SCHUR’S THEOREM AND RELATED TOPICS IN RAMSEY THEORY

by

Summer Lynne Kisner

A thesis
submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Mathematics
Boise State University

May 2013
BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Summer Lynne Kisner

Thesis Title: Schur’s Theorem and Related Topics in Ramsey Theory

Date of Final Oral Examination: 12 March 2013

The following individuals read and discussed the thesis submitted by student Summer Lynne Kisner, and they evaluated her presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Andrés E. Caicedo, Ph.D. Chair, Supervisory Committee
Marion Scheepers, Ph.D. Member, Supervisory Committee
Zach Teitler, Ph.D. Member, Supervisory Committee

The final reading approval of the thesis was granted by Andrés E. Caicedo, Ph.D., Chair, Supervisory Committee. The thesis was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.
ACKNOWLEDGMENTS

I wish to express gratitude to Dr. Andrés Caicedo. The work in this thesis would have not been possible without his patient and nurturing guidance. I would also like to thank Boise State University’s Mathematics Department for their financial support, which made this research possible. Thank you also to Dr. Marion Scheepers and Dr. Zach Teitler for agreeing to serve as committee members. I also appreciate some of my colleagues help with staying motivated during the time it took to complete this.

Lastly, I would like to thank my family. Most importantly, I would like to thank my husband for his unwavering support during this entire process.
ABSTRACT

Ramsey theory is a rich field of study and an active area of research. The theory can best be described as a combination of set theory and combinatorics; however, the arguments to prove some of its results vary across many fields. In this thesis, we survey the field of Ramsey theory highlighting three of its main theorems (Ramsey’s theorem in Chapter 2, Schur’s theorem in Chapter 4, and Van Der Waerden’s theorem in Chapter 7), paying particular attention to Schur’s theorem.

We discuss the origin (Chapter 5), proofs (Chapters 4 and 5), consequences (Chapter 6), and some generalizations (Chapter 8) of Schur’s theorem. Among generalizations we mention Rado and Szemerédi’s theorems. Special attention is also paid to upper and lower bounds for Schur numbers.

Along the way, we take a couple detours, going into areas of mathematics or history that are relevant to what we are studying. In particular Chapter 3 includes a biography of Issai Schur, and Chapter 6 discusses results related to Fermat’s Last Theorem, namely that modular arithmetic does not suffice to provide a proof, and how this can be verified as a consequence of Schur’s theorem or by using Fourier analysis.
# TABLE OF CONTENTS

ABSTRACT ................................................................. v

LIST OF FIGURES ....................................................... viii

NOTATION ................................................................. ix

LIST OF SYMBOLS ....................................................... x

1 Introduction ............................................................ 1

2 Ramsey’s Theorem ..................................................... 4
   2.1 Proof of Ramsey’s Theorem ..................................... 5
   2.2 Ramsey Numbers ................................................ 9
   2.3 Bounds for Ramsey Numbers ................................. 10

3 The Life of Issai Schur ............................................... 15
   3.1 Personal Life .................................................. 16
   3.2 Academic Life ................................................ 17
   3.3 Contributions to Mathematics ............................... 17

4 Schur’s Theorem ....................................................... 19
   4.1 Proof of Schur’s Theorem ................................. 19
   4.2 Schur Numbers ............................................... 20
   4.3 Bounds for Schur Numbers ............................... 31
5  The Origin of Schur’s Theorem ........................................... 35

6  A Detour into Fermat’s Last Theorem ................................. 39
   6.1 Fourier Analysis ......................................................... 40
   6.2 Examples ............................................................... 48

7  Van Der Waerden’s Theorem ............................................. 53
   7.1 Monochromatic Arithmetic Progressions ......................... 53
   7.2 Van Der Waerden Numbers and Their Bounds .................. 56

8  Generalizations of Schur’s Theorem .................................. 67
   8.1 Schur Quadruples ....................................................... 67
   8.2 Rado’s Theorem ........................................................ 68
   8.3 Maximum Solution-Avoiding Sets ................................. 78
   8.4 Roth and Szemerédi’s Theorems ................................. 82
   8.5 Hindman’s Theorem ................................................... 83
   8.6 Concluding Remarks .................................................. 84

BIBLIOGRAPHY ................................................................. 85

A  Calculations ................................................................. 89
LIST OF FIGURES

2.1 Frank Plumpton Ramsey, age 18. Taken from [36] page 283 ........... 4
2.2 An example that $R(4, 4) > 17$ ............................................. 13

3.1 Issai Schur, taken from [36] page 323 ................................. 15
3.2 Young Issai Schur, taken from [36] page 322 ......................... 16

4.1 Notes from Dr. Andrés Caicedo for the proof of $s(3) = 14$ ........ 31
4.2 A 2-coloring of $K_5$ .......................................................... 33

5.1 Original proof of Schur’s Theorem [35] ............................... 35

7.1 Ronald Graham presenting the check to Timothy Gowers for the proof
of this bound. Taken from [36] page 363. ......................... 65
NOTATION

- We will denote the integer interval from $a$ to $b$, that is, the set of integers
  \[ \{a, a + 1, ..., b - 1, b\} \]
  where $a < b$ are integers, as $[a, b]$.

- We will denote the set of subsets of $S$ of size $m$ as $S^{[m]}$.

- We will sometimes denote the set of all elements of $X$ that are not elements of $Y$
  as $S = X - Y$.

- If $S$ is a set of real numbers and $a$ is a real number, we will denote
  \[ \{a + s : s \in S\} \]
  as $a + S$. We will also denote \( \{as : s \in S\} \) as $aS$.

- Often we will use numbers, such as $0, 1, 2, ...$ for various “colors” simply because
  writing $0, 1, 2$ is a lot shorter to write than red, blue, and yellow.

- We will also use the abbreviation $R_r(3)$ when we mean $R(3, 3, ..., 3)$, using $r$ colors,
  where $R(a_1, ..., a_n)$ is a Ramsey number (defined in the list of symbols).

- We will write $A_n$ for the alternating group; the set of all even permutations of $n$
  numbers, which forms a group under composition.

- We will write $f''A$ to denote the pointwise image of a set $A$ under the function $f$. 
LIST OF SYMBOLS

\( \mathbb{N} \)
the set of natural numbers

\( \mathbb{Z} \)
the set of integers

\( \mathbb{Z}^+ \)
the set of positive integers

\( \mathbb{Z}/p\mathbb{Z} \)
the set \( \{ x : x \in [0, p - 1]\} \)

\( (\mathbb{Z}/p\mathbb{Z})^* \)
the set \( \{ x : x \in [1, p - 1] \text{ and } \gcd(x, p) = 1 \} \)

\( K_n \)
a complete graph on \( n \) vertices

\( s(r) \)
the Schur number of \( r \) (see Definition 1.12)

\( R(k, l) \)
the Ramsey number for \( k \) and \( l \) (see Definition 1.9)

\( R(a_1, ..., a_r) \)
the Ramsey number on \( r \) colors (see Definition 1.10)

\( w(k, r) \)
the Van der Waerden number for \( k \) and \( r \) (see Definition 7.2)

\( \left( \frac{p}{q} \right) \)
the Legendre symbol (see Definition 6.9)

\( GF(2^k) \)
the Galois field with \( 2^k \) elements (see Chapter VIII in [22] for reference)
“Complete disorder is impossible.”
Theodore S. Motzkin [27]

Ramsey theory is the study of preservation of properties under set partitions. It is a field that requires little mathematical vocabulary to pose elegant questions, but involves a wide variety of techniques in finding answers to these questions. Ramsey theory is named after Frank Plumpton Ramsey and has become an exciting area for research. See Chapter 30 in [36] for the history of the term “Ramsey Theory.”

We begin by highlighting some of the fundamental theorems in Ramsey theory and move on to the more specialized topic of Schur’s Theorem. Throughout this work, we adopt various definitions and use notation from two of the most well-known references for Ramsey Theory, mostly borrowing from Landman and Robertson [23], with occasional adaption of terms from Graham, Rothschild, and Spencer [13].

Much of Ramsey theory is based on one of the most basic principles in mathematics, the pigeonhole principle.

**Theorem 1.1 (The Generalized Pigeonhole Principle)** If a set of more than \( mn \) elements is partitioned into \( n \) sets, then some set contains more than \( m \) elements.

Before going further, let us introduce some basic notation and definitions.
Definition 1.2 A graph \( G = (V, E) \) is a set \( V \) of points, called vertices, and a set \( E \) of unordered pairs of vertices, called edges.

Definition 1.3 A subgraph \( G' = (V', E') \) of a graph \( G = (V, E) \) is a graph such that \( V' \subseteq V \) and \( E' \subseteq E \).

Definition 1.4 A complete graph on \( n \) vertices, denoted \( K_n \), is a graph on \( n \) vertices, with the property that every pair of vertices is connected by an edge. If \( V \) is the set of vertices, we also write \( K_V \) for this graph.

Definition 1.5 An edge-coloring of a graph is an assignment of a color to each edge of the graph. A graph that has been edge-colored is called a monochromatic graph if all of its edges are the same color.

Definition 1.6 An \( r \)-coloring of a set \( S \) is a function \( \chi : S \to C \), where \( |C| = r \). We also refer to \( \chi \) as a \( C \)-coloring.

Definition 1.7 A coloring \( \chi \) is monochromatic on a set \( S \) if \( \chi \) is constant on \( S \).

Now, for the theorem that sets the stage for Ramsey theory, basically, a two-dimensional refinement of the pigeonhole principle.

Theorem 1.8 (Ramsey’s Theorem for Two Colors) Let \( k, l \geq 2 \). There exists a positive integer \( R \) such that every edge-coloring of \( K_R \), with the colors red and blue, admits either a red \( K_k \) subgraph or a blue \( K_l \) subgraph.

Definition 1.9 We call the smallest number that satisfies this theorem the Ramsey number of \( k \) and \( l \) and denote it \( R(k, l) \).
Definition 1.10 In the case of \( r \) colors, where \( r > 2 \), we use the abbreviation \( R_r(3) \) for the least positive integer \( n \) such that every \( r \)-coloring of the edges of \( K_n \) admits a monochromatic \( K_3 \) subgraph. More generally, \( R(a_1,\ldots,a_r) \) denotes the least \( n \) such that any \( r \)-coloring of the edges of \( K_n \) admits a monochromatic \( K_{a_i} \) subgraph of color \( i \) for some \( i \in [1,r] \). (Note that \( R_r(3) = R(3,3,\ldots,3) \) \( r \) times.)

Finally, for the theorem that we seek to fully understand and study further.

Theorem 1.11 (Schur’s Theorem) For any \( r \geq 1 \), there exists a positive integer \( s \) such that, for any \( r \)-coloring of \( [1,s] \), there exists a monochromatic solution to \( x + y = z \).

Definition 1.12 We call the smallest number that satisfies this theorem the Schur number of \( r \), and denote it \( s(r) \).

Definition 1.13 We call the a triple, \( \{x,y,z\} \) whose elements satisfy the condition \( x + y = z \), a Schur triple.

This thesis is organized as follows: In Chapter 2, we discuss Ramsey’s theorem, the key technical tool behind the first proof we give of Schur’s Theorem 1.11, see Theorem 4.1. In Chapter 3, we sketch a brief biography of Schur. Chapter 4 is devoted to a combinatorial proof of Theorem 1.11, and to the discussion of some bounds. Chapter 5 presents Schur’s original proof. Chapter 6 discusses the connection between Schur’s theorem and Fermat’s last theorem; the main result, Theorem 6.1, is also proved via Fourier analysis. We conclude in Chapter 8 with a discussion of generalizations of Theorem 1.11, particularly, Rado’s Theorem 8.4. The key technical tool behind the proof of this result, Van der Waerden’s theorem, is discussed in Chapter 7.
CHAPTER 2

RAMSEY’S THEOREM

Ramsey’s theorem was first proved by Frank Plumpton Ramsey in 1928, see [30]. Ramsey was a bright young scholar who unfortunately died at the age of 26 in 1930. This theorem is the only contribution that he made to the field that was later called Ramsey Theory. He contributed to other fields as well, including logic, foundations of mathematics, probability, economics, decision theory, cognitive psychology, semantics and philosophy; see also [44]

Figure 2.1: Frank Plumpton Ramsey, age 18. Taken from [36] page 283
Although this theorem was proven by Ramsey, it was thanks to its rediscovery by Erdős and Szekeres that Ramsey theory started to gain popularity in the mathematical research world, as noted by Landman and Robertson in [23].

2.1 Proof of Ramsey’s Theorem

We give a proof of Ramsey’s theorem for two colors.

**Theorem 2.1 (Ramsey’s Theorem)** Let $k, l \geq 2$. There exists a least positive integer $R = R(k, l)$ such that every edge-coloring of $K_R$ with two colors, red and blue, admits either a red $K_k$ or a blue $K_l$ subgraph.

**Proof** Note that it is obvious that $R(k, 2) = k$ for all $k \geq 2$ and $R(2, l) = l$ for all $l \geq 2$. I proceed by induction on the sum $k + l$ for $k + l \geq 5$.

Let $k + l = 5$. Thus, $R(k, l) = R(3, 2) = R(2, 3) = 3$ and we are done. Now, let $k + l \geq 6$ with $k, l \geq 3$, and assume that $R(k, l - 1)$ and $R(k - 1, l)$ exist.

I now show that $R(k, l) \leq R(k - 1, l) + R(k, l - 1)$, therefore $R(k, l)$ exists. In fact, by induction this gives the upper bound

$$R(k, l) \leq \binom{k + l}{l}.$$

Let $n = R(k - 1, l) + R(k, l - 1)$ and consider an edge-coloring of $K_n$. Choose a vertex $v$ from $K_n$. Note that there are $n - 1$ edges from $v$ to other vertices. Let $A$ be the number of red edges from $v$, and let $B$ be the number of blue edges from $v$. Now, either $A \geq R(k - 1, l)$ or $B \geq R(k, l - 1)$. Else, if $A < R(k - 1, l)$ and $B < R(k, l - 1)$, then $A + B \leq (R(k - 1, l) - 1) + (R(k, l - 1) - 1) = n - 2$. This contradicts the fact that $A + B = n - 1$. 
Now, without loss of generality, we may assume $A \geq R(k - 1, l)$. Let $V$ be the set of vertices connected to $v$ by a red edge, so $|V| = A \geq R(k - 1, l)$. By our assumption, we now have that $K_V$ has either a red $K_{k-1}$ subgraph, or a blue $K_l$ subgraph. If the latter, then we are done. If the former, then by connecting each red vertex of the $K_{k-1}$ graph to $v$, we get a red $K_k$ subgraph, because $v$ was connected to all the points in $V$ by red edges. In both cases, we end up with either a blue $K_l$ subgraph or a red $K_k$ subgraph. \hfill \Box

Although we will mostly talk about the case of two colors, we also include Ramsey’s theorem for an arbitrary finite number $k$ of colors.

**Theorem 2.2** Suppose $a_1, a_2, ..., a_k$ are positive integers. Then there exists a least integer $R(a_1, a_2, ..., a_k)$ such that if $v \geq R(a_1, a_2, ..., a_k)$, then any coloring of the edges of $K_v$ with $k$ colors will contain a monochromatic $K_{a_i}$ subgraph of color $i$ for some $i \in [1, k]$.

**Proof** For $k = 2$, this is Theorem 2.1. Now, to apply induction, assume $k = n$ and that we know $R(b_1, b_2, ..., b_n)$ exists for any $b_1, b_2, ..., b_n$. We will now show that $R(a_1, ..., a_{n+1})$ exists.

**Claim 2.3** $R(a_1, ..., a_{n+1}) \leq R(a_1, R(a_2, ..., a_{n+1}))$.

**Proof** Let $M = R(a_1, R(a_2, ..., a_{n+1}))$. We will show that if $M$ is colored with colors $[1, n + 1]$, then for some $i$, there exists a monochromatic $K_{a_i}$ of color $i$. Consider an $(n + 1)$-coloring $C$ of $K_M$. Now, associate to it a new coloring, $D$ with two colors, defined as follows. If $\{\alpha, \beta\}$ is colored red (say red = 1) under the coloring of $C$, then color it red under the coloring of $D$. If $\{\alpha, \beta\}$ is not colored red under the coloring
of $\mathcal{C}$, then color it blue (say blue = 2) under the coloring of $\mathcal{D}$. Now, by definition of $M$, either we have a $K_{a_i}$ of color red or a $K_{R(a_2,\ldots,a_{n+1})}$ of color blue.

By definition of $\mathcal{D}$, if we have a $K_{a_i}$ of color red under the coloring of $\mathcal{D}$, then we have a coloring of $K_{a_i}$ of color 1 under the coloring of $\mathcal{C}$. This would mean we have shown the claim.

If we have the latter case, we have a $K_{R(a_2,\ldots,a_{n+1})}$ of color blue under the coloring of $\mathcal{D}$, but by definition of $\mathcal{D}$, we have that this $K_{R(a_2,\ldots,a_{n+1})}$ is colored with $[2, n + 1]$ colors under the coloring of $\mathcal{C}$, namely all the colors but red. But then, by induction we must have, for some $i \in [2, n + 1]$, a monochromatic $K_{a_i}$ of color $i$ within this copy of $K_{R(a_2,\ldots,a_{n+1})}$.

Claim 2.3 immediately gives the result.

Occasionally, we would also like to talk about colorings of objects that are not as easily described in geometric terms. We do this by describing colorings of sets rather than of graphs. An excellent reference for this theorem and a proof can be found in [13].

**Theorem 2.4** For all numbers $S$ and all $C$, where $C$ is a finite set of colors, there exists an $N$ such that if we color $[1, N]^{|m|}$ ($m$-sized subsets with elements from $[1, N]$) with colors from $C$, then there is a subset $B$ of $[1, N]$ with $|B| = S$ such that all the subsets of $B$ of size $m$ have the same color.

Let us look at an example. Let $m = 3$ and let $C$ and $S$ be given. We want to show that there exists an $N$ such that if we $C$-color $[1, N]^{[3]}$, then there exists a subset $B \subseteq [1, N]$ with $B$ of size $S$, and with $B^{[3]}$ monochromatic.
First, let $\chi$ be a $C$-coloring of $[1,N]^3$. To simplify notation, we write $\chi(a,b,c)$ when we mean $\chi(\{a,b,c\})$ where $a < b < c$. We will define another $C$-coloring $f_1$ of $[2,N]^2$ by $f_1(x,y) = \chi(1,x,y)$.

If we choose $N$ large enough (that is, $N \geq R_C(a)+1$ where $R_C(a)$ is as in Definition 1.10 and $a$ to be specified later), then we can find a color $i_0$ and a set $B_1 \in [2,N]^a$ with $f''_1B_1^2 = \{i_0\}$.

Now we will look inside the set $B_1$. Let $a_0 = 1$ and $a_1 = \min B_1$. We define a $C$-coloring $f_{a_1}$ on $(B_1\{a_1\})^2$, where $f_{a_1}(x,y) = \chi(a_1,x,y)$.

If we choose $a$ large enough (that is, $a \geq R_C(b)+1$ for some $b$ to be specified later), then we can find a color $i_1$ and a set $B_2 \in (B_1\{a_1\})^b$ with $f''_{a_1}B_2^2 = \{i_1\}$.

If we continue in this fashion, we see that we can get a sequence of colors

$$i_0 < i_1 < ... < i_{k-2}$$

and a sequence $a_0 < a_1 < ... < a_k$ for any $k$ we would like, with the property that $\chi(a_\alpha,a_\beta,a_\gamma) = i_\alpha$ for any $\alpha < \beta < \gamma \leq k$.

Finally, if $k - 1 \geq (S - 1)C + 1$, then some color $j$ appears as $i_\alpha$ for at least $S$ many values of $\alpha$, by the pigeonhole principle.

Then, $\{a_\alpha : i_\alpha = j\}$ has size at least $S$, and is monochromatic for $\chi$ with color $j$.

For the general case, one would proceed by induction on $m$ and instead of $R_C(a)$ we would need to use the corresponding “Ramsey numbers” for sets of size $m - 1$. See [13].
2.2 Ramsey Numbers

Below is a table of the known Ramsey numbers $R(r, s)$ collected from a survey by Stanisław Radziszowski [29], which we refer to for the original sources. We would like to note that no new values have been found since 1994.

Table 2.1: Known Ramsey Numbers, $R(r, s)$

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>14</td>
<td>18</td>
<td>23</td>
<td>28</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>9</td>
<td>18</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>9</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.3 Bounds for Ramsey Numbers

Even the famous mathematician Paul Erdős was quoted in [15] for doubting we would ever find the exact value of some of these numbers.

Suppose aliens invade the Earth and threaten to obliterate it in a year’s time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world’s best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack.

While the precise evaluation of additional Ramsey numbers seems still out of reach of even today’s computational machines, we do have bounds for some of these numbers.

Recall that $R_r(3)$ denotes $R(3,3,...,3)$, where we are using $r$ colors.

**Theorem 2.5** For $r \geq 1$, $R_r(3) \leq 3r!$.

**Proof** We prove this by induction on $r$. Note that for $r = 1$ we have $R_1(3) = 3$, so we are done. Now, for $r \geq 2$, denote the Ramsey number of $R(3, 3, 2, 3, ..., 3)$ as $R^i_r(3)$. (Clearly $R^i_r(3) = R^j_r(3)$ for any $1 \leq i < j \leq r$. The notation is meant to emphasize that the $i^{th}$ color behaves differently from the others.) First, I show the following inequality:

**Claim 2.6**

$$R_r(3) \leq \sum_{i=1}^{r} R^i_r(3).$$
Proof This is similar to the proof of Theorem 2.1. Let 
\[ m = \sum_{i=1}^{r} R_i^r(3) \]
and consider some \( r \)-coloring of the edges of \( K_m \). Choose a vertex \( v \). Let the colors be \( 1, 2, ..., r \) and for \( i = 1, 2, ..., r \) let \( C_i \) denote the set of vertices connected to \( v \) by an edge of color \( i \). Now, there exists some \( j \) such that \( |C_j| \geq R_j^r(3) \). (Otherwise, if we had \( |C_j| < R_j^r(3) \) for all \( j \), we would then have \( |C_j| \leq R_j^r(3) - 1 \). But then, \( m - 1 = \sum_{j=1}^{r} |C_j| \leq \sum_{j=1}^{r} R_j^r(3) - r = m - r \), which means that \( m - 1 \leq m - r \), and \( r \leq 1 \). But \( r \geq 2 \) by assumption.) This means that \( K_{C_j} \) must contain either a \( K_2 \) subgraph of color \( j \), or a monochromatic triangle of color \( c \) for some \( c \in \{1, 2, ..., j - 1, j + 1, ..., r\} \). If we have the case of the monochromatic triangle, we are done. Suppose we have the other case. This means that the two vertices that make up the subgraph \( K_2 \) and the vertex \( v \) create a monochromatic triangle of color \( j \). \( \Box \)

Claim 2.7 For any \( i \in [1, r] \), \( R_i^r(3) = R_{r-1}(3) \).

Proof First, I show that \( R_i^r(3) \geq R_{r-1}(3) \). Recall that \( R_i^r(3) = R(3, ..., 3, 2, 3, ..., 3) \) and \( R_{r-1}(3) = R(3, ..., 3) \). We can clearly see that \( R(3, ..., 3) \leq R(3, ..., 3, 2, 3, ..., 3) \) as any \( (r - 1) \)-coloring \( f \) of \( K_{R_i^r(3)} \) can be seen as an \( r \)-coloring \( f' \), where color \( i \) is not used. We then must have a \( K_3 \) that is monochromatic under the coloring \( f' \) in a color different from \( i \). So in fact, we would have a monochromatic triangle \( K_3 \) under the coloring \( f \).

I now want to show that \( R_i^r(3) \leq R_{r-1}(3) \). We can see by the definition of \( R_i^r(3) \), there must be an \( r \)-coloring of the edges of the subgraph \( K_{R_i^r(3) - 1} \) that avoids monochromatic triangles of color \( c \) for all \( c \in [1, r] \setminus \{i\} \), and avoids a \( K_2 \) of color \( i \).
Since we must avoid a $K_2$ of color $i$, no edge may have color $i$. But this means we actually have an $(r - 1)$-coloring of $K_{R_i(3)-1}$ that avoids monochromatic triangles. Thus, $R^i_r(3) \leq R_{r-1}(3)$. This means that in fact $R^i_r(3) = R_{r-1}(3)$. \hfill \square

Now, we can see that

$$\sum_{i=1}^{r} R^i_r(3) = rR_{r-1}(3).$$

So it follows then from Claim 2.6 that we have $R_r(3) \leq rR_{r-1}(3)$ for $r \geq 2$.

Finally, note that $R_2(3) = 6 = 3 \cdot 2!$.

But then we can see the following pattern:

- $R_3(3) \leq 3 \cdot R_2(3) = 3 \cdot 3 \cdot 2! = 3 \cdot 3!$
- $R_4(3) \leq 4 \cdot R_3(3) = 4 \cdot 3 \cdot 3! = 3 \cdot 4!$
- $R_5(3) \leq 5 \cdot R_4(3) = 5 \cdot 3 \cdot 4! = 3 \cdot 5!$
- $R_6(3) \leq 6 \cdot R_5(3) = 6 \cdot 3 \cdot 5! = 3 \cdot 6!$

... 

Which suggests $R_r(3) \leq 3r!$ for all $r$, as can be verified with a straightforward induction using the inequality $R_r(3) \leq rR_{r-1}(3)$. \hfill \square

This theorem gives us upper bounds, but in order to discover the exact value for Ramsey numbers, examples need to be found to improve the known lower bounds, and arguments of arbitrary colorings must be shown to improve the known upper bounds. There has been some recent improvements of lower bounds, as some discoveries of counterexamples have been made, we would like to mention the latest two.
In March of 2012, Geoffrey Exoo improved the previous lower bound of 35 for $R(4, 6)$ to 36 by finding 37 new examples of edge-colorings of $K_{35}$ that do not have either a monochromatic red $K_4$ or blue $K_6$ [9]. This means that the previous bounds for $R(4, 6)$ and $R(6, 4)$ of 35–41, are now 36–41.

In December of 2012, Hiroshi Fujita found a new edge-coloring of $K_{56}$ that did not admit a red $K_4$ or a blue $K_8$ [10]. This improved the previous bounds of $R(4, 8)$ and $R(8, 4)$ of 56–84 to 57–84.

For the values in Table 2.2, with the exception of $R(4, 6)$, $R(6, 4)$, $R(4, 8)$, and $R(8, 4)$, see the references listed in the survey [29].
Table 2.2: Some Known Bounds for Ramsey Numbers, $R(r, s)$

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40–43</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>36–41</td>
<td>49–61</td>
<td>57–84</td>
<td>73–115</td>
<td>92–149</td>
</tr>
</tbody>
</table>
CHAPTER 3

THE LIFE OF ISSAI SCHUR

Issai Schur’s life has been documented as being both extraordinary and tragic. Though he lived at an unfortunate time in history, his contributions to mathematics have survived. We briefly examine the man, the professor, and the mathematician responsible for one of the seminal results in Ramsey theory. Much of the details in this chapter were obtained from Soifer in the fantastic resource [36].

Figure 3.1: Issai Schur, taken from [36] page 323
3.1 Personal Life

Issai Schur was born on the 10th of January 1875 in Mogilev, Russia, what is now Belarus, unto a merchant father and a Jewish family. In 1888, Schur went to live with his older sister in Libau, Russia (what is now Latvia).

![Figure 3.2: Young Issai Schur, taken from [36] page 322](image)

As Schur was able to speak German flawlessly, in 1894 he moved to Germany to attend the University of Berlin. Berlin then became the place he called home for most of his life.

In September of 1906, Schur married a medical doctor, Regina Malka Fumkin. She was also both Jewish and a Russian immigrant. They had two children together, a son named Georg (named after Georg Frobenius, Schur’s mentor) who was born in 1907 and a daughter named Hilde who was born in 1911.

Although Schur was an incredible mathematician and held a high position at the University, with the rise of Nazi Germany, he was treated quite unjustly because of his ethnicity. After much persecution, Schur was forced to leave Germany for Switzerland, and eventually Palestine in 1939.
Sadly, once a notable professor and well-respected academic, Schur spent the rest of his life in poverty and eventually died of a heart attack on the 10th of January 1941 in Tel Aviv, Palestine, what is now Israel [43].

### 3.2 Academic Life

Being Jewish, Schur could not attend any Russian university, thus he went to live with his older sister at the age of 13. There he attended from 1888 until 1894 the Gymnasium in Libau, a school whose primary language was German. He did this in order to prepare himself to apply to a German university.

In 1894, Schur entered the University of Berlin as a doctoral student to study mathematics and physics. In 1901, Schur obtained his doctorate with Georg Frobenius as his advisor. From 1903 to 1909, Schur was lecturer at the University of Berlin. From 1909 to 1913 and again from 1916 to 1919, Schur served as the equivalent to an associate professor at the University of Berlin. Schur spent 1913 to 1916 as an assistant professor at the University of Bonn; this was the only time he spent away from the University of Berlin. In 1919, Schur was promoted to what is now the equivalent of a full professor, and remained so until he was forced to resign in 1935. Schur was a member of the Prussian Academy of Sciences before his resignation [43].

### 3.3 Contributions to Mathematics

Schur’s dissertation, *Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen*, examined rational representations of the general linear group over the complex field. Schur is best known for his work on representation theory. He was the student of Frobenius, who created the field.
He also made significant contributions in group theory, matrices, algebraic equations, Galois groups, number theory, divergent series, integral equations, and function theory; see [45]. Schur’s valuable contributions were extended by his many students, who branched out to other fields of mathematics as well. Some of his students are Richard Brauer, Alfred Brauer, Robert Frucht, Kurt Hirsch, Walter Ledermann, Bernhard Neumann, Hanna Neumann, Richard Rado, Menahem Max Schiffer, and Helmut Wielandt; see [43].

While Schur was mostly known for his work in algebra; we will be focusing on a particular theorem proven by him, Theorem 1.11, which we prove in the following chapter.
CHAPTER 4

SCHUR’S THEOREM

First established to prove a result related to Fermat’s Last Theorem, Schur’s theorem was not recognized on its own until long after its discovery. See page 301 in [36].

Recall that Schur’s Theorem (Theorem 1.11) states that for any integer \( r > 1 \), there exists a least integer \( s(r) \) such that for any \( r \)-coloring of the integer interval \([1, s(r)]\) there exists a monochromatic Schur triple, that is a monochromatic triple \( x, y, z \) such that \( x + y = z \).

Note that while the proof of this theorem gives us the existence of an explicit upper bound for \( s(r) \), it does not determine what its actual value might be.

4.1 Proof of Schur’s Theorem

**Theorem 4.1** For any \( r \geq 1 \), there exists a least positive integer \( s = s(r) \) such that for any \( r \)-coloring of \([1, s]\), there is a monochromatic solution to \( x + y = z \).

**Proof** By Ramsey’s theorem, there exists an integer \( n = R_r(3) \) such that for any \( r \)-coloring \( f \) of the edges of \( K_n \) there is a monochromatic triangle. We claim that \( s \leq n - 1 \). To see this, suppose an \( r \)-coloring of \([1, n - 1]\) is given. We can construct a coloring of \( K_n \) yielding a monochromatic solution to \( x + y = z \) as follows. Label each of the vertices of the graph \( K_n \) with the numbers 1 through \( n \). Then assign to the edge connecting a pair of vertices (the color of) the positive difference between them.
For example, the edge connecting the vertex labeled 1 and the vertex labeled 5 would be assigned the color $f(4)$. Now, by Ramsey’s theorem, we must have a triangle such that all the edges are assigned the same color. Let the vertices of this triangle be named $a < b < c$, so we know that $b - a$, $c - b$, and $c - a$ are all the same color. Finally, let $x = b - a$, $y = c - b$ and $z = c - a$, and note that $x + y = (b - a) + (c - b) = c - a = z$. Thus, since $x$, $y$, and $z$ are the same color, we have found a monochromatic solution to $x + y = z$. □

From this proof we immediately see:

**Corollary 4.2** For $r \geq 1$, $s(r) \leq R_r(3) - 1$.

### 4.2 Schur Numbers

There are very few Schur numbers known. In fact, the exact value of $s(4)$ was not established until 1931 by Baumert [3]; see also [4].

<table>
<thead>
<tr>
<th>$r$</th>
<th>$s(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
</tr>
</tbody>
</table>

**Theorem 4.3** $s(2) = 5$.

**Proof** First, we show that $s(2) \geq 5$. Consider then, without loss of generality, that 1 is colored red. If 2 is colored red, then we have the triple $\{1, 1, 2\}$, thus let 2 be colored blue. Now, if 4 is colored blue, then we have the triple $\{2, 2, 4\}$, thus let 4 be
colored red. Now, if 3 is colored red, then we have the triple \{1, 3, 4\}, thus let 3 be colored blue. We have obtained a 2-coloring \(f\) of \([1, 4]\) without monochromatic Schur triples, \(f(1) = f(4) = \text{red}\) and \(f(2) = f(3) = \text{blue}\). This means that \(s(2) \geq 5\). (We have shown more, namely, \(f\) is the only 2-coloring of \([1, 4]\) witnessing \(s(2) > 4\), up to relabeling the colors.)

Now, we show that \(s(2) \leq 5\). Consider any 2-coloring \(t\) of \([1, 5]\). If \(t \upharpoonright [1, 4] \neq f\), we are done. If \(t \upharpoonright [1, 4] = f\), note that we cannot assign 5 a color without creating a monochromatic Schur triple, so \(s(2) \leq 5\). Note that we can also use Corollary 4.2 to reach this conclusion: Since \(R_2(3) = R(3, 3) = 6\), then \(s(2) \leq 5\). Thus, \(s(2) = 5\). \(\square\)

I would now like to show a proof that \(s(3) = 14\). Along the way, we will identify all of the possible colorings of \([1, 13]\) that do not admit monochromatic Schur triples.

We suggest getting out a couple sheets of paper and a pencil. Following along with the argument will make it tremendously more clear.

**Theorem 4.4** \(s(3) = 14\).

**Proof** Suppose we have a 3-coloring of \([1, 13]\) that in colors red, yellow, and blue does not admit a monochromatic Schur triple. In order to improve readability, we adopt some *ad hoc* conventions throughout the argument: We say that coloring a number \(n\) in a given color \(i\) *admits* a Schur triple if for some numbers \(m, k, l\), we have that \(\{m, k, l\}\) is a monochromatic Schur triple. We say that \(n\) is *forced* to have color \(j\) if coloring it in any other color admits a Schur triple.

Without loss of generality, suppose 1 is colored red. This forces 2 to be colored either blue or yellow. Again, without loss of generality, let 2 be colored blue. Now we have 3 cases to consider.
Case 1: 3 is colored red.

This means that 4 is forced to be yellow (since $1 + 3 = 4$ and $2 + 2 = 4$). We now have three cases for the choice of 5.

Case 1.a: 5 is red.

This forces 8 to be blue (since 3 and 5 are red and 4 is yellow). This means that 6 is forced to be yellow (since 3 is red and 2 and 8 are blue). But then, coloring 10 would admit a monochromatic Schur triple. Thus, 5 cannot be colored red.

Case 1.b: 5 is blue.

Now, we look at what would happen when coloring 6. Note that it cannot be colored red, as 3 is red.

Case 1.b.i: 6 is blue.

This forces 8 to be red (since 2 and 6 are blue and 4 is yellow). But then 11 is forced to be yellow (since 3 and 8 are red and 5 and 6 are blue). But this means that coloring 7 would admit a monochromatic Schur triple. 6 is not blue.

Case 1.b.ii: 6 is yellow.

This forces 10 to be red (5 is blue and 4 and 6 are yellow). Then, 7 is forced to be yellow (since 3 and 10 are red and 2 and 5 are blue). But then, 11 is forced to be blue (since 1 and 10 are red and 4 and 7 are yellow). But this means that coloring 13 would admit a monochromatic Schur triple. Thus, 6 cannot be yellow either. This means that if 5 is colored blue, 6 cannot be colored without admitting a monochromatic Schur triple. Thus, 5 cannot be colored blue.

Case 1.c: 5 is yellow.

Now we have that either 8 or 10 must be red (since 8 and 10 cannot be yellow because 4 and 5 are, and 8 and 10 cannot both be blue because 2 is). Either way, if 8
or 10 is colored red, we have that 9 must be colored blue (since 1 and either 8 or 10 is red and 4 and 5 are yellow). Now we have two cases, 7 can be either red or yellow (but not blue since 2 and 9 are blue).

**Case 1.c.i:** 7 is red.

This forces 10 to be blue (since 3 and 7 are red and 5 is yellow). This means that any coloring of 8 would admit a monochromatic Schur triple.

**Case 1.c.ii:** Thus, 7 is not red, and must be colored yellow.

Now, we have the following (partial) coloring:

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Yellow</td>
<td>4 5</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Note now that 11 is forced to be red (since 2 and 9 are blue and 4 and 7 are yellow), and then 10 is forced to be blue (since 1 and 11 are red and 5 is yellow). But this means that any coloring of 8 would admit a monochromatic Schur triple. Thus, we see that 5 cannot be colored yellow either.

**Case 2:** Since assuming 3 was colored red lead to 5 not being able to be colored without admitting a monochromatic Schur triple, we see that 3 cannot be colored red. So now we assume 3 is colored yellow.

Note that 4 can be either red or yellow (since 2 is blue).

**Case 2.a:** 4 is red.

We have then that 5 can either be blue or yellow (since 1 and 4 are red).

**Case 2.a.i:** 5 is blue.

This forces 6 to either be red or blue (since 3 is yellow).
Case 2.a.i.α: 6 is red.

This forces 10 to be yellow (since 4 and 6 are red and 5 is blue). But then, 7 cannot be colored without admitting a monochromatic Schur triple (since 1 and 6 are red, 2 and 5 are blue, and 3 and 10 are yellow). Thus, 6 is not red.

Case 2.a.i.β: 6 is blue.

This forces 8 to be yellow (since 4 is red and 2 and 6 are blue). Then, 11 is forced to be red (since 5 and 6 are blue and 3 and 8 are yellow). This forces 10 to be yellow (since 1 and 11 are red and 5 is blue). But this means that 7 cannot be colored without creating a monochromatic Schur triple (since 4 and 11 are red, 2 and 5 are blue, and 3 and 10 are yellow). Thus, 6 cannot be colored blue. This means, since coloring 6 would admit a a monochromatic Schur triple, 5 cannot be colored blue.

Suppose 5 is yellow, then 8 is forced to be blue (since 4 is red and 3 and 5 are yellow). This means that 10 is forced to be red (since 2 and 8 are blue and 5 is yellow). Thus, 6 cannot be colored without admitting a monochromatic Schur triple, since 4, 10 are red, 2, 8 are blue, and 3 is yellow. Thus, 5 cannot be yellow either. Since 5 cannot be colored without admitting a monochromatic Schur triple, 4 cannot be colored red.

Case 2.b: 4 is yellow.

Note 6 should either be red or blue (since 3 is yellow).

Case 2.b.i: 6 is red.

This forces 7 to be blue (since 1 and 6 are red and 3 and 4 are yellow). Then, 5 is forced to be yellow (since 1 and 6 are red and 2 and 7 are blue). This means that 9 is forced to be red (since 2 and 7 are blue and 4 and 5 are yellow). This forces 8 to be blue (since 1 and 9 are red and 4 is yellow). But this means that 10 cannot
be colored without admitting a monochromatic Schur triple (since 1 and 9 are red, 2 and 8 are blue, and 5 is yellow). Thus, 6 is not red.

**Case 2.b.ii:** 6 is blue.

This forces 8 to be red (since 2 and 6 are blue and 4 is yellow). This means that 7 is forced to be blue (since 1 and 8 are red and 3 and 4 are yellow). This means that 9 is forced to be yellow (since 1 and 8 are red and 2 and 7 are blue). But then 5 is forced to be red (because 2 and 7 are blue and 4 and 9 are yellow). Now, 13 cannot be colored without admitting a monochromatic Schur triple (since 5 and 8 are red, 6 and 7 are blue, and 9 and 4 are yellow). This means 6 cannot be blue either.

Since 6 cannot be colored without admitting a monochromatic Schur triple, 4 cannot be yellow either.

**Case 3:** Since 4 cannot be colored without admitting a monochromatic Schur triple, 3 cannot be yellow.

Since we have seen 3 cannot be red or yellow, 3 is forced to be colored blue.

At last, we have 3 numbers colored, and we know if we are to have a coloring of $[1, 13]$, it must look like this so far:

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now, we consider 4. Note that since 2 is blue, 4 cannot be colored blue. This means we have two cases, either 4 is colored yellow or red.

**Case 3.a:** Suppose now that 4 is colored yellow.

We see that we can color 5 either red or yellow (but not blue since 2 and 3 are
blue).

**Case 3.a.i:** 5 is red.

This forces 6 to be yellow (since 1 and 5 are red and 3 is blue), so 10 is forced to be blue (since 5 is red and 4 and 6 are yellow). But then 12 is forced to be red (since 2 and 10 are blue and 6 is yellow). This forces 7 to be yellow (since 5 and 12 are red and 3 and 10 are blue.) Therefore, 13 cannot be colored without admitting a monochromatic Schur triple (since 1 and 12 are red, 3 and 10 are blue, and 6 and 7 are yellow). This means 5 cannot be colored red.

**Case 3.a.ii:** 5 is yellow.

Now, we should color 6 red or yellow.

**Case 3.a.ii.α:** 6 is red.

We can either color 8 red or blue.

**Case 3.a.ii.α.1:** 8 is red.

This forces 9 to be blue (since 1 and 8 are red and 4 and 5 are yellow). Therefore, we have that 7 is forced to be yellow (since 1 and 8 are red and 2 and 9 are blue). But then, 12 cannot be colored without admitting a monochromatic Schur triple (since 6 is red, 3 and 9 are blue, and 5 and 7 are yellow). This means 8 cannot be red.

**Case 3.a.ii.α.2:** 8 is blue.

We have then that 10 is forced to be red (since 2 and 8 are blue and 5 is yellow). This forces 9 to be blue (since 1 and 10 are red and 4 and 5 are yellow) and this forces 11 to be yellow (since 1 and 10 are red and 2 and 9 are blue). But then 7 cannot be colored without admitting a monochromatic Schur triple (since 1 and 6 are red, 2 and 9 are blue, and 4 and 11 are yellow). This means 8 cannot be blue either. Since coloring 6 red lead to not being able to color 8 without admitting a monochromatic
Schur triple, 6 cannot be colored red.

**Case 3.a.ii.β:** 6 is colored yellow.

Note now that 9 can either be colored red or blue.

**Case 3.a.ii.β.1:** 9 is red.

This forces 10 to be blue (since 1 and 9 are red and 5 is yellow). But then 8 cannot be colored without admitting a monochromatic Schur triple (since 1 and 9 are red, 2 and 10 are blue, and 4 is yellow). This means 9 cannot be red.

**Case 3.a.ii.β.2:** 9 is blue.

This forces 11 to be red (since 2 and 9 are blue and 4 and 5 are yellow). This forces 10 to be blue (since 1 and 11 are red and 5 is yellow). But then 12 cannot be colored without admitting a monochromatic Schur triple (since 1 and 11 are red, 2 and 10 are blue, and 6 is yellow). This means 9 cannot be blue either.

So we have that coloring 6 yellow admitted a monochromatic Schur triple in each case.

**Case 3.b:** So we see that 4 must be colored red, giving us the coloring of:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We can now clearly see that 5 must be colored yellow, so we have:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>
Now note that because 3 is colored blue, 6 should either be red or yellow.

**Case 3.b.i:** 6 is red.

This forces 10 to be blue (because 4 and 6 are red and 5 is yellow). But then 7 is forced to be yellow (because 1 and 6 are red and 3 and 10 are blue). But then 12 cannot be colored without admitting a monochromatic Schur triple (since 6 is red, 2 and 10 are blue, and 5 and 7 are yellow). This means 6 cannot be red.

**Case 3.b.ii:** This means that 6 must be colored yellow, and any coloring of \([1, 13]\) must start like this:

\[
\begin{array}{c|ccc}
\text{Red} & 1 & 4 \\
\hline
\text{Blue} & 2 & 3 \\
\hline
\text{Yellow} & 5 & 6 \\
\end{array}
\]

Note that since 5 is yellow, 10 is forced to be either red or blue.

**Case 3.b.ii.α:** 10 is colored blue.

This forces 8 to be yellow (since 4 is red and 2 and 10 are blue). But this forces 13 to be red (since 10 and 3 are blue and 5 and 8 are yellow). But then we cannot color 12 without admitting a monochromatic Schur triple (since 1 and 13 are red, 2 and 10 are blue, and 6 is yellow.) This means that 10 cannot be colored blue.

**Case 3.b.ii.β:** Thus, 10 must be colored red and we have any coloring of \([1, 13]\) must start like this:

\[
\begin{array}{c|ccc}
\text{Red} & 1 & 4 & 10 \\
\hline
\text{Blue} & 2 & 3 \\
\hline
\text{Yellow} & 5 & 6 \\
\end{array}
\]

From here we can see that 11 must be colored blue (since 1 and 10 are red and 5 and 6 are yellow):
Now, we can see that 8 must be colored yellow (since 4 is red and 3 and 11 are blue):

<table>
<thead>
<tr>
<th>Red</th>
<th>1 4 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2 3 11</td>
</tr>
<tr>
<td>Yellow</td>
<td>5 6</td>
</tr>
</tbody>
</table>

We can also see that 9 must be colored yellow (since 1 and 10 are red and 2 and 11 are blue):

<table>
<thead>
<tr>
<th>Red</th>
<th>1 4 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2 3 11</td>
</tr>
<tr>
<td>Yellow</td>
<td>5 6 8 9</td>
</tr>
</tbody>
</table>

And, 13 must be colored red (since 2 and 11 are blue and 5 and 8 are yellow):

<table>
<thead>
<tr>
<th>Red</th>
<th>1 4 10 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2 3 11</td>
</tr>
<tr>
<td>Yellow</td>
<td>5 6 8 9</td>
</tr>
</tbody>
</table>

Lastly, we see that 12 must be colored blue (since 1 and 13 are red and 6 is yellow):

<table>
<thead>
<tr>
<th>Red</th>
<th>1 4 10 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2 3 11 12</td>
</tr>
<tr>
<td>Yellow</td>
<td>5 6 8 9</td>
</tr>
</tbody>
</table>

But now we can see that coloring 7 either red, yellow, or blue would not admit a
monochromatic Schur triple. Thus, we have precisely 3 possible colorings of \([1, 13]\) that do not admit a monochromatic Schur triple, namely one for each color of 7.

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Blue</th>
<th>Yellow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 4 7 10 13</td>
<td>2 3 11 12</td>
<td>5 6 8 9</td>
</tr>
<tr>
<td>Red</td>
<td>1 4 10 13</td>
<td>2 3 7 11 12</td>
<td>5 6 8 9</td>
</tr>
<tr>
<td>Blue</td>
<td>2 3 7 11 12</td>
<td>2 3 7 11 12</td>
<td>5 6 7 8 9</td>
</tr>
<tr>
<td>Yellow</td>
<td>5 6 8 9</td>
<td>5 6 7 8 9</td>
<td></td>
</tr>
</tbody>
</table>

Therefore, we have three examples that \(s(3) \geq 14\). Now, note that none of these colorings can be extended to a coloring of \([1, 14]\) without admitting a monochromatic Schur triple. Thus, \(s(3) = 14\).

I believe that it is now obvious why so few Schur numbers have been found. It is extremely tedious to check all the cases for these colorings, and it only becomes worse and computationally unfeasible as we consider more colors.

This demonstrates that trying to find these numbers “by hand,” as we might say, is just not going to work.
In 1931 Baumert [3] used a computer to find that \( s(4) = 45 \).

Here is an example of one of the coloring that he found demonstrating \( s(4) > 44 \).

<table>
<thead>
<tr>
<th>Red</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>26</th>
<th>28</th>
<th>40</th>
<th>42</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>18</td>
<td>21</td>
<td>24</td>
<td>27</td>
<td>33</td>
<td>37</td>
<td>38</td>
<td>43</td>
</tr>
<tr>
<td>Yellow</td>
<td>4</td>
<td>6</td>
<td>13</td>
<td>20</td>
<td>22</td>
<td>23</td>
<td>25</td>
<td>30</td>
<td>32</td>
<td>39</td>
<td>41</td>
</tr>
<tr>
<td>Green</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>29</td>
<td>31</td>
<td>34</td>
<td>35</td>
<td>36</td>
</tr>
</tbody>
</table>

### 4.3 Bounds for Schur Numbers

As the exact value of so few Schur numbers is known, our focus naturally shifts to bounds. This has been where most of the research lies on this topic. I now discuss some of these bounds.
We proved in Corollary 4.2 that $s(r) \leq R_r(3) - 1$. From this, and Theorem 2.5, we get:

**Corollary 4.5** For $r \geq 1$, $s(r) \leq 3r! - 1$.

We can now look at a lower bound. Theorem 4.6 and its proof, due to Schur, can be found in [35].

**Theorem 4.6** For $r \geq 1$, $s(r) \geq \frac{3^r + 1}{2}$.

**Proof** Let $n \geq 1$. Assume we have a coloring $\chi: [1, n] \to [1, r]$ such that there is no monochromatic Schur triple. We define an $(r + 1)$-coloring $\hat{\chi}: [1, 3n + 1] \to [1, r + 1]$ that extends $\chi$ as follows:

For $x \in [1, n]$, let $\hat{\chi}(x) = \chi(x)$.

For $x \in [n + 1, 2n + 1]$, let $\hat{\chi}(x) = r + 1$.

For $x \in [2n + 2, 3n + 1]$, let $\hat{\chi}(x) = \chi(x - (2n + 1))$.

Now, we claim that $[1, 3n + 1]$ does not contain a monochromatic Schur triple, under $\hat{\chi}$. Assume otherwise, and let $\{x, y, z\}$ be a monochromatic Schur triple, with $x \leq y < z$. Consider the color $r + 1$. Since $z = x + y \geq 2x \geq 2(n + 1) > 2n + 1$, our triple cannot be of color $r + 1$.

Consider the color $j \neq r + 1$. Since $\hat{\chi}$ is identical to $\chi$, we have that $\{x, y, z\} \subseteq [1, n]$ cannot be a monochromatic Schur triple of color $j$.

Now, since $2(2n + 2) > 3n + 1$, it is not possible to have $x, y \in [2n + 2, 3n + 1]$, thus $[2n + 2, 3n + 1]$ does not contain a monochromatic Schur triple of color $j$ for $x$.

This means that if we have a monochromatic Schur triple of color $j$, it must be the case that $x \in [1, n]$ and $y \in [2n + 2, 3n + 1]$. 

But if \( \{x, y, z\} \) is such a Schur triple of color \( j \), then by taking \( y' = y - (2n + 1) \) and \( z' = z - (2n + 1) \), we see that \( \{x, y', z'\} \) is a Schur triple of color \( j \) for \( \chi \), a contradiction.

So, it must be the case that if \( s(r) \geq n + 1 \), then \( s(r + 1) \geq 3n + 2 \), and we have \( s(r + 1) \geq 3s(r) - 1 \). Now, we can do induction on \( r \). Note that \( s(1) = 2 \geq \frac{3^1 + 1}{2} \).

Assume that \( r > 1 \) and \( s(r) \geq \frac{3^r + 1}{2} \). Because we have shown \( s(r + 1) \geq 3s(r) - 1 \), we have that \( s(r + 1) = 3s(r) - 1 \geq 3\left(\frac{3^r + 1}{2}\right) - 1 = \frac{3^{r+1} + 1}{2} \). Thus, we have shown that for all \( r \), \( s(r) \geq \frac{3^r + 1}{2} \).

\( \square \)

We have seen in Corollary 4.2 that Ramsey numbers and Schur numbers are related. We would like to recall the proof of \( s(2) = 5 \) in Theorem 4.3 and the coloring of the \( K_5 \) graph that showed \( R(3, 3) > 5 \) in Figure 4.2.
We can see that the coloring in the argument for the Schur number demonstrates exactly how we might color a graph in order to obtain a counterexample for a Ramsey number.
CHAPTER 5

THE ORIGIN OF SCHUR’S THEOREM

Schur’s original proof of Theorem 1.11, which preceded Ramsey’s theorem, was published in 1917; see [35]. It was introduced as a lemma in a paper meant to improve some results of Dickson. An English translation of his original proof appeared much later in [24]. I now prove Schur’s Theorem 1.11 again, this time closely following the original proof of Schur.

Proof Suppose we have an $m$-coloring of $[1, N]$ without monochromatic Schur triples, let the color that is used most frequently be denoted $Z_1$. Let us denote the numbers that it colors as $x_1, x_2, \ldots, x_{n_1}$, where $x_1 < x_2 < \ldots < x_{n_1}$.

Figure 5.1: Original proof of Schur’s Theorem [35]
Note that $N \leq n_1 m$, and that there are $n_1 - 1$ differences

$$x_2 - x_1, x_3 - x_1, \ldots, x_{n_1} - x_1.$$  

Note that these differences are all integers in $[1, N]$. If any of them have color $Z_1$, we would have $x_i$, $x_1$ and $x_i - x_1$ all of color $Z_1$ and $(x_i - x_1) + x_1 = x_i$, creating a monochromatic Schur triple. Since we are considering a coloring that avoids monochromatic Schur triples, we know that none of these differences is colored with $Z_1$. This means that we have an $(m - 1)$-coloring of these differences.

Now, let $Z_2$ be the color that is used the most to color these differences.

If $Z_2$ colors the $n_2$ differences

$$x_\alpha - x_1, x_\beta - x_1, \ldots$$

then we can see that $n_1 - 1 \leq n_2 (m - 1)$.

Now, consider the $n_2 - 1$ differences

$$x_\beta - x_\alpha = (x_\beta - x_1) - (x_\alpha - x_1), \quad x_\gamma - x_\alpha = (x_\gamma - x_1) - (x_\alpha - x_1), \quad \ldots$$

and that all these differences are integers in $[1, N]$. If any of these differences, say $x_\tau - x_\alpha$, were to be colored with $Z_1$, we would have the monochromatic Schur triple, $x_\alpha, x_\tau - x_\alpha, x_\tau$. Similarly, if any of these differences were to be colored with $Z_2$, we would have for some $i > \alpha$ that $x_\alpha - x_1, x_i - x_1,$ and $x_i - x_\alpha$ all have color $Z_2$ and $(x_i - x_\alpha) + (x_\alpha - x_1) = (x_i - x_1)$, creating a monochromatic Schur triple. Since we are coloring to avoid monochromatic Schur triples, we know that none of these differences is colored with $Z_1$ or $Z_2$, thus we have a coloring with $m - 2$ colors.
Let \(n_3\) be the number that is used most frequently in coloring these differences. Then, we can see that \(n_2 - 1 \leq n_3(m - 2)\).

Continuing in this way, we can see that we will have:

\[
n_\mu - 1 \leq n_{\mu+1}(m - \mu)
\]

for \(\mu = 1, 2, 3, ..., m_1\), where \(m_1\) is such that \(n_{m_1} = 1\). Now, from the inequalities we have that

\[
\frac{n_\mu}{(m - \mu)!} \leq \frac{1}{(m - \mu)!} + \frac{n_{\mu+1}}{(m - \mu - 1)!}
\]

for \(\mu = 1, 2, ..., n_1\). Taking the sum of these inequalities gives

\[
\frac{n_1}{(m - 1)!} \leq \frac{1}{(m - 1)!} + \frac{1}{(m - 2)!} + ... + \frac{1}{(m - m_1)!} < e.
\]

Hence, \(N \leq n_1m \leq m!e\). In particular, any \(m\)-coloring of \([1, k]\) admits a monochromatic Schur triple whenever \(k > m!e\).

Now, we can see that this proof not only gives existence of a solution, it gives a bound that actually improves on Corollary 4.2.

**Corollary 5.1** For \(r \geq 1\), \(s(r) \leq \lceil r!e \rceil\).

Note that \(\frac{3^4+1}{2} = 41\) and \(\lceil 4!e \rceil = 66\) while \(s(4) = 45\). In general, it is expected that the lower bound \(\frac{3^n+1}{2}\) from Theorem 4.6 is closer to the actual value of \(s(r)\) than the upper bound \(s(r) \leq \lceil r!e \rceil\) from Corollary 5.1.

In [41], the authors use Corollary 5.1 to show the following bound.

**Theorem 5.2** \(s(n) \geq \frac{1}{2}(89 \cdot 3^{n-4} + 1)\).
They prove a stronger result as well. Define $g(l)$ to be the smallest number of colors needed to color $[1, l]$ without admitting a monochromatic Schur triple; that is, if $s(n - 1) < l + 1 \leq s(n)$, then $g(l) = n$. The following three results come from [41].

**Lemma 5.3** For $l$ sufficiently large, $g(l) < \log l$.

**Theorem 5.4** For all positive integers $m$ and $k$,

$$s(km + g(k \cdot s(n) - k)) - 1 \geq (2s(n) - 1)^k - 1.$$ 

**Corollary 5.5** For all sufficiently large $n$ and $c$ an absolute constant,

$$s(n) > 89n^{2 - c\log n} + 1.$$ 

The relationship between Schur numbers and Ramsey numbers is noted in [41] as well. For the proofs of these results, see Chapter 2 of Part III of [41]. Also, for more details about the discovery of the relationship between Schur numbers and Ramsey numbers, see references of [41].
CHAPTER 6

A DETOUR INTO FERMAT’S LAST THEOREM

As mentioned in Chapter 4, Schur’s theorem was first established in order to prove a result related to Fermat’s Last Theorem. Here we explain this relation.

In 1637 (there is some debate about the date; see [25]), Fermat scribbled into the margins of his copy of *Arithmetica* by Diophantus, that

It is impossible for a cube to be the sum of two cubes, a fourth power to be the sum of two fourth powers, or in general for any number that is a power greater than the second to be the sum of two like powers. I have discovered a truly marvelous demonstration of this proposition that this margin is too narrow to contain.

See [42] and [25]. After his death, this note became one of the most famous conjectures of mathematics. We now know that it is unlikely that Fermat had a proof with the elementary tools he had available, but it is interesting to wonder if it would have even been possible.

After many attempts without success, mathematicians decided to consider “local versions” of Fermat’s Last Theorem. Theorem 6.1 shows that there is no proof of Fermat’s Last Theorem that uses only modular arithmetic.

Note that if there are integers $x, y, z$ satisfying $x^n + y^n = z^n$, then for every prime $p$, they also solve the congruence equation: $x^n + y^n \equiv z^n \pmod{p}$, and if $p$ is large
enough then \( p \not| xyz \). Schur’s Theorem was used to show that this congruence has a non-trivial solution as long as \( p \) is sufficiently large; see Theorem 6.1 below.

This demonstrates that if one were to try and prove Fermat’s Last Theorem, by necessity, tools beyond modular arithmetic must be involved, and a “local” analysis, that is, modulo some \( p \), cannot succeed.

### 6.1 Fourier Analysis

I now show two proofs of the result indicated above. The first is combinatorial and follows from Theorem 1.11 as in Schur’s original approach. The second uses Fourier analysis on the additive group \( \mathbb{Z}/p\mathbb{Z} \), and for this the reader may follow the outline in [46].

**Theorem 6.1** Let \( n > 1 \). Then, for all primes \( p > s(n) \), the congruence

\[
x^n + y^n \equiv z^n \pmod{p}
\]

has a solution in the integers, such that \( p \) does not divide \( xyz \).

**Proof** Recall from Definition 1.12 that \( s(n) \) is the least positive integer \( s \) such that, for any \( n \)-coloring of \([1, s]\), there exists a monochromatic solution to \( x + y = z \).

Let \( p > s(n) \) be a prime. Recall that \( (\mathbb{Z}/p\mathbb{Z})^* = \{1, 2, \ldots, p - 1\} \). Now, let

\[
G_n = S = \{x^n \pmod{p} : x \in (\mathbb{Z}/p\mathbb{Z})^*\}.
\]

Note that \( (\mathbb{Z}/p\mathbb{Z})^* \) forms a group under multiplication modulo \( p \) and \( S \) forms a subgroup of \( (\mathbb{Z}/p\mathbb{Z})^* \). We can now write \( (\mathbb{Z}/p\mathbb{Z})^* \) as a disjoint union of cosets,
\[(\mathbb{Z}/p\mathbb{Z})^* = \bigcup_{i=1}^{k} a_i S\]

where \( k = \frac{n}{\gcd(n, p-1)} \). Now, we define a \( k \)-coloring of \((\mathbb{Z}/p\mathbb{Z})^*\) by coloring \( t \in (\mathbb{Z}/p\mathbb{Z})^*\) with color \( j \) if and only if \( t \in a_j S \). Since \( k \leq n \) and \( p - 1 \geq s(n) \), we have a monochromatic Schur triple. That is, there exist \( \alpha, \beta, \gamma \in (\mathbb{Z}/p\mathbb{Z})^* \) such that \( \alpha + \beta \equiv \gamma \pmod{p} \) and, for some \( i, \alpha, \beta, \gamma \in a_i S \). This means there exists \( x, y, z \in (\mathbb{Z}/p\mathbb{Z})^* \) such that \( a_i x^n + a_i y^n \equiv a_i z^n \pmod{p} \). Thus, we have that \( x, y, z \in (\mathbb{Z}/p\mathbb{Z})^* \) and \( x^n + y^n \equiv z^n \pmod{p} \). \( \square \)

We now discuss the Fourier analytic approach. First, we let \( \omega \) denote the \( p \)-th root of unity \( e^{\frac{2\pi i}{p}} \), where \( p \) is a prime number. Below, \( n \) is fixed.

For \( k \in \mathbb{Z}/p\mathbb{Z} \), we denote by \( S_k \) the exponential sum as \( \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{k x n} = \sum_{x=0}^{p-1} e^{2\pi i k x n / p} \). Of course, \( S_0 = p \). We are mostly interested in \( S_k \) for \( k \neq 0 \pmod{p} \). We use the notation \( \#(G) \) to denote the cardinality of the set \( G \), and \( |G| \) to denote the norm of the complex number \( G \).

**Lemma 6.2** If \( a \in (\mathbb{Z}/p\mathbb{Z})^* \), then \( S_k = S_{ka^n} \).

**Proof** Suppose \( a \in (\mathbb{Z}/p\mathbb{Z})^* \). Then, by the definitions of \( S \) and of \( \omega \), we have that

\[
S_{ka^n} = \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{ka^n x^n / p} = \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{k(ax)^n / p}.
\]

But \( x \mapsto ax \) is a bijection of \( \mathbb{Z}/p\mathbb{Z} \) with itself, because \( a \neq 0 \). Thus, we have

\[
\sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{k(ax)^n / p} = \sum_{b \in \mathbb{Z}/p\mathbb{Z}} \omega^{kb^n / p} = S_k.
\]

This shows that \( S_{ka^n} = S_k \). \( \square \)
Let

$$T_m = \sum_{k=0}^{p-1} \omega^{km}$$

be the sum of the $m$-th powers of the $p$-th roots of unity.

**Lemma 6.3** $T_m = p$ when $m \equiv 0 \pmod{p}$.

**Proof** Let $m \equiv 0 \pmod{p}$. Then, we have

$$T_m = \sum_{k=0}^{p-1} e^{2\pi ikm/p} = \sum_{k=0}^{p-1} e^{2\pi ikt}$$

for some $t$.

Now we use the fact that $e^{2\pi i} = 1$.

$$T_m = \sum_{k=0}^{p-1} e^{2\pi ikt} = \sum_{k=0}^{p-1} (1)^k = \sum_{k=0}^{p-1} 1 = p.$$

$\square$

**Lemma 6.4** $T_m = \sum_{k=0}^{p-1} e^{2\pi km/p} = 0$ when $m \not\equiv 0 \pmod{p}$.

**Proof** Let $m \not\equiv 0 \pmod{p}$. We use the fact that

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1}$$

when $x \neq 1$.

Note that we can use this because if we let $x = \omega^m = e^{2\pi im/p}$, we see that $x \neq 1$ since $m \not\equiv 0 \pmod{p}$. Then,

$$T_m = \sum_{k=0}^{p-1} e^{2\pi km/p} = \sum_{k=0}^{p-1} (e^{2\pi im/p})^k = \frac{(e^{2\pi im/p})^{(p-1)} - 1}{e^{2\pi im/p} - 1}$$

$$= \frac{(e^{2\pi im/p})^{p-1} - 1}{e^{2\pi im/p} - 1} = \frac{e^{2\pi im} - 1}{e^{2\pi im/p} - 1} = \frac{e^{2\pi i} m - 1}{e^{2\pi im/p} - 1} = 0.$$

$\square$
We can use the previous two lemmas to prove Lemma 6.5.

**Lemma 6.5**

\[
\sum_{k \in \mathbb{Z}/p\mathbb{Z}} |S_k|^2 = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \sum_{k \in \mathbb{Z}/p\mathbb{Z}} e^{2\pi i k \left( \frac{x^n - y^n}{p} \right)} = pN,
\]

where \(N\) is the number of solutions to the equation \(x^n = y^n\) within the field \(\mathbb{Z}/p\mathbb{Z}\).

**Proof** First, we see that

\[
|S_k|^2 = \left| \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{kx^n} \right|^2 = \left( \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{kx^n} \right) \left( \sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{-kx^n} \right) = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \omega^k (x^n - y^n).
\]

This means we have

\[
\sum_{k \in \mathbb{Z}/p\mathbb{Z}} |S_k|^2 = \sum_{k \in \mathbb{Z}/p\mathbb{Z}} \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \omega^k (x^n - y^n) = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \sum_{k \in \mathbb{Z}/p\mathbb{Z}} \omega^k (x^n - y^n) = pN.
\]

Now, let \(N\) denote \(#\{ (x, y) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} : x^n = y^n \}\). From the facts shown above,

\[
\sum_{k \in \mathbb{Z}/p\mathbb{Z}} |S_k|^2 = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \sum_{k \in \mathbb{Z}/p\mathbb{Z}} \omega^k (x^n - y^n) = Np.
\]

With \(N\) as in Lemma 6.5, we have the following upper bound:

**Lemma 6.6** \(N \leq 1 + n(p - 1)\).
Proof Recall that $N = \#(\{(x, y) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} : x^n = y^n\})$. Note that when $x = 0$, we have only one solution to the equation $x^n = y^n$, namely when $y = 0$. Now, if we fix $x \in (\mathbb{Z}/p\mathbb{Z})^*$, we see that we have at most $n$ solutions $y$ to the equation $x^n = y^n$. We have $p - 1$ choices for $x$. This means that $N \leq 1 + n(p - 1)$. □

With $G_n$ as in the proof of Theorem 6.1, we have the following lower bound.

**Lemma 6.7** $\#(G_n) \geq \frac{p - 1}{n}$.

**Proof** Recall that $G_n = \{a^n : a \in (\mathbb{Z}/p\mathbb{Z})^*\}$. Given $x \in (\mathbb{Z}/p\mathbb{Z})^*$, we have $x^n \in G_n$. We partition $(\mathbb{Z}/p\mathbb{Z})^*$ into $\#(G_n)$ sets $\{x \in (\mathbb{Z}/p\mathbb{Z})^* : x^n = t\}$, where $t$ varies over $G_n$. We have $\#(\{x \in (\mathbb{Z}/p\mathbb{Z})^* : x^n = t\}) \leq n$, thus $\#((\mathbb{Z}/p\mathbb{Z})^*) \leq \#(G_n)n$. But $\#((\mathbb{Z}/p\mathbb{Z})^*) = p - 1$, thus, $\#(G_n) \geq \frac{p - 1}{n}$. □

Now we need the following estimate on the sums $S_k$ for our second proof of Theorem 6.1.

**Lemma 6.8** If $k \in (\mathbb{Z}/p\mathbb{Z})^*$, 

$$\left| \sum_{x=0}^{p-1} e^{2\pi i \frac{kx^n}{p}} \right| \leq \sqrt{2np^{\frac{1}{2}}}.$$ 

**Proof** By the previous lemmas, we have the following equalities and inequalities.

$$\#(G_n) \cdot |S_k|^2 = \sum_{a^n \in G_n} |S_{ka^n}|^2 \leq \sum_{t \in (\mathbb{Z}/p\mathbb{Z})^*} |S_t|^2$$

from Lemma 6.2.

$$\sum_{t \in (\mathbb{Z}/p\mathbb{Z})^*} |S_t|^2 \leq Np$$
from Lemma 6.5.

\[ Np \leq p(1 + n(p - 1)) \]

from Lemma 6.6. We can also see that \( p(1 + n(p - 1)) \leq np^2 \).

This means that we have

\[ |S_k|^2 \leq \frac{np^2}{\#(G_n)}. \]

And from Lemma 6.7, we can see that

\[ \frac{np^2}{\#(G_n)} \leq \frac{np^2}{\frac{p-1}{n}} \leq 2n^2 p. \]

Thus, we have

\[ |\sum_{x=0}^{p-1} e^{2\pi i \frac{kx^p}{p}}| \leq \sqrt{2np^2}. \]

\[ \square \]

We are finally ready for the second proof of Theorem 6.1. In fact, we show that if \( p \) is a prime number and \( p \geq 32n^6 \), then there are nontrivial solutions to the equation \( x^n + y^n \equiv z^n \pmod{p} \).

**Proof** Let \( M \) denote the number of ordered triples \((x, y, z)\) in \((\mathbb{Z}/p\mathbb{Z})^3\) such that \( x^n + y^n = z^n \). Note that this happens precisely when \( x^n + y^n - z^n = 0 \pmod{p} \). But we have shown when \( m \equiv 0 \pmod{p} \), then \( T_m = p \), and when \( m \not\equiv 0 \pmod{p} \), then \( T_m = 0 \). Thus, we have that when \((x, y, z)\) is a solution to the equation

\[ x^n + y^n - z^n = 0 \pmod{p}, \]
then \( \frac{1}{p} \sum_{k=0}^{p-1} e^{2\pi ik(x^n+y^n-z^n)} = 1 \) and when it is not, then \( \frac{1}{p} \sum_{k=0}^{p-1} e^{2\pi ik(x^n+y^n-z^n)} = 0 \). We now have a nice way to keep track of the number of solutions,

\[
M = \sum_{x,y,z \in \mathbb{Z}/p\mathbb{Z}} \frac{1}{p} \sum_{k=0}^{p-1} e^{2\pi ik(x^n+y^n-z^n)}.
\]

We would now like to analyze this number to show that for \( p \) large enough, we are guaranteed to have a nontrivial solution. Note the following equalities:

\[
M = \sum_{x,y,z \in \mathbb{Z}/p\mathbb{Z}} \frac{1}{p} \sum_{k=0}^{p-1} e^{2\pi ik(x^n+y^n-z^n)} = \sum_{k=0}^{p-1} \sum_{x \in \mathbb{Z}/p\mathbb{Z}} e^{2\pi ikx^n} \sum_{y \in \mathbb{Z}/p} e^{2\pi iky^n} \sum_{z \in \mathbb{Z}/p} e^{2\pi ik(-z^n)} = \sum_{k=0}^{p-1} S_k \overline{S_k} = \frac{1}{p} \left( S_0^2 S_0 + \sum_{k=1}^{p-1} S_k^2 \overline{S_k} \right) = p^2 + \frac{1}{p} \sum_{k=1}^{p-1} S_k^2 \overline{S_k}.
\]

Since \( M - p^2 \) is a real number, we know that

\[
\frac{1}{p} \sum_{k=1}^{p-1} S_k^2 \overline{S_k}
\]

is also real. This means that

\[
\frac{1}{p} \sum_{k=1}^{p-1} S_k^2 \overline{S_k} \geq -\frac{1}{p} \sum_{k=1}^{p-1} |S_k^2 \overline{S_k}| = -\frac{1}{p} \sum_{k=1}^{p-1} |S_k^2 S_k| = -\frac{1}{p} \sum_{k=1}^{p-1} |S_k|^3
\]

\[
\geq -\frac{1}{p} \sum_{k=1}^{p-1} (\sqrt{2np^2})^3 = -\frac{p-1}{p} (\sqrt{2np^2})^3 \geq -(\sqrt{2np^2})^3,
\]

where the previous to last inequality follows from Lemma 6.8.
We have that $M \geq p^2 - (\sqrt{2np^2})^3$. Now, if we assume $p \geq 32n^6$, then we can see the following\(^1\):

\[
p \geq 32n^6 \iff \frac{1}{4}p \geq 8n^6 \iff \frac{1}{4}p^4 \geq 8n^6p^3 \iff \frac{1}{2}p^2 \geq 2\frac{1}{2}n^3p^\frac{3}{2} \iff p^2 - 2\frac{3}{2}n^3p^\frac{3}{2} \geq \frac{1}{2}p^2 \iff p^2 - (\sqrt{2np^2})^3 \geq \frac{1}{2}p^2.
\]

We have that,

\[
M \geq \frac{1}{2}p^2.
\]

Lastly, I would like to show that as long as $p \geq 32n^6$, we have that $1 + 3np < \frac{1}{2}p^2$.

Note that $p \geq 32n^6$ is true if and only if $\frac{1}{32} > n$ is true (the inequality is strict, since $p$ is a prime number, so it cannot equal $32n^6$). Also, $\frac{1}{32} > n$ is true if and only if $7\frac{1}{32}p > 7np$ is true. But, we can see that $\frac{1}{2}p^2 > 7\frac{1}{32}p$ (with room to spare\(^2\)) and $7np = np + 6np > 2 + 6np > 1 + 3np$. Thus, $\frac{1}{2}p^2 > 1 + 3np$.

Now, we can count the number of trivial solutions and see that it is less than $1 + 3np$. Trivial solutions to the equation $x^n + y^n = z^n \pmod{p}$ are those where $p$ divides $xyz$. Note that when $x = y = z = 0$, we have a trivial solution. So we have at least 1. Now, we need to count those where exactly one of $x, y,$ or $z$ is zero (because if any two are zero the last will be forced to be zero.) This means we have three cases. The number of solutions where $x = 0$ is exactly the number of pairs with $y^n = z^n \pmod{p}$ with $y, z \neq 0$. But this is the same as the number of pairs $(tz, z)$ where $t^n = 1$ and $z \neq 0$. There are $p - 1$ choices for $z$, and at most $n$ solutions for the equation $t^n = 1$. This means there is at most $n(p - 1)$ solutions when $x = 0$.

By symmetry, the same bound applies to the number of solutions when $y = 0$, and

\(^1\)Note that, at this point, \([46]\) claims without justification that the bound $p \geq 16n^6$ suffices.

\(^2\)We have $p \geq 32n^6 \geq 32 \cdot 2^6 = 2^{11}$. Note that $2^9 > 7$, so $2^{54} > 7^6$ and $2^{55} > 7^6 \cdot 2$. But $2^{55} \leq p^5$, so $p^5 > 7^6 \cdot 2$. This gives us $p^5 > 7 \cdot 2^7$, or $\frac{1}{2}p^2 > 7\frac{1}{32}p$. 
by a similar argument, it also applies to the number of solutions when \( z = 0 \) (here, we consider the equation \( t^n = -1 \) instead of \( t^n = 1 \)).

Thus, we have the number of trivial solutions is at most \( 1 + 3n(p - 1) \). But this is less than the \( 1 + 3np \). This means we have \( M \geq \frac{1}{2}px \geq 1 + 3np > 1 + 3n(p - 1) \), therefore we have that there is at least one nontrivial solution. This means that from some point on, namely when \( p \geq 32n^6 \), there must be a nontrivial solution to the equation \( x^n + y^n \equiv z^n \pmod{p} \). □

We can see that the Fourier analytic proof gave much better results than combinatorial one using Ramsey numbers. If we look back, we can also see that at times we were quite generous with some inequalities, meaning that the bound we obtained can be improved. This leads us to believe that this would be a better approach to trying to find bounds for other results as well.

### 6.2 Examples

We consider nontrivial solutions of the equation \( x^n + y^n \equiv z^n \pmod{p} \) for fixed \( n \), as \( p \) varies. We would now like to work out a couple of examples in detail.

Consider \( n = 1 \). This means we look at the equation \( x + y \equiv z \pmod{p} \). Note that when \( p = 2 \), we never have a nontrivial solution because \( 1 + 1 \equiv 0 \pmod{2} \). Also, if \( p > 2 \), we always have a nontrivial solution, namely \( 1 + 1 \equiv 2 \pmod{p} \).

Consider \( n = 2 \). This means we look at the equation \( x^2 + y^2 \equiv z^2 \pmod{p} \). Note that when \( p = 2 \), we never have a nontrivial solution because \( 1^2 + 1^2 \equiv 0 \pmod{2} \). Also, if \( p = 3 \), we never have a nontrivial solution because \( 1^2 + 1^2 \equiv 0 \pmod{3} \) and \( 2^2 \equiv 1 \pmod{3} \), thus \( 1^2 + 1^2 \equiv 1^2 + 2^2 \equiv 2^2 + 2^2 \equiv 2 \not\equiv z^2 \pmod{3} \) for any nonzero \( z \pmod{3} \). If \( p = 5 \), \( 1^2 \equiv 4^2 \equiv 1 \) and \( 2^2 \equiv 3^2 \equiv 4 \), and we also never have a nontrivial
solution. Finally, for $p \geq 7$, we always have a nontrivial solution because $3^2 + 4^2 \equiv 5^2 \pmod{p}$.

Consider $n = 3$. This means we will look at the equation $x^3 + y^3 \equiv z^3 \pmod{p}$. Note again, if $p = 2$, we have no nontrivial solutions. Now we can look at a couple of primes to see what is going on.

When $p = 3$, we have
\[
1^3 \equiv 1 \\
2^3 \equiv 2
\]
We have the nontrivial solution $1^3 + 1^3 \equiv 2^3$.

When $p = 5$, we have
\[
1^3 \equiv 1 \\
2^3 \equiv 3 \\
3^3 \equiv 2 \\
4^3 \equiv 4
\]
We have a nontrivial solution.

When $p = 7$, we have
\[
1^3 \equiv 1 \\
2^3 \equiv 1 \\
3^3 \equiv 6 \\
4^3 \equiv 1 \\
5^3 \equiv 6 \\
6^3 \equiv 6
\]
We do not have a nontrivial solution.
We can see that when we have all nonzero numbers are cubic residues (as when \( p = 5 \)), then we have solutions.

Now, we would like to mention a basic result in number theory.

**Definition 6.9** Given a prime \( q \) and an integer \( p \), the Legendre Symbol \( \left( \frac{p}{q} \right) \) has value 0 when \( q \) divides \( p \), 1 when \( p \) is a quadratic residue modulo \( q \), that is, when there is some \( x \) such that \( x^2 \equiv p \pmod{q} \), and \(-1\) if \( p \) is not a quadratic residue modulo \( q \).

**Theorem 6.10 (Laws of Quadratic Reciprocity)** Let \( p \) and \( q \) be distinct odd primes, and \( a \) and \( b \) be integers. Then,

1. \( \left( \frac{-1}{q} \right) = (-1)^{\frac{q-1}{2}}. \)
2. \( \left( \frac{2}{q} \right) = (-1)^{\frac{q^2-1}{8}}. \)
3. \( \left( \frac{p}{q} \right) = (-1)^{\frac{p-1}2 \frac{q-1}2 \left( \frac{q}{p} \right)}. \)
4. \( \left( \frac{ab}{q} \right) = \left( \frac{a}{q} \right) \left( \frac{b}{q} \right). \)

See [19] for a proof.

**Theorem 6.11** If \( p \equiv -1 \pmod{3} \), then for all \( x \), there exists a \( y \) such that \( y^3 \equiv x \pmod{p} \).

**Proof** Let \( p > 3 \) be a prime (the case \( p = 2 \) being trivial), and work modulo \( p \). Note that it is enough to show that \( f : y \mapsto y^3 \) is injective, because we have a finite set, namely \( [0, p - 1] \), being mapped to itself. So, if \( f \) is injective, it is also surjective.

Since \( y = 0 \) if and only if \( y^3 = 0 \), we may assume \( y \neq 0 \).
Suppose that \( y_1, y_2 \neq 0, y_1 \neq y_2 \) and \( y_1^3 = y_2^3 \). This means we have

\[
\frac{y_1^3}{y_2^3} = \left( \frac{y_1}{y_2} \right)^3 = 1.
\]

But if \( \frac{y_1}{y_2} = 1 \), we would have \( y_1 = y_2 \). I will show that it must be the case that if \( x^3 = 1 \) then \( x = 1 \). This will prove that \( y \mapsto y^3 \) is injective.

Note that \( x^3 = 1 \) if and only if \((x - 1)(x^2 + x + 1) = 0\). This means either \( x - 1 = 0 \) or \( x^2 + x + 1 = 0 \). I will show that \( x^2 + x + 1 = 0 \) has no solutions when \( p \equiv -1 \) (mod 3).

We can see by using the Laws of Quadratic Reciprocity (Theorem 6.10), that

\[
\left( \frac{-3}{p} \right) = \left( \frac{-1}{p} \right) \left( \frac{3}{p} \right) = (-1)^{\frac{p-1}{2}} (-1)^{\frac{p-1}{2} \cdot \frac{3-1}{2}} \left( \frac{3}{p} \right).
\]

But then we see that

\[
\left( \frac{-3}{p} \right) = \left( \frac{p}{3} \right) = p \pmod{3}
\]

Thus, when \( p \equiv -1 \) (mod 3), we have that \(-3\) is not a quadratic residue.

This means that our equation \( x^2 + x + 1 \equiv 0 \pmod{p} \) has no solutions in this case, because if \( x^3 + x + 1 = 0 \) then \( 4x^3 + 4x + 4 = 0 \) or \( (2x + 1)^2 + 3 = 0 \), so \(-3\) is a square. Thus, we have that \( x - 1 = 0 \) must hold if \( x^3 = 1 \).

This shows that \( y \mapsto y^3 \) is injective and we have proved the theorem. \( \square \)

**Corollary 6.12** We always have nontrivial solutions to the equation \( x^3 + y^3 \equiv z^3 \pmod{p} \) when \( p \equiv -1 \) (mod 3).
Unfortunately, for \( n = 3 \) where \( p \not\equiv -1 \pmod{3} \) and for \( n > 3 \) we have not found anything quite as nice as Corollary 6.12, but we have been able to find some results using a computer search; see Table 6.1 below.

In Lecture 12 of [31], we see that in 1909 Dickson improved the bound for when the congruence \( x^n + y^n \equiv z^n \pmod{p} \) has a solution mod \( p \) to \( \mathcal{N} = (n-1)^2(n-2)^2 + 6n - 2 \), meaning that for \( p > \mathcal{N} \), we are guaranteed a solution to the congruence. This can be proved by an extension of the Fourier analytic argument from the previous section; see [31].

We were also curious about \( p \leq \mathcal{N} \). For \( n \) from 4 to 10, we ran a program that would check when we have solutions to \( x^n + y^n \equiv z^n \pmod{p} \) for \( p \leq \mathcal{N} \).

Table 6.1: Computer search for primes below \( \mathcal{N} \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \mathcal{N} )</th>
<th>primes below ( \mathcal{N} ) that do not admit a solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>20</td>
<td>2, 7, 13</td>
</tr>
<tr>
<td>4</td>
<td>58</td>
<td>2, 3, 5, 13, 17, 41</td>
</tr>
<tr>
<td>5</td>
<td>172</td>
<td>2, 11, 41, 71, 101</td>
</tr>
<tr>
<td>6</td>
<td>434</td>
<td>2, 3, 5, 7, 13, 19, 43, 61, 97, 157, 277</td>
</tr>
<tr>
<td>7</td>
<td>940</td>
<td>2, 29, 71, 113, 491</td>
</tr>
<tr>
<td>8</td>
<td>1810</td>
<td>2, 3, 5, 13, 17, 41, 113</td>
</tr>
<tr>
<td>9</td>
<td>3188</td>
<td>2, 7, 13, 19, 37, 73, 181, 523, 577</td>
</tr>
<tr>
<td>10</td>
<td>5242</td>
<td>2, 3, 5, 11, 31, 41, 71, 101, 281, 401, 1181</td>
</tr>
</tbody>
</table>

For the code that produced these results, please see Appendix A.
CHAPTER 7

VAN DER WAERDEN’S THEOREM

Van der Waerden’s theorem is a key result in Ramsey theory, perhaps the most responsible for the development of the field. Schur was first to conjecture the result that Van der Waerden proved (and published in 1927) [13].

7.1 Monochromatic Arithmetic Progressions

We state Van der Waerden’s theorem in terms of colors and natural numbers.

**Theorem 7.1 (Van der Waerden)** There exists a least positive integer \( w(k, r) \) such that any \( r \)-coloring of \([1, w(k, r)]\) admits a monochromatic arithmetic progression of length \( k \).

For Van der Waerden’s original proof, see [39]; for his own expository account of the proof, see [40]; and, for a more modern proof, see Terence Tao’s presentation in [38].

We show an example, following an argument similar to the one in [13]. It demonstrates how, when dealing with 2 colors, it suffices to color the numbers up to 325 in order to ensure an arithmetic progression of length 3. This example can be easily adapted to an inductive proof of Theorem 7.1 for any number of colors and any length of arithmetic progression.
Suppose we have a 2-coloring of $[1, 325]$, and we color each of the numbers either red or blue. We now break these 325 numbers up into blocks of length 5. The choice of 5 will become more clear later. This means we have 65 blocks, $B_1 = [1, 5], B_2 = [6, 10], \ldots, B_{65} = [321, 325]$. Note that there are $2^5 = 32$ ways to color each block with 2 colors, therefore by the pigeonhole principle, if we take the first 33 blocks, two of them must be colored identically.

Suppose that $B_i$ and $B_j$ are these two identically colored blocks, where $i < j$. Note that $B_{2j-i}$ is also defined (that is, $2j - i \leq 65$), which is why we chose 65 blocks ($65 = 2 \cdot 32 + 1$).

Now we can see that if the two identically colored blocks are colored so that they contain a monochromatic arithmetic progression of length 3, then we are done, so suppose otherwise.

Let $n_1, n_2, n_3, n_4, n_5$ enumerate in order the numbers in the block $B_a$, so for example, $2_4$ is the 4th number in $B_2$. Recall that $B_2 = [6, 10]$, so $2_4 = 9$.

Within the numbers $i_1, i_2, i_3$ of block $B_i$, two of them must have the same color, say $i_a$ and $i_{a+d}$, and note that $i_{a+2d}$ is also defined (meaning $a + 2d \leq 5$), which is why our blocks have length 5 (because $5 = 2 \cdot 2 + 1$). Note that in $B_j$, the numbers $j_a, j_{a+d}$ are also colored the same way. Assume that they are all of color red.

Now, let us look at block $B_{2j-i}$. Note that this is the same distance away from $B_j$ as $B_j$ was from $B_i$. Let $k = 2j - i$.

Note that the number $k_a$ must be blue or else we would have a monochromatic arithmetic progression of length 3, namely $\{i_a, j_a, k_a\}$, where $i_a - j_a = k_a - j_a = 5(j-i)$. Similarly, $k_{a+d}$ must be colored blue. Recall that the element $i_{a+2d}$ must be of color blue or else we would have an arithmetic progression in the block $B_i$. This means that when we consider the element $k_{a+2d}$ in $B_k$, it must be colored red,
else we would have a monochromatic arithmetic progression of length 3, namely 
\{i_{a+2d}, j_{a+2d}, k_{a+2d}\}. But then we have a monochromatic arithmetic progression of 
length 3, namely \{i_a, j_{a+d}, k_{a+2d}\}. Thus, we have found that if we color \([1, 325]\) with 
just two colors, we must have an monochromatic arithmetic progression of length 3; 
that is, \(w(3, 2) \leq 325\).

We can see that this argument (commonly referred to as the color-focusing tech-
nique) will generalize with an induction on both the number of colors and the length 
of the arithmetic progression. We also note that 325 was quite a large number for 
just 2 colors and an arithmetic progression of length only 3. In fact, \(w(3, 2) = 9\), as 
shown in Theorem 7.5.

Let us look at how our example might change if we wanted an arithmetic pro-
gression of length 3 with 3 colors. These examples demonstrate how the inductive 
argument proceeds.

In the first example, we knew that within the first three numbers of any block, we 
would get two of the same color. We needed to be able to extend the block to get an 
arithmetic progression of length three, so the blocks needed to be of length 5.

Now we are coloring with 3 colors, so we consider the first 4 numbers in the block. 
Within the first 4 we have to have two of the same color. Now we need to extend 
the block to ensure we can have a monochromatic arithmetic progression of length 3. 
Thus, we need the blocks to be of length 7 (because 7 = 3 \cdot 2 + 1).

There are \(3^7\) ways of coloring each of these blocks, so we need to have \((2 \cdot 3^7 + 1) = 
4375\) blocks to ensure that the block that we named \(B_k\) in the previous example will 
be defined.

That means, if we are to follow in a similar fashion as before, coloring with three 
colors, and trying to find an arithmetic progression of length three, we need 4375
blocks of length 7. Thus, we need to color at most 30625 numbers to find an arithmetic progression of length 3 when coloring with 3 colors, $w(3, 3) \leq 30625$. In reality, $w(3, 3) = 27$; see Section 7.2 for a discussion of bounds and known values.

Proceeding with the examples, if we were to increase the length of arithmetic progressions rather than the number of colors, the numbers that were of the form $2n+1$ (that determined the length of the blocks) would now have to be defined in terms of the Van der Waerden number for the previous length of arithmetic progressions. This means that the argument hinted at by the examples proceeds by a double induction: That is, we argue by induction on the length of the arithmetic progressions and, for each fixed length, we argue by induction on the number of colors.

For more proofs and information about this result, please see [12], [13], [14], [23], or references within.

### 7.2 Van Der Waerden Numbers and Their Bounds

**Definition 7.2** The Van der Waerden number for $r$ and $k$, denoted $w(k, r)$, is the least positive integer such that for every $r$-coloring of $[1, w(k, r)]$ there exists a monochromatic arithmetic progression of length $k$.

The known values of the function $w(k, r)$ are shown below, in Table 7.1.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9 (see [8])</td>
<td>35 (see [8])</td>
<td>178 (see [37])</td>
<td>1132 (see [20])</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27 (see [8])</td>
<td>293 (see [21])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>76 (see [5])</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As so few values of $w(r, k)$ are known, we naturally move to bounds for these numbers. In 1968 (see [6]), Berlekamp proved the following result.

**Theorem 7.3 (Berlekamp)** If $k$ is prime, then $w(k + 1, 2) > k(2^k - 1)$.

**Proof** Let $GF(2^k)$ denote the finite field with $2^k$ elements; see [22, Chapter VIII]. $GF(2^k)$ contains an element, say $\alpha$, that is primitive. That is, $\alpha$ generates the multiplicative group $GF(2^k)^*$. This means we can write every element of $GF(2^k)^*$ as $\alpha^i$ for some $i$. Now, fix a basis $v_1, v_2, ..., v_k$ for $GF(2^k)$ as a vector space over $\mathbb{Z}/2\mathbb{Z}$. We can write

$$\alpha^i = a_{1i}v_1 + a_{2i}v_2 + \ldots + a_{ki}v_k \quad (7.1)$$

where $a_{ij} \in \mathbb{Z}/2\mathbb{Z}$ for each $i$ and $1 \leq j \leq k$.

Now we color the numbers represented by powers of $\alpha$ by writing them in the expanded form 7.1 described above and coloring the element red if the coefficient of $v_1$ is 0, and coloring it blue if the coefficient of $v_1$ is 1. This gives us two sets of elements; set $R$ will represent the ones colored red, and set $B$ will represent the ones colored blue.

$$R = \{ i : a_{1i} = 0, 1 \leq i \leq k(2^k - 1) \},$$

$$B = \{ i : a_{1i} = 1, 1 \leq i \leq k(2^k - 1) \}.$$

We claim that this is a 2-coloring of $[1, k(2^k - 1)]$ that does not have a monochromatic arithmetic progression of length $k + 1$.

To show this, assume otherwise, so for some $a$ and $d$, \{a, a+d, a+2d, ..., a+kd\} $\subseteq R$ or \{a, a+d, a+2d, ..., a+kd\} $\subseteq B$.

Let $\beta = \alpha^a$ and $\gamma = \alpha^d$. Note we know
\[1 \leq a < a + kd \leq k(2^k - 1),\]
so we know \(\frac{a}{k} + \frac{kd}{k} \leq 2^k - 1\),
so we know \(\frac{a}{k} + d \leq 2^k - 1\),
so we know \(d < 2^k - 1\).

Thus, \(\gamma \neq 1\). This means we know \(\beta, \beta\gamma, \ldots, \beta\gamma^k\), which are \(\alpha^a, \alpha^{a+d}, \alpha^{a+2d}, \ldots, \alpha^{a+kd}\), have the same coefficient of \(v_1\) when written in the expanded form 7.1.

**Case 1:** First, assume \(\{a, a + d, a + 2d, \ldots, a + kd\} \subseteq R\), so for each of \(\beta, \beta\gamma, \ldots, \beta\gamma^{k-1}\), we have that \(v_1 = 0\). These are \(k\) vectors in a \((k - 1)\)-dimensional space because \(v_1 = 0\) for each of them. This means they are not linearly independent. Thus, there exists some \(a_0, \ldots, a_{k-1} \in \mathbb{Z}/2\mathbb{Z}\) such that

\[
\sum_{i=0}^{k-1} a_i (\beta \gamma^i) = 0
\]
and not all \(a_i\) are 0. Then, we have,

\[
\sum_{i=0}^{k-1} a_i \gamma^i = 0.
\]

But we have seen that \(\gamma \neq 1\) and we know \(\gamma \neq 0\). Since \(\gamma \in GF(2^k)\), the degree of \(\gamma\) over \(\mathbb{Z}/2\mathbb{Z}\) must divide \(k\). Since \(k\) is prime, \(\gamma\) must have degree precisely \(k\) over \(\mathbb{Z}/2\mathbb{Z}\). This is a contradiction.

Note that this argument shows in fact that \(R\) does not contain a progression of length \(k\) when \(d < 2^k - 1\).

**Case 2:** Now, assume \(\{a, a + d, a + 2d, \ldots, a + kd\} \subseteq B\). Then, for \(\beta, \beta\gamma, \ldots, \beta\gamma^k\), \(v_1 = 1\), so \(\beta(\gamma - 1), \beta(\gamma^2 - 1), \ldots, \beta(\gamma^k - 1)\) all have \(v_1 = 0\) and all lie in a \((k - 1)\)-
dimensional space. As before, this means that

$$\sum_{i=0}^{k} a_i[\beta(\gamma^i - 1)] = 0$$

for some $a_i \in \mathbb{Z}/2\mathbb{Z}$ with at least one $a_i \neq 0$. Dividing by $\beta(\gamma - 1)$, we find that $\gamma$ satisfies a polynomial of degree at most $k - 1$. Again, this is a contradiction.

Thus, there cannot be a monochromatic arithmetic progression of length $k + 1$ of either color. Therefore, $w(k + 1, 2) > k(2^k - 1)$. □

For clarity, we would now like to include an example that follows this proof. Let $k = 2$. First, we look at the field with $2^2$ elements, $GF(2^2)$. Here is the Cayley table for the additive group:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>$\alpha$</th>
<th>$\alpha^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>$\alpha$</td>
<td>$\alpha^2$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\alpha^2$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\alpha$</td>
<td>$\alpha^2$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha^2$</td>
<td>$\alpha^2$</td>
<td>$\alpha$</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Note that 1 and $\alpha$ form a basis for the field $GF(2^2)^*$ considered as a vector space over $\mathbb{Z}/2\mathbb{Z}$. Let $v_1 = 1$ and $v_2 = \alpha$. This means we can write the powers of $\alpha$ from 1 to $2(2^2 - 1)$ as follows:
\[
\alpha = 0 \cdot v_1 + 1 \cdot v_2 \\
\alpha^2 = 1 \cdot v_1 + 1 \cdot v_2 \\
\alpha^3 = 1 \cdot v_1 + 0 \cdot v_2 \\
\alpha^4 = 0 \cdot v_1 + 1 \cdot v_2 \\
\alpha^5 = 1 \cdot v_1 + 1 \cdot v_2 \\
\alpha^6 = 1 \cdot v_1 + 0 \cdot v_2
\]

Now we have the sets \( R = \{1, 4\} \) and \( B = \{2, 3, 5, 6\} \). Thus, we see that if we color the numbers in \( R \) red and the numbers in \( B \) blue, then we get a 2-coloring of \([1, 2(2^2 - 1)]\) that does not have a monochromatic arithmetic progression of length 3. This shows that \( w(3, 2) > 2(2^2 - 1) = 6 \).

Berlekamp [6] in fact proved a stronger result:

**Theorem 7.4 (Berlekamp)** If \( k \) is prime, then \( w(k + 1, 2) > k2^k \).

**Proof** We continue the argument in the proof of Theorem 7.3. The key is to choose a specific basis \( v_1, \ldots, v_k \), namely

\[
v_1 = 1, v_2 = 1 + \alpha, \ldots, v_{\frac{k+1}{2}} = 1 + \alpha^{\frac{k-1}{2}},
\]

and

\[
v_{\frac{k+3}{2}} = 1 + \alpha^{-1}, v_{\frac{k+5}{2}} = 1 + \alpha^{-2}, \ldots, v_k = 1 + \alpha^{-\frac{k-1}{2}}.
\]

(When \( k = 2 \), this is just \( v_1 = 1, v_2 = 1 + \alpha^{-1} = 1 + \alpha^2 = \alpha \). Similar adjustments apply in what follows.)
Let us begin by verifying that the $v_i$ are linearly independent. Otherwise, we obtain an equation of the form

$$\sum_{j=-\frac{k-1}{2}}^{\frac{k-1}{2}} a_j \alpha^j = 0,$$

where the $a_j$ are in $\mathbb{Z}/2\mathbb{Z}$ and at least one is nonzero. But this equation is equivalent to

$$\sum_{j=0}^{k-1} b_j \alpha^j = 0$$

for some $b_j \in \mathbb{Z}/2\mathbb{Z}$, not all equal to zero. This contradicts that $\alpha$ is primitive.

Now that we know that the $v_i$ are independent, we continue the argument in the proof of Theorem 7.3, explaining how $k$ additional numbers can be added to the partition described there without adding large monochromatic sets in arithmetic progression. The proof of Theorem 7.3 now gives us a partition $[1, k(2^k - 1)] = R \cup B$ without monochromatic arithmetic progressions of length $k + 1$. Note that the conclusions are not altered if we replace this interval with $[0, k(2^k - 1) - 1]$, the point being that $\alpha^0 = \alpha^{k(2^k-1)} = 1 = 1 \cdot v_1$, so $R$ does not change, $B$ becomes $(B \setminus \{k(2^k-1)\}) \cup \{0\}$, and the argument in Case 2 is unaffected. To simplify notation, we keep calling the sets $R$ and $B$.

Note that

$$C = \left[0, \frac{k-1}{2}\right] \subseteq B,$$  \hfill (7.2)

since

$$\alpha^t = 1 + (1 + \alpha^t) = 1 \cdot v_1 + 1 \cdot v_{t+1}$$

for $1 \leq t \leq \frac{k-1}{2}$, and $\alpha^0 = 1 \cdot v_1$. 

Similarly,

\[ D = \left[ k(2^k - 1) - \frac{k - 1}{2}, k(2^k - 1) - 1 \right] \subseteq B, \]  

(7.3)
since

\[ \alpha^{k(2^k - 1) - t} = \alpha^{-t} = 1 + (1 + \alpha^{-t}) = 1 \cdot v_1 + 1 \cdot v_{\frac{k+1}{2} + t} \]

for \( 1 \leq t \leq \frac{k-1}{2} \).

Let \( A = [-\frac{k-1}{2}, k(2^k - 1) + \frac{k-1}{2}] \). Now we color the elements of \( A \) red or blue, according to whether they belong to

\[ R' = R \cup R_1 \cup R_2 \]

or to

\[ B' = A \setminus R' = B, \]

where

\[ R_1 = \left[ -\frac{k - 1}{2}, -1 \right] \]

and

\[ R_2 = \left[ k(2^k - 1), k(2^k - 1) + \frac{k - 1}{2} \right]. \]

We claim that this coloring does not contain monochromatic arithmetic progressions of size \( k + 1 \). (Of course, this gives the result, as the corresponding coloring of \([1, k2^k]\) can be obtained by shifting \( A, R', B' \) to the right.)

If there is such a progression \( a, a + d, \ldots, a + kd \), it must be in \( R' \), by the proof of Theorem 7.3, since \( B' = B \). Also, it must meet \( R_1 \cup R_2 \). There are now three cases:

**Case 1:** The progression contains one element of \( R_1 \) and one of \( R_2 \).

This is impossible, because then \( a \in R_1 \) and \( a + kd \in R_2 \), so for some \( i \in R_2 \) and
$j \in R_1$, we have that $k$ divides $i - j$. But this is not the case.

**Case 2:** The progression contains at least 2 elements of $R_1$, or 2 elements of $R_2$.

This is also impossible, since it implies that at least one term of the progression must be in $C \cup D = [0, \frac{k-1}{2}] \cup [k(2^k - 1) - \frac{k-1}{2}, k(2^k - 1) - 1] \subseteq B$, as indicated in 7.2 and 7.3 above.

**Case 3:** The progression contains at most one element of $R_1 \cup R_2$. Then it has precisely $k$ elements in $R$, by Theorem 7.3, say $c, c + d, \ldots, c + (k - 1)d$. But this means that $d \geq 2^k - 1$, since otherwise the same argument as in Case 1 in the proof of Theorem 7.3 applies. It follows that $kd \geq k(2^k - 1)$. But since the remaining term of the progression is in $R_1 \cup R_2$, and since the progression is disjoint from $C \cup D$, this is impossible as well.

We continue the example when $k = 2$ started above. As indicated in the proof of Theorem 7.4, the basis we use is $v_1 = 1, v_2 = 1 + \alpha^{-1} = 1 + \alpha^2 = \alpha$. This means that $R = \{1, 4\}$ and $B = \{2, 3, 5, 6\}$ as before, except that we replace 6 with 0, so $B = \{0, 2, 3, 5\}$.

With notation as in the proof of Theorem 7.4, $R_1 = \{-1\}$ and $R_2 = \{6\}$, so $R' = \{-1, 1, 4, 6\}$ and $B' = B = \{0, 2, 3, 5\}$. Shifting this by 2 units gives us the coloring of $[1, 8]$ where 1, 3, 6, 8 are red, and 2, 4, 5, 7 are blue. Observe that neither color has 3 terms in arithmetic progression.

The above shows that $w(3, 2) \geq 9$. This is actually optimal:

**Theorem 7.5** $w(3, 2) = 9$.

**Proof** It remains to show that $w(3, 2) \leq 9$. It seems easiest to proceed by an analysis of cases: Consider a 2-coloring of $[1, 9]$, and let us see what restrictions
the coloring must satisfy in order to avoid arithmetic triples. We must conclude that it is impossible to have such a coloring. The key is to consider 4, 5, 6. They cannot all be the same color, but two of them must be. Let us call that color A, and let B be the other one.

**Case 1:** 4, 6 ∈ A.

This is the easiest case to eliminate, since 2, 5, 8 must then be an arithmetic triple of color B: Consider, respectively, the triples 2, 4, 6, and 4, 5, 6, and 4, 6, 8. If we place any of 2, 5, 8 ∈ A, one of these three arithmetic triples ends up in A.

**Case 2:** 4, 5 ∈ A.

Then, 3, 6 ∈ B (consider, respectively, the triples 3, 4, 5 and 4, 5, 6), so 9 ∈ A (consider 3, 6, 9), but then 1, 7 ∈ B (consider 1, 5, 9 and 5, 7, 9), so 2, 8 ∈ A (consider 1, 2, 3 and 6, 7, 8). We now see this case cannot happen either, because the triple 2, 5, 8 is in A.

**Case 3:** 5, 6 ∈ A.

This is really the same as Case 2, by symmetry. (In this case, 4, 7 ∈ B, so 1 ∈ A, so 3, 9 ∈ B, so 2, 8 ∈ A, and we see that the triple 2, 5, 8 is in A.) □

We also have a lower bound for non-prime k,

**Theorem 7.6** \(w(2, k) > \left\lceil \left(\frac{2^k}{2ek}\right)(1 + o(1)) \right\rceil\)

The proof of this uses probabilistic methods and can be found in [13].

The study of Van der Waerden numbers and the rate of growth of the function \(w(r, k)\) is very popular among researchers in Ramsey Theory. We only mention a few remarks.

One of the best upper bounds known to date was proven by Timothy Gowers in 1998; see [36].
**Theorem 7.7 (Gowers)** For any positive integer $k$, $w(2, k) \leq 2^{2^{2k}+9}$.

![Figure 7.1: Ronald Graham presenting the check to Timothy Gowers for the proof of this bound. Taken from [36] page 363.](image)

This result is a significant improvement of a 1988 upper bound of Shelah (see [34]), itself a significant improvement of the bound produced by the argument sketched in Section 7.1.

Improvements of these bounds are of great interest, and Ronald Graham has offered a $1000 prize for a proof of this conjecture:

**Conjecture 7.8** For any $k$, $w(2, k) < 2^{k^2}$.

As a suggestion of how these numbers grow, we mention the following conjecture by Erdős.

**Conjecture 7.9** $\lim_{k \to \infty} \frac{w(2, k)}{2^k} = \infty$.

See [13], [23], and [34] and the references therein for a detailed discussion of these numbers. Also, for more information on bounds and conjectures having to do with Van der Waerden numbers, see [36, Chapter 35.7].
Here is a table from the survey [16] on some of the known Van der Waerden numbers and lower bounds. See [16] for references.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9</td>
<td>35</td>
<td>178</td>
<td></td>
<td>&gt;1131</td>
<td>&gt;3703</td>
<td>&gt;11495</td>
<td>&gt;41265</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>&gt;292</td>
<td>&gt;1209</td>
<td>&gt;8886</td>
<td></td>
<td>&gt;43855</td>
<td>&gt;238400</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>76</td>
<td>&gt;1048</td>
<td>&gt;17705</td>
<td>&gt;91331</td>
<td></td>
<td>&gt;393469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>&gt;125</td>
<td>&gt;2254</td>
<td>&gt;24045</td>
<td>&gt;246956</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>&gt;207</td>
<td>&gt;9778</td>
<td>&gt;63473</td>
<td>&gt;633981</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2: Known Van Der Waerden Numbers, $w(r, k)$ and Lower Bounds
CHAPTER 8

GENERALIZATIONS OF SCHUR’S THEOREM

8.1 Schur Quadruples

It is a natural question whether Schur’s theorem can be generalized to structures other than the positive integers. In [26], McCutcheon colors the elements of the symmetric group $S_n$. Instead of searching for triples, he searches for quadruples of the form \{x, y, xy, yx\} ($S_n$ is non-abelian, so $xy$ and $yx$ could \emph{a priori} be different). Such a quadruple is called a \emph{Schur quadruple}.

**Theorem 8.1 (McCutcheon)** Let $r \in \mathbb{N}$. There exists $n = n(r)$ such that for any $r$-coloring of the alternating group $A_n$, there is a monochromatic Schur quadruple.

Recall that the elements of $A_n$ are the even permutations of $[1, n]$. In particular, all 3-cycles $(ijk)$ are in $A_n$.

**Proof** By Ramsey’s Theorem 2.1, we can choose an $n$ such that for any $r$-coloring of the 3-element subsets of $[1, n]$ there is a 4-element set whose 3-element subsets form a monochromatic family.

Now suppose we are given an $r$-coloring of $A_n$ with color set $i$ where each $C_i$ is a different color.

$$A_n = \bigcup_{i=1}^{r} C_i,$$
we define an $r$-coloring $[1,n]^3 = \bigcup_{m=1}^r D_m$ as follows: If the permutation $(ijk)$ is in $C_m$, where $i < j < k$, we put $\{i,j,k\} \in D_m$.

Now, by choice of $n$, there exists a 4-element set $\{a,b,c,d\}$ with $a < b < c < d$, all of whose 3-element subsets have the same color $D_y$. That is,

$$\{\{a,b,c\}, \{a,b,d\}, \{a,c,d\}, \{b,c,d\}\} \subseteq D_y.$$ 

By definition of $D_y$, this gives us that

$$E = \{(abc), (abd), (acd), (bcd)\} \in C_y.$$ 

Observe that $(abc)(acd) = (bcd)$ and $(acd)(abc) = (abd)$, so $E$ has the form $\{x, y, xy, yx\}$, where $(abc) = x$ and $(acd) = y$. This shows that any $r$-coloring of the alternating group $A_n$ admits a monochromatic Schur quadruple, whenever $n$ is sufficiently large, as indicated. \hfill \square

### 8.2 Rado’s Theorem

Schur’s Theorem naturally leads us to ask which (linear) equations have monochromatic solutions under finite colorings of the positive integers, or equivalently, which equations admit monochromatic solutions under finite colorings of $\mathbb{Z}^+$, the nonzero integers. Schur’s theorem corresponds to the equation

$$x + y - z = 0.$$
Richard Rado, a student of Schur’s, turned his attention to this general setting, considering arbitrary coefficients and an arbitrary number of variables. He determined exactly when equations of the form

$$\sum_{i=1}^{k} c_i x_i = 0$$

are guaranteed to have monochromatic solutions under any finite coloring of the positive integers. In the 1930’s, his results appeared in a series of papers as well as in his Ph.D. dissertation [28]; see also [23].

First, recall that a \textit{(linear) homogeneous} equation is any equation of the form

$$\sum_{i=1}^{k} c_i x_i = 0$$

where each $c_i \in \mathbb{Z}$ is a nonzero constant and each $x_i$ is a variable.

\textbf{Definition 8.2} Let $S$ be a linear homogeneous equation and let $r \geq 1$. We say that $S$ is $r$-regular if for every $r$-coloring of $\mathbb{Z}^+$ there is a monochromatic solution to $S$. If $S$ is $r$-regular for all $r \geq 1$, we say that $S$ is regular.

For the proof of Rado’s Theorem, we need the following compactness lemma.

\textbf{Lemma 8.3} If there is a monochromatic solution to

$$\sum c_i x_i = 0$$

under any finite coloring of $\mathbb{Z}^+$, then for all $r$, there exists a number $n(r)$ such that for all $r$-colorings of $[1, n(r)]$ there is a monochromatic solution to
\[ \sum c_i x_i = 0. \]

**Proof** Suppose, for the sake of contradiction, that for some fixed \( r \) and for all \( n \), we have at least one function, \( f : [1, n] \rightarrow [1, r] \), without a monochromatic solution to the equation \( \sum_{i=1}^k c_i x_i = 0 \). Let the set of all such functions for a given \( n \) (those without a monochromatic solution) be named \( F_n \). Note that \( F_n \) is finite, and in fact \( |F_n| \leq r^n \). By assumption, \( F_n \) is nonempty.

Let \( n_1 = 1 \) and choose \( f_m \in F_m \) for all \( m \geq n_1 \). We see that \( f_1 \mid [1, n_1], f_2 \mid [1, n_1], \ldots \) are all in \( F_{n_1} \). But note that there are infinitely many functions \( f_i \), and \( F_{n_1} \) is finite.

Let \( g_{n_1} \in F_{n_1} \) be such that for infinitely many \( m \), \( f_m \mid [1, n_1] = g_{n_1} \). Let \( A_{n_1} \) be the set of all these \( m \).

Suppose we have defined \( n_1 < n_2 < \ldots < n_k \), we have chosen functions \( g_{n_1} \in F_{n_1}, \ldots, g_{n_k} \in F_{n_k} \), and infinite sets \( A_1, \ldots, A_k \) with \( \min A_j > n_j \) and such that for all \( m \in A_j \), \( f_m \mid [1, n_j] = g_{n_j} \), for all \( j \) and \( A_1 \supseteq A_2 \supseteq \ldots \supseteq A_k \).

Now, let \( n_{k+1} = \min A_k \). Then, \( f_i \mid [1, n_{k+1}] \in F_{n_{k+1}} \) for all \( i \in A_k \). Since \( F_{n_{k+1}} \) is finite but \( A_k \) is infinite, there must be a function \( g_{n_{k+1}} \in F_{n_{k+1}} \) and an infinite set \( A_{k+1} \subseteq A_k \) with \( \min A_{k+1} > n_{k+1} \) and \( f_i \mid [1, n_{k+1}] = g_{n_{k+1}} \) for all \( i \in A_{k+1} \). This completes the recursive construction.

For all \( k \), \( A_k \supseteq A_{k+1} \), the functions \( g_{n_k} \) are compatible (meaning \( g_{n_j} \subseteq g_{n_k} \), whenever \( j \leq k \)) and \( n_k < n_{k+1} \). Let \( G = \bigcup_k g_{n_k} \), so \( G \) is an \( r \)-coloring of \( \mathbb{Z}^+ \) without a monochromatic solution to \( \sum_{i=1}^k c_i x_i = 0 \). But this is a contradiction to the assumption of our lemma. Therefore, for any \( r \), there exists an \( n \) such that for any \( r \)-coloring of \([1, n]\), we have a monochromatic solution to \( \sum_{i=1}^k c_i x_i = 0 \). \( \square \)

We could have used a similar argument to prove Schur’s Theorem. However, our knowledge of Schur numbers would have only gone as far as existence. Of course we
wanted to know more about Schur numbers and their bounds, so we needed a more elaborate proof. The existence of the number \( n(r) \) is all we will need in the proof of Rado’s Theorem, therefore this proof of existence will suffice.

Now we are ready to prove Rado’s Theorem.

**Theorem 8.4 (Rado’s Single Equation Theorem)**

Let \( k \geq 2 \), and let \( c_i \in \mathbb{Z}, 1 \leq i \leq k \), be constants. Then,

\[
\sum_{i=1}^{k} c_i x_i = 0
\]

is regular if and only if there exists a nonempty \( D \subseteq \{c_i : 1 \leq i \leq k\} \) such that

\[
\sum_{d \in D} d = 0.
\]

**Proof** First, given \( k \geq 2 \) and constants \( c_i \in \mathbb{Z}, 1 \leq i \leq k \), we prove that if there exists a nonempty \( D \subseteq \{c_i : 1 \leq i \leq k\} \) such that \( \sum_{d \in D} d = 0 \), then

\[
\sum_{i=1}^{k} c_i x_i = 0 \quad (8.1)
\]

is regular. Assume that \( \sum_{d \in D} d = 0 \) and assume without loss of generality, that \( c_1 > 0 \) and that \( D = \{c_1, c_2, ..., c_m\} \), with \( m = |D| \) maximal. If \( m = k \), then we have

\[
\sum_{d \in D} d = \sum_{i=1}^{k} c_i = 0.
\]

This would mean we could take \( x_i = 1 \) for \( 1 \leq i \leq k \) and we would have the monochromatic solution
\[
\sum_{i=1}^{k} x_i c_i = \sum_{i=1}^{k} c_i = 0.
\]

So we assume that \( m < k \) and let \( s = c_{m+1} + c_{m+2} + \ldots + c_k \). Note that \( s \neq 0 \) (by the maximality of \( m \)). Now, we impose the additional restrictions that \( x_2 = x_3 = \ldots = x_m \) and \( x_{m+1} = x_{m+2} = \ldots = x_k \). The equation (8.1) now becomes

\[
c_1 x_1 + x_2 (c_2 + c_3 + \ldots + c_m) + x_{m+1} (c_{m+1} + \ldots + c_k) = 0.
\]

But since \( c_1 + \ldots + c_m = 0 \), equation (8.1) becomes

\[
c_1 (x_1 - x_2) + sx_{m+1} = 0. \tag{8.2}
\]

We now begin an induction on \( r \). For \( r = 1 \), any positive integers \( x_1 \) and \( x_2 \) such that \( x_2 - x_1 = s \) together with \( x_{m+1} = c_1 \) will be a monochromatic solution to (8.2).

Assume that \( r \geq 2 \) and that the result holds for \( r - 1 \). For each \( t \leq r - 1 \), let \( n(t) \) be the least positive integer such that for every \( t \)-coloring of \([1, n(t)]\) there is a monochromatic solution to equation (8.1). (Note that this exists by inductive hypothesis; see Lemma 8.3.)

Let \( b = \sum_{i=1}^{k} |c_i| \) and write \( n \) for \( n(r-1) \). Let \( \chi \) be an \( r \)-coloring of the interval \([1, w(n+1, r)b]\), where \( w(n+1, r) \) is the Van der Waerden number defined in Section 7.2.

Our goal is now to show that \( \chi \) admits a monochromatic solution to equation (8.2). This means we want to find \( x_1, x_2, \) and \( x_{m+1} \) such that \( \chi(x_1) = \chi(x_2) = \chi(x_{m+1}) \), and \( c_1 (x_1 - x_2) + sx_{m+1} = 0 \).

Recall that \( 0 \neq s = \sum_{i=m+1}^{k} c_i \) and that \( b = \sum_{i=1}^{k} |c_i| \), thus \( 1 \leq |s| < b \).

Given \( 1 \leq l \leq b \), let \( \chi_l \) be the coloring of \([1, w(n+1, r)]\) given by \( \chi_l(i) = \chi(li) \).
This gives us for each $l$ a $\chi_l$-monochromatic set

$$\{\alpha, \alpha + d, \alpha + 2d, \ldots, \alpha + nd\},$$

where each of the elements is less than or equal to $w(n + 1, r)$. So we have that $l\alpha, l\alpha + ld, \ldots, l\alpha + lnd$ are monochromatic under the coloring of $\chi$ and all are less than or equal to $l \cdot w(n + 1, r)$. This means that they are all in $[1, w(n + 1, r)b]$.

Now, let $l = |s|$ and $a = l\alpha$, and we have for some $d \geq 1$ a monochromatic set

$$\{a, a + d|s|, a + 2d|s|, \ldots, a + nd|s|\}$$

with $a + nd|s| \leq w(n + 1, r)b$.

If there exists $j \in [1, n]$ such that $\chi(jdc_1) = \chi(a)$, then by letting $|x_2 - x_1| = jd|s|$ and $x_{m+1} = jdc_1$, we have a monochromatic solution to equation (8.2). If, on the other hand, for all $j \in [1, n]$ we have $\chi(jdc_1) \neq \chi(a)$, then

$$\{dc_1, 2dc_1, \ldots, ndc_1\} = dc_1[1, n]$$

is colored with only $r - 1$ colors. This means we have a monochromatic solution to equation (8.2).

Thus, in either case, we are finished with the first implication.

Now, we prove if

$$\sum_{i=1}^{k} c_i x_i = 0$$

is regular, then there exists a nonempty $D \subseteq \{c_i : 1 \leq i \leq k\}$ such that
\[ \sum_{d \in D} d = 0. \]

We will prove this by proving the contrapositive, so assume we have \( c_1, c_2, ..., c_k \) such that no nonempty subset of \( D \) sums to 0.

Now, we will show that for some \( r \) there is an \( r \)-coloring of \( \mathbb{Z}^+ \) without monochromatic solutions to equation (8.1).

First, choose a prime \( p \) such that, for any \( C \subseteq \{ c_i : 1 \leq i \leq k \} \), we have that \( p \) does not divide \( \sum_{c \in C} c \). Note that we can do this because there are only finitely many choices for \( C \). We claim that \( r = p - 1 \) works. For this, we will define a \((p - 1)\)-coloring \( \chi : \mathbb{Z}^+ \to [1, p - 1] \) as follows:

Given \( i \in \mathbb{Z}^+ \), let \( s \) be largest such that \( p^s \mid i \), so \( i = p^s j \) where \( j \not\equiv 0 \pmod{p} \), and define \( \chi(i) = j \pmod{p} \).

To show that \( \chi \) does not admit a monochromatic solution to equation (8.1), assume that we have one. Let \( R = \{ d_1, d_2, ..., d_k \} \) be a monochromatic solution under \( \chi \), and call \( m \) its color, so \( 1 \leq m \leq p - 1 \), and for each \( d_i \) there are numbers \( s_i, k_i \) such that \( d_i = p^{s_i}(pk_i + m) \). Let \( s = \min\{s_1, \ldots, s_k\} \). We have that

\[
0 = \sum_{i=1}^{k} c_i d_i = p^s \sum_{i=1}^{k} c + ip^{s_i - s}(pk_i + m),
\]

and therefore

\[
0 = \sum_{i=1}^{k} c_i p^{t_i}(pk_i + m),
\]

where \( t_i = s_i - s \). Note that \( 0 \leq t_i \) for all \( i \), with equality for at least one \( i \). Modulo \( p \), this equation becomes

\[
0 \equiv m \sum_{i=1}^{k} p^{t_i} c_i \pmod{p}.
\]
Since $p$ is prime and $p \nmid m$, this gives us that

$$p \mid \sum_{i \in [1,k], t_i = 0} c_i.$$  

Call $C = \{c_i : i \in [1, k], t_i = 0\}$, and note that $C \neq \emptyset$. But $p$ was chosen such that

$$p \nmid \sum_{c_i \in C} c_i.$$  

This is a contradiction. □

We have already considered examples of the first direction in this proof, through our study of Schur's theorem. Since the second direction of the proof was constructive, it would be nice to show an example as well.

Consider the equation

$$x + 2y - 4z = 0. \quad (8.3)$$

We have $c_1 = 1, c_2 = 2, c_3 = -4$, and no nonzero subset of $\{c_1, c_2, c_3\}$ sums to 0:

$$c_1 + c_2 + c_3 = -1,$$
$$c_1 + c_3 = -3,$$
$$c_2 + c_3 = -2,$$
$$c_1 + c_2 = 3.$$  

Since $p = 5$ is a prime that does not divide any of these sums, the proof of Theorem 8.4 gives us that the coloring $\chi : \mathbb{Z}^+ \to [1, 4]$ given by

$$\chi(5^kj) = j \mod 5,$$
where \(5 \nmid j\), is a 4-coloring of \(\mathbb{Z}^+\) without monochromatic solutions of the equation (8.3), that is, equation (8.3) is not 4-regular.

It is natural to wonder what the least \(r\) is such that equation (8.3) is not \(r\)-regular. This is addressed in the following theorem, a particular case of results from [1].

**Theorem 8.5** The equation \(x + 2y - 4z = 0\) is 2-regular over \(\mathbb{Z}^+\), but not 3-regular.

**Proof** To see that equation (8.3) is 2-regular, note that

\[
4 + 2 \cdot 2 = 4 \cdot 2, \\
2 + 2 \cdot 3 = 4 \cdot 2, \text{ and} \\
4 + 2 \cdot 4 = 4 \cdot 3,
\]

so if a coloring avoids monochromatic solutions of equation (8.3), then \(2, 3, 4\) must all have different colors, and the coloring requires at least 3 colors.

To see that equation (8.3) is not 3-regular, we define a coloring \(\chi : \mathbb{Z}^+ \to [0, 2]\) as follows: Given \(i \in \mathbb{Z}^+\), write \(i = 2^oj\) where \(2 \nmid o\), and let \(\chi(i) = j \pmod{3}\).

We claim that no solution of equation (8.3) is monochromatic under \(\chi\); suppose that \(x, y, z\) is a solution and \(\chi(x) = \chi(y) = \chi(z) = j\), so there are natural numbers, \(x_1, y_1, z_1\), and \(x_2, y_2, z_2\), possibly 0, such that

\[
x = 2^{3x_1+j}(2x_2 + 1), \quad y = 2^{3y_1+j}(2y_2 + 1), \quad z = 2^{3z_1+j}(2z_2 + 1),
\]

and

\[
2^{3x_1+j}(2x_2 + 1) + 2 \cdot 2^{3y_1+j}(2y_2 + 1) = 4 \cdot 2^{3z_1+j}(2z_2 + 1). \quad (8.4)
\]

As in the proof of Theorem 8.4, let \(t = \min\{x_1, y_1, z_1\}\), and set
\[ x_3 = x_1 - t, \quad y_3 = y_1 - t, \quad z_3 = z_1 - t, \]

so \( x_3, y_3, z_3 \geq 0 \), with equality for at least one of them.

Dividing equation (8.4) by \( 2^{3t+j} \), we obtain

\[
dfrac{2^{3x_3}(2x_2 + 1) + 2 \cdot 2^{3y_3}(2y_2 + 1)}{2^{3t+j}} = \dfrac{4 \cdot 2^{3z_3}(2z_2 + 1)}{2^{3t+j}}.
\]

We reduce equation (8.5) modulo 2 to obtain

\[
2^{3x_3} \equiv 0 \pmod{2};
\]

that is, \( x_3 > 0 \). Similarly, considering equation (8.5) modulo 4 now gives us

\[
2 \cdot 2^{3y_3} \equiv 0 \pmod{4};
\]

so \( y_3 > 0 \) as well. But then, considering equation (8.5) modulo 8 finally gives us

\[
0 \equiv 4 \cdot 2^{3z_3} \pmod{8};
\]

so \( z_3 > 0 \). This is a contradiction. \( \square \)

We would like to mention one more example. E.G. Straus showed (see [32]) that any 3-coloring of \([1, 54]\) would admit a monochromatic solution to the equation \( x + y = 3z \), and that it is possible to color \([1, 53]\) without any monochromatic solution.

Here is the example that he found. Let \([1, 53]\) be colored as follows:
\[ R = \{x : x \equiv 1 \pmod{3}\} \cup \{x : x \equiv 3 \pmod{9}\} \]
\[ B = \{x : x \equiv 2 \pmod{3}\} \cup \{9, 27, 36\} \]
\[ Y = \{6, 15, 18, 24, 33, 42, 45, 51\} \]

This is a 3-coloring of [1, 53] that does not admit any monochromatic solution to the equation \(x + y = 3z\). This means that \(x + y = 3z\) is not 3-regular in [1, 53] but it is 3-regular in \(\mathbb{Z}^+\); it is not 4-regular by the argument of Theorem 8.4.

We can now see that the following generalization of Schur’s Theorem (Theorem 1.11) follows directly from Theorem 8.4.

**Theorem 8.6** Assume \(r \geq 1\) and, for \(1 \leq i \leq r\), assume that \(k_i \geq 3\), then there exists a least positive integer \(S = S(k_1, k_2, ..., k_r)\) such that for every \(r\)-coloring of [1, \(S\)] there is a monochromatic solution of color \(j \in [1, r]\) to the equation \(x_1 + x_2 + ... + x_{k_j-1} = x_{k_j}\) where \(x_1, x_2, ..., x_{k_j}\) are all variables.

Much less is known about non-regular equations. We refer the reader to [1], [2], and [11].

**8.3 Maximum Solution-Avoiding Sets**

In this and the following section, we would like to mention some theorems that have to do with particular properties of the sets that are created when we seek to avoid solutions to equations. The first theorem has to do with the size and elements of the sets that avoid Schur triples, and the second theorem has to do with the number of times a particular color is used. See [7] for more information on this theorem.

**Theorem 8.7** A maximum subset of [1, \(n\)] containing no solutions to \(x + y = z\) (not necessarily distinct) has size \(\lceil \frac{n}{2} \rceil\). If \(n \geq 3\) is odd there are precisely two maximum
subsets: the odd integers less than or equal to \(n\) and \(\{x \in \mathbb{Z} : \frac{n+1}{2} \leq x \leq n\}\). If \(n \geq 4\) is even there are at least three maximum subsets: the two maximum subsets for the odd number \(n - 1\) and \(\{x \in \mathbb{Z} : \frac{n+1}{2} \leq x \leq n\}\). For even \(n \geq 10\), these three are the only ones. For smaller even numbers, \(\{1, 4\}, \{2, 5, 6\}, \{1, 4, 6\}, \) and \(\{2, 3, 7, 8\}\) are the only additional ones.

For this proof, a sum-free subset is a subset of the interval \([1, n]\) that does not contain a triple \(\{x, y, z\}\) such that \(x + y = z\). A sum-free subset is maximum if no sum-free subset of \([1, n]\) has size larger than \(|X|\).

**Proof** First, we will show that a maximum sum-free subset always has size \(\lceil \frac{n}{2} \rceil\).

Let \(X \subset [1, n]\) be a maximum sum-free subset. Let \(m\) be the largest element of \(X\). Then, for \(i = 1, 2, 3, ..., \lceil \frac{m}{2} \rceil\), at most one of \(i\) or \(m - i\) can be in \(X\). Else \(\{i, m - i, m\}\) would be in \(X\). But then, the size of \(X\) must be at most the number of values of \(i\) plus one, because \(m \in X\). Thus, \(|X| \leq \lceil \frac{m}{2} \rceil\). (Note that if \(m\) is even, \(\frac{m}{2}\) cannot be in \(X\).) Now, since \(m \leq n\), \(|X| \leq \lceil \frac{n}{2} \rceil\).

We know there exist maximum sum-free subsets of \(X\) such that \(|X| \geq \lceil \frac{n}{2} \rceil\). For example, if \(n\) is even, then we could take \(X\) to be the odd numbers from 1 to \(n - 1\), so, \(|X| = \lceil \frac{n}{2} \rceil\). Since \(n\) was arbitrary, this shows that maximum sum-free subsets always have size \(\lceil \frac{n}{2} \rceil\).

For the rest of the proof, we will characterize the maximum sum-free subsets, which naturally breaks into two cases, depending on the parity of \(n\).

**Case 1:** \(n\) is odd.

To show that any maximum sum-free subset \(X \in [1, n]\) must be either the set of odd integers less than or equal to \(n\) or the set \(\lceil \frac{n+1}{2}, n\rceil\), we will assume \(X\) is neither of the two, and arrive at a contradiction.
Let \( n \geq 5 \) be the least odd integer such that there exists a maximum sum-free subset \( X \in [1, n] \) that is neither the set of odd integers less than or equal to \( n \), nor the set \( [\frac{n+1}{2}, n] \).

Note that \( n \in X \), or else \( |X| \leq n - 1 < \left\lceil \frac{n}{2} \right\rceil \). Now suppose that \( n - 1 \notin X \). Note that \( |X \cap [1, n - 2]| = \left\lceil \frac{n-2}{2} \right\rceil \), then by the minimality of \( n \), either \( X \) is the set of all odd integers less than or equal to \( n \) (which is not the case by assumption), or \( X = [\frac{n-1}{2}, n - 2] \cup \{n\} \). But this is a contradiction since \( \{\frac{n-1}{2}, \frac{n+1}{2}, n\} \in X \) and \( \frac{n-1}{2} + \frac{n+1}{2} = n \). Thus, \( n - 1 \) and \( n \) are both in \( X \), which means 1 and \( \frac{n-1}{2} \) are not in \( X \).

Let \( G \) be the graph vertex set \( V \), where

\[
V = \left\{ v_i : i \in \left[ 2, \frac{n-3}{2} \right] \cup \left[ \frac{n+1}{2}, n - 2 \right] \right\},
\]

so \( V = n - 4 \), and where \( \{v_i, v_j\} \) is an edge of \( G \) if and only if \( i + j = n \) or \( i + j = n - 1 \).

Note that \( G \) has an edge from \( v_{n-2} \) to \( v_2 \), and from \( v_2 \) to \( v_{n-3} \), and from \( v_{n-3} \) to \( v_3 \), and so on. This means \( G \) is the path \( v_{n-2}, v_2, v_{n-3}, \ldots, v_{\frac{n-3}{2}}, v_{\frac{n+1}{2}} \).

Every element of \( X \) other than \( n \) and \( n - 1 \) corresponds to a vertex in \( G \). Note that no two adjacent vertices of \( G \) can both be in \( X \). Since we know \( X \) is maximal, and must have \( \left\lceil \frac{n}{2} \right\rceil \) elements (including \( n - 1 \) and \( n \), and not including 1, and \( \frac{n-1}{2} \)), we have to choose \( \frac{n-3}{2} \) vertices from \( G \). In order to be able to choose this many vertices from \( G \), without choosing any two adjacent, we must choose the vertices with indices \( [\frac{n+1}{2}, n - 2] \), so \( X = [\frac{n+1}{2}, n] \).

**Case 2:** \( n \) is even.

For the case where \( n = 4, 6, \) or 8, it is easy to check that the theorem is correct. Let \( n \geq 10 \) be the smallest even integer such that there exists a maximum sum-free
subset $X$ of $[1, n]$ not contained in $\left[\frac{n}{2} + 1, n\right]$, and is not one of the two maximum sum-free subsets of $[1, n - 1]$. First, we will show that $n - 1$ and $n$ must both be in $X$.

Suppose $n \notin X$. Then, we are in the odd case, and $X$ must be one of the two maximum sum-free subsets of $[1, n - 1]$, by the previous case, against our assumption. Thus, we have $n \in X$. Now suppose that $n - 1 \notin X$. We let $Y = X \cap [1, n - 2]$, so that $|Y| = \frac{n - 2}{2}$. $Y$ cannot be the odd integers less than or equal to $n - 3$ because $3 + (n - 3) = n$ and $\{3, n - 3, n\}$ would be a subset of $X$. Note that $Y$ cannot be $[\frac{n - 2}{2}, n - 3]$ or $[\frac{n}{2}, n - 2]$ either because $\frac{n}{2} \notin X$. This means that $Y$ cannot be any of the three maximum subsets for even $n$ that we described in the theorem. Thus, by minimality of $n$ we would have to have for $n = 10$, $Y = \{2, 3, 7, 8\}$, which is a contradiction because then $\{2, 8, 10\} \subset X$. Thus, as with the odd case, $n - 1 \in X$.

Now, let $H$ be the graph with vertex set $V$, where

$$V = \left\{ v_i : i \in \left[2, \frac{n - 2}{2}\right] \cup \left[\frac{n + 2}{2}, n - 2\right]\right\},$$

so $|V| = n - 4$, and where $\{v_i, v_j\}$ is an edge of $H$ if and only if $i + j = n$, $i + j = n - 1$, or $\{i, j\} = \{n - 2, \frac{n - 2}{2}\}$.

Similarly to the odd case, $H$ has an edge from $v_{n-2}$ to $v_2$, and from $v_2$ to $v_{n-3}$, and from $v_{n-3}$ to $v_3$, and so on. But $H$ also has an edge from $v_{\frac{n-2}{2}}$ to $v_{n-2}$. This means $H$ is the cycle $v_{n-2}, v_2, v_{n-3}, ..., v_{\frac{n+2}{2}}, v_{n-2}$.

Every element of $X$ other than $n$ and $n - 1$ corresponds to a vertex of $H$. Note that no two adjacent vertices of $H$ can both be in $X$. Since we know $X$ is maximal, and must have $\frac{n}{2}$ elements (including $n - 1$ and $n$, and not including $1$, and $\frac{n}{2}$), we have to choose $\frac{n-4}{2}$ vertices from $H$. In order to be able to choose this many vertices
from $H$, without choosing any two adjacent, we must choose either the vertices with indices $[\frac{n+2}{2}, n-2]$, or $[2, \frac{n-2}{2}]$. But we see that if $n \geq 10$, we cannot choose the vertices with indices $[2, \frac{n-2}{2}]$ because it will contain 2 and 4. Therefore, $X = [\frac{n+1}{2}, n]$. 

\[ \Box \]

8.4 Roth and Szemerédi’s Theorems

Van der Waerden’s Theorem 7.1 assures us that for any given $k$ and $r$, if $n$ is sufficiently large, then any $r$-coloring of $[1, n]$ must contain a monochromatic arithmetic progression of length $k$.

**Theorem 8.8 (Roth’s Theorem)** For any $\delta > 0$ and any $r$, there is an $N$ such that

if $n > N$ and $[1, n]$ is $r$-colored, then there are nontrivial monochromatic solutions to the equation $x + y = 2z$ in any color used with density more than $\delta$, where the density of a set $A$ is defined by $\frac{|A|}{n}$.

Note the when $k$ is 3, this corresponds to the nontrivial monochromatic solution to the equation $x + y = 2z$, where the nontriviality means that $x \neq y$.

Szemerédi’s theorem (see [36]) generalizes Roth’s Theorem 8.8 to arithmetic progressions of arbitrary length. This is how the latest upper bounds on Van der Waerden numbers discussed in Chapter 7 have been found.

It follows from Szemerédi’s result that the maximum size of a subset of $[1, n]$ avoiding arithmetic progressions of a given length approaches 0, as a function of $n$, and the determination of the rate at which this happens is an active area of research.
Let $f^*(n, 2)$ denote the maximum size subset of $[1, n]$ containing no three-term arithmetic progressions. Roth in [33] showed that:

$$f^*(n, 2) = O\left(\frac{n}{\log \log n}\right).$$

Some of the most recent bounds for $f^*(n, 2)$ are as follows, where the $c_i$ are absolute constants:

$$ne^{-c_1 \sqrt{\log n}} < f^*(n, 2) < \frac{c_2 n}{(\log n)^{c_3}}.$$  

The lower bound is due to Salem and Spencer, while the upper bound was proved by Heath-Brown and Szemerédi; see [7].

The situation is again different once we examine $x + y = 3z$ or $x + y = kz$ for $k > 2$, since these equations are no longer regular. Although not much is known, the paper [7] studies the case $k = 3$ in some detail. Here are two of the results from [7]:

**Theorem 8.9** Let $T_n$ be a subset of $[1, n]$ of maximum size such that $x + y = 3z$ has no solutions with $x, y, z \in T_n$ ($x, y, z$, not necessarily distinct). If $n \neq 4$, then $|T_n| = \left\lceil \frac{n}{2} \right\rceil$.

**Theorem 8.10** If $n \geq 23$ and $T_n$ is a subset of maximum size of $[1, n]$ having no solutions to $x + y = 3z$, then $T_n$ is the set of all odd integers less than or equal to $n$.

### 8.5 Hindman’s Theorem

Lastly, we would like to state a far-reaching generalization of Schur’s Theorem 1.11: Hindman’s Theorem (see [17]).
Definition 8.11 Given a set $A \subseteq \mathbb{N}$, let $\mathcal{FS}(A)$ be the set of sums of finite non-empty subsets of $A$. Also, let $\mathcal{FP}(A)$ be the set of products of finite non-empty subsets of $A$.

Theorem 8.12 (Hindman, 1979) If $\mathbb{N}$ is finitely colored, then there exist infinite subsets $A \subseteq \mathbb{N}$ and $B \subseteq \mathbb{N}$ such that $\mathcal{FS}(A)$ and $\mathcal{FP}(B)$ are monochromatic.

In the same paper that his theorem first appeared, Neil Hindman asked the following:

Question 8.13 Suppose that the positive integers are partitioned into finitely many pieces,

$$\mathbb{N} = A_1 \cup A_2 \cup \ldots \cup A_n.$$ 

Must there be integers $x, y$ such that $x, y, x + y$ and $xy$ belong to the same $A_i$?

This question is still open; see for example Question 3 in [18].

8.6 Concluding Remarks

In closing, we hope that this survey of Ramsey theory, with a particular interest in Schur’s Theorem has been enlightening and that the reader feels an appreciation for these theorems and the means by which they have been proven. We also hope that it has become clear that although most of the results mentioned have been found, Ramsey theory remains an active area of research.
BIBLIOGRAPHY


http://mathoverflow.net/questions/33289 (version: 2012-02-04)


[35] Issai Schur. *Über die Kongruenz $x^m + y^m \equiv z^m$ (mod $p$)*, Jahresbericht der Deutschen Mathematiker-Vereinigung 25 (1917), 114–117.


[38] Tao, Terence. Short story on Arithmetic Ramsey Theory
http://www.math.ucla.edu/~tao/preprints/acnt.html
(version: 2012-10-01)


http://www-history.mcs.st-andrews.ac.uk/Biographies/Schur.html
Retrieved on January 10th 2012

http://www-history.mcs.st-andrews.ac.uk/Biographies/Ramsey.html
Retrieved on January 13th 2012

http://www.encyclopedia.com
Retrieved on January 11th 2012

[46] Notes from 2005 REU on Ramsey Theory, available at
With the help of Kameryn Williams, we wrote a loop in MAPLE that would find the moduli beneath the $n^4$ bound for which we have a solution to the equation $x^n + y^n \equiv z^n \pmod{p}$. Here is our MAPLE code. The output follows in blue.

**Loop to find when $x^n + y^n = z^n \pmod{p}$ has solutions**

```maple
> with(numtheory);
> pset := map(proc (x) options operator, arrow; end proc, [seq(i, i = 1 .. 10)]);
primes := {};
for n from 3 to 10 do
Bound := n^4 - 6n^3 + 13n^2 - 6n + 2;
for i to Bound do
primes := primes \cup \{nextprime(i)\}
end do;
for p in primes do
for x from 1 to p-1 do
for y from 1 to p-1 do
X := x^n mod p;
Y := y^n mod p;
```
\[ z := \text{simplify}((X + Y)^{\frac{1}{n}} \mod p); \]

\[ \text{if type}(z, \text{integer}), \text{and } z \neq 0 \text{ then} \]
\[ \text{pset}[n-1] := \{p\} \cup \text{pset}[n-1]; \]
\[ \text{break} \]
\[ \text{end if} \]
\[ \text{end if} \]
\[ \text{end do}; \]

\[ \text{if type}(z, \text{integer}) \text{ and } z \neq 0 \text{ then} \]
\[ \text{break} \]
\[ \text{end if} \]
\[ \text{end do}; \]

\#printf("Primes that work for n=%d \%a\n", n, pset[n-1]);
printf("Primes that don't work for n=%d \%a\n", nprimes \ pset[n-1]))
end do;

Primes that don’t work for n=3 \{2, 7, 13\}
Primes that don’t work for n=4 \{2, 3, 5, 13, 17, 41\}
Primes that don’t work for n=5 \{2, 11, 41, 71, 101\}
Primes that don’t work for n=6 \{2, 3, 5, 7, 13, 19, 43, 61, 97, 157, 277\}
Primes that don’t work for n=7 \{2, 29, 71, 113, 491\}
Primes that don’t work for n=8 \{2, 3, 5, 13, 17, 41, 113\}
Primes that don’t work for n=9 \{2, 7, 13, 19, 37, 73, 181, 523, 577\}
Primes that don’t work for n=10 \{2, 3, 5, 11, 31, 41, 71, 101, 281, 401, 1181\}