda_{m+1}) \in \mathbb{Z}^{(m+1)} \mid \ \textit{some } \lambda_i \neq 0\textit{, the first non-zero } \lambda_i = c, \\ |\lambda_j| \leq p \textit{ for all other } \lambda_j\}.$$

Definition 4.5.11. *An (m, p, c)-set, generated by the set of positive integers, $\{y_1, y_2, \ldots, y_{m+1}\}$, is a set S of the form*

$$S = \left\{ \sum_{i=1}^{m+1} \lambda_i y_i \,\middle|\, (\lambda_1, \ldots, \lambda_{m+1}) \in \mathbb{Z}_{m,p,c} \right\}.$$

Example. A $(1, 2, 2)$-set generated by y_1 and y_2 is

$$S = \left\{ \sum_{i=1}^{2} \lambda_i y_i \,\middle|\, (\lambda_1, \lambda_2) \in \mathbb{Z}_{1,2,2} \right\},$$

where

$$\mathbb{Z}_{1,2,2} = \{(2,0), (2,1), (2,-1), (2,2), (2,-2), (0,2)\}.$$

That is, $S = \{2y_1, 2y_1 + y_2, 2y_1 - y_2, 2y_1 + 2y_2, 2y_1 - 2y_2, 2y_2\}$.

Example. A $(2, 2, 1)$-set generated by y_1, y_2 and y_3 is

$$S = \left\{ \sum_{i=1}^{3} \lambda_i y_i \,\middle|\, (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{Z}_{2,2,1} \right\},$$

where

$$\begin{aligned}\mathbb{Z}_{2,2,1} =\{&(1,-2,-2),(1,-2,-1),(1,-2,0),(1,-2,1),(1,-2,2),\\ &(1,-1,-2),(1,-1,-1),(1,-1,0),(1,-1,1),(1,-1,2),\\ &(1,0,-2),(1,0,-1),(1,0,0),(1,0,1),(1,0,2),\\ &(1,1,-2),(1,1,-1),(1,1,0),(1,1,1),(1,1,2),\\ &(1,2,-2),(1,2,-1),(1,2,0),(1,2,1),(1,2,2),\\ &(0,1,-2),(0,1,-1),(0,1,0),(0,1,1),(0,1,2),\\ &(0,0,1)\}.\end{aligned}$$

If the generators, $y_1, y_2, \ldots, y_{m+1}$, of an (m,p,c)-set, S, satisfy $y_1 \gg y_2 \gg \cdots \gg y_{m+1} \geq 0$ then $S \subseteq \mathbb{N}$.

Theorem 4.5.12. *Let $m, p, c \in \mathbb{N}$. If $\mathbb{N}$ is coloured with a finite number of colours there exists a monochromatic (m,p,c)-set, S.*

Proof. We will prove this theorem by induction on m.

For the case in which $m = 1$ we refer to Corollary 4.3.2. We first take an r-colouring of $\mathbb{N}$, say χ. We now define an r-colouring of $\mathbb{N}$, say $\bar{\chi}$, in which we colour x with the colour given to cx under the χ colouring, for all $x \in \mathbb{N}$. A general $(1,p,c)$-set takes the form

$$S = \{cy_1, cy_1 + y_2, cy_1 - y_2, \ldots, cy_1 + py_2, cy_1 - py_2, cy_2\},$$

where

$$\mathbb{Z}_{1,p,c} = \{(c,0),(c,1),(c,-1),\ldots,(c,p),(c,-p),(0,c)\}.$$

Setting $k = p+1$ and $s = c$ in Corollary 4.3.2 we may deduce that under the $\bar{\chi}$ colouring there exists a monochromatic set of the form $\{a+\lambda d \,|\, |\lambda| \leq p\} \cup \{cd\}$. From the definition of our $\bar{\chi}$ colouring, the set $\{ca + \lambda cd \,|\, |\lambda| \leq p\} \cup \{c^2 d\}$ must also be monochromatic under the χ colouring. We may now set $a = y_1$ and $cd = y_2$ and we have shown that the monochromatic set $\{cy_1 + \lambda y_2 \,|\, |\lambda| \leq p\} \cup \{cy_2\}$ can be found in any r-colouring of $\mathbb{N}$. Since this set is a $(1,p,c)$-set we have shown that in any r-colouring of $\mathbb{N}$ there must exist a monochromatic $(1,p,c)$-set for any $p, c \in \mathbb{N}$.

We now assume that if $\mathbb{N}$ is coloured with a finite number of colours there exists a monochromatic (m,p,c)-set and we must prove that an $(m+1,p,c)$-set exists.

We may take $\mathfrak{F}$ to be the family of (m,p,c)-sets. $\mathfrak{F}$ is regular from the induction hypothesis. $\mathfrak{F}$ is homogeneous since multiplying the elements of an (m,p,c)-set by some $a \in \mathbb{N}$ will produce an (m,p,c)-set. The generators of the new (m,p,c)-set will be $ay_1, ay_2, \ldots, ay_{m+1}$, where previously they were $y_1, y_2, \ldots, y_{m+1}$. Therefore, from Corollary 4.5.9 if $\mathbb{N}$ is coloured with a finite number of colours there must exist a monochromatic $(m+1,p,c)$-set S. Indeed, we replace the M in Corollary 4.5.9 with p, d now becomes y_{m+1} and A represents our (m,p,c)-set. $\square$

To complete the proof of Rado's Theorem we now prove Theorem 4.5.13, the converse of Lemma 4.5.5.

Theorem 4.5.13. *If C satisfies the Columns condition then the system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$.*

Proof. We prove this theorem by producing a monochromatic parametric solution to the system $C\mathbf{x} = \mathbf{0}$. We will produce a solution of the form

$$\mathbf{x} = \mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \cdots + \mathbf{a}_n y_n,$$

where $\mathbf{a}_i \in \mathbb{Z}^m$ and $y_j \in \mathbb{Z}^+$. We then will show that in any finite colouring of $\mathbb{N}$ there exist $y_1, y_2, \ldots, y_n \in \mathbb{N}$ such that $\mathbf{x} \in \mathbb{N}^m$, and such that $\mathbf{x}$ is a monochromatic solution to $C\mathbf{x} = \mathbf{0}$. We note that the matrix C satisfies the Columns condition, and assume that C has n partition sets. We now label the columns of the matrix C as $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{k_n}$, and the n partition sets as $\mathcal{C}_1 = \{\mathbf{c}_1, \ldots, \mathbf{c}_{k_1}\}$, $\mathcal{C}_2 = \{\mathbf{c}_{k_1+1}, \ldots, \mathbf{c}_{k_2}\}, \ldots, \mathcal{C}_n = \{\mathbf{c}_{k_{n-1}+1} + \cdots + \mathbf{c}_{k_n}\}$. That is,

$$\begin{aligned}
\mathbf{c}_1 + \cdots + \mathbf{c}_{k_1} &= \mathbf{0}, \\
\mathbf{c}_{k_1+1} + \cdots + \mathbf{c}_{k_2} &= \alpha_1 \mathbf{c}_1 + \cdots + \alpha_{k_1} \mathbf{c}_{k_1}, \\
&\vdots \\
\mathbf{c}_{k_{n-1}+1} + \cdots + \mathbf{c}_{k_n} &= \gamma_1 \mathbf{c}_1 + \cdots + \gamma_{k_{n-1}} \mathbf{c}_{k_{n-1}}.
\end{aligned}$$

We now define the vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n \in \mathbb{Q}^{k_n}$ as follows,

$$\begin{aligned}
\mathbf{x}^1 &= (x_1^1, x_2^1, \ldots, x_{k_1}^1, x_{k_1+1}^1, \ldots, x_{k_2}^1, \ldots, x_{k_{n-1}+1}^1, \ldots, x_{k_n}^1)^T, \\
&= (1, 1, \ldots, 1, 0, \ldots, 0, \ldots, 0, \ldots, 0)^T, \\
\mathbf{x}^2 &= (x_1^2, x_2^2, \ldots, x_{k_1}^2, x_{k_1+1}^2, \ldots, x_{k_2}^2, \ldots, x_{k_{n-1}+1}^2, \ldots, x_{k_n}^2)^T, \\
&= (-\alpha_1, -\alpha_2, \ldots, -\alpha_{k_1}, 1, \ldots, 1, \ldots, 0, \ldots, 0)^T, \\
&\vdots \\
\mathbf{x}^n &= (x_1^n, x_2^n, \ldots, x_{k_1}^n, x_{k_1+1}^n, \ldots, x_{k_2}^n, \ldots, x_{k_{n-1}+1}^n, \ldots, x_{k_n}^n)^T, \\
&= (-\gamma_1, -\gamma_2, \ldots, -\gamma_{k_1}, -\gamma_{k_1+1}, \ldots, -\gamma_{k_2}, \ldots, 1, \ldots, 1)^T.
\end{aligned}$$

Note that each of these vectors, $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n$, is a solution to the system $C\mathbf{x} = \mathbf{0}$. We multiply these vectors by an appropriate $c \in \mathbb{N}$, to produce n integer solutions. We call these n solutions $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. Thus

$$\begin{aligned}
\mathbf{a}_1 &= c\mathbf{x}^1 = (c, c, \ldots, c, 0, \ldots, 0, \ldots, 0, \ldots, 0)^T, \\
\mathbf{a}_2 &= c\mathbf{x}^2 = (-c\alpha_1, -c\alpha_2, \ldots, -c\alpha_{k_1}, c, \ldots, c, \ldots, 0, \ldots, 0)^T, \\
&\vdots \\
\mathbf{a}_n &= c\mathbf{x}^n = (-c\gamma_1, -c\gamma_2, \ldots, -c\gamma_{k_1}, -c\gamma_{k_1+1}, \ldots, -c\gamma_{k_2}, \ldots, c, \ldots, c)^T.
\end{aligned}$$

We are now able to produce our parametric solution to $C\mathbf{x} = \mathbf{0}$,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k_1} \\ x_{k_1+1} \\ \vdots \\ x_{k_2} \\ \vdots \\ x_{k_{n-1}} \\ \vdots \\ x_{k_n} \end{pmatrix} = \begin{pmatrix} c \\ c \\ \vdots \\ c \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} y_1 + \begin{pmatrix} -c\alpha_1 \\ -c\alpha_2 \\ \vdots \\ -c\alpha_{k_1} \\ c \\ \vdots \\ c \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} y_2 + \cdots + \begin{pmatrix} -c\gamma_1 \\ -c\gamma_2 \\ \vdots \\ -c\gamma_{k_1} \\ -c\gamma_{k_1+1} \\ \vdots \\ -c\gamma_{k_2} \\ \vdots \\ c \\ \vdots \\ c \end{pmatrix} y_n.$$

Indeed, $\mathbf{x}$ is a solution to $C\mathbf{x} = \mathbf{0}$, since it is a linear combination of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, which are all solutions to $C\mathbf{x} = \mathbf{0}$.

Clearly the first non-zero coefficient of each of the parametric solutions will be c. For example, $x_{k_2} = cy_2 + \ldots - c\gamma_{k_2}y_n$, here the coefficient of y_2 is c.

To prove that we can find a monochromatic solution of this form in any finite colouring of $\mathbb{N}$ we set p to be the largest modulus of an entry from the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ and $m = n - 1$. From Theorem 4.5.12, for any finite colouring of $\mathbb{N}$ we can find a monochromatic (m, p, c)-set, $S = \{\sum_{i=1}^{n} \lambda_i y_i \,|\, (\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}_{n-1,p,c}\}$. S must contain the parametric solution. Indeed, the modulus of the maximum coefficient of each y_i in our parametric solution is at most p, there are n terms y_i, and every leading non-zero coefficient in the parametric solution is c. Therefore the system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$. □

This completes the proof of Rado's Theorem: *a system $C\mathbf{x} = \mathbf{0}$ is regular over $\mathbb{N}$ if and only if C satisfies the Columns condition.*

We now give an example of how a particular parametric solution, used in the proof of Theorem 4.5.13, is found.

Example. Take the system $C\mathbf{x} = \mathbf{0}$ where

$$C = \begin{pmatrix} 2 & -2 & 0 & 2 & 0 & 0 & 1 \\ 3 & -2 & -1 & -9 & 12 & 0 & 4 \\ 1 & -1 & 0 & -1 & 2 & 0 & 1 \\ 0 & 1 & -1 & -10 & 10 & 2 & 0 \end{pmatrix},$$

giving the set of equations

$$\begin{aligned} 2x_1 - 2x_2 \quad\quad\quad + 2x_4 \quad\quad\quad\quad\quad + 1x_7 &= 0, \\ 3x_1 - 2x_2 - 1x_3 - 9x_4 + 12x_5 \quad\quad + 4x_7 &= 0, \\ 1x_1 - 1x_2 \quad\quad\quad - 1x_4 + 2x_5 \quad\quad + 1x_7 &= 0, \\ 1x_2 - 1x_3 - 10x_4 + 10x_5 + 2x_6 \quad\quad &= 0. \end{aligned} \tag{4.5.1}$$

We have

$$
\begin{aligned}
\mathbf{c}_1 &= (2,3,1,0)^T,\\
\mathbf{c}_2 &= (-2,-2,-1,1)^T,\\
\mathbf{c}_3 &= (0,-1,0,-1)^T,\\
\mathbf{c}_4 &= (2,-9,-1,-10)^T,\\
\mathbf{c}_5 &= (0,12,2,10)^T,\\
\mathbf{c}_6 &= (0,0,0,2)^T,\\
\mathbf{c}_7 &= (1,4,1,0)^T.
\end{aligned}
$$

The matrix, C, satisfies the Columns condition. The partitions of C are $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$, $\{\mathbf{c}_4, \mathbf{c}_5\}$ and $\{\mathbf{c}_6, \mathbf{c}_7\}$, since

$$
\begin{aligned}
\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 &= \mathbf{0},\\
\mathbf{c}_4 + \mathbf{c}_5 &= \mathbf{c}_1,\\
\mathbf{c}_6 + \mathbf{c}_7 &= \frac{1}{2}\mathbf{c}_1 + \frac{1}{2}\mathbf{c}_3 + \frac{1}{4}\mathbf{c}_5.
\end{aligned}
$$

Therefore we have solutions to the system,

$$
\begin{aligned}
\mathbf{x}^1 &= (1,1,1,0,0,0,0)^T,\\
\mathbf{x}^2 &= (-1,0,0,1,1,0,0)^T,\\
\mathbf{x}^3 &= (-\frac{1}{2},0,-\frac{1}{2},0,-\frac{1}{4},1,1)^T.
\end{aligned}
$$

We multiply these solution by 4 so that they are all integer solutions, making

$$
\begin{aligned}
\mathbf{a}_1 &= (4,4,4,0,0,0,0)^T,\\
\mathbf{a}_2 &= (-4,0,0,4,4,0,0)^T,\\
\mathbf{a}_3 &= (-2,0,-2,0,-1,4,4)^T,
\end{aligned}
$$

where $c = 4$. So our parametric solution is

$$
\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{pmatrix} = \begin{pmatrix} 4\\ 4\\ 4\\ 0\\ 0\\ 0\\ 0 \end{pmatrix} y_1 + \begin{pmatrix} -4\\ 0\\ 0\\ 4\\ 4\\ 0\\ 0 \end{pmatrix} y_2 + \begin{pmatrix} -2\\ 0\\ -2\\ 0\\ -1\\ 4\\ 4 \end{pmatrix} y_3. \tag{4.5.2}
$$

Since p is to be the largest modulus of an entry from the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $m = n - 1$ we have $p = 4$ and $m = 2$. In any finite colouring of $\mathbb{N}$ there is a monochromatic $(2,4,4)$-set, S, and this must contain a solution as in (4.5.2) and so a solution to the system (4.5.1). Indeed, $S = \{4y_1 - 4y_2 - 4y_3, 4y_1 - 4y_2 - 3y_3, \ldots, 4y_1 - 4y_2 + 3y_3, 4y_1 - 4y_2 + 4y_3, \ldots, 4y_1 - 3y_2 + 4y_3, \ldots, 4y_1 + 4y_2 + 4y_3, 4y_2 - 4y_3, \ldots, 4y_2 + 4y_3, \ldots, 4y_3\}$. In particular $4y_1 -$

$4y_2 - 2y_3, 4y_1, 4y_1 - 2y_3, 4y_2, 4y_2 - 1y_3, 4y_3 \in S$ and since $4y_1 - 4y_2 - 2y_3 = x_1, 4y_1 = x_2, 4y_1 - 2y_3 = x_3, 4y_2 = x_4, 4y_2 - 1y_3 = x_5, 4y_3 = x_6 = x_7$ we have that $x_1, x_2, x_3, x_4, x_5, x_6, x_7 \in S$. S is a monochromatic set so $\mathbf{x}$ is a monochromatic solution.

4.6 Consistency of finite linear systems

Definition 4.6.1. *Consider two regular systems, $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$. So in any finite colouring of $\mathbb{N}$ there exists a monochromatic vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$ and in every finite colouring of $\mathbb{N}$ there exists a monochromatic vector $\mathbf{y}$ such that $B\mathbf{y} = \mathbf{0}$. These two systems are said to be consistent if in every finite colouring of $\mathbb{N}$ there exist monochromatic vectors $\mathbf{x}$ and $\mathbf{y}$, which have the same colour, such that $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$.*

Example. We can easily see that

$$A = \begin{pmatrix} 1 & -1 & -2 \\ 2 & -2 & -4 \\ -3 & 3 & 6 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 0 & 1 & -1 & 2 \\ 1 & 0 & -1 & 3 \\ -1 & 2 & -1 & 1 \end{pmatrix}$$

satisfy the Columns condition and so by Rado's Theorem are regular. Therefore, in any any finite colouring of $\mathbb{N}$ we can find a monochromatic vector $\mathbf{x}$ and a monochromatic vector $\mathbf{y}$ such that $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$. However we can only say these two systems are consistent if we can always find monochromatic vectors $\mathbf{x}$ and $\mathbf{y}$, which have the same colour, such that $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$. It is completely analogous to say the systems are consistent if we can find a monochromatic vector $\mathbf{z}$ such that $C\mathbf{z} = \mathbf{0}$ where

$$C = \begin{pmatrix} \begin{pmatrix} & A & \end{pmatrix} & \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{pmatrix} & B & \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -2 & 0 & 0 & 0 & 0 \\ 2 & -2 & -4 & 0 & 0 & 0 & 0 \\ -3 & 3 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 & 3 \\ 0 & 0 & 0 & -1 & 2 & -1 & 1 \end{pmatrix}.$$

We can see C satisfies the Columns condition. Let $\mathbf{c}_1, \ldots, \mathbf{c}_7$ be the columns of C, we take the partitions to be $\mathcal{C}_1 = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_4, \mathbf{c}_5, \mathbf{c}_6\}$ and $\mathcal{C}_2 = \{\mathbf{c}_3, \mathbf{c}_7\}$. $\sum_{i=1}^{2} \mathbf{c}_i + \sum_{i=4}^{6} \mathbf{c}_i = 0$ and $\mathbf{c}_3 + \mathbf{c}_7 = 2\mathbf{c}_2 + \mathbf{c}_4 - 2\mathbf{c}_6$ so C satisfies the Columns condition and by Rado's Theorem is regular. Since C is regular we can certainly find a monochromatic vector, $\mathbf{z}$, in any r-colouring of $\mathbb{N}$ such that $C\mathbf{z} = \mathbf{0}$ and so $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$ are consistent.

Using Rado's Theorem we will now formally prove that finite linear systems are always consistent.

Theorem 4.6.2. *Any two finite regular linear homogeneous systems are consistent.*

Proof. We take two regular linear homogeneous systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{y} = \mathbf{0}$. From Rado's Theorem we can see that these matrices satisfy the Columns condition. Let $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m$ be the partition of the columns of A and $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$ the partition of the columns of B. Let the matrix C be the diagonal sum of A and B, that is $C = \left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$. We can now define a partition of the columns of C which satisfies the Columns condition. The first partition set of C is $\mathcal{C}_1 = \mathcal{A}_1 \cup \mathcal{B}_1$, the sum of the columns in this partition set must sum to zero since the columns in $\mathcal{A}_1$ and $\mathcal{B}_1$ sum to zero. The remaining partition sets are defined as $\mathcal{C}_2 = \mathcal{A}_2, \mathcal{C}_3 = \mathcal{A}_3, \ldots, \mathcal{C}_m = \mathcal{A}_m, \mathcal{C}_{m+1} = \mathcal{B}_2, \mathcal{C}_{m+2} = \mathcal{B}_3, \ldots, \mathcal{C}_{m+n-1} = \mathcal{B}_n$. It is clear that the sum of the vectors in each of these partitions sets can be written as a linear combination of the vectors in the partitions sets of lower index as they were in A and B. Therefore C is regular and A and B are consistent. □

4.7 The Finite Sums Theorem

The Finite Sums Theorem is also known as Folkman's Theorem, Sanders' Theorem and the Folkman-Rado-Sanders Theorem. The reason for the differing names is that various proofs, published and unpublished, have been produced by Sanders, Rado and Folkman. Sanders' proof was published in 1969 and Rado's in 1970.

To understand the Finite Sums Theorem Theorem we must first define a sum-set.

Definition 4.7.1. *Given a set $S \subseteq \mathbb{N}$,*

$$\mathcal{S}(S) = \left\{ \sum_{s \in S} z_s s \;\middle|\; z_s = 0 \text{ or } 1 \text{ for all } s \in S \text{ and } z_s = 1 \text{ for some non-zero finite number of } s \in S \right\}.$$

We call $\mathcal{S}(S)$ the sum-set for the set S.

Example. If $S = \{2, 4, 6, 8\}$ then $\mathcal{S}(S) = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$. If $S = \{2, 4, 5\}$ then $\mathcal{S}(S) = \{2, 4, 5, 6, 7, 9, 11\}$.

We state the Finite Sums Theorem below as Theorem 4.7.2.

Theorem 4.7.2. *In any r-colouring of $\mathbb{N}$ there exists some k-element set, S, such that $\mathcal{S}(S)$ is monochromatic.*

Proof. We shall first prove that the monochromatic set, $\mathcal{S}(S)$, exists for $k = 4$. We take the matrix

$$C = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

C satisfies the Columns condition, indeed, labeling C's columns $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{15}$, we see that $\mathbf{c}_4 + \sum_{i=9}^{15} \mathbf{c}_i = \mathbf{0}$, $\mathbf{c}_1 + \mathbf{c}_5 + \mathbf{c}_7 + \mathbf{c}_8 = -\mathbf{c}_9 - \mathbf{c}_{12} - \mathbf{c}_{14} - \mathbf{c}_{15}$, $\mathbf{c}_2 + \mathbf{c}_6 = -\mathbf{c}_5 - \mathbf{c}_8 - \mathbf{c}_{10} - \mathbf{c}_{12} - \mathbf{c}_{13} - \mathbf{c}_{15}$ and $\mathbf{c}_3 = -\sum_{i=6}^{8} \mathbf{c}_i - \mathbf{c}_{11} - \sum_{i=13}^{15} \mathbf{c}_i$. That is, the partition sets of C are $\mathcal{C}_1 = \{\mathbf{c}_4, \mathbf{c}_9, \mathbf{c}_{10}, \mathbf{c}_{11}, \mathbf{c}_{12}, \mathbf{c}_{13}, \mathbf{c}_{14}, \mathbf{c}_{15}\}$, $\mathcal{C}_2 = \{\mathbf{c}_1, \mathbf{c}_5, \mathbf{c}_7, \mathbf{c}_8\}$, $\mathcal{C}_3 = \{\mathbf{c}_2, \mathbf{c}_6\}$ and $\mathcal{C}_4 = \{\mathbf{c}_3\}$. From Rado's Theorem we may deduce that since C satisfies the Columns condition it is a regular matrix. Since C is regular there must exist some monochromatic vector, $\mathbf{x}$, such that $C\mathbf{x} = \mathbf{0}$. The sum-set, $\mathcal{S}(S)$, consist of all the elements of the vector $\mathbf{x}$. Indeed, $C\mathbf{x} = \mathbf{0}$ may also be written

$$\begin{aligned}
x_1 + x_2 \qquad\qquad\quad &= x_5, \\
x_2 + x_3 \qquad\quad &= x_6, \\
x_1 \qquad + x_3 \qquad\quad &= x_7, \\
x_1 + x_2 + x_3 \qquad\quad &= x_8, \\
x_1 \qquad\qquad\quad + x_4 &= x_9, \\
x_2 \qquad + x_4 &= x_{10}, \\
x_3 + x_4 &= x_{11}, \\
x_1 + x_2 \qquad + x_4 &= x_{12}, \\
x_2 + x_3 + x_4 &= x_{13}, \\
x_1 \qquad + x_3 + x_4 &= x_{14}, \\
x_1 + x_2 + x_3 + x_4 &= x_{15}.
\end{aligned}$$

From this set of equations we can see that, since $C\mathbf{x} = \mathbf{0}$ is a regular system, not only will the elements, x_1, x_2, x_3 and x_4 be monochromatic but also every sum in the sum-set, $\mathcal{S}(x_1, x_2, x_3, x_4)$.

The proof for a general k is clearly an extension of the example given here for $k = 4$. We must simply produce the matrix which represents the sums in the sum-set, these matrices will all have the Columns condition, since they will be of the same form as C. $\square$

5 Hindman's Theorem

In Neil Hindman's paper, *Finite sums from sequences within cells of a partition of* $\mathbb{N}$ [17], he proved the theorem which would later become known as Hindman's Theorem. Hindman's theorem states that *if* $\mathbb{N}$ *is coloured with some finite number of colours there must exist some infinite set* $S \subseteq \mathbb{N}$ *such that* $\mathcal{S}(S)$ *is monochromatic.* $\mathcal{S}(S)$ is defined as in the Finite Sums Theorem, in Definition 4.7.1.

To prove Hindman's Theorem we will use ultrafilters. Therefore, we will begin by defining filters and ultrafilters and studying some of their properties. Throughout this chapter we shall denote the complement of a set A by $\bar{A}$.

5.1 Filters and ultrafilters

Definition 5.1.1. *A filter on* $\mathbb{N}$ *is a collection* $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$ *such that*

- $\emptyset \notin \mathcal{F}$ *and* $\mathbb{N} \in \mathcal{F}$,
- *if* $A \in \mathcal{F}$ *and* $A \subset B$ *then* $B \in \mathcal{F}$,
- *if* $A, B \in \mathcal{F}$ *then* $A \cap B \in \mathcal{F}$.

Example. Some examples of filters are given by $\{A \subset \mathbb{N} \,|\, 1 \in A\}$, $\{A \subset \mathbb{N} \,|\, 1, 2 \in A\}$, $\{A \subset \mathbb{N} \,|\, B \subset A\}$, for some fixed $B \subset \mathbb{N}$. Another example is $\{A \subset \mathbb{N} \,|\, \bar{A}$ is finite $\}$, this is called the *cofinite filter* and denoted by $\mathcal{F}_{cof}$.

Definition 5.1.2. *An ultrafilter is a maximal filter.*

Lemma 5.1.3. *Every filter is contained in some ultrafilter.*

Proof. We shall prove that every filter is contained in a maximal filter using Zorn's Lemma. Zorn's Lemma states that *every partially ordered set, P, has a maximal element if P has the Zorn Property.* P has the the Zorn property if *for every chain C in P there is an upper bound $p \in P$ of C.*

For some filter $\mathcal{F}_1$ we take $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \cdots$ to be any chain of filters containing $\mathcal{F}_1$, and must find an upper bound of the chain. We will show that $\mathcal{F} = \bigcup_{i=1}^{\infty} \mathcal{F}_i$ is an upper bound of the chain. We first check that $\mathcal{F}$ is a filter. Clearly $\emptyset \notin \mathcal{F}$ since $\emptyset \notin \mathcal{F}_i$ for all i. Since $\mathbb{N} \in \mathcal{F}_i$ for all i, $\mathbb{N} \in \bigcup_{i=1}^{\infty} \mathcal{F}_i = \mathcal{F}$. If $A \in \mathcal{F}$ and $A \subset B$ then $A \in \mathcal{F}_i$ for some i, and since $\mathcal{F}_i$ is a filter $B \in \mathcal{F}_i$, but then $B \in \mathcal{F}$. If $A, B \in \mathcal{F}$ then $A \in \mathcal{F}_i$ for some i and $B \in \mathcal{F}_j$ for some j. If $i = j$ we are done, since $\mathcal{F}_i$ is a filter so $A \cap B \in \mathcal{F}_i$. If $i < j$ then since $\mathcal{F}_i \subseteq \mathcal{F}_j$ we must have that $A \in \mathcal{F}_j$, and again since $\mathcal{F}_j$ is a filter $A \cap B \in \mathcal{F}_j$. If $i > j$ then we must have that $B \in \mathcal{F}_i$, and $A \cap B \in \mathcal{F}_i$. Therefore $\mathcal{F}$ is a filter since is satisfies the definition of a filter given in Definition 5.1.1. Therefore the set of all filters containing $\mathcal{F}_1$ has a maximal element. A maximal filter is an ultrafilter, so every filter is contained in some ultrafilter. $\square$

Lemma 5.1.4. *$\mathcal{U}$ is an ultrafilter if and only if for all $A \subset \mathbb{N}$ either $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$.*

Proof. We shall prove this lemma using a proof by contradiction. For some ultrafilter, $\mathcal{U}$, we assume there exists a set A such that $A, \bar{A} \notin \mathcal{U}$. From the maximality of $\mathcal{U}$ there must exist some set $B \in \mathcal{U}$ such that $B \cap A = \emptyset$ and $B \cap \bar{A} = \emptyset$. It is then clear that $B \cap (A \cup \bar{A}) = \emptyset$ and $A \cup \bar{A} = \mathbb{N}$ but then $B \cap (A \cup \bar{A}) = B \cap \mathbb{N} = \emptyset$ so $B = \emptyset$ and $\emptyset \in \mathcal{U}$. This is a contradiction since the empty set is defined not to be in any filter and $\mathcal{U}$ is a filter. So one of A and $\bar{A}$ must be in $\mathcal{U}$. We cannot have that both A and $\bar{A}$ are subsets of a filter since if $A, \bar{A} \in \mathcal{F}$ then from the definition of a filter $A \cap \bar{A} = \emptyset \in \mathcal{F}$. The empty set is defined not to be in any filter. Therefore, either $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$.

We have shown that if $\mathcal{U}$ is an ultrafilter then for all $A \subset \mathbb{N}$ either $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$. We must now show that if for all sets $A \subset \mathbb{N}$ either $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$ then $\mathcal{U}$ is an ultrafilter.

We take some filter $\mathcal{U}$ which is not an ultrafilter and therefore is not a maximal filter. We assume, for all $A \subset \mathbb{N}$ either $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$. Since $\mathcal{U}$ is not a maximal filter there must exist some filter $\mathcal{V}$ such that $\mathcal{U} \subsetneqq \mathcal{V}$ and so there must exist some set, say B, such that $B \notin \mathcal{U}$ but $B \in \mathcal{V}$. From our assumption we have that if $B \notin \mathcal{U}$ then $\bar{B} \in \mathcal{U}$, however, $\mathcal{U} \subsetneqq \mathcal{V}$ so $\bar{B} \in \mathcal{V}$. This is a contradiction since both B and $\bar{B}$ cannot both be in $\mathcal{V}$ from our assumption. Therefore $\mathcal{U}$ must be a maximal filter, an ultrafilter. $\square$

Proposition 5.1.5. *For every ultrafilter, $\mathcal{U}$ we have that $A_1 \cup A_2 \in \mathcal{U}$ if and only if $A_1 \in \mathcal{U}$ or $A_2 \in \mathcal{U}$.*

Proof. We assume that $A_1 \cup A_2 \in \mathcal{U}$ but $A_1, A_2 \notin \mathcal{U}$. Since $\mathcal{U}$ is an ultrafilter from Lemma 5.1.4 we must have $\bar{A}_1, \bar{A}_2 \in \mathcal{U}$. Taking $A_3 = \bar{A}_1 \cap \bar{A}_2$ we have that $A_3 \in \mathcal{U}$ but, using De Morgan's laws, $\bar{A}_3 = A_1 \cup A_2$, and we already have that $A_1 \cup A_2 \in \mathcal{U}$ so $\bar{A}_3 \in \mathcal{U}$. Since $\mathcal{U}$ is an filter, $\bar{A}_3$ and A_3 cannot both be in $\mathcal{U}$. We have a contradiction and so $A_1 \in \mathcal{U}$ or $A_2 \in \mathcal{U}$. From the definition of an ultrafilter we can see that if $A_1 \in \mathcal{U}$ or $A_2 \in \mathcal{U}$ then since $A_1 \subset A_1 \cup A_2$ and $A_2 \subset A_1 \cup A_2$ we have $A_1 \cup A_2 \in \mathcal{U}$. $\square$

Lemma 5.1.6. *Suppose that $\mathcal{U}$ is an ultrafilter and $A_1, A_2, \ldots, A_n$ are subsets of $\mathbb{N}$. Then the subsets of $\mathbb{N}$, $A_1, A_2, \ldots, A_n$, are all in $\mathcal{U}$ if and only if $A_1 \cap A_2 \cap \cdots \cap A_n \in \mathcal{U}$.*

Proof. We shall prove the case for two sets, A_1 and A_2, and then three sets, A_1, A_2 and A_3, the general case follows similarly.

From the definition of a filter we have, if $A_1, A_2 \in \mathcal{U}$ then $A_1 \cap A_2 \in \mathcal{U}$. If $A_1 \cap A_2 \in \mathcal{U}$ then since $A_1 \cap A_2 \subset A_1$ and $A_1 \cap A_2 \subset A_2$ we may say, again from the definition of a filter, $A_1, A_2 \in \mathcal{U}$.

Again, from the definition of a filter we have, if $A_1, A_2, A_3 \in \mathcal{U}$ then $A_1 \cap A_2 \in \mathcal{U}$, but then $(A_1 \cap A_2), A_3 \in \mathcal{U}$ so we must have that $A_1 \cap A_2 \cap A_3 \in \mathcal{U}$. If $A_1 \cap A_2 \cap A_3 \in \mathcal{U}$ then since $A_1 \cap A_2 \cap A_3 \subset A_1$, $A_1 \cap A_2 \cap A_3 \subset A_2$ and $A_1 \cap A_2 \cap A_3 \subset A_3$ we may say, again from the definition of a filter, $A_1, A_2, A_3 \in \mathcal{U}$. $\square$

Definition 5.1.7. $\tilde{n} = \{A \subset \mathbb{N} \mid n \in A\}$ *is called the principal ultrafilter at* n.

Example. The principal ultrafilter at 1 is $\{A \subset \mathbb{N} \mid 1 \in A\}$.

There is clearly a bijective mapping between the elements of $\mathbb{N}$ and the principal ultrafilters. Under this mapping $n \in \mathbb{N}$ is mapped to the principal ultrafilter, $\tilde{n}$. This is a bijective mapping since any principal ultrafilter is mapped onto by exactly one element of $\mathbb{N}$. This allows us, in some circumstances, to use n and $\tilde{n}$ interchangeably.

Proposition 5.1.8. *If* $A_1, \dots, A_n$ *are subsets of* $\mathbb{N}$ *and* $\mathcal{U}$ *is an ultrafilter then* $A_1 \cup \cdots \cup A_n \in \mathcal{U}$ *if and only if there exists some* $i \in \{1, 2, \dots, n\}$ *such that* $A_i \in \mathcal{U}$.

Proof. We set $A = A_1 \cup \cdots \cup A_n$ and assume that $A \in \mathcal{U}$. It is clear that $A = A_1 \cup (A_2 \cup \cdots \cup A_n) \in \mathcal{U}$ and so Proposition 5.1.5 implies $A_1 \in \mathcal{U}$ or $A_2 \cup \cdots \cup A_n \in \mathcal{U}$. If $A_1 \notin \mathcal{U}$ then $A_2 \cup \cdots \cup A_n \in \mathcal{U}$. We may now write $A_2 \cup \cdots \cup A_n = A_2 \cup (A_3 \cup \cdots \cup A_n)$. Again, Proposition 5.1.5 implies that $A_2 \in \mathcal{U}$ or $A_3 \cup \cdots \cup A_n \in \mathcal{U}$. We can continue in this way. Since A consists of the union of a finite number of sets, A_i, we know that at some stage there must exist some $i \in \{1, \dots, n\}$ such that $A_i \in \mathcal{U}$.

If there exists some $i \in \{1, \dots, n\}$ such that $A_i \in \mathcal{U}$ then from the definition of a filter, since $A_i \subset (A_1 \cup \cdots \cup A_n)$, we must have that $A_1 \cup \cdots \cup A_n \in \mathcal{U}$. $\square$

Corollary 5.1.9. *If* $\mathcal{U}$ *is an ultrafilter and* A *is a finite set such that* $A \in \mathcal{U}$ *then there exists some* $n \in A$ *such that* $\mathcal{U} = \tilde{n}$.

Proof. Since $A = \{n_1, n_2, \dots, n_k\} = \{n_1\} \cup \{n_2\} \cup \cdots \cup \{n_k\}$, from Proposition 5.1.8 we have that for some $i \in \{1, 2, \dots, n\}$, $\{n_i\} \in \mathcal{U}$. However, since $\{n_i\} \in \mathcal{U}$, we must have, from the definition of a filter, that for every set $B \subset \mathbb{N}$ such that $n_i \in B$, $B \in \mathcal{U}$. Moreover, if $B \subseteq \mathbb{N}$ and $n_i \notin B$ then $B \notin \mathcal{U}$ since otherwise, from the definition of a filter, $B \cap \{n_i\} = \emptyset \in \mathcal{U}$. Thus $\mathcal{U} = \tilde{n}_i$. $\square$

Proposition 5.1.10. *Any ultrafilter,* $\mathcal{U}$*, containing the cofinite filter,* $\mathcal{F}_{cof}$*, is non-principal.*

Proof. We assume $\mathcal{U}$ is a principal ultrafilter containing the $\mathcal{F}_{cof}$. Then $\tilde{n} = \mathcal{U}$ for some $n \in \mathbb{N}$ and from the definition of $\tilde{n}$, $\{n\} \in \mathcal{U}$. The complement of $\{n\}$ is $\mathbb{N} \setminus \{n\}$, an infinite set, so $\{n\} \notin \mathcal{F}_{cof}$ and $\mathbb{N} \setminus \{n\} \in \mathcal{F}_{cof}$. $\mathcal{F}_{cof} \subset \mathcal{U}$ and so $\mathbb{N} \setminus \{n\} \in \mathcal{U}$. $\mathcal{U}$ is a filter and thus $\{n\}$ and its complement, $\mathbb{N} \setminus \{n\}$, cannot both be in $\mathcal{U}$, since otherwise $\{n\} \cap \mathbb{N} \setminus \{n\} = \emptyset \in \mathcal{U}$. We have a contradiction and so $\mathcal{U}$ must be a non-principal ultrafilter. $\square$

Proposition 5.1.11. *If* $\mathcal{U}$ *is a non-principal ultrafilter then* $\mathcal{F}_{cof} \subseteq \mathcal{U}$.

Proof. We assume there exists some set A where $A \notin \mathcal{U}$ and $\bar{A}$ is finite. Since $\mathcal{U}$ is an ultrafilter Lemma 5.1.4 implies that $\bar{A} \in \mathcal{U}$. Since $\bar{A}$ is a finite set and $\bar{A} \in \mathcal{U}$, Proposition 5.1.9 implies that there exists some $n \in \bar{A}$ such that $\tilde{n} = \mathcal{U}$. This is a contradiction since $\mathcal{U}$ is a non-principal ultrafilter. Therefore, $A \in \mathcal{U}$ for every set A whose complement, $\bar{A}$, is finite. This is the definition of the cofinite filter. So $\mathcal{F}_{cof} \subseteq \mathcal{U}$. $\square$

Proposition 5.1.10 and Proposition 5.1.11 together imply that $\mathcal{F}_{cof} \subseteq \mathcal{U}$ if and only if $\mathcal{U}$ is a non-principal ultrafilter.

5.2 A topology on $\beta\mathbb{N}$

We must now produce a topological space on the set of all ultrafilters on $\mathbb{N}$. This will enable us to prove results concerning the set of all ultrafilters on $\mathbb{N}$ which we will need in the proof of Hindman's Theorem.

Definition 5.2.1. *The set of all ultrafilters on* $\mathbb{N}$ *is called* $\beta\mathbb{N}$.

We define a topology on $\beta\mathbb{N}$ by taking basic open sets of the form $C_A = \{\mathcal{U} \in \beta\mathbb{N} \mid A \in \mathcal{U}\}$ for $A \subseteq \mathbb{N}$.

Lemma 5.2.2. *Let* $A_1, A_2, \ldots, A_n$ *be subsets of* $\mathbb{N}$. *Then* $C_{A_1} \cap C_{A_2} \cap \cdots \cap C_{A_n} = C_{A_1 \cap A_2 \cap \cdots \cap A_n}$, *a basic open set.*

Proof. We shall prove the case for two basic open sets, C_{A_1} and C_{A_2}, and then three basic open sets, C_{A_1}, C_{A_2} and C_{A_3}, the general case follows similarly.

$C_{A_1} \cap C_{A_2} = \{\mathcal{U} \in \beta\mathbb{N} \mid A_1 \in \mathcal{U}\} \cap \{\mathcal{U} \in \beta\mathbb{N} \mid A_2 \in \mathcal{U}\} = \{\mathcal{U} \in \beta\mathbb{N} \mid A_1 \cap A_2 \in \mathcal{U}\} = C_{A_1 \cap A_2}$. Indeed, this holds since from Lemma 5.1.6 we have that $A_1, A_2 \in \mathcal{U}$ if and only if $A_1 \cap A_2 \in \mathcal{U}$. So the set of ultrafilters containing A_1 and A_2 is precisely the set of ultrafilters containing $A_1 \cap A_2$. So $C_{A_1} \cap C_{A_2} = C_{A_1 \cap A_2}$. Now, $(C_{A_1} \cap C_{A_2}) \cap C_{A_3} = C_{A_1 \cap A_2} \cap C_{A_3}$, since $C_{A_1 \cap A_2}$ is a basic open set we may say $C_{A_1 \cap A_2} \cap C_{A_3} = C_{A_1 \cap A_2 \cap A_3}$. Clearly we may generalise this proof for any finite number of basic open sets. □

The union of a collection of basic open sets in the topology, $C_{A_1}, C_{A_2}, C_{A_3} \ldots$, is open since they are open sets. That is, $C_{A_1} \cup C_{A_2} \cup C_{A_3} \ldots$ is an open set for any basic open sets $C_{A_1}, C_{A_2}, C_{A_3} \ldots$. The intersection of finitely many basic open sets is a basic open itself. Therefore the open sets in $\beta\mathbb{N}$ are the basic open sets and their unions.

Proposition 5.2.3. $\beta\mathbb{N} \setminus C_A = C_{\bar{A}}$.

Proof. As proved in Proposition 5.1.4 for any set $A \subset \mathbb{N}$, every ultrafilter must contain A or $\bar{A}$. So $\beta\mathbb{N} \setminus C_A$, the set of all ultrafilters which do not contain A, must be the set of all ultrafilters which do contain $\bar{A}$, the set $C_{\bar{A}}$. □

The complement of an open set is, by definition, a closed set and so our basic open sets are clopen.

Proposition 5.2.4. *Let* $A_1, A_2, \ldots, A_n$ *be subsets of* $\mathbb{N}$. *Then* $C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n} = C_{A_1 \cup A_2 \cup \cdots \cup A_n}$ *and* $\beta\mathbb{N} \setminus \{C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n}\} = C_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n}$.

Proof. $C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n}$ is the set of ultrafilters which contain at least one of the sets $A_1, A_2, \ldots, A_n$. From Proposition 5.1.8, an ultrafilter contains A_i for some $i \in \{1, 2, \ldots, n\}$ if and only if it contains $A_1 \cup A_2 \cup \cdots \cup A_n$. Therefore the set of ultrafilters $C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n}$ is equal to the set of ultrafilters $C_{A_1 \cup A_2 \cup \cdots \cup A_n}$. Therefore the set $\beta\mathbb{N} \setminus \{C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n}\} =$

$\beta\mathbb{N} \setminus C_{A_1 \cup A_2 \cup \cdots \cup A_n}$. We may set $A = A_1 \cup A_2 \cup \cdots \cup A_n$ and then from Proposition 5.2.3 we have $\beta\mathbb{N} \setminus C_{A_1 \cup A_2 \cup \cdots \cup A_n} = C_{\overline{A_1 \cup A_2 \cup \cdots \cup A_n}}$, but then, from De Morgan's Laws, we have that $\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n$. Therefore $C_{\overline{A_1 \cup A_2 \cup \cdots \cup A_n}} = C_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n}$, and so $\beta\mathbb{N} \setminus \{C_{A_1} \cup C_{A_2} \cup \cdots \cup C_{A_n}\} = C_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n}$. □

We note that Proposition 5.2.4 implies that a finite union of basic open sets in also basic open set.

Definition 5.2.5. *A subset, A, of a topological space, X, is dense if every non-empty open subset of X intersects A.*

Lemma 5.2.6. *$\mathbb{N}$ is dense in $\beta\mathbb{N}$.*

Proof. We recall that every element of $\mathbb{N}$ may be identified with a principal ultrafilter. Consider any basic open set, C_A, where $A \neq \emptyset$, then for every $n \in A$, we have $A \in \tilde{n}$ and thus $\tilde{n} \in C_A$. Since $\tilde{n} \in C_A$ we have that C_A intersects $\mathbb{N}$. For any $n \in I$ where $I \in \{A_1, A_2, A_3, \ldots\}$ we have $\tilde{n} \in C_I$ so $\tilde{n} \in C_{A_1} \cup C_{A_2} \cup C_{A_3} \ldots$. Therefore the union of any basic open sets intersect $\mathbb{N}$. So the basic open sets in $\beta\mathbb{N}$ and the union of any number of those sets both intersect $\mathbb{N}$. So every non-empty open set in $\beta\mathbb{N}$ intersects $\mathbb{N}$ and $\mathbb{N}$ is dense in $\beta\mathbb{N}$. □

Definition 5.2.7. *A neighbourhood of a point, x, in a topological space, X, is any open set containing x.*

Definition 5.2.8. *An element, x, of a topological space, X, is said to be an isolated point, if there exists a neighbourhood of x, say U, such that $X \cap U = \{x\}$. That is, the one-point set $\{x\}$ is open in X.*

Lemma 5.2.9. *Every point of $\mathbb{N}$ is isolated in $\beta\mathbb{N}$.*

Proof. For any element $n \in \mathbb{N}$ we can take

$$\begin{aligned} C_{\{n\}} &= \{\mathcal{U} \in \beta\mathbb{N} \mid \{n\} \in \mathcal{U}\}, \\ &= \{\tilde{n}\}, \end{aligned}$$

where $\tilde{n}$ is the principal ultrafilter at n. $C_{\{n\}}$ is an open set and contains only $\tilde{n}$, so the definition of an isolated point is satisfied. Since n was an arbitrary element of $\mathbb{N}$, every point in $\mathbb{N}$ must be isolated in $\beta\mathbb{N}$. □

Lemma 5.2.10. *For a topological space, X, and a collection of closed sets, $(F_i)_{i \in I}$,*

$$\bigcap_{i \in I} F_i = \emptyset \Leftrightarrow (U_i)_{i \in I} \text{ is an open cover of } X$$

where $U_i = X \setminus F_i$.

Proof. We first assume that $\bigcap_{i\in I} F_i = \emptyset$. By De Morgan's laws we have

$$\bigcup_{i\in I}(X \setminus F_i) = X \setminus \bigcap_{i\in I} F_i.$$

Now, since $U_i = X \setminus F_i$ and $\bigcap_{i\in I} F_i = \emptyset$

$$\bigcup_{i\in I}(X \setminus F_i) = X \setminus \emptyset, \text{ and so}$$
$$\bigcup_{i\in I} U_i = X.$$

Thus the collection of open sets, $(U_i)_{i\in I}$, is an open cover of X.

We now assume the collection of open sets, $(U_i)_{i\in I}$, is an open cover of X and we must prove that $\bigcap_{i\in I} F_i = \emptyset$. Again by De Morgan's law,

$$\bigcup_{i\in I}(X \setminus F_i) = X \setminus \bigcap_{i\in I} F_i \text{ , therefore}$$
$$\bigcup_{i\in I} U_i = X \setminus \bigcap_{i\in I} F_i.$$

Since $(U_i)_{i\in I}$ is an open cover of X we must have $\bigcup_{i\in I}(U_i) = X$ but then $X = X \setminus \bigcap_{i\in I} F_i$ so we must have $\bigcap_{i\in I} F_i = \emptyset$. $\square$

Definition 5.2.11. *A collection, $(F_i)_{i\in I}$, of closed sets of a topological space, X, has the finite intersection property if the intersection of any finite number of these sets is non-empty. That is, if for any finite $J \subset I$ we have that*

$$\bigcap_{j\in J} F_j \neq \emptyset.$$

Definition 5.2.12. *A topological space, X, is compact if and only if every open cover of X has a finite subcover.*

Lemma 5.2.13. *A topological space is compact if and only if for every collection of closed sets, $(F_i)_{i\in I}$, satisfying the finite intersection property we have $\bigcap_{i\in I} F_i \neq \emptyset$.*

Proof. For this proof we must show that our lemma is equivalent to the definition of compactness given in Definition 5.2.12. We shall do this by proving that for every collection of closed sets, $(F_i)_{i\in I}$, satisfying the finite intersection property we have $\bigcap_{i\in I} F_i \neq \emptyset$ if and only if every open cover of X has a finite subcover.

We first assume that for any collection of closed sets, $(F_i)_{i\in I}$, satisfying the finite intersection property we have $\bigcap_{i\in I} F_i \neq \emptyset$. We now take an open cover of X and assume that this open cover does not have a finite subcover. We take $(U_i)_{i\in I}$ to be the collection of sets in the open cover of X. We may take any finite subcollection of open sets from $(U_i)_{i\in I}$, say $(U_j)_{j\in J}$ where $J \subset I$. Since this collection of open sets cannot be a cover of X we must have that there exists some element of X, say x, which is not in $\bigcup_{j\in J} U_j$. But then, by De Morgan's laws, we have $x \in \bigcap_{j\in J} F_j$, where $F_j = X \setminus U_j$. This is true of any finite subcollection of open sets from $(U_i)_{i\in I}$ so $\bigcap_{j\in J} F_j \neq \emptyset$ for all finite J such that $J \subset I$. Therefore $(F_i)_{i\in I}$ has the finite intersection property, but then from our assumption $\bigcap_{i\in I} F_i \neq \emptyset$, and Lemma 5.2.10 implies that $(U_i)_{i\in I}$ is not an open cover of X. We have a contradiction and so there must exist some finite open subcover of X in every open cover of X.

This completes the proof that if, for every collection of closed sets, $(F_i)_{i\in I}$, satisfying the finite intersection property we have $\bigcap_{i\in I} F_i \neq \emptyset$ then every open cover of X has a finite subcover. We must now prove the converse of this statements.

We now assume that every open cover of X has a finite subcover of X. Considering any collection of closed set $(F_i)_{i\in I}$, with the finite intersection property, we assume that $\bigcap_{i\in I} F_i = \emptyset$. Since our collection of closed set $(F_i)_{i\in I}$ has the finite intersection property we have that $\bigcap_{j\in J} F_j \neq \emptyset$ for all finite $J \subset I$. Since $\bigcap_{i\in I} F_i = \emptyset$, from Lemma 5.2.10 we have that $(U_i)_{i\in I}$ forms an open cover of X, where $U_i = X \setminus F_i$. However, again from Lemma 5.2.10, we have that $(U_j)_{j\in J}$ does not form a finite open subcover of X for all finite $J \subset I$. This is a contradiction since every open cover of X has a finite subcover of X. Therefore $\bigcap_{i\in I} F_i \neq \emptyset$ for every collection of closed sets, $(F_i)_{i\in I}$, satisfying the finite intersection property in a compact topological space. □

Definition 5.2.14. *Consider a topological space, X. If a and b are elements of this topological space we call them separated by neighbourhoods if we can find two neighbourhoods, A and B, of a and b respectively such that $A \cap B = \emptyset$.*

Definition 5.2.15. *We call a topological space, X, a Hausdorff space if any two distinct points of X can be separated by neighborhoods.*

We shall later prove results about general compact Hausdorff spaces. Using Theorem 5.2.16 we will be able to use these results to deduce results about $\beta\mathbb{N}$.

Theorem 5.2.16. *$\beta\mathbb{N}$ is a compact Hausdorff space.*

Proof. First we take two distinct ultrafilters, $\mathcal{U}$ and $\mathcal{V}$. Since $\mathcal{U}$ and $\mathcal{V}$ are distinct there must exist some set A such that $A \in \mathcal{U}$ and $A \notin \mathcal{V}$. $\mathcal{V}$ is an ultrafilter and $A \notin \mathcal{V}$ so Lemma 5.1.4 implies $\bar{A} \in \mathcal{V}$. Since $A \in \mathcal{U}$ we have $\mathcal{U} \in C_A$ and since $\bar{A} \in \mathcal{V}$ we have $\mathcal{V} \in C_{\bar{A}}$. The two neighbourhoods, C_A and $C_{\bar{A}}$, of $\mathcal{U}$ and $\mathcal{V}$ respectively are disjoint since no ultrafilter can contain both A and $\bar{A}$. So $\beta\mathbb{N}$ is a Hausdorff space.

We must now prove that $\beta\mathbb{N}$ is compact. To do this we must prove that for every family of closed sets, say $(F_i)_{i\in I}$, with the finite intersection property we have $\bigcap_{i\in I} F_i \neq \emptyset$. We assume that each F_i is basic, that is, $F_i = C_{A_i}$. Recall that Lemma 5.2.2 implies $C_{A_{i_1}} \cap C_{A_{i_2}} \cap C_{A_{i_3}} \cap \cdots \cap C_{A_{i_n}} = C_{A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n}}$ for all $i_1, i_2, \ldots, i_n \in I$.

We define a filter,

$$\mathcal{F} = \{A \subset \mathbb{N} \,|\, A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n} \subset A \quad \text{for some } i_1, i_2, i_3 \ldots, i_n \in I \text{ and some } n \in \mathbb{N}\}.$$

We now prove that $\mathcal{F}$ is indeed a filter. From our assumption, the family of closed sets, $(F_i)_{i\in I} = (C_{A_i})_{i\in I}$, has the finite intersection property so $C_{A_{i_1}} \cap C_{A_{i_2}} \cap C_{A_{i_3}} \cap \cdots \cap C_{A_{i_n}} = C_{A_{i_1}\cap A_{i_2}\cap A_{i_3}\cap\cdots\cap A_{i_n}} \neq \emptyset$ for any $i_1, i_2, i_3 \ldots, i_n \in I$. Since $C_{A_{i_1}\cap A_{i_2}\cap A_{i_3}\cap\cdots\cap A_{i_n}} \neq \emptyset$ we must have that $A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n} \neq \emptyset$ and since $A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n} \subset A$ clearly $A \neq \emptyset$ and $\emptyset \notin \mathcal{F}$. $A_i \subset \mathbb{N}$ for every $i \in I$ so from the definition of $\mathcal{F}$, $\mathbb{N} \in \mathcal{F}$.

If $A \in \mathcal{F}$ and $A \subset B$ then $A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n} \subset A$ for some $i_1, i_2, i_3 \ldots, i_n \in I$ thus $A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap \cdots \cap A_{i_n} \subset B$ and so $B \in \mathcal{F}$.

If $A, B \in \mathcal{F}$ then $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n} \subset A$ for some $i_1, i_2 \ldots, i_n \in I$ and $A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_n} \subset B$ for some $j_1, j_2 \ldots, j_n \in I$. Since $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n} \cap A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_n} \subset A \cap B$ and $i_1, i_2 \ldots, i_n, j_1, j_2 \ldots, j_n \in I$ we have $A \cap B \in \mathcal{F}$.

We have shown that $\emptyset \notin \mathcal{F}$ and $\mathbb{N} \in \mathcal{F}$, if $A \in \mathcal{F}$ and $A \subset B$ then $B \in \mathcal{F}$ and if $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$. Therefore $\mathcal{F}$ is a filter since is satisfies the definition of a filter given in Definition 5.1.1.

From Lemma 5.1.3 we may take $\mathcal{W}$, an ultrafilter extending $\mathcal{F}$. Since $A_i \in \mathcal{F}$ for all $i \in I$ and $\mathcal{F} \subseteq \mathcal{W}$ we have $A_i \in \mathcal{W}$ for all $i \in I$ so $\mathcal{W} \in C_{A_i}$ for all $i \in I$. Therefore $\mathcal{W} \in \bigcap_{i\in I} C_{A_i}$ and $\bigcap_{i\in I} C_{A_i} \neq \emptyset$ so Lemma 5.2.13 implies that $\beta\mathbb{N}$ is compact. □

In order to prove Hindman's Theorem we must define $\mathcal{U} + \mathcal{V}$ for two ultrafilters $\mathcal{U}$ and $\mathcal{V}$. Therefore we turn our attention to ultrafilter quantifiers which will enable us to manage the notation more easily.

5.3 Ultrafilter quantifiers

Definition 5.3.1. *$(\forall_{\mathcal{U}} x)p(x)$ if and only if $\{x \in \mathbb{N} \,|\, p(x)\} \in \mathcal{U}$ where $p(x)$ is some statement with variable x.*

The following lemma demonstrates that $(\forall_{\mathcal{U}} x)$ and $(\forall_{\mathcal{V}} y)$ are not always commutative.

Lemma 5.3.2. *For two non-principal ultrafilters $\mathcal{U}$ and $\mathcal{V}$*

$$(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)x < y \neq (\forall_{\mathcal{V}} y)(\forall_{\mathcal{U}} x)x < y.$$

Proof. First we note $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)x < y$ if and only if $\{x \in \mathbb{N} \,|\, (\forall_{\mathcal{V}} y)x < y\} \in \mathcal{U}$ if and only if $\{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x < y\} \in \mathcal{V}\} \in \mathcal{U}$. Comparatively $(\forall_{\mathcal{V}} y)(\forall_{\mathcal{U}} x)x < y$ if and only if $\{y \in \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)x < y\} \in \mathcal{V}$ if and only if $\{y \in \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, x < y\} \in \mathcal{U}\} \in \mathcal{V}$. So to prove that

$$(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)x < y \neq (\forall_{\mathcal{V}} y)(\forall_{\mathcal{U}} x)x < y$$

we must show

$$\{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x < y\} \in \mathcal{V}\} \in \mathcal{U} \neq \{y \in \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, x < y\} \in \mathcal{U}\} \in \mathcal{V},$$

when $\mathcal{U}$ and $\mathcal{V}$ are non-principal ultrafilters. For any natural number there exists an infinite number of larger natural numbers. Therefore, if we choose any element, say $x \in \mathbb{N}$, then the set $\{y \in \mathbb{N} \,|\, x < y\} \in \mathcal{V}$ since the complement of this set is $\{z \in \mathbb{N} \,|\, x \geq z\}$, which is clearly finite. Since x was an arbitrary natural number we must have that $\{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x < y\} \in \mathcal{V}\} = \mathbb{N}$. Therefore $\{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x < y\} \in \mathcal{V}\} \in \mathcal{U}$. Conversely, if we choose any element, say $y \in \mathbb{N}$, then the set $\{x \in \mathbb{N} \,|\, x < y\} \notin \mathcal{U}$. Indeed, any set in $\mathcal{U}$ must be infinite, and the set $\{x \in \mathbb{N} \,|\, x < y\}$ is clearly finite. Since y was an arbitrary natural number we must have that $\{y \in \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, x < y\} \in \mathcal{U}\} = \emptyset$. However, since $\mathcal{V}$ is an ultrafilter, $\emptyset \notin \mathcal{V}$ so $\{y \in \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, x < y\} \in \mathcal{U}\} \notin \mathcal{V}$. Therefore $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x < y$ is not equal to $(\forall_{\mathcal{V}} y)(\forall_{\mathcal{U}} x) x < y$. □

To be able to effectively use ultrafilter quantifiers we must first prove some results about them.

Proposition 5.3.3. *Take an ultrafilter, say $\mathcal{U}$, and statements $p(x)$ and $q(x)$. Then*

1. $(\forall_{\mathcal{U}} x)(p(x)$ *and* $q(x)) \Leftrightarrow ((\forall_{\mathcal{U}} x)p(x)$ *and* $(\forall_{\mathcal{U}} x)q(x))$,
2. $(\forall_{\mathcal{U}} x)(p(x)$ *or* $q(x)) \Leftrightarrow ((\forall_{\mathcal{U}} x)p(x)$ *or* $(\forall_{\mathcal{U}} x)q(x))$,
3. *If* $(\forall_{\mathcal{U}} x)(p(x))$ *does not hold then* $(\forall_{\mathcal{U}} x)($*not* $p(x))$.

Proof. Take $A = \{x \in \mathbb{N} \,|\, p(x)\}$ and $B = \{x \in \mathbb{N} \,|\, q(x)\}$.

Part *1.* now states that $A \cap B \in \mathcal{U}$ if and only if $A \in \mathcal{U}$ and $B \in \mathcal{U}$. This was proved in Lemma 5.1.6.

Part *2.* now states that $A \cup B \in \mathcal{U}$ if and only if $A \in \mathcal{U}$ or $B \in \mathcal{U}$. This was proved in Proposition 5.1.5.

Part *3.* now states that if $A \notin \mathcal{U}$ then $\bar{A} \in \mathcal{U}$. This was proved in Lemma 5.1.4. □

Definition 5.3.4. *For $\mathcal{U}, \mathcal{V} \in \beta\mathbb{N}$ we define an addition on the topological space $\beta\mathbb{N}$ as*

$$\mathcal{U} + \mathcal{V} = \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x + y \in A\}.$$

We may also write $\mathcal{U} + \mathcal{V}$ in a less elegant form,

$$\begin{aligned}\mathcal{U} + \mathcal{V} &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x + y \in A\}, \\ &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, (\forall_{\mathcal{V}} y) x + y \in A\} \in \mathcal{U}\}, \\ &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in A\} \in \mathcal{V}\} \in \mathcal{U}\}.\end{aligned}$$

Example. $\tilde{3} + \tilde{5} = \tilde{8}$.

$$\begin{aligned}\tilde{3} + \tilde{5} &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in A\} \in \tilde{5}\} \in \tilde{3}\}, \\ &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, y \in A - x\} \in \tilde{5}\} \in \tilde{3}\}, \\ &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, A - x \in \tilde{5}\} \in \tilde{3}\}, \\ &= \{A \subset \mathbb{N} \,|\, 3 \in \{x \in \mathbb{N} \,|\, A - x \in \tilde{5}\}\}, \\ &= \{A \subset \mathbb{N} \,|\, A - 3 \in \tilde{5}\}, \\ &= \{A \subset \mathbb{N} \,|\, 5 \in A - 3\}, \\ &= \{A \subset \mathbb{N} \,|\, 8 \in A\}, \\ &= \tilde{8}.\end{aligned}$$

We may generalise this example for the addition of any two principal ultrafilters.

Lemma 5.3.5. $\widetilde{n}_1 + \widetilde{n}_2 = \widetilde{n_1 + n_2}$.

Proof.

$$\begin{aligned}
\widetilde{n}_1 + \widetilde{n}_2 &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in A\} \in \widetilde{n}_2\} \in \widetilde{n}_1\}, \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, y \in A - x\} \in \widetilde{n}_2\} \in \widetilde{n}_1\}, \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, A - x \in \widetilde{n}_2\} \in \widetilde{n}_1\}, \\
&= \{A \subset \mathbb{N} \,|\, n_1 \in \{x \in \mathbb{N} \,|\, A - x \in \tilde{n}_2\}\}, \\
&= \{A \subset \mathbb{N} \,|\, A - n_1 \in \tilde{n}_2\}, \\
&= \{A \subset \mathbb{N} \,|\, n_2 \in A - n_1\}, \\
&= \{A \subset \mathbb{N} \,|\, n_1 + n_2 \in A\}, \\
&= \widetilde{n_1 + n_2}.
\end{aligned}$$

□

Lemma 5.3.6. $\mathcal{U} + \mathcal{V}$ *is an ultrafilter.*

Proof. To prove that $\mathcal{U}+\mathcal{V}$ is an ultrafilter we must check it satisfies Definition 5.1.1, the definition of a filter, and that if $A \notin \mathcal{U} + \mathcal{V}$ then $\bar{A} \in \mathcal{U} + \mathcal{V}$ for all $A \subset \mathbb{N}$. That is, we must check that $\mathcal{U} + \mathcal{V}$ is a filter and satisfies the conditions given in Lemma 5.1.4 for it to be an ultrafilter.

Since $\mathcal{U}$ and $\mathcal{V}$ are ultrafilters we have $\emptyset \notin \mathcal{U}, \mathcal{V}$ and $\mathbb{N} \in \mathcal{U}, \mathcal{V}$. We first show $\emptyset \notin \mathcal{U} + \mathcal{V}$ and $\mathbb{N} \in \mathcal{U} + \mathcal{V}$. If $\emptyset \in \mathcal{U} + \mathcal{V}$ then there must exist some $x \in \mathcal{U}$ and $y \in \mathcal{V}$ such that $x + y = \emptyset$. No two natural numbers add together to give the empty set so $\emptyset \notin \mathcal{U} + \mathcal{V}$. $\mathbb{N} \in \mathcal{U} + \mathcal{V}$ since $\mathbb{N} \in \mathcal{U}, \mathcal{V}$. Indeed, for any $x \in \mathbb{N}$ we have $\{y \in \mathbb{N} \,|\, x + y \in \mathbb{N}\} = \mathbb{N} \in \mathcal{V}$. Thus $\{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in \mathbb{N}\} \in \mathcal{V}\} = \mathbb{N} \in \mathcal{U}$ and so $\mathbb{N} \in \mathcal{U} + \mathcal{V}$.

We now write $\mathcal{U} + \mathcal{V}$ in a different form,

$$\begin{aligned}
\mathcal{U} + \mathcal{V} &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in A\} \in \mathcal{V}\} \in \mathcal{U}\}, \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, y \in A - x\} \in \mathcal{V}\} \in \mathcal{U}\}, \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, A - x \in \mathcal{V}\} \in \mathcal{U}\}.
\end{aligned}$$

We must now show that if $A \in \mathcal{U} + \mathcal{V}$ and $A \subset B$ then $B \in \mathcal{U} + \mathcal{V}$. For ease of notation we define the sets $D_A = \{x \in \mathbb{N} \,|\, A - x \in \mathcal{V}\}$ and $D_B = \{x \in \mathbb{N} \,|\, B - x \in \mathcal{V}\}$. From the above it is clear that for any $A \in \mathcal{U} + \mathcal{V}$ we have $D_A \in \mathcal{U}$. We already have that $A \subset B$, therefore $A - x \subset B - x$. Since $\mathcal{V}$ is a filter it follows that $B - x \in \mathcal{V}$ whenever $A - x \in \mathcal{V}$. Thus $D_A \subseteq D_B$, and since $D_A \in \mathcal{U}$ and $\mathcal{U}$ is a filter we must have that $D_B \in \mathcal{U}$. We now have that $B \in \mathcal{U} + \mathcal{V}$, so we have shown if $A \in \mathcal{U} + \mathcal{V}$ and $A \subset B$ then $B \in \mathcal{U} + \mathcal{V}$.

We must now show that if $A, B \in \mathcal{U} + \mathcal{V}$ then $A \cap B \in \mathcal{U} + \mathcal{V}$. We have $A, B \in \mathcal{U} + \mathcal{V}$ so $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x + y \in A$ and $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x + y \in B$, then from

Proposition 5.3.3.1 we have $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)(x+y \in A$ and $x+y \in B)$ so clearly $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x+y \in A \cap B$, and $A \cap B \in \mathcal{U}+\mathcal{V}$.

Finally, if $A \notin \mathcal{U}+\mathcal{V}$ then $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x+y \in A$ does not hold. From Proposition 5.3.3.3 we have $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x+y \notin A$ and so $(\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x+y \in \bar{A}$. This shows that if $A \notin \mathcal{U}+\mathcal{V}$ then $\bar{A} \in \mathcal{U}+\mathcal{V}$. □

Lemma 5.3.7. *The operation $+$ on $\beta\mathbb{N}$ is associative.*

Proof. For three ultrafilters $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ we must prove $(\mathcal{U}+\mathcal{V})+\mathcal{W} = \mathcal{U}+(\mathcal{V}+\mathcal{W})$. We first set $B_A = \{t \in \mathbb{N} \,|\, (\forall_{\mathcal{W}} z) t+z \in A\}$. Then

$$\begin{aligned}(\mathcal{U}+\mathcal{V})+\mathcal{W} &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}+\mathcal{V}} t)(\forall_{\mathcal{W}} z) t+z \in A\},\\ &= \{A \subset \mathbb{N} \,|\, \{t \in \mathbb{N} \,|\, (\forall_{\mathcal{W}} z) t+z \in A\} \in \mathcal{U}+\mathcal{V}\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y) x+y \in B_A\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)(\forall_{\mathcal{W}} z) x+y+z \in A\}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathcal{U}+(\mathcal{V}+\mathcal{W}) &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}+\mathcal{W}} s) x+s \in A\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}+\mathcal{W}} s) s \in A-x\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)\{s \in \mathbb{N} \,|\, s \in A-x\} \in \mathcal{V}+\mathcal{W}\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x) A-x \in \mathcal{V}+\mathcal{W}\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)(\forall_{\mathcal{W}} z) y+z \in A-x\},\\ &= \{A \subset \mathbb{N} \,|\, (\forall_{\mathcal{U}} x)(\forall_{\mathcal{V}} y)(\forall_{\mathcal{W}} z) x+y+z \in A\}.\end{aligned}$$

Therefore $(\mathcal{U}+\mathcal{V})+\mathcal{W} = \mathcal{U}+(\mathcal{V}+\mathcal{W})$. □

Definition 5.3.8. *If $\star$ is a binary operation on a set S, then $\star$ is left-continuous if $f(x) = y \star x$ is a continuous mapping for all $y \in S$.*

Lemma 5.3.9. *Addition on $\beta\mathbb{N}$ is left continuous. That is for every fixed $\mathcal{V}$,*

$$\begin{aligned}f : \beta\mathbb{N} &\to \beta\mathbb{N} \times \beta\mathbb{N}\\ \mathcal{U} &\mapsto \mathcal{U}+\mathcal{V}\end{aligned}$$

is a continuous mapping.

Proof. For the mapping to be continuous the preimage of any open set in $\mathcal{U}+\mathcal{V}$ must be an open set in $\mathcal{U}$. Clearly, it is sufficient to consider any basic open set C_A in $\mathcal{U}+\mathcal{V}$ and show that $f^{-1}(C_A)$ is an open set in $\mathcal{U}$. We can see that $f^{-1}(C_A) = \{\mathcal{U} \in \beta\mathbb{N} \,|\, A \in \mathcal{U}+\mathcal{V}\}$. Now, $A \in \mathcal{U}+\mathcal{V}$ if and only if $\{x \in \mathbb{N} \,|\, (\forall_{\mathcal{V}} y) x+y \in A\} \in \mathcal{U}$. We can now define the set $B = \{x \in \mathbb{N} \,|\, (\forall_{\mathcal{V}} y) x+y \in A\}$. Then $A \in \mathcal{U}+\mathcal{V}$ if and only if $B \in \mathcal{U}$. Thus $f^{-1}(C_A) = \{\mathcal{U} \in \beta\mathbb{N} \,|\, B \in \mathcal{U}\} = C_B$. C_B is an open set in $\mathcal{U}$, therefore for every fixed $\mathcal{V}$, f is a continuous mapping. □

Definition 5.3.10. *If $\star$ is a binary operation on a set S, an element, x, of S is an idempotent for $\star$ if $x \star x = x$.*

Before going on to prove Hindman's Theorem we must use our knowledge of ultrafilters and the topology on $\beta\mathbb{N}$ to find an ultrafilter in $\beta\mathbb{N}$ which is idempotent under addition. We shall first prove an idempotent element exists for a general Hausdorff space, X, of the form of $\beta\mathbb{N}$ we may then apply this result to $\beta\mathbb{N}$. We do this using good subsets.

5.4 Good subsets

Throughout this section X is a non-empty compact Hausdorff space, and $+ : X \times X \to X$ is an associative left-continuous binary operation on X.

Definition 5.4.1. *A subset, M, of X is good if it is compact, non-empty and satisfies $M + M \subseteq M$.*

Lemma 5.4.2. *There exists a minimal good subset of X.*

Proof. We shall prove that a minimal good subset of X exists using Zorn's Lemma. Zorn's Lemma states that *every partially ordered set, P, has a maximal element if P has the Zorn Property.* P has the the Zorn property if *for every chain C in P there is an upper bound $p \in P$ of C.* We are concerned with finding a minimal element and so we take the reverse order of the set.

We take $M_1 \supseteq M_2 \supseteq M_3 \supseteq \cdots$ to be the chain of good sets and must find a lower bound of the chain. Under the reverse order this chain would be $M_1 \subseteq M_2 \subseteq M_3 \subseteq \cdots$ and we would need to have a upper bound for Zorn's lemma to hold, it is completely equivalent to find a lower bound in our ordering.
Claim. $M = \bigcap_{i=1}^{\infty} M_i$ is a lower bound of the chain.

X is defined to be a compact Hausdorff space. $M_1, M_2, M_3, \ldots$ are all good sets, and so are all compact in X. Compact subsets of Hausdorff spaces must be closed so all the sets M_i, for $i \in \mathbb{N}$, are closed. Since M is the intersection of these closed sets M is also closed. M is a closed subset of X, a compact space and so M is compact.

Since each M_i is a subset of M_{i-1} for $i \geq 2$, every element of M_i must also be in M_{i-1}. Therefore any intersection of some finite collection of these sets cannot be empty since it must contain at least all the elements of M_j where M_j is the smallest set in the intersection. Thus the collection of closed sets, $(M_i)_{i\in\mathbb{N}}$, satisfies the finite intersection property and so, by Theorem 5.2.13, the intersection of all the M_i must be non-empty. We now have $M \neq \emptyset$.

Finally $M + M = \{m_1 + m_2 \,|\, m_1, m_2 \in \bigcap_{i=1}^{\infty} M_i\}$ and any two points in $\bigcap_{i=1}^{\infty} M_i$ must both be in every M_i. So from the definition of a good set $m_1 + m_2$ must also be in every M_i, and therefore in $\bigcap_{i=1}^{\infty} M_i$. So $M + M \subseteq M$

Since M is compact, non-empty and $M + M \subseteq M$ it is a good set, so M is the lower bound. Therefore, by Zorn's Lemma, there must exist a minimal good set. □

Lemma 5.4.3. *Let M be a minimal good set (which must exist by Lemma 5.4.2) and take $x \in M$. Then the set $M + x$ is good.*

Proof. $M + x$ is the image of M under the continuous function,

$$\begin{aligned} f : X &\to X \\ m &\mapsto m + x. \end{aligned}$$

We already have that $+$ is an associative left-continuous binary operation on X, so f is a continuous mapping. The image of a compact space under a continuous mapping is compact. Since M is compact and $M + x$ is the image of M under a continuous mapping, $M + x$ is compact.

$M + x$ is non-empty since M is non-empty. That is, there exists some element, say a, in M. This element will map to an element of $M + x$, in particular a will map to $a + x$, therefore $M + x$ is non-empty.

From the definition of M we have $M + M \subseteq M$. But $x \in M$ and so $M + x \subseteq M$. We can now see that $M + x$ satisfies $(M + x) + (M + x) \subseteq M + x$ since $(M + x) + (M + x) = (M + x + M) + x \subseteq (M + M) + x \subseteq M + x$.

This proves that the set $M + x$ is good since it is compact, non-empty and satisfies $(M + x) + (M + x) \subseteq M + x$. □

M is a good set and $x \in M$ so we have that $M + x \subseteq M$. We have now shown that $M + x$ is a good set, but M is the minimal good set so we must have $M + x = M$.

Theorem 5.4.4. *There exists some $x \in X$ such that $x + x = x$. That is, there exists an idempotent element of X.*

Proof. Since $M + x = M$ there must exist some element of M, say y, such that $y + x = x$. We take a set $Y = \{y \in M \mid y + x = x\}$. This set is the inverse image of $\{x\}$ under the function

$$\begin{aligned} g : X &\to X \\ y &\mapsto y + x. \end{aligned}$$

We already have that $+$ is an associative left-continuous binary operation on X, so g is a continuous mapping. Therefore, Y is closed, since it is the inverse image of a singleton under a continuous function.

A closed subspace of a compact space is compact, therefore since X is a compact Hausdorff space and Y is a closed subspace of X, Y is also compact.

For two elements in Y, say y_1 and y_2, we have $(y_1 + y_2) + x = y_1 + (y_2 + x) = y_1 + x = x$. So $y_1 + y_2 \in Y$.

We now have that Y is compact, non-empty and $Y + Y \subseteq Y$, so Y is good. We may now conclude that $Y = M$ since $M \subseteq Y$ as M is the minimal good set and $Y \subseteq M$ from the definition of Y. Since $Y = M$ we must have that $x \in Y$. Thus we have an element x such that $x + x = x$, as required. □

We are now ready to prove Hindman's Theorem. First we shall clarify exactly what the results we have proved imply.

The results for the general compact Hausdorff space, X, may now be applied to our compact Hausdorff space, $\beta\mathbb{N}$. In particular Theorem 5.4.4 may be applied to $\beta\mathbb{N}$. In Lemma 5.3.7 and Lemma 5.3.9 we showed that $+$ is an associative left-continuous binary operation on $\beta\mathbb{N}$. Therefore we may conclude that there exists some idempotent element of $\beta\mathbb{N}$, that is, there exists some $\mathcal{U}$ in $\beta\mathbb{N}$ such that $\mathcal{U} + \mathcal{U} = \mathcal{U}$.

5.5 Hindman's Theorem

Theorem 5.5.1. *If $\mathbb{N}$ is coloured with some finite number of colours there must exist some infinite set $S \subseteq \mathbb{N}$ such that $\mathcal{S}(S)$ is monochromatic.*

Proof. As shown above there must exist an idempotent ultrafilter $\mathcal{U}$ in $\beta\mathbb{N}$. If $\mathcal{U}$ is the principal ultrafilter $\tilde{n}$ we must have that $\tilde{n} = \tilde{n} + \tilde{n} = \widetilde{2n}$, a contradiction. So $\mathcal{U}$ is a non-principal idempotent ultrafilter. We first show that in a colouring of $\mathbb{N}$ with some finite number of colours, $r_1, r_2, r_3, \dots, r_n$ there must exist some infinitely large monochromatic set A_1 in $\mathcal{U}$. Indeed, if $\mathbb{N}$ is n-coloured then $R_1 \cup R_2 \cup \cdots \cup R_n = \mathbb{N}$ where R_i is the set of all r_i coloured natural numbers. From Proposition 5.1.8 there must exist some $i \in \{1, 2, \dots, n\}$ such that $R_i \in \mathcal{U}$ since $R_1 \cup R_2 \cup \cdots \cup R_n = \mathbb{N} \in \mathcal{U}$. Since $\mathcal{U}$ is a non-principal ultrafilter Proposition 5.1.9 implies that any set in $\mathcal{U}$ cannot be finite. Therefore the set R_i in $\mathcal{U}$, must be infinite, so we have our desired monochromatic infinite set in $\mathcal{U}$. We call this set A_1. We may write $\mathcal{U} + \mathcal{U}$ as

$$\begin{aligned}
\mathcal{U} + \mathcal{U} &= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, x + y \in A\} \in \mathcal{U}\} \in \mathcal{U}\} \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, \{y \in \mathbb{N} \,|\, y \in A - x\} \in \mathcal{U}\} \in \mathcal{U}\} \\
&= \{A \subset \mathbb{N} \,|\, \{x \in \mathbb{N} \,|\, A - x \in \mathcal{U}\} \in \mathcal{U}\}.
\end{aligned}$$

Since $A_1 \in \mathcal{U}$ and $\mathcal{U} = \mathcal{U} + \mathcal{U}$ we must have that $A_1 \in \mathcal{U} + \mathcal{U}$. Therefore there must exist some $x_1 \in A_1$ such that $A_1 - x_1 \in \mathcal{U}$. Indeed, otherwise $\{x \in \mathbb{N} \,|\, A_1 - x \in \mathcal{U}\} \subseteq \bar{A}_1$ since $A_1 \cup \bar{A}_1 = \mathbb{N}$. That is, for any $x \in \mathbb{N}$ such that $A_1 - x \in \mathcal{U}$ we must have $x \in \bar{A}_1$ since $x \notin A_1$. A contradiction, since then $A_1, \bar{A}_1 \in \mathcal{U}$, which from the definition of a filter cannot happen. We now define $A_2 = A_1 \cap (A_1 - x_1)$, since $A_1 \in \mathcal{U}$ and $A_1 - x_1 \in \mathcal{U}$, we must have from the definition of a filter that $A_2 \in \mathcal{U}$. Therefore $A_2 \in \mathcal{U} + \mathcal{U}$. As with A_1, there must exist some $x_2 \in A_2$ such that $A_2 - x_2 \in \mathcal{U}$. We may continue in this fashion, defining $A_3 = A_2 \cap (A_2 - x_2)$. This produces an infinite sequence of nested sets, $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$ and elements, $x_i \in A_i$ such that $A_i - x_i \supseteq A_{i+1}$. Therefore $A_i \supseteq A_{i+1} + x_i$ and $A_i \supseteq A_j + x_i$ for all $j \geq i + 1$. We have thus found the monochromatic set $S = \{x_i \,|\, i \in \mathbb{N}\}$. We will now show that $\mathcal{S}(S)$ is monochromatic. To do this we consider a subset of S, the natural numbers $x_{i_1}, x_{i_2}, \dots, x_{i_n}$, where $i_1 < i_2 < \cdots < i_{n-1} < i_n$.

Since $A_{i_n} + x_{i_{n-1}} \subseteq A_{i_{n-1}}$ we have $x_{i_n} + x_{i_{n-1}} \in A_{i_{n-1}}$, but then since $A_{i_{n-1}} + x_{i_{n-2}} \subseteq A_{i_{n-2}}$ we have $(x_{i_n} + x_{i_{n-1}}) + x_{i_{n-2}} \in A_{i_{n-2}}$. We may continue in this way until we reach A_{i_1}, where we have $x_{i_n} + x_{i_{n-1}} + x_{i_{n-2}} + \cdots + x_{i_1} \in A_{i_1} \subseteq A_1$. This shows that $\mathcal{S}(S)$ is in A_1 and is therefore monochromatic. So we have shown that if $\mathbb{N}$ is coloured with some finite number of colours there must exist some infinite set $S \subseteq \mathbb{N}$ such that $\mathcal{S}(S)$ is monochromatic. $\square$

6 Conclusion

We began by introducing Ramsey's numbers and giving examples of ways in which some of these smaller numbers may be found. We then went on to prove Ramsey's Theorem, initially for graphs, which we were able to prove in two ways, and then, more generally, for sets. Next we proved Van der Waerden's Theorem. An interesting example enabled us to see that Van der Waerden's Theorem cannot be extended to infinite arithmetic progressions.

We then gave the proof of Rado's Theorem which provided a characterisation of the homogeneous systems to which a monochromatic solution can be found in any finite colouring of the natural numbers. This result enabled us to prove that any two finite regular linear homogeneous systems are consistent. We also proved the Finite Sums Theorem using Rado's Theorem.

Finally we proved Hindman's Theorem which is a natural extension of the Finite Sums Theorem. In order to prove this theorem we had to prove many quite abstract results. After understanding ultrafilters and their quantifiers we were able to prove results about a topology on the set of ultrafilters.

The origins of the theorems proved here are far from common. The intensions of the mathematicians were, in general, not advancing the field of Ramsey Theory. Ramsey himself came across the field which now bares his name while attempting to solve problems of decidability in logic. Ramsey died aged only twenty six in 1930. A year later it was proved that the problem which he had been working on is in fact impossible to solve. So Ramsey is now famous for proving a theorem he didn't need while trying prove something which he couldn't prove. Schur was trying to solve Fermat's last theorem over finite fields. Van der Waerden became interested in his theorem when Baudet, a student at Göttingen, asked if it could be proved. The academics at Göttingen had been struggling to find a proof before Van der Waerden produced his.

Ramsey Theory began to expand in the second half of the twentieth century and is now recognised as an interesting and expanding area of mathematics. In 1983 Frank Harary [18] wrote about Ramsey Theory, *unsolved problems abound, and additional interesting open questions arise faster than solutions to the existing problems.* This remains true, indeed, there are many problems which are natural and easy to understand but have not been solved.

An example of one of these open problems follows directly from Chapter 4. We proved in Section 4.6 that any two finite regular linear homogeneous systems are consistent. A characterisation of the infinite regular linear homogeneous systems which are consistent has not been produced. There are results which show that certain types of infinite matrices are consistent and certain other types are inconsistent. The proofs of these results can be found in papers by Leader and Russell, [8] and [9]. However, the characterisation in general is far from complete.

Another open question is an extension of Schur's Theorem. *If* $\mathbb{N}$ *is 2-coloured, does there always exists* $x, y \in \mathbb{N}$ *such that* $x, y, x + y$ *and* xy *are all monochromatic?*

Another problem are the bounds for the Ramsey, and Van der Waerden numbers which are not generally known with any accuracy. These bounds are unlikely to be improved upon without some deeper mathematical understanding. For example $43 \leq R(5,5) \leq 49$, however to prove that $R(5,5) \neq 43$ by drawing every possible 2-colouring of K_{43} would require us to draw $2^{\binom{43}{2}} = 2^{903}$ graphs. As we take larger Ramsey and Van der Waerden numbers the bounds become less informative. Some of these open questions have prizes attached to them. For example, in 1998 Ronald Graham offered \$1000 to anyone who could prove or disprove that $W(k,2) < 2^{k^2}$.

7 Sources

The book *Ramsey Theory* [1] has been my main source, however, in some chapters I referred to other sources.

Throughout Chapter 2, Ramsey's Theorem, I mainly referred to the book *Ramsey Theory* [1], however, I also referred to *Modern Graph Theory* [2], *A Friendly Introduction to Graph Theory* [3], *Asymptotic Bounds for Classical Ramsey Numbers* [4] and *Combinatorics: set systems, hypergraphs, families of vectors and combinatorial probability* [5].

In Chapter 3, Van der Waerden's Theorem, I mainly referred to *Ramsey Theory* [1], however, I have also referred to *Three Pearls of Number Theory* [7].

In Chapter 5, Hindman's Theorem, I referred to *Ramsey Theory* [1]. However my main resource was *Partition Regular Equations* [8]. I also referred to *Topology* [11] and *Combinatorial Topology* [12] in Chapter 5 for formal topological definitions.

Rudiments of Ramsey Theory [6] and *Ramsey Theory on the Integers* [10] provided some of the historical information which was included.

Appendix

A Bounding $W(3,3)$

We give an example here of how the value of $W(3,3)$ may be bounded. In Proposition 3.1.1 we showed how the value for $W(3,2)$ may be bounded, this example uses an extension of the method used in that proposition.

Example. $W(3,3) \leq 7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)$

The set of integers, $\{1,2,\ldots,7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\}$, are first divided into $2\cdot 3^{7(2\cdot 3^7+1)}+1$ blocks, each of order $7(2\cdot 3^7+1)$. That is, $\{1,2,\ldots,7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\} = \{1,\ldots,7(2\cdot 3^7+1)\}\cup\{7(2\cdot 3^7+1)+1,\ldots,14(2\cdot 3^7+1)\}\cup\ldots\cup\{7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)})+1,\ldots,7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\}$. We then label these blocks $B_1, B_2, \ldots, B_{2\cdot 3^{7(2\cdot 3^7+1)}+1}$ where $B_1 = \{1,\ldots,7(2\cdot 3^7+1)\}, B_2 = \{7(2\cdot 3^7+1)+1,\ldots,14(2\cdot 3^7+1)\},\ldots,B_{2\cdot 3^{7(2\cdot 3^7+1)}+1} = \{7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)})+1,\ldots,7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\}$. Since there are only $3^{7(2\cdot 3^7+1)}$ ways in which any of the blocks may be 3-coloured, by the pigeonhole principle, at least two of the first $3^{7(2\cdot 3^7+1)}+1$ blocks must be coloured in the same way. We call these blocks B_a and B_{a+d}. Each block is of order $7(2\cdot 3^7+1)$ so we may split each into subblocks, $SB_{i,j}$, each of order 7, where i denotes the block which the subblock is in and y denotes its position within that block. For example, $SB_{1,2} = \{8,9,\ldots,14\}$. Since each subblock is coloured with only three colours there are only 3^7 ways in which a subblock may be coloured. So again by the pigeonhole principle we must have that at least two of the first 3^7+1 subblocks must be identically coloured. We call the two identically coloured subblocks of B_a, $SB_{a,\alpha}$ and $SB_{a,\alpha+\delta}$. We now label each element of $\{1,2,\ldots,7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\}$, $b_{x,y,z}$, where x denotes the block which the element is in, y denotes the subblock the element is in, in the block B_x, and z denotes the elements position within $SB_{x,y}$. For example $7(2\cdot 3^7+1)+2$ is labeled $b_{2,1,2}$. Since each subblock is coloured with only three colours, again by the pigeonhole principle, we can say that at least two of the first four elements in each subblock must be monochromatic. We call the two monochromatic elements of $SB_{a,\alpha}$, $b_{a,\alpha,\varphi}$ and $b_{a,\alpha,\varphi+\psi}$. The next term in this arithmetic progression must be in $SB_{a,\alpha}$ since $b_{a,\alpha,\varphi}$ and $b_{a,\alpha,\varphi+\psi}$ are both in its first four terms, we call the next term $b_{a,\alpha,\varphi+2\psi}$. If $b_{a,\alpha,\varphi+2\psi}$ is the same colour as $b_{a,\alpha,\varphi}$ and $b_{a,\alpha,\varphi+\psi}$ we would have our monochromatic 3-term arithmetic progression, so we assume $b_{a,\alpha,\varphi+2\psi}$ is a different colour. Since $SB_{a,\alpha}$ and $SB_{a,\alpha+\delta}$ are identically coloured the elements $b_{a,\alpha,\varphi}, b_{a,\alpha,\varphi+\psi}, b_{a,\alpha+\delta,\varphi}$ and $b_{a,\alpha+\delta,\varphi+\psi}$ must be monochromatic. However this also means that $b_{a,\alpha,\varphi+2\psi}$ and $b_{a,\alpha+\delta,\varphi+2\psi}$ must be monochromatic. We now turn to the next subblock in the arithmetic progression of subblocks, $SB_{a,\alpha}$ and $SB_{a,\alpha+\delta}$, this subblock must be in B_a since $SB_{a,\alpha}$ and $SB_{a,\alpha+\delta}$ are both in the first 3^7+1 subblocks of B_a. We label this subblock $SB_{a,\alpha+2\delta}$ and turn to the element $b_{a,\alpha+2\delta,\varphi+2\psi}$. If $b_{a,\alpha,\varphi+2\psi}, b_{a,\alpha+\delta,\varphi+2\psi}$ and $b_{a,\alpha+2\delta,\varphi+2\psi}$ are monochromatic then we have a monochromatic 3-term arithmetic progression, so we assume $b_{a,\alpha+2\delta,\varphi+2\psi}$ is a different colour. Equally we must assume that $b_{a,\alpha+2\delta,\varphi+2\psi}$ is a different colour to $b_{a,\alpha,\varphi}$ and $b_{a,\alpha+\delta,\varphi+\psi}$ otherwise, again, we have a monochromatic 3-term arithmetic progression.

We noted earlier that there must exist a block, B_{a+d}, coloured identically to B_a. We can also see that since B_a and B_{a+d} were both taken from the first $3^{7(2\cdot 3^7+1)} + 1$ blocks, the next block in the arithmetic progression of blocks, B_{a+2d}, is in the first $2 \cdot 3^{7(2\cdot 3^7+1)} + 1$ blocks. Therefore we turn to the element $b_{a+2d,\alpha+2\delta,\varphi+2\psi}$. $b_{a+2d,\alpha+2\delta,\varphi+2\psi}$ must be the same colour as one of $b_{a,\alpha,\varphi}$, $b_{a,\alpha,\varphi+2\psi}$ or $b_{a,\alpha+2\delta,\varphi+2\psi}$ since these are each coloured using a different colour in our 3-colouring. Therefore we either have the monochromatic 3-term arithmetic progression $b_{a,\alpha,\varphi}, b_{a+d,\alpha+\delta,\varphi+\psi}, b_{a+2d,\alpha+2\delta,\varphi+2\psi}$ or $b_{a,\alpha,\varphi+2\psi}, b_{a+d,\alpha+\delta,\varphi+2\psi}, b_{a+2d,\alpha+2\delta,\varphi+2\psi}$ or
$b_{a,\alpha+2\delta,\varphi+2\psi}, b_{a+d,\alpha+2\delta,\varphi+2\psi}, b_{a+2d,\alpha+2\delta,\varphi+2\psi}$. Since one of these monochromatic 3-term arithmetic progressions must exist in the set of integers, $\{1, 2, \ldots, 7(2 \cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)\}$, we must have that $W(3,3) \leq 7(2\cdot 3^7+1)(2\cdot 3^{7(2\cdot 3^7+1)}+1)$.

References

[1] R. Graham, B. Rothschild and J. Spencer, *Ramsey Theory*, Wiley-Interscience, New York. 1980.

[2] B. Bollobás, *Modern Graph Theory*, Springer Science, New York. 1998.

[3] F. Buckley and M. Lewinter, *A Friendly Introduction to Graph Theory*, Pearson Education LTD., London. 2003, p. 297–304.

[4] J. Winn, *Asymptotic Bounds for Classical Ramsey Numbers*, Polygonal, Washington. 1988.

[5] B. Bollobás, *Combinatorics: set systems, hypergraphs, families of vectors and combinatorial probability*, Cambridge University Press, Cambridge. 1986, p. 155–159.

[6] R. Graham, Rudiments of Ramsey Theory, *Conference Board of Mathematical Science.* **45** (1983).

[7] A. Khinchin, *Three Pearls of Number Theory*, Graylock Press, Rochester, New York. 1948, p. 11–14

[8] I. Leader, Partition Regular Equations, *Surveys in Combinatorics.* **307** (2003), p. 309–329.

[9] I. Leader and P. Russell, *Consistency for Partition Regular Equations*, submitted. (2003).

[10] B. Landman and A. Robertson, *Ramsey Theory on the Integers*, STML **24**, American Mathematical Society, Providence, Rhode Island. 2003.

[11] J. Hocking and G. Young, *Topology*, Addison-Wesley Publishing Company, Inc. Reading, Massachusetts. 1961.

[12] P. Alexandrow, *Combinatorial Topology*, Dover Publications, Inc. Mineola, New York. 1998.

[13] F. Ramsey, On a Problem of Formal Logic, *Proc. London Math. Soc.* **30** (1928), p. 264–286.

[14] B. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Archief voor Wiskunde* **15** (1927), p. 212–216.

[15] R. Rado, Studien zur Kombinatorik, *Math Zeit* **36** (1933), p. 242–280.

[16] I. Schur, Über die Kongruenz $x^m + y^m \equiv z^m \mod p$, *Jber. Deutsch. Math.-Verlein.* **25** (1916), p. 114–117.

[17] N. Hindman, Finite sums from sequences within cells of a partition of $\mathbb{N}$, *Journal of Combinatorial Theory, Ser. A.* **17** (1974), p. 1–11.

[18] F. Harary, A tribute to Frank P Ramsey, 1903-1930, *J. Graph Theory* **7** (1983), p. 1–7.