SUPER FIVE OF RAMSEY THEORY

JIALIN YANG

Abstract. This paper aims to give a friendly introduction to five of the "Super Six" [1] fundamental theorems in Ramsey Theory, the study of how order may arise from disorder. We begin with Ramsey's theorem, its generalizations, and a survey of known examples and bounds. We then consider how order arises in colorings of n-dimensional cubes, the natural numbers, and solutions to linear systems of equations. These occurrences are described by the Hales-Jewett, Van der Waerden's, and Rado's theorem, respectively.

Contents

1. Introduction 1
2. Ramsey's Theorem 2
3. The Hales-Jewett Theorem 6
4. Van der Waerden's Theorem 10
5. Schur's Theorem 11
6. Rado's Theorem 12
Acknowledgments 14
References 14

1. Introduction

Theodore Motzkin once said, "Complete disorder is impossible." This quote is a standard staple of papers on Ramsey Theory and is perhaps best thought of as a mathematician's excuse for his office space. Ramsey Theory studies how ordered structures may emerge from disorderly structures that grow extremely large. Although Ramsey Theory studies chaos, many of the theorems themselves are precise and elegant.

In Section 2, we begin with stating definitions needed for Ramsey's theorem and then prove Ramsey's Theorem Abridged and Unabridged. These theorems concern ordered structures that emerge from coloring graphs and hypergraphs. In Section 3, we prove the Hales-Jewett Theorem, which concerns coloring vertices of higher dimensional cubes. We also state the Graham-Leeb-Rothchild Theorem, the only member of the "Super Six" we do not prove. In Section 4, we prove Van der Waerden's theorem, which concerns coloring the natural numbers and finding arbitrarily long arithmetic progressions of a single color. In Sections 5 and 6, we prove Schur's Theorem and Rado's Theorem, respectively, both of which concern finding monochromatic solutions to equations.

Before we begin our journey into Ramsey Theory, we make a seemingly obvious statement that will be incredibly useful going forward.

Lemma 1.1 (Pigeonhole principle)**.** *If m pigeons roost in n holes and $m > n$, then at least two pigeons must share a hole.* □

For us, pigeons may be colorings and holes may be edges. Problems in Ramsey Theory often ask how many elements must be in a set for a given property to be true. We can, therefore, say that pigeons and holes encompass the fundamental spirit of Ramsey Theory.

2. Ramsey's Theorem

Notation 2.1. For $n \in \mathbb{N}$, denote the set $\{1, 2, \ldots, n\}$ as $[n]$.

Notation 2.2. For any set N and $k \in \mathbb{N}$, denote the set $\{M : M \subset N, |M| = k\}$ as $[N]^k$. Furthermore, in the special case where $N = [n]$ for some $n \in \mathbb{N}$, we remove the second pair of brackets and use $[n]^k$. Thus, $[n]^k$ is the set of subsets of $[n]$ with k elements.

A graph is a set $G = (V, E)$ such that V is a set of elements called *vertices* and E is a set of elements called *edges* that are sets of two distinct vertices, that is, unordered pairs of vertices. A complete graph is a graph such that every pair of distinct vertices has a unique edge connecting them. We note that a complete graph with n vertices can be represented as $[n]^2$.

Definition 2.3. Let S be any set. We say a map

$$\chi : S \to [r]$$

is an *r-coloring of S*. We say the *color of $s \in S$* is $\chi(s)$. We say $T \subset S$ is *monochromatic* if χ is constant on T.

Definition 2.4. The *Ramsey function* $R(l_1, \ldots, l_r)$ denotes the minimal n such that, for every r-coloring χ of $[n]^2$, there exists i, $1 \leq i \leq r$, and a set $T \subset [n]$, $|T| = l_i$, such that $[T]^2$ is colored i.

In the context of graphs, the Ramsey function $R(l_1, \ldots, l_r)$ denotes the minimal n such that if the edges of a complete graph with n vertices are colored with r colors, then for some $1 \leq i \leq r$, there exists a complete subgraph (its vertices and edges are subsets of the original graph) with l_i vertices and edges all colored i.

Example 2.5. We show $R(3, 3) = 6$. Observe we may color a regular pentagon so the five exterior edges are all red and the five interior diagonals are all blue. Since no monochromatic triangles are formed, we conclude that $R(3, 3) > 5$.

Proof. To show $R(3, 3) \leq 6$, we consider 6 vertices labeled by A, B, C, D, E, F. In this example, we call the vertices "people." Given a 2-coloring

$$\chi : [\{A, B, C, D, E, F\}]^2 \to \{1, 2\},$$

we say $S, T \in \{A, B, C, D, E, F\}$ "know" each other if $\chi(\{S, T\}) = 1$ and S, T "do not know" each other if $\chi(\{S, T\}) = 2$.

By the pigeonhole principle, A must either know 3 people or not know 3 people. Without loss of generality, suppose A knows B, C, and D. If any pair of B, C, or D know each other, say B and C, then $\chi(\{A, B\}) = \chi(\{A, C\}) = \chi(\{B, C\})$. Otherwise, B, C, and D mutually do not know each other, in which case we have also formed a mutual coloring, that is $\chi(\{B, C\}) = \chi(\{C, D\}) = \chi(\{B, D\})$. □

Remark 2.6. If $l_i \leq l'_i$ for all $1 \leq i \leq r$, then $R(l_1, \ldots, l_r) \leq R(l'_1, \ldots, l'_r)$. To illustrate, $R(4,4)$ must have the property $R(4,4) \geq R(3,3)$ because if 4 people mutually know each other, then we may consider just 3 of them who must also mutually know each other.

Lemma 2.7. *$R(l_1, l_2)$ is finite for all positive integers l_1, l_2.*

Proof. Let $l = \max(l_1, l_2)$. We will show that $2^{2l-1} \geq R(l_2, l_1)$. Let χ be a 2-coloring of $[2^{2l-1}]^2$. We define the sets S_i and positive integers x_i, $i \geq 1$, as the following:

(1) $S_1 = [2^{2l-1}]$.
(2) Having chosen S_i, select $x_i \in S_i$ arbitrarily.
(3) Having chosen x_i, denote

$$T_{i_j} = \{u \in S_i \colon \chi(\{x_i, u\}) = j\}, \qquad j = 1, 2.$$

Define S_{i+1} to be the set with larger cardinality of T_{i_1}, T_{i_2}.

We note that $|T_{i_1}| + |T_{i_2}| = |S_i| - 1$, so $|S_{i+1}| \geq \frac{|S_i|-1}{2}$.

Because $|S_1| = 2^{2l-1}$, this process will produce $x_1, \ldots, x_{2l-1}$. Define a new 2-coloring $\chi^* \colon \{x_1, \ldots, x_{2l-1}\} \to \{1, 2\}$ as $\chi^*(x_i) = j$ if $\chi(\{x_i, y\}) = j$ for all $y \in S_{i+1}$. Then, by the pigeonhole principle, we can choose $\{y_1, \ldots, y_l\} \subset \{x_1, \ldots, x_{2l-1}\}$ to be a monochromatic set under χ^*.

Consider $y_i, y_j \in \{y_1, \ldots, y_l\}$ with $i < j$. We know that $y_j \in S_j \subset S_{i+1}$, which means that $\chi(\{y_i, y_j\}) = \chi^*(y_i) = \chi^*(y_j)$ by construction. It follows that $\{y_1, \ldots, y_l\}$ is our monochromatic set under χ. □

There are many proofs for Lemma 2.7 [1, Theorem 1]. We have chosen our method because the proof for Ramsey's Theorem Unabridged generalizes this approach.

The classical proof uses double induction to show $R(l_1 - 1, l_2) + R(l_1, l_2 - 1) \geq R(l_1, l_2)$. We give a quick sketch of the proof here. The base case is obvious by $R(2, n) = R(n, 2) = n$. Consider a complete graph with $G = R(l_1 - 1, l_2) + R(l_1, l_2 - 1)$ vertices. Given a 2-coloring $\chi \colon [G] \to \{1, 2\}$, denote the colors $1, 2$ as *red* and *blue*, respectively. Let $v \in [G]$ be an arbitrary vertex. Take N to be the set of all vertices u such that the edge connecting v and u is colored red, that is, $\chi(\{v, u\}) = 1$. Take M to be the same for blue edges. Because $|G| = |N| + |M| + 1$, we have either $|N| \geq R(l_1, l_2 - 1)$ or $|M| \geq R(l_1 - 1, l_2)$. If $|N| \geq R(l_1, l_2 - 1)$ holds, then we have found our red monochromatic complete graph with l_1 vertices in N. Otherwise, $|M| \geq R(l_1 - 1, l_2)$ holds, so we have found our blue monochromatic complete graph with l_2 vertices in M. Thus, our double induction is complete.

The earliest bounds for $R(n, n)$ were

$$\frac{1}{e\sqrt{2}} n 2^{\frac{n}{2}} < R(n, n) \leq \binom{2n-2}{n-1}.$$

The upper bound is obtained by the double induction proof for Lemma 2.7. A summary of the lower bound proof is given as follows.

Theorem 2.8. *$R(n, n)$ has a lower bound of $\frac{1}{e\sqrt{2}} n 2^{\frac{n}{2}}$.*

Proof. There are $2^{\binom{m}{2}}$ distinct 2-colorings of $[m]^2$. The number of colorings that contain n vertices with mutual monochromatic edges is at most

$$\binom{m}{n} \frac{2^{\binom{m}{n}}}{2^{\binom{n}{2}+1}}.$$

It follows that there cannot exist n vertices with mutual monochromatic edges if

$$2^{\binom{m}{2}} > \binom{m}{n} 2^{\binom{m}{n} - \binom{n}{2} + 1}.$$

This inequality is true when $m > \frac{1}{e\sqrt{2}} n 2^{\frac{n}{2}}$ [2]. □

Details of the original proof can be found in [3]. The current best bounds are

$$\frac{\sqrt{n}}{2} n 2^{\frac{n}{2}} < R(n,n) < n^{-\frac{1}{2} + \frac{c}{\log n}} \binom{2n-2}{n-1},$$

which is sadly little progress in the intervening fifty years [2, Section 2.3.1].

Theorem 2.9 (Ramsey's Theorem Abridged)**.** *$R(l_1, \dots, l_r)$ is finite for all positive integers $l_1, \dots, l_r$.*

Proof. We will prove this theorem by inducting on r. We have shown our base case of $r = 2$ in Lemma 2.7. We move onto the inductive step. Let $l_1, \dots, l_r$ be positive integers. We claim that

$$R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r)) \geq R(l_1, \dots, l_r).$$

From our base case, we know that $R(l_{r-1}, l_r)$ is finite. From our inductive step, we know that $R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))$ is finite.

Fix an r-coloring χ of $[R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))]^2$. We define an $(r-1)$-coloring χ^* of $[R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))]^2$ by

$$\chi^*(\{x,y\}) = \begin{cases} i & \text{if } \chi(\{x,y\}) = i, 1 \leq i \leq r-2 \\ r-1 & \text{if } \chi(\{x,y\}) = i, i \geq r-1. \end{cases}$$

By the induction hypothesis applied to χ^*, we consider two cases.

1. Suppose there is i, $1 \leq i \leq r-2$, and a set $T \subset [R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))]$, $|T| = l_i$, such that $[T]^2$ is colored i under χ^*. By construction of χ^*, this set is also monochromatic under χ.
2. If the first case does not apply, then there is $T \subset [R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))]$, $|T| = R(l_{r-1}, l_r)$, such that $[T]^2$ is colored $r-1$. Consider the coloring of T under χ. By the definition of χ^*, $\chi([T]^2) = \{r-1, r\}$. By the construction of $R(l_{r-1}, l_r)$, there exists i, $r-1 \leq i \leq r$, and a set $T' \subset T \subset [R(l_1, \dots, l_{r-2}, R(l_{r-1}, l_r))]$, such that $[T']^2$ is colored i under χ. We take this to be our monochromatic set. □

Example 2.10. We show that $R(3,3,3) = 17$.

Proof. We will show that a 3-colored graph without a monochromatic triangle can be at most 16 vertices. Let n be the number of vertices in the graph and χ be a 3-coloring of $[n]$. Denote the colors as *red, green,* and *blue*. Let $v \in [n]$. Take S to be the set of all vertices u such that the edge connecting v and u is colored red, that is, $\chi(\{v,u\})$ is red. Because we do not want to form a monochromatic triangle, no two vertices in S can have a red edge, so any edge connecting two vertices in r must be green or blue. Because $R(3,3) = 6$, the cardinality of S must be less than 6, or a monochromatic triangle will be formed. It follows that S has at most 5 elements. We can repeat this process for the vertices with a green or blue edge connecting to v. Taken altogether, the entire graph can have at most $5 + 5 + 5 + 1 = 16$ vertices, 5 for each color and 1 for v itself. Therefore, we have shown $R(3,3,3) \leq 17$.

To show $R(3,3,3) > 16$, we can color a complete graph with 16 vertices using three colors without any monochromatic triangles. A beautiful diagram can be found in [4]. □

Only two nontrivial Ramsey numbers using three or more colors have been found, the other being $R(3,3,4) = 30$ [5, Section 6].

Definition 2.11. For a positive integer $k \geq 1$, the *Ramsey function* $R_k(l_1, \dots, l_r)$ denotes the minimal n such that, for every r-coloring χ of $[n]^k$, there exists i, $1 \leq i \leq r$, and a set $T \subset [n]$, $|T| = l_i$, such that $[T]^k$ is colored i.

This definition generalizes Definition 2.4. If k is not given, it is assumed to be 2.

Theorem 2.12 (Ramsey's Theorem Unabridged)**.** *$R_k(l_1, \dots, l_r)$ is finite for all positive integers $l_1, \dots, l_r$.*

Proof. We will prove this theorem by inducting on k. For our base case of $k = 1$, we know that

$$1 + \sum_{i=1}^{r} (l_i - 1) \geq R_1(l_1, \dots, l_r).$$

By the pigeonhole principle, there must be be a color i that is used l_i times.

We move onto the induction step. Let $l = \max(l_1, \dots, l_r)$. We claim that

$$2r^c \geq R_k(l_1, \dots, l_r), \qquad \text{where } c = \sum_{i=k-1}^{t-1} \binom{i+1}{k+1}, t = R_{k-1}(\underbrace{l, \dots, l}_{r}).$$

Let $n = 2r^c$. Let χ be an r-coloring of $[n]^k$. Select $a_1, \dots, a_{k-2} \in [n]$ arbitrarily. Define $a_i, S_i, i \geq k-2$ as

(1) $S_{k-2} = [n] \setminus \{a_1, \dots, a_{k-2}\}$.
(2) Having chosen S_i, select $a_{i+1} \in S_i$ arbitrarily.
(3) Having chosen a_{i+1}, split $S_i \setminus \{a_{i+1}\}$ into equivalence classes by $x \equiv y$ if and only if for every $T \subset \{a_1, \dots, a_{i+1}\}, |T| = k-1,$ we have

$$\chi(T \cup \{x\}) = \chi(T \cup \{y\}).$$

From $\{a_1, \dots, a_{i+1}\}$, there are $\binom{i+1}{k-1}$ ways to select T. It follows that there exist at most $r^{\binom{i+1}{k-1}}$ classes. We set S_{i+1} to be the largest of those classes.

For all $i \geq 1$, we have that $S_{i+1} \subset S_i \setminus \{a_{i+1}\}$ and $r^{\binom{i+1}{k-1}}|S_{i+1}| \geq (|S_i| - 1)$. This suggests that we need to define n such that $|S_{k-2}| = n - (k-2)$ and the recursion $|S_{i+1}| \geq (|S_i| - 1)r^{-\binom{i+1}{k-1}}$ satisfies $|S_{t-1}| \geq 1$, that is, a_t is defined.

Consider $A = \{a_1, \dots, a_t\}$. Define χ^* to be an r-coloring of $[A]^{k-1}$ by

$$\chi^*(\{a_{i_1}, \dots, a_{i_{k-1}}\}) = \chi(\{a_{i_1}, \dots, a_{i_{k-1}}, a_s\})$$

for $1 \leq i_1 < i_2 < \dots < i_{k-1} < s \leq t$. If $a_{i_{k-1}} = t$, we define χ^* arbitrarily. By the construction of t, there exists $\{b_1, \dots, b_l\} \subset \{a_1, \dots, a_t\}$ such that χ^* maps all $k-1$ cardinality subsets of $\{b_1, \dots, b_l\}$ to a single color, denoted as *red*. By the construction of χ^*, for all $1 \leq i_1 < i_2 < \dots < i_{k-1} < i_k \leq l$, we have

$$\chi(\{b_{i_1}, ..., b_{i_{k-1}}, b_{i_k}\} = \chi^*(\{b_{i_1}, ..., b_{i_{k-1}}\}),$$

which means every k cardinality subset of $\{b_1, \dots, b_l\}$ will also be colored red. Thus, we have found our monochromatic subset. □

There is only one known value for Ramsey numbers of hypergraphs, which is $R_3(4,4) = 13$ [5, Section 7.1]. To show $R_3(4,4) \geq 13$, we know there exist 434714 different 2-colorings of $[12]^3$, none of which contain our desired monochromatic structure. We opt not to list out all 434714 colorings in our paper. A computer evaluation in 1991 improved the upper bound from 15 to 13, so we can conclude $R_3(4,4) = 13$. In general, we know that [5, Section 7.3]

$$2^{cn^2} < R_3(n,n) < 2^{2^n}.$$

Further details about the Ramsey numbers of hypergraphs can be found in [5], but such calculations are beyond the scope of this paper.

3. The Hales-Jewett Theorem

Definition 3.1. We say a set

$$C_t^n \coloneqq \{(x_1, \dots, x_n) \colon x_i \in [t]\}$$

is the *n-cube over t* elements.

Definition 3.2. We call a *line in C_t^n* a set of points $\{\mathbf{x}_1, \dots, \mathbf{x}_t\} \subset C_t^n$ with $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ such that, for all $j, 1 \leq j \leq n$,

$$x_{1j} = x_{2j} = \cdots = x_{t,j} \qquad \text{or} \qquad x_{sj} = s \text{ for } s \in [t]$$

and the latter condition occurs for at least one j.

Notation 3.3. We will sometimes represent $(x_1, x_2, \dots, x_n)$ as $x_1 x_2 \dots x_n$ for conciseness. Given two points $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$, we may also write $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m) = x_1 \dots x_n y_1 \dots y_m$.

Example 3.4. For $t = 4, n = 3$, the set $\{111,212,313,414\}$ is a line in C_4^3. Our definition, however, differs from the geometric interpretation of a line, as $\{13, 22, 31\}$ is not a line in C_3^2.

In 1987, Saharon Shelah found a new proof for the Hales-Jewett Theorem. We will show Shelah's proof in this paper. It is helpful to provide some additional definitions to assist with the proof.

Definition 3.5. Given a point $x = (x_1, \dots, x_n)$, we call n the *length* of x.

Definition 3.6. We call a set of two points $a, b \in C_t^n$ *neighbors* if they differ at exactly one coordinate and that coordinate in a, b is 1 and 2, respectively.

Example 3.7. The points $(3,3,1,3)$ and $(3,3,2,3)$ are neighbors, but $(3,3,3,3)$ is not a neighbor to either.

Definition 3.8. A point $x = (x_1, \dots, x_n)$ is called a *root* if $x_i \in (\{[t] \cup \{*\})$ and at least one $x_i = *$. Given a root τ, for $s \in [t]$, we denote $\tau(s)$ as the result of replacing the $*$ in the root with s.

Example 3.9. Let $\tau = (1, *, 2, *, 3)$ be a root in C_3^5. Then, we have $\tau(2) = (1,2,2,2,3) = 12223$ and $\tau(3) = (1,3,2,3,3) = 13233$.

Remark 3.10. Any root τ can represent a line by substituting $*$ for the corresponding position in the line. For example, the root $\tau = (1, *, 2, *, 3)$ in C_3^5 represents the line $\{11213, 12223, 13233\}$.

Definition 3.11. Two or more roots $\tau_1, \tau_2, \ldots, \tau_n$ may be concatenated to form a *concatenated root* $\tau = (\tau_1, \tau_2, \ldots, \tau_n) = \tau_1\tau_2 \ldots \tau_n$. Given this concatenated root, we say that $\tau(x)$ where $x = (x_1, x_2, \ldots, x_n)$ is equal to $(\tau_1(x_1), \tau_2(x_2), \ldots, \tau_n(x_n))$.

Theorem 3.12 (Hales-Jewett)**.** *For all $r, t \in \mathbb{N}$, there exists a natural number $N' = HJ(r,t)$ such that, for $N \geq N'$, if the set C_t^N is r-colored, there exists a monochromatic line.*

Proof. We will prove this theorem by inducting on t. For our base case, we have that $HJ(r,1) = 1$. A point of length 1 forms a monochromatic line by itself. Assume that $n = HJ(r, t-1)$ exists. We want to show $HJ(r,t)$ exists.

We define the numbers N_i, $0 \leq i \leq n$ as the following:

(1) $N_0 = 0$.
(2) $N_i = r^{t^n + \sum_{j=0}^{i-1} N_j}$.

Let $N = N_1 + \cdots + N_n = \displaystyle\sum_{i=1}^{n} r^{t^n + \sum_{j=1}^{i-1} N_j}$.

Claim. Given an r-coloring χ of C_t^N, there exists a concatenated root $\tau = \tau_1 \ldots \tau_n$ such that the following holds.

(1) For $1 \leq i \leq n$, the length of τ_i is N_i. This implies that the length of τ is N.
(2) The root τ has the property that for any neighbors $a, b \in C_t^N$, we have that $\chi(\tau(a)) = \chi(\tau(b))$.

Proof of Claim. We proceed by backwards induction. Assume we have found roots $\tau_{i+1}, \ldots, \tau_n$. We can choose τ_n with a simple application of the pigeonhole principle. Define $P_i = (x_{i1}, ..., x_{iN_n})$, for $0 \leq i \leq n$, by

$$x_{ij} = \begin{cases} 1, & j \leq i \\ 2, & j > i. \end{cases}$$

Given that $N_n + 1 > r$, by the pigeonhole principle, we know two of these points P_a, P_b must have the same color under χ if the points from the first coordinate to the $(N - N_n)^{\text{th}}$ coordinate are given. Define

$$\tau_n = \underbrace{1 \ldots 1}_{a}\underbrace{* \cdots *}_{b-a}\underbrace{2 \ldots 2}_{n-b}.$$

This root is length N_n and satisfies $\chi(\tau(a)) = \chi(\tau(b))$ if a, b differs on the n^{th} coordinate.

We want to show τ_i exists.

Let $L_{i-1} = \displaystyle\sum_{j=1}^{i-1} N_j$ be the length of $\tau_1 \ldots \tau_{i-1}$. For all $0 \leq k \leq N_i$, let

$$W_k = \underbrace{1 \ldots 1}_{k}\underbrace{2 \ldots 2}_{N_i - k}.$$

For all $0 \leq k \leq N_i$, let the r-coloring χ_k of $C_t^{L_{i-1}+n-i}$ be defined as

$$\chi_k(x_1 \ldots x_{L_{i-1}} y_{i+1} \ldots y_n) = \chi(\tau_1(x_1) \ldots \tau_{L_{i-1}}(x_{L_{i-1}}) W_k \tau_{i+1}(y_{i+1}) \ldots \tau_n(y_n)).$$

Then, we have defined $N_i + 1$ number of colorings $\chi_0, \ldots, \chi_{N_i}$. For all $1 \leq i \leq n$, we have that there are a total of $t^{L_{i-1}+n-i}$ points in $C_t^{L_{i-1}+n-i}$, which means there can only exist $r^{t^{L_{i-1}+n-i}}$ distinct colorings of $C_t^{L_{i-1}+n-i}$. Because $r^{t^{L_{i-1}+n-i}} \leq$

$r^{t^{L_{i-1}+n}} = N_i$, by the pigeonhole principle, two colorings χ_s and χ_k must be identical. Without loss of generality, assume $s < k$. Let $\tau_i = \underbrace{1\dots1}_{s}\underbrace{*\cdots*}_{k-s}\underbrace{2\dots2}_{N_i-k}$. We want to show this root satisfies the conditions of the claim.

(1) By construction, the length of τ_i is N_i.
(2) We have that

$$\tau_i(1) = \underbrace{1\dots1}_{s}\underbrace{1\dots1}_{k-s}\underbrace{2\dots2}_{N_i-k} = W_k$$

and

$$\tau_i(2) = \underbrace{1\dots1}_{s}\underbrace{2\dots2}_{k-s}\underbrace{2\dots2}_{N_i-k} = W_s.$$

Let $a = (a_1, \dots, a_{i-1}, 1, a_{i+1}, \dots, a_n)$ and $b = (a_1, \dots, a_{i-1}, 2, a_{i+1}, \dots, a_n)$ be neighbors that differ in the i^{th} coordinate.

We have that

$$\begin{aligned}
\chi(\tau(a)) &= \chi(\tau_1(a_1)\dots\tau_{i-1}(a_{i-1})\tau_i(1)\tau_{i+1}(a_{i+1})\dots\tau_n(a_n)) \\
&= \chi(\tau_1(a_1)\dots\tau_{i-1}(a_{i-1})W_k\tau_{i+1}(a_{i+1})\dots\tau_n(a_n)) \\
&= \chi_k(a_1\dots a_{i-1}a_{i+1}\dots a_n) \\
&= \chi_s(a_1\dots a_{i-1}a_{i+1}\dots a_n) \\
&= \chi(\tau_1(a_1)\dots\tau_{i-1}(a_{i-1})W_s\tau_{i+1}(a_{i+1})\dots\tau_n(a_n)) \\
&= \chi(\tau_1(a_1)\dots\tau_{i-1}(a_{i-1})\tau_i(2)\tau_{i+1}(a_{i+1})\dots\tau_n(a_n)) \\
&= \chi(\tau(b)).
\end{aligned}$$

Therefore, by backwards induction, we have shown that the claim is true.

We now move onto the rest of the proof. Recall that $n = HJ(r, t-1)$ exists. We want to show that $HJ(r,t) \leq N$. From our claim, there exists a root τ of length N such that $\chi(\tau(a)) = \chi(\tau(b))$ for any two neighbors $a, b \in C_t^N$.

Let χ^* be an r-coloring of C_{t-1}^N be defined as $\chi^*(x) = \chi(\tau(x))$. By the definition of n, we know that there exists a root $V = V_1 \dots V_n$ such that, for $1 \leq i \leq n$, $V_i \in ([t-1] \cup \{*\})$, V_i is length 1, and the line $\{V(1), \dots, V(t-1)\} \in C_{t-1}^N$ is monochromatic under χ^*.

We want to show $\tau(V) = \tau_1(V_1)\dots\tau_n(V_n)$ is a root. Because V is a root, there exists some $V_i = *$. It follows that some $\tau_i(V_i)$ contains a $*$, so $\tau(V)$ is a root.

We will prove that $\{\tau(V(1)), \dots, \tau(V(t))\}$ is our monochromatic line. Consider $\tau(V(i)), \tau(V(i+1)) \in \{\tau(V(1)), \dots, \tau(V(t))\}$ for $1 \leq i < t$. If $\tau(V)$ only has one $*$ and therefore differ in only one coordinate, then by the construction of τ, we know $\chi(\tau(V(i)) = \chi(\tau(V(i+1))$. If $\tau(V)$ contains more than one $*$, then we can still show they are the same color. For example, if $\tau(V)$ contains three $*$ we have that

$$\begin{aligned}
\chi(\tau(V(i))) &= \chi(...i...i...i...) = \chi(...(i+1)...i...i...) \\
&= \chi(...(i+1)...(i+1)...i...) = \chi(...(i+1)...(i+1)...(i+1)...) \\
&= \chi(\tau(V(i+1))).
\end{aligned}$$

Thus, we have shown that a monochromatic line must exist given any arbitrary r-coloring of C_t^N. □

In their original paper, A. W. Hales and R. I. Jewett looked at how the Hales-Jewett Theorem could be applied to generalizations of tic-tac-toe. The original game, played by two players on a 3×3 grid, is known to end in a draw (though perhaps not against one's less attentive friends). We may think of the original game as two players coloring C_3^2 with their own color. A player wins when he or she obtains an entire line. By the Hales-Jewett Theorem, if N is large enough, then (borrowing terminology from [1, Section 2.3]) the "r-person, N-dimensional, t-in-a-row tic-tac-toe " cannot end in a draw. This is because there will always exist a line under our more restrictive definition. Finite games with perfect information that do not end in a draw must always have a winning strategy, and in symmetrical games, the winning strategy always belongs to the first player [6]. Thus, the first player must always have a winning strategy for a large enough multidimensional tic-tac-toe board. The exact dimension when the first player always has a winning strategy on a r-person, t-in-a-row tic-tac-toe is not known, but we have an upper bound with $HJ(r, t)$.

The exact values given by the Hales-Jewett Theorem are difficult to find. The only known non-trivial values where $r > 1$ are $HJ(2, 2) = 2$ and $HJ(2, 3) = 4$ [6]. Additionally, the upper bounds of the Hales-Jewett Theorem are incredibly large. How large?

We now take a brief detour to explore how much Shelah's proof improved the upper bound of the Hales-Jewett Theorem. For convenience, we only consider when $r = 2$. It is necessary to first define the Ackermann hierarchy, a sequence of rapidly growing functions. For the sake of consistency and humor, we use the exact language found in [1, Section 2.7, "EEEEENORMOUS UPPER BOUNDS"].

We define the sequence of functions $f_1, f_2, ...$ by the following:

(1) $f_1(x) = 2x$.
(2) $f_2(x) = 2^x$. We notice that we can derive f_2 from f_1 by starting at 1 and applying f_1 a total of x times.
(3) $f_3(x) = 2^{2^{\cdot^{\cdot^{2}}}}$ where there are x number of twos. We obtained f_3 by starting at 1 and applying f_2 a total of x times.
(4) $f_{i+1} = f_i^{(x)}(1)$, where $f^{(x)}$ denotes the x^{th} iteration of f.

The sequence can also be defined by

(1) $f_{i+1}(1) = 2$,
(2) $f_{i+1}(x + 1) = f_i(f_{i+1}(x))$.

This sequence of functions, as the reader may have noticed, grows extremely quickly. For context, $f_3(5) = 2^{65536}$ is a number with $20,000$ digits. We call f_4 the WOW function because this "fanciful description comes from trying to grasp the magnitude of $f_4(5)$ – a tower of twos of size 65536 – what can we say but "oh wow" [1, Section 2.7]!

We now define the the Ackermann function, denoted by f_ω or ACKERMANN. This function is defined as

$$f_\omega(x) = \text{ACKERMANN}(x) = f_x(x).$$

Observe that for any $x \geq n$, then $\text{ACKERMANN}(x) = f_x(x) \geq f_n(x)$. It follows that the Ackermann functions grows faster than f_n, for all $n \in \mathbb{N}$. We say that a function g has a level i (including $i = \omega$) if there exist $c', c'' > 0$ and x' such that if $x > x'$,

$$f_i(c'x) < g(x) < f_i(c''x).$$

For $i = 1, 2, 3, 4, \omega$, we use the terms linear, exponential, towerian, wowzer, and *ackermanic*, respectively, to describe the level of g.

We are now ready to describe Shelah's breakthrough. All previous proofs of the Hales-Jewett theorem *gave an ackermanic function as an upper bound.* Shelah's proof yields a *wowzer* upper bound. For such a fundamental improvement, Shelah was awarded $500 by the senior author of *Ramsey Theory* [1].

***

The next theorem of the "Super Six" is the Graham-Leeb-Rothchild Theorem, otherwise known as the Affine Ramsey Theorem, which can be proven as a result of the Hales-Jewett Theorem. As the name would suggest, this theorem can be seen as Ramsey's Theorem for spaces. We omit this proof because it is highly technical. Details can be found in [1, Section 2.4, Theorem 9].

Theorem 3.13. *Let F be a finite field. For all positive integers r, t, k there exists a positive integer n_0 such that the following holds for $n > n_0$. Let V be an n-dimensional vector space over F. Color the t-dimensional subspaces of V with r colors. Then there exists an k-dimensional subspace of V all of whose t-dimensional subspaces have the same color.*

4. Van der Waerden's Theorem

Although Van der Waerden's Theorem came chronologically before the Hales-Jewett Theorem, we can obtain Van der Waerden's Theorem as a corollary to the Hales-Jewett Theorem.

Van der Waerden originally set out to prove a simpler theorem than what he ultimately proved.

Theorem 4.1. *If $\mathbb{N}$ is 2-colored, then there exist monochromatic arithmetic progressions of arbitrary length.*

We make two modifications to the original theorem. First, instead of only 2 colors, we allow for any positive integer r number of colors. Second, we consider the first natural number such that if the number is r-colored, then at least one color contains an arithmetic progression of l terms. This leads us to a modified statement.

Theorem 4.2 (Van der Waerden)**.** *For all $l, r \in \mathbb{N}$, there exists a natural number $W(l, r)$ such that if $[W(l, r)]$ is r-colored, then there exists a monochromatic arithmetic progression of length l.*

Proof. Let $n = HJ(r, t)$. Take $W(l, r) = nt + 1$. Let χ be an r-coloring of $[W(l, r)]$. Let $f : C_t^n \to [W(l, r)]$ by

$$f(x_1, \dots, x_n) = x_1 + \cdots + x_n.$$

The last point of $f(C_t^n)$ is $\sum_{i=1}^{n} t = nt$, which is less than $W(l, r)$. Then, we can define r-coloring χ^* of C_t^n as $\chi^*(x) = \chi(f(x))$.

By the Hales-Jewett Theorem, there exists a monochromatic line $\{\tau(1), \dots, \tau(t)\}$ in C_t^n. For $1 \le i < t$, the difference $f(\tau(i + 1)) - f(\tau(i))$ is the number of $*$ in τ.

Because this is a constant, we can take $f(\tau(1)), \ldots, f(\tau(t))$ to be our monochromatic arithmetic progression of length t. $\square$

For any $r \in \mathbb{N}$, we may take $r+1$ to satisfy the conditions $W(2,r)$ by the pigeonhole principle. Let us consider the first nontrivial example, where $l = 3, r = 2$.

Example 4.3. We show $W(3,2) \leq 325$.

Proof. We want to find a monochromatic arithmetic progression of length 3. Let χ be a 2-coloring of $[325]$. We call a set $B_i = \{5i + j \mid i \in [5]\}$ a "block." We can write $[325]$ as the union of 65 blocks, that is,

$$[325] = B_1 \cup \cdots \cup B_{65} = \{1,2,3,4,5\} \cup \cdots \cup \{321, 322, 323, 324, 325\}.$$

Because there are only two colors, there are at most $2^5 = 32$ distinct colorings of a single block. By the pigeonhole principle, two blocks in the first 33 blocks B_{b_1}, B_{b_2} must have the same coloring. This means that $\chi(5b_1 + i) = \chi(5b_2 + i)$ for all $1 \leq i \leq 5$. Consider $5b_1 + 1, 5b_1 + 2, 5b_1 + 3$. By the pigeonhole principle, two of the numbers must have the same color. Denote these numbers as $5b_1 + a_1, 5b_1 + a_2$ where $a_1 < a_2$ and denote their color as *red*. Let $a_3 = 2a_2 - a_1$ and consider $5b_1 + a_3$.

(1) If $5b_1 + a_3$ is red, then we have found our red monochromatic arithmetic progression $5b_1 + a_1, 5b_1 + a_2, 5b_1 + a_3$.
(2) If $5b_1 + a_3$ is red, denote its color as *blue*. Because B_{b_1} and B_{b_2} have identical colorings, we know that $5b_2 + a_3$ is also blue. Let $b_3 = 2b_2 - b_1$. We observe that $5b_3 + a_3 \leq 325$.
  (a) If $5b_3 + a_3$ is red, then we have found our red monochromatic arithmetic progression $5b_1 + a_1, 5b_2 + a_2, 5b_3 + a_3$.
  (b) If $5b_3 + a_3$ is blue, then we have found our blue monochromatic arithmetic progression $5b_1 + a_3, 5b_2 + a_3, 5b_3 + a_3$. $\square$

The original proof of Van der Waerden's Theorem involved a double induction on l, the length of the arithmetic progression, and r, the number of colors. The idea is to divide the set of integers into equal sized blocks of consecutive integers and apply the induction hypothesis to the blocks. Details can be found in [1, Section 2.1, Theorem 2].

5. Schur's Theorem

Theorem 5.1 (Schur)**.** *For all $r \in \mathbb{N}$, there exists a natural number $S(r)$ such that if $[S(r)]$ is r-colored, there exist $x, y, z \in [S(r)]$ having the same color such that*

$$x + y = z.$$

Proof. Let $r \in \mathbb{N}$. Let $S(r)$ be such that

$$S(r) = R(\underbrace{3, \ldots, 3}_{r}).$$

Let χ be an r-coloring of $S(r)$. Let χ^* be a new r-coloring defined as $\chi^*(i,j) = \chi(|i-j|)$. By construction of $S(r)$, we know there exist $a, b, c \in [S(r)]$ with $a > b > c$ such that $\chi^*(a,b) = \chi^*(b,c) = \chi^*(a,c)$. Let $x = a - b, y = b - c, z = a - c$. Then, we have that $x + y = z$ and $\chi(x) = \chi(y) = \chi(z)$. $\square$

Theorem 5.2. *For all $k, r, s \in \mathbb{N}$, there exists $n = n(k, r, s)$ such that if $[n]$ is r-colored, there exist $a, d \in [n]$ such that*

$$\{a, a + d, a + 2d, \ldots, a + kd\} \cup \{sd\}$$

is monochromatic.

Proof. We will prove this theorem by inducting on r. In our base case, we take $n(k, 1, s) = \max(k + 1, s)$. Because every element in $[n(k, 1, s)]$ is the same color, we will have a monochromatic set $\{a, a + d, a + 2d, \ldots, a + kd\} \cup \{sd\}$ where $a = 1, d = n(k, 1, s) - 1$.

We proceed to the inductive step. Assume $n(k, r - 1, s)$ exists. By Theorem 4.2 (Van der Waerden), let $W(t, r) \in \mathbb{N}$ such that if $W(t, r)$ is r-colored, there exists a t length arithmetic progression. Take

$$n(k, r, s) = sW(kn(k, r - 1, s), r).$$

Fix an r-coloring of $n(k, r, s)$. By the definition of $W(kn(k, r - 1, s), r)$ we know there exists a monochromatic set $\{a + id' : 0 \le i \le kn(k, r - 1, s)\}$. We denote this set's color as *mustard yellow* (incidentally, this happens to be Nikolay's favorite color). We have two cases.

(1) If there exists $j \in \mathbb{N}$ such that $1 \le j \le n(k, r - 1, s)$ and sdj' is mustard yellow, then we take $d = jd'$. It follows that all elements of the set
$$\{a, a + d, a + 2d, \ldots, a + kd\} \cup \{sd\}$$
are colored mustard yellow.

(2) If, for all j, $1 \le j \le n(k, r - 1, s)$, sdj' is not mustard yellow, then we know that the set $\{sd'j : 1 \le j \le n(k, r - 1, s)\}$ is colored with $r - 1$ colors (mustard yellow is not included). By the definition of $n(k, r - 1, s)$, we know there exist some a', d such that
$$\{a', a' + d, a' + 2d, \ldots, a' + kd\} \cup \{sd\}$$
is monochromatic.

By induction, $n(k, r, s)$ exists for all $r \in \mathbb{N}$. This proves the theorem. □

The following is immediate from the preceding theorem.

Corollary 5.3. *For all $k, r, s \in \mathbb{N}$, there exists $n = n(k, r, s)$ such that if $[n]$ is r-colored, there exist $a, d \in [n]$ such that*

$$\{a + \lambda d : \lambda \in \mathbb{Z}, |\lambda| \le k\} \cup \{sd\}$$

is monochromatic.

6. Rado's Theorem

Let $S = S(x_1, \ldots, x_n)$ be a linear system of equations in the variables $x_1, \ldots, x_n$.

Definition 6.1. Let $r \in \mathbb{N}$. We say S is *r-regular* if, given any r-coloring of $\mathbb{N}$, there exist $x_1, \ldots, x_n \in \mathbb{N}$ such that $x_1, \ldots, x_n$ satisfy S and are monochromatic.

Definition 6.2. We say S is *regular* if S is r-regular for all $r \in \mathbb{N}$.

Theorem 6.3 (Rado)**.** *The equation*

$$c_1x_1 + \ldots_n x_n = 0, \qquad c_i \in \mathbb{Z}$$

is regular if and only if the elements of a nonempty subset of $\{c_1, \ldots, c_n\}$ sum to 0.

Proof. We will show this theorem in both directions.

($\Longleftarrow$) Assume there exists $k \le n$ such that $c_1 + \cdots + c_k = 0$. If $k = n$, then we take $x_1 = \cdots = x_n = 1$ to be our monochromatic solution. Therefore, we assume $k < n$. Let $A = \gcd(c_1, \ldots, c_k)$, $B = c_{k+1} + \cdots + c_n$. We note that if $B = 0$, we have $k = n$, so we can assume $B \neq 0$. Let $s = \frac{A}{\gcd(A,b)}$. Because $\frac{Bs}{A} \in \mathbb{Z}$, there exists $t \in \mathbb{Z}$ such that

$$At + Bs = 0.$$

By Bézout's identity [7], we know that there exist $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$ such that

$$c_1\lambda_1 + \cdots + c_k\lambda_k = At.$$

Let $k = \max(|\lambda_1|, \ldots, |\lambda_k|)$. By Corollary 5.3, there exist $a, d \in \mathbb{N}$ such that $\{a + \lambda d \colon \lambda \in \mathbb{Z}, |\lambda| \le k\} \cup \{sd\}$ is a monochromatic set. We take our solution to be

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

Then, we have that

$$c_1x_1 + \cdots + c_kx_k = c_1(a + \lambda_1) + \cdots + c_k(a + \lambda_k) = 0 \cdot a + At = At.$$

We also have

$$c_{k+1}x_{k+1} + \cdots + c_nx_n = Bsd.$$

It follows that we have our monochromatic solution.

($\Longrightarrow$) We will show this direction by proving the contrapositive. Let $c_1, \ldots, c_n$ be such that no nonempty subset of $\{c_1, \ldots, c_n\}$ sums to 0. We will define a coloring of $\mathbb{Q} \setminus \{0\}$ so that the equation has no monochromatic solution.

Let $p > 0$ be a prime such that p does not divide the sum of any nonempty subset of $\{c_1, \ldots, c_n\}$. Consider an arbitrary $q \in \mathbb{Q} \setminus \{0\}$. We know that q may be uniquely expressed as

$$q = \frac{p^j a}{b}, \qquad j \in \mathbb{Z}, a \in \mathbb{Z}, b \in \mathbb{N}, \gcd(a, b) = 1, p \nmid ab.$$

Define a $(p-1)$-coloring χ of $\mathbb{Q} \setminus \{0\}$ by

$$\chi(q) = \frac{a}{b} \pmod{p}$$

where $\frac{a}{b} \pmod{p} = x$ if and only if $b \cdot x \pmod{p} = 1$. Note that $\chi(x) = \chi(y)$ implies $\chi(\alpha x) = \chi(\alpha y)$ for all $\alpha \in \mathbb{Q} \setminus \{0\}$. Assume there exist $y_1, \ldots, y_n \in \mathbb{Q} \setminus \{0\}$ that is a monochromatic solution. Then, we note that $\mu y_1, \ldots, \mu y_n$ is also a monochromatic solution for all $\mu \in \mathbb{Q} \setminus \{0\}$. Then, we can take $x_1, \ldots, x_n \in \mathbb{Z}, \gcd(x_1, \ldots, x_n) = 1$ to be a monochromatic solution.

We can reorder $x_1, \ldots, x_n$ such that

$$p \nmid x_i, 1 \le i \le k \qquad \text{and} \qquad p \mid x_i, k < i \le n.$$

We know that $k \ge 1$ because $x_1, \ldots, x_n$ are relatively prime. We reduce the equation modulo p to obtain

$$\sum_{i=1}^{n} \bar{c}_i \bar{x}_i = \bar{0} \qquad \text{in } \mathbb{Z}/p\mathbb{Z}$$

where $\bar{a}$ represents the residue class $a \pmod p$. By construction, we have that $\bar{x_i} = \bar{0}$ for $k < i \leq n$. Because $\{x_1, \dots, x_n\}$ is monochromatic, we know that $\bar{x_i}$, $1 \leq i \leq k$ are equal. Thus, we obtain

$$\sum_{i=1}^{n} \bar{c_i}\bar{x_i} = \sum_{i=1}^{k} \bar{c_i}\bar{x_i} = \left(\sum_{i=1}^{n} \bar{c_i}\right) \bar{x_1}.$$

Since $\bar{x_1} \neq \bar{0}$ and p is prime, we have $\sum_{i=1}^{n} \bar{c_i} = 0$, which contradicts our assumptions. Thus, we have shown the contrapositive. □

Rado's Theorem can be generalized to an arbitrary system of linear equations. We will state the theorem here but omit the proof because it is rather technical. Details can be found in [1, Section 3.3, Theorem 5].

Definition 6.4. A matrix $C = (c_{ij})$ satisfies the *columns condition* if one can order the column vectors $\mathbf{c}_1, \dots, \mathbf{c}_n$ and find $1 \leq k_1 < k_2 < \cdots < k_t = n$ such that, setting

$$\mathbf{A}_i = \sum_{j=k_{i-1}+1}^{k_i} \mathbf{c}_j$$

we have:

(1) $\mathbf{A}_1 = 0$.
(2) For $1 \leq i < t$, $\mathbf{A}_i$ may be expressed as a linear combination of $\mathbf{c}_1, \dots, \mathbf{c}_{k_{i-1}}$.

Theorem 6.5 (Rado's Theorem Complete)**.** *The system of equations $Cx = 0$ is regular on N if and only if C satisfies the columns condition.*

Acknowledgments

I would like to thank my mentor, Nikolay Grantcharov, for guiding me through mathematical research and for being a living legend. I would like to thank my friends who have kept me company during lockdown for their support. Finally, I would thank Peter May for his efforts in organizing and running the REU.

References

[1] Ronald Graham, Bruce Rothschild, and Joel Spencer. Ramsey Theory. John Wiley & Sons, Inc. 1980.
[2] Fan Chung and Ronald Graham. Erdõs on Graphs : His Legacy of Unsolved Problems. http://www.math.ucsd.edu/~fan/epbk/ramsey.pdf.
[3] Paul Erdos. Some Remarks on the Theory of Graphs. https://projecteuclid.org/download/pdf_1/euclid.bams/1183510596.
[4] Wikipedia. Ramsey's Theorem. https://en.wikipedia.org/wiki/Ramsey%27s_theorem.
[5] Stanisław P. Radziszowski. Small Ramsey Numbers. https://www.combinatorics.org/ojs/index.php/eljc/article/view/DS1/pdf.
[6] Marcus Näslund. The Hales-Jewett theorem and its application to further generalisations of m, n, k-games. http://uu.diva-portal.org/smash/get/diva2:631339/FULLTEXT01.pdfd.
[7] Eric Weisstein. "Bézout's Identity." https://mathworld.wolfram.com/BezoutsIdentity.html.