Ramsey Theory: From Finite to Infinite

J.J. Marshall
Department of Mathematics
Washington University in St. Louis

March 2011

ABSTRACT

In this paper we present a survey of some results in Ramsey Theory, ranging from classic theorems that form the basis of the field to recent results that apply combinatorial techniques to the theory of Banach Spaces.

First we will present classical Ramsey Theory, culminating with the full version of Ramsey's Theorem in the finite case. From there we will move to some extensions of Ramsey's result, presenting results regarding the induced monochromatic substructures of colorings of the natural numbers and of vector spaces. We will also consider in more detail the subject of partition regular equations, in which we present the idea of the chromatic number of a linear equation or a system of linear equations (i.e., the smallest number of colors needed to find a coloring of $\mathbb{N}$ without a monochromatic solution).

Next we will present a discussion of how the results in finite Ramsey Theory carry over (or fail to carry over) to the infinite case, examining finite and infinite colorings of countable and uncountable sets. Just as finite Ramsey Theory lends itself to results on discrete spaces, infinite Ramsey Theory produces many results on continuous spaces, and we will explore some of these on both $\mathbb{R}^n$ and general metric spaces.

Finally, we will present an application of Ramsey-theoretic combinatorial techniques to a class of Banach spaces. Somewhat surprisingly, this application was recently used to answer an open question of Banach in Hilbert space analysis.

To Greta and Aerys, for whom all of the following is trivial.

TABLE OF CONTENTS

1 Introduction

1.1 Motivation, History 5

1.2 Plan of Action 5

1.3 An Example 6

1.4 Preliminaries 8

2 Finite Ramsey Theory

2.1 Colorings 8

2.2 The Arrow Notation 10

2.3 Ramsey Numbers 11

2.4 Ramsey's Theorem, Finite Version 13

3 Further Results in Finite Ramsey Theory

3.1 Chromatic Numbers, Schur's Theorem, Rado's Theorem 17

3.2 Arithmetic Progressions and van der Waerden's Theorem 20

3.3 Proof of Rado's Theorem 24

3.4 Generalization to Systems of Equations 29

3.5 Additional Results 30

4 Infinite Ramsey Theory

4.1 Preliminaries 31

4.2 Some Graph Theory and the Infinite Ramsey Theorems 33

4.3 The Compactness Principle 40

4.4 Larger Cardinals 41

4.5 Infinite Colorings 47

5 Continuous Ramsey Theory

5.1 Two Examples in $\mathbb{R}^n$ 49

5.2 The Happy Ending Problem 51

5.3 Application to Topology 52

6 Ramsey Theory in Banach Spaces

6.1 Banach, Hilbert Spaces and Banach's Question 54

6.2 Connection to Ramsey Theory 55

6.3 Application to Banach's Conjecture 63

7 Odds and Ends

7.1 Semigroups 67

7.2 The Paris-Harrington Theorem 68

Acknowledgements 70

References 71

1 Introduction

1.1 Motivation, History

Results in Ramsey Theory stem from the idea that "complete disorder is impossible." Ramsey Theory can be described as including all results that impose structure. The idea is that if a mathematical "object" such as a set, graph or vector space is "large" enough, under any partition of the object there will be one part which maintains some structure.

We will explore many variations on this theme. Most of our discussion will surround the idea of a "coloring" of the subsets of a set. Specifically, we will color (i.e., partition) all subsets of a set X that contain exactly k elements and then attempt to find a large subset of X all of whose k-element subsets are the same color. Ramsey's Theorem–the beginning of the field–gives us the guarantee that such monochromatic sets exist, as long as our starting set X is sufficiently big.

Frank Plumpton Ramsey published the first finite and infinite results of this nature in 1930. Ramsey was a promising young mathematician, philosopher and economist, having studied under Bertrand Russell, Ludwig Wittgenstein and John Maynard Keynes, giants in their respective fields. He authored several landmark papers in economics that are still part of Keynesian economic theory. Ramsey died in 1930, the same year his paper was published, at the age of just 26 [HOF].

The paper in which we find Ramsey's Theorems was titled "On a Problem of Formal Logic," and as such, the results he presented and upon which Ramsey Theory was later built were used mainly as technical results for a lengthier discussion of first-order logic rather than standing alone for their relevance in combinatorial set theory. It was not until later in the century that Paul Erdős and some of his many colleagues seized these results and began to expound them into a true field of their own.

1.2 Plan of Action

What is fascinating about Ramsey Theory is that it extends in some fashion to almost every branch of mathematics; in the remainder of the paper we will try to illustrate this. The Ramsey Theorems themselves are simply results in set theory but they have a very natural and direct application to graph theory. Coincidentally, graph-theoretic techniques will be used in several proofs in our discussion of infinite Ramsey Theory.

In Section 2 we will present the classic buildup of the finite Ramsey's Theorems. Readers who have studied some combinatorial set theory (or graph theory) will find this section to be mostly a review. In Section 3 we will consider one of the main results in Ramsey Theory, van der Waerden's Theorem on monochromatic arithmetic progressions in the natural numbers. This represents an application of Ramsey Theory to number theory. We will also discuss the theorems of Schur and Rado on monochromatic solutions to linear equations; the proofs here rely heavily on number theory. When we extend to systems of equations we naturally enter the field of linear algebra, in which we get another of the classic Ramsey-theoretic results on colorings of the subspaces of vector spaces.

After we conclude our work with finite Ramsey results, in Section 4 we will move on to the infinite cases, starting with the infinite version of Ramsey's Theorem from 1930, which Ramsey proved in the countable case, and continuing on to an analysis of the uncountable cases as well. This will lead into a discussion of some Ramsey results on continuous spaces in Section 5, including some of geometric and topological interest.

In Section 6 we will present a recent result of W.T. Gowers, who has been awarded the Fields medal for his work linking combinatorics and analysis. Gowers uses a Ramseyian proof technique to prove a result about infinite dimensional vector spaces that can be used to answer a question in Hilbert space analysis that had been initially posed by Banach and had remained open for many decades.

Finally, in Section 7 we will wrap up the discussion by presenting an application of Ramsey Theory to algebra and a connection with mathematical logic. Our goal is to impress on the reader how useful and versatile the Ramsey theorems and techniques truly are, and to show how they pop up in almost every area of mathematics.

1.3 An Example

Discussions of Ramsey Theory traditionally begin with a specific example, which we have modified slightly:

Example 1.1 At the Happy Pines Country Club, members congregate in the pro shop, waiting to form groups of three with which to set out on the golf course. The rule for forming parties is this: any two of the three members must either be acquainted or be strangers. (We assume acquaintance is symmetric and these are the only two possibilities; i.e., it is a binary relation.) What is the minimum number of people needed in the shop to guarantee a suitable threesome?

Solution We first note that the required number is greater than 5, for if we have five people (named A, B, C, D, E), consider the following acquaintance schedule:

Person	Knows	Does Not Know
A	B,C	D,E
B	A,D	C,E
C	A,E	B,D
D	B,E	A,C
E	C,D	A,B

A thorough inspection reveals that if any three of the five individuals are chosen, there is one of the three who knows exactly one of the other two. So it is possible to have five people present yet no suitable party.

We claim that with 6 people present, there must be a suitable threesome that can set out on the course. The argument is classic in its simplicity, and similar arguments drive the brunt of finite Ramsey Theory.

Say there are 6 people, namely A, B, C, D, E, F. Consider person A. There are five others, so he must either be acquainted with three of the five, or not be acquainted with three of the five (for $2 + 2 < 5$). Without loss of generality, we can assume that A knows B, C and D. Now, if any pair among B, C, D is acquainted, then that pair along with A forms the desired threesome (all of whom know one another). If no pair among B, C, D is acquainted, then B, C, D are mutual strangers and can therefore form a party. □

The same question can be asked when the parties are made up of 4 people rather than 3, and then for 5, etc. The minimum number required to guarantee the existence of a suitable foursome is 18; for any $n \geq 5$, the exact answer is not known. Similarly, the question can be extended to situations in which there are more than two possible ways to describe the level of acquaintance between two individuals. If there are three such ways–say, the two are strangers, they hate each other or they like each other–and a party of three must be formed in a similar way as above (with every pair among the party sharing the same relation) , it can be shown that with 17 people, it is certain that a suitable threesome exists. (We will prove this later.) With more than 3 "levels," the exact answer is, again, unknown. The fact that the exact answer is known for so few cases is a testament to the algorithmic complexity of the question, and of Ramsey Theory in general.

1.4 Preliminaries

The example above serves nicely as a warmup for more general Ramsey Theory, a setting in which the question will be sufficiently formalized. Before we embark, however, we need to agree on some notation.

For the purposes of this paper let the natural numbers be exactly the positive integers. That is, $\mathbb{N} = \{1, 2, \ldots\}$. If $n \in \mathbb{N}$, then we define $[n]$ to be the set consisting of the first n natural numbers. More explicitly, $[n] \equiv \{1, 2, \ldots, n\}$. For any set X (finite or infinite), we will use $|X|$ to denote the cardinality of X.

Definition 1.2 If X is a finite set, then $[X]_k$ is defined to be the set of all subsets of X containing exactly k elements. That is,

$$[X]_k \equiv \{Y \subseteq X : |Y| = k\}.$$

We will often refer to the elements of $[X]_k$ as the *k-subsets* of X. When $X = [n]$ for some $n \in \mathbb{N}$ (as defined above), we will abuse the notation slightly and write $[n]_k$ to denote the collection of k-subsets of $[n]$, instead of the more tedious $[[n]]_k$. (Though set theory enthusiasts will note that $n = \{0, 1, \ldots, n-1\}$, so it is not too much of a stretch.)

2 Finite Ramsey Theory

2.1 Colorings

We begin with a discussion of colorings of a finite set.

Definition 2.1 Let X be a finite set and let $r \in \mathbb{N}$. An *r-coloring* of X is a function

$$\chi : X \to [r].$$

If $Y \subseteq X$, we say Y is *monochromatic* (or *monochromatic under* χ or t-*monochromatic*) if $\chi(y) = t$ for all $y \in Y$, which is to say that $\chi[Y] = \{t\}$. In other words, Y is monochromatic if χ is constant on Y.

Our main goal throughout this section is to find "large" monochromatic subsets. Eventually we will be more concerned with coloring the k-subsets of X than with coloring the elements of X themselves, but we start with one of the most basic theorems in all of mathematics.

Theorem 2.2 (Pigeonhole Principle) Let $|X| = n$. Then if $m < n$ and χ is an m-coloring of X, there exists a monochromatic subset $Y \subseteq X$ with $|Y| \geq 2$.

Proof Suppose, to the contrary, that no such subset Y exists. For each $1 \leq i \leq m$, let M_i be the set of points of color i, or the set $\{x \in X : \chi(x) = i\}$. Then we can write $X = \displaystyle\bigcup_{i=1}^{m} M_i$ and thus $|X| = \displaystyle\sum_{i=1}^{m} |M_i|$ since the sets M_i are pairwise disjoint. But by assumption, $|M_i| \leq 1$ for each i, so $|X| \leq m < n$, a contradiction. $\square$

The Pigeonhole Principle, so presented, seems almost too trivial to bother with. But it is important for two reasons. When presented in the context of an axiomatic set-theoretic construction of the natural numbers, it can be understood as a crucial result linking the mathematical formality of numbers and the natural notion of quantity. See [END]. This is not a paper on axiomatic set theory, but we nonetheless present it here with deserved fanfare, for Ramsey Theory as a whole is often seen as its generalization.

It is important to note that colorings of sets are "preserved" under bijections. In this sense, if there is a bijection between two sets A and B, then any coloring of A is equivalent to a coloring of B. We can think of this equivalence as a "weak" isomorphism between colorings of A and colorings of B, where it is only important that $|A| = |B|$. Because of this we can, in proving the subsequent theorems, restrict our attention to colorings of $[n]_k$, but without hesitance generalize, when needed, to colorings of $[A]_k$ for any set A with $|A| = n$.

2.2 The Arrow Notation

Theorem 2.2 gives us a precise solution to the question of how big a set must be in order to guarantee that two elements share a color under any r-coloring. We want to extend this notion and find the (minimal) size for a set X that guarantees, under any r-coloring of $[X]_k$, that there is a subset $Y \subseteq X$ all of whose k-subsets are the same color.

Before we continue, we introduce some additional notation. If $n, k, a_1, \ldots, a_r$ are natural numbers we will write

$$n \longrightarrow (a_1, a_2, \ldots, a_r)_k$$

as shorthand for the following condition: if the elements of $[n]_k$ (that is, the k-subsets of $[n]$) are r-colored, then, for some $1 \leq i \leq r$ there is a subset $Y \subseteq [n]$ such that $|Y| = a_i$ and all the k-subsets of Y are colored i. Or, in other words, $[Y]_k$ is i-monochromatic. If $a_1 = a_2 = \ldots = a_r = a$, we will instead write

$$n \longrightarrow (a)^r_k.$$

By convention, when either k or r is equal to 2, it is usually omitted. The negation of such a statement will unsurprisingly be written as

$$n \nrightarrow (a_1, a_2, \ldots, a_r)_k$$

(or some conventional equivalent). The reader should confirm that to say $n \nrightarrow (a_1, a_2, \ldots, a_r)_k$ is to say that there exists an r-coloring of $[n]_k$ such that for each $1 \leq i \leq r$, there is no i-monochromatic subset $[Y]_k$ with $|Y| = a_i$.

With this new notation in hand, we can restate the Pigeonhole Principle as follows: if $m < n$, then $n \longrightarrow (2)^m_1$.

In introducing the notation, we noted the equivalence between saying that the k-subsets of Y are all of the same color and saying that $[Y]_k$ is monochromatic. The following definition melds these together:

Definition 2.3 Let $\chi : [X]_k \to [r]$ be a coloring. If either of the above conditions (and thus both) holds of a subset $Y \subseteq X$, we say Y is *homogeneous* under the given coloring. We may say Y is *t-homogeneous* if $\chi[[Y]_k] = \{t\}$.

It is important to note that in the case where the numbers a_i are distinct, the size of the "desired" homogeneous subset is dependent on color. But when they are not distinct, the color becomes less important. The extreme is when $a_1 = \ldots = a_r = a$; in this case we simply seek a homogeneous subset of size a and we do not care for which color this subset is homogeneous..

The following lemma states some obvious, yet useful facts.

Lemma 2.4 Suppose that $n \longrightarrow (a_1, \ldots, a_r)_k$. The following hold.

a) If $b_i \leq a_i$ for $1 \leq i \leq r$, then $n \longrightarrow (b_1, \ldots, b_r)_k$.

b) If $N \geq n$, then $N \longrightarrow (a_1, \ldots, a_r)_k$.

c) Let $\sigma \in S_r$ be a permutation. Then $n \longrightarrow (a_{\sigma(1)}, \ldots, a_{\sigma(r)})_k$.

d) $n \longrightarrow (a_1, \ldots, a_r, k)_k$.

Proof

a) Under any r-coloring of $[n]_k$ there is a subset $Y \subseteq [n]$ with $|Y| = a_i$ and Y homogeneous for color i. Choose b_i points from Y and call this set Y'. Every k-subset of Y' is a k-subset of Y, so Y' is homogeneous for color i.

b) A coloring of $[N]_k$ induces a coloring of $[n]_k$, namely the "restriction" coloring $\chi|_{[n]_k}$. Look at the latter coloring and we at once have the existence of the desired homogeneous set.

c) Suppose $n \nrightarrow (a_{\sigma(1)}, \ldots, a_{\sigma(r)})_k$. Then there is a coloring $\chi : [n]_k \to [r]$ such that for each $1 \leq i \leq r$, there is no i-homogeneous subset Y with $|Y| = a_{\sigma(i)}$. We form a "re-coloring" χ^* such that

$$\chi^*(\{x_1, \ldots, x_k\}) = \sigma^{-1}(\chi(\{x_1, \ldots, x_k\})).$$

That is, everything colored j under χ is colored $\sigma^{-1}(j)$ under χ^*. A set is homogeneous under χ iff it is homogeneous under χ^*, so we have that for each i, there is no i-homogeneous subset A with $|A| = a_{\sigma^{-1}(\sigma(i))} = a_i$, so $n \nrightarrow (a_1, \ldots, a_r)_k$.

d) If an $(r+1)$-coloring of $[n]_k$ "uses" the color $r+1$, which is to say that $\chi^{-1}[\{r+1\}] \neq \emptyset$, then there is a subset of size k with that color, and this is a homogeneous set of size k. Otherwise, it is an r-coloring, which by hypothesis produces a homogeneous subset of the requisite size. $\square$

2.3 Ramsey Numbers

Suppose that for some $n \in \mathbb{N}$ we have $n \longrightarrow (a_1, \ldots, a_r)_k$. Since there are only a finite number of r-colorings of $[n-1]_k$, we can (in theory) check to see if $(n-1) \longrightarrow (a_1, \ldots, a_r)_k$. If it does, we can "check" for $n-2$, and so on. Eventually we reach a number j such that $j \longrightarrow (a_1, \ldots, a_r)_k$, but $(j-1) \not\rightarrow (a_1, \ldots, a_r)_k$.

Definition 2.5 The *Ramsey number* $R_k(a_1, \ldots, a_r)$ is the smallest positive integer n such that $n \longrightarrow (a_1, \ldots, a_r)_k$. That is,

$$R_k(a_1, \ldots, a_r) \equiv \min\{n \in \mathbb{N} : n \longrightarrow (a_1, \ldots, a_r)_k\}.$$

In the case where $a_1 \ldots = a_r = a$, we will often write $R_k(a; r)$. As before, when $k = 2$, we omit it and write $R(a_1, \ldots, a_r)$.

Thus, in Example 1.1 we showed both that $5 \not\rightarrow (3,3)$ and that $6 \longrightarrow (3,3)$. So we, in fact, proved that $R_2(3,3) = R(3,3) = R(3;2) = 6$.

The Ramsey numbers for $k = 2$ (and specifically when $r = 2$ as well) are of greatest interest to many, because they have a very simple and natural interpretation in graph theory. Recall that a graph G is an ordered pair (V, E) where V is any set (but typically assumed for convenience to be $[n]$ for some $n \in \mathbb{N}$) and $E \subseteq [V]_2$. The points in V are the *vertices*, and the elements of E (which are 2-subsets of V) are the *edges*. When $E = [V]_2$, G is called the *complete graph on* $|V|$ *vertices*. Where $|V| = n$, this graph is denoted K_n.

If we take a subset $V' \subseteq V$, the *induced subgraph* on V' is the graph with vertex set V' and edge set

$$E' \equiv \{\{x, y\} \in E : x, y \in V'\}.$$

If the induced subgraph on $V' \subseteq V$ has an empty edge set, then V' is an *independent set* of vertices in V. We denote the independent set of size n as I_n.

Hence, when $k = 2$, the Ramsey number $R(a, b)$ is the minimum number n of vertices needed so that if the edges of the complete graph K_n are 2-colored (say with red and blue), there is guaranteed either to be a subgraph K_a all of whose edges are red or a subgraph K_b with edges all blue. Of course, we need not only consider complete graphs here where the edges are colored. If we take any graph on a vertex set V of size n, we can think of the edges in E as red and the "possible" edges that are not in E (i.e., $[V]_2 - E$) as blue. This is a structure-preserving bijection between the set of all graphs on n vertices and the set of all 2-colorings of the graph K_n. Thus we have the following result.

Proposition 2.6 The minimum number n of vertices needed to guarantee that every graph on n vertices has either K_a or I_b as an induced subgraph is exactly the Ramsey number $R(a, b)$.

We also mentioned that very few of the Ramsey numbers have been calculated precisely. For $r = k = 2$ (and ignoring the simple cases where either $a_1 \leq 2$ or $a_2 \leq 2$, the exact values are known only for $R(3, k) = R(k, 3)$ with $k \leq 9$, for $R(4, 4)$, and for $R(5, 4) = R(4, 5)$. For a full account of the known bounds for other Ramsey numbers, see [RAD].

When $r \geq 3$, there is only one non-trivial Ramsey number known.

Example 2.7 $R(3, 3, 3) = 17$.

As in Example 1.1, to show this to be true we need to demonstrate both that $17 \longrightarrow (3, 3, 3)$ and that $16 \not\rightarrow (3, 3, 3)$. For the latter fact it is necessary to provide a 3-coloring of K_{16} which has no monochromatic K_3 subgraph (K_3 is simply a triangle). Two such colorings exist; for the details we refer the reader to [GG].

To show that $17 \longrightarrow (3, 3, 3)$ we use a similar argument to the one employed in Example 1.1. Fix a vertex $x \in [17]$ and consider any 3-coloring (red, blue, green) of K_{17}. The vertex x shares edges with 16 others, so by the Pigeonhole Principle at least 6 of these must be of the same color ($5+5+5 < 15$). Without loss of generality, say that this color is red and let $x_1, \ldots, x_6$ be the 6 vertices that share a red edge with x.

If the edge $\{x_i, x_j\}$ is red for some i, j with $1 \leq i < j \leq 6$, then $\{x, x_i, x_j\}$ is a red triangle, and we are done. Otherwise, all the edges among $x_1, \ldots, x_6$ are blue or green, so we have an induced 2-coloring of a K_6 subgraph, and since $6 \longrightarrow (3, 3)$ (indeed, $R(3, 3) = 6$), there is a monochromatic triangle among those edges. □

In the preceding example it seemed that the Ramsey number $R(3; 3)$ was a property inherent to graphs. But this is not true–the Ramsey numbers are properties of *sets*. Ramsey Theory is not embedded in graph theory. However, for $k = 2$ it is easier to "see" the results when we consider their application to graphs.

2.4 Ramsey's Theorem, Finite Version

We have defined Ramsey numbers generally but have not proven that they exist. We will rectify this problem now. For two reasons, we will treat the case $k = 2$ separately. First, this case has a special graph-theoretic interpretation, as we have discussed above. Also, it gives us an opportunity to present both methods of proving the finite version(s) of Ramsey's theorem. The proofs both explicitly present a "large enough" N and therefore are of use when trying to provide upper bounds on the Ramsey numbers. Both techniques utilize induction; one inducts over the "smaller" Ramsey numbers of a fixed dimension, and the other inducts over the Ramsey numbers of smaller dimensions.

Theorem 2.8 Let $a_1, \ldots, a_r \in \mathbb{N}$. The Ramsey number $R(a_1, \ldots, a_r)$ exists.

Proof The proof proceeds by induction on the sum $a_1 + \ldots + a_r$. The base case is where $a_1 = \ldots = a_r = 2$. It is clear that $R(2; r) = 2$; with 2 elements there is a 2-set which is homogeneous for its color.

For the induction step, we assume the existence of $R(a_1, \ldots, a_j - 1, \ldots, a_r)$ for each $1 \leq j \leq r$. Let $N = 2 + \sum_{j=1}^{r} (R(a_1, \ldots, a_j - 1, \ldots, a_r) - 1)$.

Let $\chi : [N]_2 \to [r]$ be any r-coloring. We fix $x \in [N]$ and partition $[N] - \{x\}$ as follows: for each $1 \leq j \leq r$, define

$$S_j \equiv \{y \in [N] : \chi(\{x, y\}) = j\}.$$

We claim that for some k, $1 \leq i \leq r$, we have

$$|S_k| \geq R(a_1, \ldots, a_i - 1, \ldots, a_r).$$

If not, we have $[N] = \{x\} \cup S_1 \cup \ldots \cup S_r$, so

$$\begin{aligned}
|[N]| = N &= 1 + |S_1| + \ldots + |S_r| \\
&\leq 1 + (R(a_1 - 1, a_2, \ldots, a_r) - 1) + \ldots + (R(a_1, \ldots, a_{r-1}, a_r - 1) - 1) \\
&= 1 + \sum_{j=1}^{r} (R(a_1, \ldots, a_j - 1, \ldots, a_r) - 1)
\end{aligned}$$

which contradicts our choice of N. (Here we have tacitly applied the Pigeonhole Principle.)

For this i, we have a subset $T \subseteq S_i$ that is either homogeneous for color $m \neq i$ with $|T| = a_m$ or for color i with $|T| = a_i - 1$. In the former case, we are done. Otherwise we note that for each $y \in T$ we have $\chi(\{x, y\}) = i$ so $T \cup \{x\}$ is i-homogeneous with $|T \cup \{x\}| = |T| + 1 = a_k$. □

As promised, we now present and prove the more general version of Theorem 2.8, which handles the case $k > 2$ as well.

Theorem 2.9 (Finite Ramsey's Theorem) For each $k \geq 1$ and for any $a_1, \ldots, a_r \geq 1$, the Ramsey number $R_k(a_1, \ldots, a_r)$ exists.

Proof We proceed by induction on k.

The base case is where $k = 1$. If we choose $n = 1 + \sum_{i=1}^{r}(a_i - 1)$, it is clear that $n \longrightarrow (a)_1^r$. If not, we would have a coloring $\chi : [n] \to [r]$ such that each of the color classes, $1 \leq i \leq r$, had size at most $a_i - 1$, implying that $n \leq \sum_{i=1}^{r}(a_i - 1)$. (In fact, the n we have chosen is precisely the Ramsey number $R_1(a_1, \ldots, a_r)$. Where $a_1 = \ldots = a_r = 2$, we simply have the Pigeonhole Principle.)

For the induction step, we assume that $R_{k-1}(a_1, \ldots, a_r)$ exists. It suffices (by Lemma 2.4) to show the existence of $R_k(a; r)$, where $a \equiv \max\{a_1, \ldots, a_r\}$, since any upper bound on the Ramsey number proves its existence. Let $\zeta = R_{k-1}(a; r)$ and let $n = 2r^{\alpha}$, where $\alpha = \sum_{i=k-1}^{\zeta-1} \binom{i+1}{k-1}$. (The choice of such n will become apparent later.) We fix a coloring $\chi : [n]_k \to [r]$, and define inductively a sequence $x_1, x_2, \ldots, x_\zeta$ as follows. Let $x_1, x_2, \ldots, x_{k-2} \in [n]$ be arbitrary (but distinct) points and let $S_{k-2} = [n] - \{x_1, \ldots, x_{k-2}\}$. Then, we define the remaining x_i for $k - 1 \leq i \leq \zeta$ and S_i for $k - 1 \leq i \leq \zeta - 1$:

- Once S_i has been chosen, let $x_{i+1} \in S_i$ be arbitrary.
- After selecting x_{i+1}, we partition $S_i - \{x_{i+1}\}$ into equivalence classes with the relation R, where yRz iff for each $T \subseteq \{x_1, \ldots, x_{i+1}\}$ with $|T| = k - 1$ it holds that $\chi(T \cup \{y\}) = \chi(T \cup \{z\})$. That is, yRz iff y and z are "identical" with regard to color when forming k-sized sets with the members of the sequence we are defining.
- Let S_{i+1} be the largest of the equivalence classes.

For now we will take it for granted that we can, indeed, define our sequence all the way to x_ζ. (See below the proof for why this holds.) Having defined the sequence $\mathbf{x} = \{x_1, \dots, x_\zeta\}$, we next define a coloring $\chi^* : [\mathbf{x}]_{k-1} \to [r]$ as follows: if $\{x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}\} \subseteq \mathbf{x}$, with $i_1 < i_2 < \dots < i_{k-1}$, then

$$\chi^*(\{x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}\}) = \chi(\{x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_\beta\})$$

where $i_{k-1} < \beta \leq j$. Of course, there are a number of issues to deal with. First, we must check that $\chi*$ is well-defined; that is, we need to show that the choice of β does not affect the "color." In other words, we need that

$$\chi(\{x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_\beta\}) = \chi(\{x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_\gamma\})$$

for any choices of $i_{k-1} < \beta, \gamma \leq \zeta$. To see this, note that $x_\beta \in S_{\beta-1}$ and $x_\gamma \in S_{\gamma-1}$, and both $S_{\beta-1}, S_{\gamma-1}$ are subsets of $S_{i_{k-1}}$, which was defined so that all of its elements "agree" under χ when joined with $(k-1)$-subsets of $\{x_1, \dots, x_{i_{k-1}}\}$.

The other issue is when $i_{k-1} = \zeta$; in this case we let $\chi^*(\{x_{i_1}, \dots, x_{i_{k-1}}\})$ be arbitrary.

By our choice of ζ, the coloring χ^* induces a homogeneous set $\{y_1, \dots, y_a\} \subseteq \mathbf{x}$, all of whose $(k-1)$-subsets are the same color (say, color q). Choose any $\{y_{m_1}, \dots, y_{m_k}\} \subseteq \{y_1, \dots, y_a\}$ with $1 \leq m_1 < \dots < m_k \leq a$ and

$$\chi(\{y_{m_1}, \dots, y_{m_{k-1}}, y_{m_k}\}) = \chi^*(\{y_{m_1}, \dots, y_{m_{k-1}}\}) = q,$$

so $\{y_1, \dots, y_a\}$ is the required homogeneous set. □

We did not demonstrate in the proof why our process of defining the sequence $\{x_1, \dots, x_\zeta\}$ does not terminate after selecting some x_ξ for $\xi < \zeta$. We claim, though, that the process we have defined will, indeed, allow us to choose ζ elements for the sequence. The equivalence classes we used are defined by $(k-1)$-sized subsets of $\{x_1, \dots, x_{i+1}\}$, so it follows that there are (at most) $r^{\binom{i+1}{k-1}}$ such classes. Thus, when we select S_{i+1} to be the largest class, we can be assured that

$$|S_{i+1}| \geq \frac{|S_i - \{x_{i+1}\}|}{r^{\binom{i+1}{k-1}}} = \frac{|S_i| - 1}{r^{\binom{i+1}{k-1}}}.$$

We omit the details, but this is precisely why we made our choice of $n = 2r^\alpha$. We eventually get to the stage where we have the set $S_{\zeta-1}$, and repeated applications of the inequality above guarantee that $|S_{\zeta-1}| \geq 2-\sigma$, where σ is some partial geometric series less than 1, so $|S_{\zeta-1}| \geq 1$ and we can select x_ζ.

A remark on the technique used in this proof is in order, as it is typical throughout Ramsey Theory. With the result in hand for the case $k-1$, to show it holds for k we construct a coloring of the k-subsets of some set such that the color depends only on $k-1$ of the elements in any k-subset. Then we use the inductive hypothesis to obtain a homogeneous set in the $k-1$ case which transfers immediately to the case k.

We could have proven Theorem 2.9 by using a recursive technique similar to the one used in the proof of Theorem 2.8. As noted above, though, this type of strategy is a very important lesson of Ramsey Theory, and in particular it will help us prove the infinite version of this theorem in Section 4. Similarly, we could have used this technique to prove Theorem 2.8, but we would be remiss if we did not present both methods.

Ramsey's Theorem provides a powerful tool in confirming the existence of "nice" substructures of sufficiently large mathematical structures. If a coloring can be defined on a set based on some property, we automatically know there is a *finite* N such that, if the size of the set is at least N, there is guaranteed to be a monochromatic substructure–and, if we have defined the coloring in a clever manner, this monochromaticity will translate into some meaningful property. As long as i) the size of the desired substructure is known and ii) we have freedom to choose arbitrarily large (though finite) universes to work in, we can produce the substructure we want. The proof we have employed for Theorem 2.9 itself is a perfect paradigm for this strategy.

In the next section, we explore some of the major extensions and applications of the finite version of Ramsey's Theorem.

3 Further Results in Finite Ramsey Theory

3.1 Chromatic Numbers, Schur's Theorem, Rado's Theorem

The idea of the Ramsey theorems that we encountered in the previous section was that if we start with a large enough structure, for any partition of that structure there is a substructure of the original with some desired property that falls entirely within one part of the partition. In this section we will describe how this template has been used to provide various results on discrete spaces. We begin with the subject of chromatic numbers of equations.

Definition 3.1 Let $c_1, \ldots, c_n \in \mathbb{Z}$. The *chromatic number* of the equation

$$c_1 x_1 + \ldots c_n x_n = 0,$$

denoted $\nu(c_1, \ldots, c_n)$, is the smallest positive integer r such that there exists a coloring $\chi : \mathbb{N} \to [r]$ with no monochromatic solution.

Note that we are coloring the natural numbers (not the integers) and searching there for solutions to the equation. (Also note that we have gone back to coloring elements of a set rather than coloring its subsets.) Obviously, if for each $1 \leq i \leq n$ we have $c_i \geq 1$, then $\nu(c_1, \ldots, c_n) = 1$. It is also clear that the order of the numbers c_i does not affect the chromatic number of the collection, so that we may be better off writing $\nu(\{c_1, \ldots, c_n\})$. We will not do this, but we shall keep in mind that for any permutation $\sigma \in S_n$ we have $\nu(c_1, \ldots, c_n) = \nu(c_{\sigma(1)}, \ldots, c_{\sigma(n)}) = \nu(-c_1, \ldots, -c_n)$.

Example 3.2 We try to find $\nu(1, -2)$. We want a minimal coloring χ (in terms of the fewest number of colors needed) for $\mathbb{N}$ such that if x, y satisfy the equation

$$x - 2y = 0$$

then $\chi(x) \neq \chi(y)$.

It is clear that $\nu(1, -2) > 1$, since a chromatic number of an equation is 1 if and only if there are no solutions to that equation in $\mathbb{N}$; we claim that $\nu(1, -2) = 2$. We define a coloring as follows. For each $n \in \mathbb{N}$, we can write n in terms of primes $p_1, \ldots, p_m$ as

$$n = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}$$

where each $r_i \geq 1$. If each $p_i > 2$, we can extend this to

$$n = 2^{r_0} p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}$$

where $r_0 \geq 0$. Otherwise let r_0 be the power of 2 in the prime factorization of n and define

$$\chi(n) = \begin{cases} \text{red} & \text{if } r_0 \text{ even} \\ \text{blue} & \text{if } r_0 \text{ odd.} \end{cases}$$

Then if $x - 2y = 0$ we have $x = 2y$, so $\chi(x) \neq \chi(y)$. Thus we have shown that $\nu(1, -2) = 2$. □

Note that the same technique can be used to show that $\nu(p, -q) = 2$ whenever $\gcd(p, q) = 1$.

To this point we have not discussed the possibility that there might be $c_1, \ldots, c_n$ such that no finite chromatic number $\nu(c_1, \ldots, c_n)$ exists. This would be the case if any only if for any positive r, every r-coloring of $\mathbb{N}$ had a monochromatic solution.

Definition 3.3 Let $c_1, \ldots, c_n \in \mathbb{Z}$. If $\nu(c_1, \ldots, c_n)$ does not exist then we say the equation

$$c_1x_1 + \ldots + c_nx_n = 0$$

is *partition regular.*

Note that if $c_1 = \ldots = c_n = 0$ we get a trivial partition regular equation. But beyond that, it is not clear how one would demonstrate that a particular equation is partition regular; indeed, it is not obvious that any nontrivial such equations exist. However, Ramsey's theorem gives us the machinery to prove their existence.

Theorem 3.4 (Schur's Theorem) Nontrivial partition regular equations exist. In particular, the equation

$$x + y = z$$

is partition regular.

Proof Let $r \in \mathbb{N}$ and let $\chi : \mathbb{N} \to [r]$ be any coloring.

We choose $N \geq R(3; r)$ (which exists by Theorem 2.8), and define a coloring $\chi^* : [N] \to [r]$ by

$$\chi^*(\{i, j\}) = \chi(|i - j|)$$

for any $1 \leq i, j \leq N$ with $i \neq j$.

By our choice of N, there is a monochromatic triangle under χ^* (i.e., there are $i, j, k \in [N]$ with $i > j > k$ and such that $\chi^*(\{i, j\}) = \chi^*(\{i, k\}) = \chi^*(\{j, k\})$). We have, therefore, that

$$\chi(i - j) = \chi(i - k) = \chi(j - k).$$

Let $x = i - j$, let $y = j - k$, and let $z = i - k$, so $\chi(x) = \chi(y) = \chi(z)$. Further,

$$\begin{aligned} x + y &= (i - j) + (j - k) \\ &= i - k \\ &= z. \end{aligned}$$

□

It is worthwhile to note that this theorem uses Ramsey's Theorem in a very direct way, relying on the existence of $R(3; r)$. Because of this, we didn't actually need all of $\mathbb{N}$ in order to find our monochromatic solution. We could have rephrased our statement of the theorem as "for any r, there exists an N such that any coloring $\chi : [N] \to [r]$ yields a monochromatic solution to $x + y = z$.

Our goal is to work toward a powerful result which will allow us to classify completely the partition regular equations. We state the theorem now, but it will take a good deal of work, including a lengthy (yet interesting) digression, to get us there.

Theorem 3.5 (Rado's Theorem) The equation

$$c_1 x_1 + \ldots + c_n x_n = 0$$

is partition regular if and only if there is a nonempty subset $C \subseteq \{c_1, \ldots, c_n\}$ such that

$$\sum_{c \in C} c = 0.$$

We are not ready to prove Rado's Theorem yet. After the next discussion, we will be.

3.2 Arithmetic Progressions and van der Waerden's Theorem

In the context of Ramsey theory, the main idea of this section is that if we color $[N]$ (for large enough N) then we can find an arbitrarily long arithmetic progression that is monochromatic. Let us make this more precise.

Definition 3.6 An *arithmetic progression* of length l is a set

$$\{a, a+d, a+2d, \ldots, a+(l-1)d\} \subseteq \mathbb{N}$$

for some $a, d \in \mathbb{N}$.

If χ is a coloring of the natural numbers, a monochromatic arithmetic progression is one contained entirely in the set $\chi^{-1}[\{i\}]$ for some color i. The following is one of the classic theorems of Ramsey theory.

Theorem 3.7 (van der Waerden's Theorem) Let $\chi : \mathbb{N} \to [2]$ be a coloring, and let $l \in \mathbb{N}$. Then there is a monochromatic arithmetic progression of length l.

Just as with Ramsey's Theorem, it is natural to extend results on 2-colorings to hold also for r-colorings for any $r \in \mathbb{N}$. Similarly, we saw above in Theorem 3.4 that although the general result was for coloring *all* of $\mathbb{N}$, we really only needed a large (but finite) N to find the desired monochromatic structure. We modify van der Waerden's Theorem in the analogous way.

Theorem 3.8 Let $l, r \in \mathbb{N}$. There exists $N \in \mathbb{N}$ such that if $\chi : [N] \to [r]$ is a coloring, there is a monochromatic arithmetic progression of length l.

It is surprising, but if we (once again) strengthen Theorem 3.8 to provide an even *more* general result, we wind up with a theorem that is easier to prove, and one from which Theorem 3.8 immediately follows.

Let $l, m \in \mathbb{N}$. We define the *l-equivalence classes* of the set $\{0, 1, \ldots, l\}^m$ (the Cartesian product of $\{0, 1, \ldots, l\}$ with itself m times) to be those formed by the relation R where $(x_1, \ldots, x_m)R(y_1, \ldots, y_m)$ iff the two have the same number of leading l entries and both have no l appearing anywhere else. More formally, if $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ then $\mathbf{x}R\mathbf{y}$ iff for each $1 \leq i \leq m, x_i, y_i \neq l$ or if there is some $1 \leq k \leq m$ such that $x_i = y_i = l$ for $i \leq k$ and $x_i, y_i \neq l$ for $k < i \leq m$.

For positive integers l, m, let the "proposition" $\Sigma(l, m)$ be the following:

> For any r, there exists $N = N(l, m, r)$ for which it holds that if $\chi : [N] \to [r]$ is a coloring, then there exist $a, d_1, \ldots, d_m \in \mathbb{N}$ such that
>
> $$\chi(a + \sum_{i=1}^{m} x_i d_i) = \chi(a + \sum_{i=1}^{m} y_i d_i)$$

if $\mathbf{x}R\mathbf{y}$. That is, $\chi(a + \sum_{i=1}^{m} x_i d_i)$ is constant on members $\mathbf{x}$ of an l-equivalence class.

First note that $\Sigma(1,1)$ is trivially true. It requires only that there exist a, d with $\chi(a + xd) = \chi(a + yd)$ if $\{x\}, \{y\}$ are in the same 1-equivalence class. This happens if $x = y = 0$ or $x = y = 1$, so any a and d will do.

Lemma 3.10 If $\Sigma(l,k)$ holds for all $1 \le k \le m$, then $\Sigma(l, m+1)$ holds.

Proof Let $r \in \mathbb{N}$ and set $M = N(l,m,r)$ and $M' = N(l,1,r^M)$. Our claim is that $N(l,m+1,r) = MM'$, so let $\chi : [MM'] \to [r]$ be arbitrary.

We define a coloring $\chi^* : [M'] \to [r^M]$ where the vertices $\{1, \ldots, M'\}$ can be thought of as "blocks" of length M. That is, we divide our MM' points into M' blocks of M points each, and color each block based on the color of each of its M points under χ. The actual colors are not important; what is important is that two blocks i, j are the same color under χ^* iff $\chi(iM + q) = \chi(jM + q)$ for each $1 \le q \le M$. Each block has M points, and each is one of r colors, so r^M is precisely the number of colors we need to use in order to achieve this property.

By our choice of M', there exist constants a', d' such that

$$\chi^*(a' + xd') = \chi^*(a' + yd')$$

if $\{x\}$ and $\{y\}$ are l-equivalent–meaning if both are l (which is a trivial case) or if both $x, y \in \{0, 1, \ldots, l-1\}$. (That is, there is a monochromatic arithmetic progression of "blocks" under χ^*; see the proof of Theorem 3.8 below.) We then look at the a'th block of M points. By our choice of M (and a translation of $[M]$ to $\{a'M + 1, \ldots, a'M + M\}$) we have constants $a, d_1, \ldots, d_m$ such that $\chi(a + \sum_{i=1}^{m} x_i d_i)$ is constant on members $\mathbf{x}$ of an l-equivalence class (and all possible sums fall within that block).

We need constants $a^*, d_1^*, \ldots, d_m^*, d_{m+1}^*$ such that $\chi(a^* + \sum_{i=1}^{m+1} x_i d_i^*)$ is constant on l-equivalence classes of sequences of length $(m+1)$. We take $a^* = a'M + a$, $d_i^* = d_i$ for $1 \le i \le m$ and set $d_{m+1}^* = d'M$. If $\mathbf{x}, \mathbf{y} \in \{0, 1, \ldots, l\}^{m+1}$ are l-equivalent then we get the desired monochromaticity for the first m coordinates from within the block, and the last coordinate is handled by the monochromatic progression of identical blocks. So $\Sigma(l, m+1)$ holds. $\square$

Lemma 3.11 If $\Sigma(l, m)$ holds for all $m \in \mathbb{N}$, then $\Sigma(l+1, 1)$ holds.

Proof Let $r \in \mathbb{N}$. We claim that $N(l+1, 1, r) = N(l, r, r)$, so let

$$\chi : [N(l, r, r)] \to [r]$$

be arbitrary. Therefore, by assumption, there exist constants $a, d_1, \ldots, d_r$ such that $\chi(a + \sum_{i=1}^{r} x_i d_i)$ is constant on l-equivalence classes.

Next, we examine the colors of the points

$$a, a + ld_1, a + l(d_1 + d_2), \ldots, a + l(d_1 + \ldots + d_r).$$

Each is one of r colors and there are $r+1$ points, so by the Pigeonhole Principle two of these share the same color. That is, there exist $j, k \in \{0, 1, \ldots, r\}$ with $j < k$ such that

$$\chi(a + l\sum_{i=1}^{j} d_i) = \chi(a + l\sum_{i=1}^{k} d_i).$$

We take $a' = a + l\sum_{i=1}^{j} d_i$ and $d' = \sum_{i=j+1}^{k} d_i$. Then the $(l+1)$-equivalence classes are $\{l+1\}$ and $\{0, 1, \ldots, l\}$. For the first class, the fact that $\chi(a' + xd')$ is constant on the class is trivial. In the second case, we have

$$\chi(a' + xd') = \chi((a + l\sum_{i=1}^{j} d_i) + x(\sum_{i=j+1}^{k} d_i)).$$

That this is constant for $x \in \{0, 1, \ldots, l-1\}$ follows from the inductive hypothesis, that this is constant on $\{0, l\}$ follows from the application of the pigeonhole principle above. Thus it is constant on $\{0, \ldots, l\}$ and we have $\Sigma(l+1, 1)$. $\square$

Theorem 3.12 $\Sigma(l, m)$ holds for all l, m.

Proof We know $\Sigma(1,1)$ to be true. Lemmas 3.10 and 3.11 do the rest of the work by double induction. □

Now we have (the strengthened version of) van der Waerden's Theorem.

Proof (of Theorem 3.8) We show that Theorem 3.8 follows from $\Sigma(l,1)$.

$\Sigma(l,1)$ implies that for any r, there exists $N \in \mathbb{N}$ such that any coloring $\chi : [N] \to [r]$ satisfies

$$\chi(a + dx) = \chi(a + dy)$$

for some a, d and any $x, y \in \{0, 1, \ldots, l-1\}$. Thus the set

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

is monochromatic. It is an arithmetic progression of length l. □

Corollary 3.13 If $\mathbb{N}$ is r-colored, there is an arbitrarily long monochromatic arithmetic progression.

Proof Once the desired length l is chosen, restrict the coloring down from $\mathbb{N}$ to $N(l,1,r)$. □

3.3 Proof of Rado's Theorem

To prove Rado's Theorem (Theorem 2.5), we will need a small result which strengthens Theorem 3.8 mildly. Note that a consequence of Theorem 3.8 is that for any l, r there is a minimal $W = W(l,r)$ such that every r-coloring of $[W]$ yields a monochromatic arithmetic progression of length l. (Back in Section 2, in analogy with Ramsey numbers, we may have called this a *van der Waerden number*.)

Lemma 3.14 For any $l, m, r \in \mathbb{N}$ there exists $M = M(l,m,r) \in \mathbb{N}$ such that for any coloring $\chi : [N] \to [r]$, there exist $a, d \in \mathbb{N}$ for which

$$\{a, a+d, a+2d, \ldots, a+ld\} \cup \{md\}$$

is monochromatic.

Proof We proceed by induction on r. The base case is where $r = 1$; here, there is only one color, so we take $a = d = 1$ and make sure we have enough numbers to color red by letting $N = \max\{l+1, m\}$.

For the induction step, assume the existence of $M(l,m,r)$. The claim is that $M(l,m,r+1) = m \cdot W(l \cdot M(l,m,r), r+1)$. Let

$$\chi : [m \cdot W(l \cdot M(l,m,r), r+1)] \to [r+1]$$

be a coloring. Looking at χ just on the set $[W(l \cdot M(l,m,r), r+1)]$, we find a monochromatic arithmetic progression of length $l \cdot M(l,m,r)$, namely a monochromatic set

$$\{a + id : 0 \le i \le l \cdot M(l,m,r)\}.$$

Say this set is colored blue.

If there is a j, $1 \le j \le M(l,m,r)$ such that $\chi(mdj)$ is blue as well, then

$$\{a, a+jd, a+2jd, \ldots, a+ljd\} \cup \{mjd\}$$

forms the desired monochromatic set. If not, then χ is in fact an r-coloring on the set $\{mjd : 1 \le j \le M(l,m,r)\} = md[M(l,m,r)]$. So by inductive hypothesis there is a monochromatic set

$$\{a', a'+d', a'+2d', \ldots, a'+ld'\} \cup \{md'\} \subseteq [M(l,m,r)],$$

which means that χ is monochromatic on the set

$$\{a'md, a'md + d'md, a'md + 2d'md, \ldots, a'md + ld'md\} \cup \{m^2dd'\}.$$

Set $a^* = a'md$ and $d^* = d'md$ and this becomes

$$\{a^*, a^*+d^*, a^*+2d^*, \ldots, a^*+ld^*\} \cup \{md^*\}.$$

□

In particular, we will use the following case of Lemma 3.14.

Corollary 3.15 For any $l, m, r \in \mathbb{N}$ there exists $M = M(l,m,r)$ such that for any coloring $\chi : [N] \to [r]$, there exists $a, d \in \mathbb{N}$ for which

$$\{a + jd : -l \le j \le l\} \cup \{md\}$$

is monochromatic.

Proof By the preceding lemma we can find a big enough M as to provide a monochromatic progression $\{a' + jd : 0 \leq j \leq 2l\} \cup \{md\}$ under any r-coloring. We take our a to be $a' + ld$ and this monochromatic set turns into

$$\{a + jd : -l \leq j \leq l\} \cup \{md\}.$$

□

We now have all the machinery we need to prove Rado's Theorem, the main result of this section. Recall from number theory that for any integers a, b, it holds that $ab = \gcd(a, b) \cdot \text{lcm}(a, b)$ where gcd and lcm are the greatest common divisor and least common multiple functions, respectively. Also recall that if $d = \gcd(a_1, \ldots, a_n)$, then there exist integers $c_1, \ldots, c_n$ such that $d = c_1a_1 + \ldots + c_na_n$.

Proof (of Theorem 3.5)

($\Leftarrow$) Let C be a nonempty subset of $\{c_1, \ldots, c_n\}$ with $\sum C = 0$. Without loss of generality, we assume $C = \{c_1, \ldots, c_k\}$; otherwise, just reorder the variables. Let $r \in \mathbb{N}$ and let $\chi : \mathbb{N} \to [r]$ be any coloring. We need to show that

$$c_1x_1 + \ldots + c_nx_n = 0$$

has a monochromatic solution.

Note that if $k = n$, we can take $x_1 = \ldots = x_n = 1$ as our monochromatic solution, so assume $k < n$. Also, if $c_{k+1} + \ldots + c_n = 0$, the same choice of $x_1 = \ldots = x_n = 1$ works, so assume that $\sum_{i=k+1}^{n} c_i \neq 0$. Set $A = \gcd(c_1, \ldots, c_k)$ and set $B = c_{k+1} + \ldots + c_n$, and let $p = \dfrac{A}{\gcd(A, B)}$. (We can do this because $b \neq 0$.)

Note that $Bp = \dfrac{AB}{\gcd(A, B)} = \text{lcm}(A, B) = Aq'$ for some integer q'. Let $q = -q'$ and we have that

$$Aq + Bp = 0.$$

We can also find integers $j'_1, \ldots, j'_k$ such that $A = j'_1c_1 + \ldots + j'_kc_k$; set $j_i = j'_iq$ and we have

$$c_1j_1 + \ldots + c_kj_k = Aq.$$

Let $\beta = \max\{|j_1|, \ldots, |j_k|\}$. Then by Corollary 3.15 there exist $a, d \in \mathbb{N}$ such that $\{a + jd : -\beta \leq j \leq \beta\} \cup \{pd\}$ is monochromatic. For each $1 \leq i \leq n$, let

$$x_i = \begin{cases} a + j_i d & \text{if } 1 \leq i \leq k \\ pd & \text{if } k < i \leq n. \end{cases}$$

Then

$$\begin{aligned} c_1x_1 + \ldots + c_nx_n = \sum_{i=1}^{n} c_i x_i &= \sum_{i=1}^{k} c_i(a + j_i d) + \sum_{i=k+1}^{n} c_i pd \\ &= a \sum_{i=1}^{k} c_i + d \sum_{i=1}^{k} c_i j_i + pd \sum_{i=k+1}^{n} c_i \\ &= a(0) + dAq + pdB \\ &= d(Aq + Bp) \\ &= 0 \end{aligned}$$

so we have our desired monochromatic solution to the equation.

($\Rightarrow$) We assume that every nonempty subset C of $\{c_1, \ldots, c_n\}$ satisfies $\sum C \neq 0$, and we show that there is a coloring for which

$$c_1x_1 + \ldots + c_nx_n = 0$$

has no monochromatic solution. We will construct a coloring of $\mathbb{Q}^* = \mathbb{Q} - \{0\}$, the nonzero rationals. If we can show there is no monochromatic solution in $\mathbb{Q}^*$, we will certainly know that there is no monochromatic solution in $\mathbb{N}$.

Any rational $q \in \mathbb{Q}^*$ can be written as $q = \dfrac{a'}{b'}$ where $a' \in \mathbb{Z}, b' \in \mathbb{N}$ and $\gcd(a, b) = 1$. If $p \in \mathbb{N}$ is any prime, we can factor it out of a' and b' and write $q = \dfrac{p^d a}{b}$ for some $d \in \mathbb{Z}$. We have $d = 0$ if p is not in the prime factorization of a' or b'; in any case, p is not in the prime factorization of a or b. For any integer λ, let $[\lambda]_p \in \{0, 1, \ldots, p-1\}$ be λ *modulo* p (not to be confused with the set of p-subsets of $\{1, \ldots, \lambda\}$). Then any rational $\dfrac{a}{b}$ can be considered as $[ab^{-1}]_p$ where b^{-1} is the multiplicative inverse of b in the field $\mathbb{Z}_p$.

We define a coloring $\chi_p : \mathbb{Q}^* \to [p-1]$ as follows: for $q \in \mathbb{Q}^*$, write $q = \dfrac{p^d a}{b}$ and let

$$\chi_p(q) = [a]_p[b]_p^{-1}.$$

Note that $0 < \chi(q) < p$ because $\mathbb{Z}_p$ is a field and neither a nor b is a multiple of p (so neither is equal to 0 modulo p).

We claim that if p is a prime that does not divide the sum of any nonempty subset of $\{c_1, \ldots, c_n\}$, then the coloring χ_p defined above has no monochromatic solution to the equation $\sum_{i=1}^{n} c_i x_i = 0$. For each nonempty subset $C \subseteq \{c_1, \ldots, c_n\}$, there is a prime p_C that does not divide the sum $\sum C$; we can define p_C to be the smallest prime bigger than every prime in the factorization of $\sum C$. There are $2^n - 1$ such subsets C, so we can take $p = \max\{p_C : C \subseteq \{c_1, \ldots, c_n\}, C \neq \emptyset\}$.

Seeking a contradiction, assume $c_1 x_1' + \ldots + c_n x_n' = 0$ with $\{x_1', \ldots, x_n'\} \subseteq \mathbb{Q}^*$ monochromatic. By the definition of χ_p, if $\chi_p(q_1) = \chi_p(q_2)$ then $\chi_p(mq_1) = \chi_p(mq_2)$ for any $m \in \mathbb{Q}^*$. So we can find $\mu \in \mathbb{Q}^*$ such that $\mu x_1', \ldots, \mu x_n'$ are integers, such that $\{\mu x_1', \ldots, \mu x_n'\}$ is monochromatic, and such that

$$c_1 \mu x_1' + \ldots + c_n \mu x_n' = 0.$$

Set $x_i = \mu x_i'$ for $1 \leq i \leq n$. We can assume without loss of generality that $\gcd(x_1, \ldots, x_n) = 1$. (If $x_i = \frac{p^{d_i} a_i}{b_i}$, we can take $\mu = \prod_{i=i}^{n} b_i$ and then "scale down" by the gcd if it is greater than 1.)

Now reorder $x_1 \ldots, x_n$ such that p does not divide x_i for $1 \leq i \leq k$ and p does divide x_i for $k < i \leq n$. Note that $k = n$ is a possibility here, but $k = 0$ is not, since we are assuming $x_1, \ldots, x_n$ are relatively prime. We have that $\sum_{i=1}^{n} c_i x_i = 0$; it follows then that

$$\sum_{i=1}^{n} [c_i]_p [x_i]_p = [0]_p$$

in the field $\mathbb{Z}_p$. Recognizing that $[x_i]_p = 0$ for $k + 1 \leq i \leq n$, we have

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

Since $\{x_1, \ldots, x_k\}$ is monochromatic, we have $[x_1]_p = \ldots = [x_k]_p$ so this can be reduced to

$$[0]_p = [x_1]_p \sum_{i=1}^{k} [c_i]_p.$$

We have assumed that p does not divide x_1, so $[x_1]_p \neq [0]_p$ and thus

$$\sum_{i=1}^{k} [c_i]_p = [0]_p,$$

contradicting the fact that p does not divide the sum of any nonempty subset of $\{c_1, \ldots, c_n\}$. Thus there is no monochromatic solution even in $\mathbb{Q}^*$, so there is certainly no monochromatic solution in the coloring χ_p restricted to $\mathbb{N}$. □

3.4 Generalization to Systems of Equations

We defined chromatic number and examined partition regularity for single equations only. But we can extend these definitions and results to systems of equations. The system

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\
&\vdots \\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0
\end{aligned}$$

can be expressed in matrix notation, where A is the $m \times n$ matrix with $(A)_{ij} = a_{ij}$, as

$$A\mathbf{x} = \mathbf{0},$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{0}$ are $n \times 1$ column vectors.

Definition 3.16 Let A be an $m \times n$ matrix with integer entries. The *chromatic number* of the system $A\mathbf{x} = \mathbf{0}$ is the smallest positive integer r such that there exists a coloring $\chi : \mathbb{N} \to [r]$ with no monochromatic set $\{x_1, \ldots, x_n\}$ for which $\mathbf{x} = (x_1, \ldots, x_n)$ solves the system. If the chromatic number does not exist (because there is a monochromatic solution under any r-coloring for every $r \in \mathbb{N}$, then the system $A\mathbf{x} = \mathbf{0}$ is *partition regular*.

While we will not prove the details here, we present the analogous results about partition regular systems. In the case of a single equation, the key characteristic was whether a nonempty subset of the coefficients summed to 0. Here, the following condition is the key.

Definition 3.17 Let A be an $m \times n$ matrix with columns $\mathbf{A_1}, \ldots, \mathbf{A_n}$. If it is possible to reorder the columns of A as $\mathbf{A'_1}, \ldots, \mathbf{A'_n}$ and find $\lambda_1, \ldots, \lambda_t$ with $1 \leq \lambda_1 < \lambda_2 < \ldots < \lambda_t = n$ in a way such that when we define $n \times 1$ column vectors $\mathbf{B_i}$ as

$$\mathbf{B_i} = \sum_{j=\lambda_{i-1}-1}^{\lambda_i} \mathbf{A'_j},$$

it holds that $\mathbf{B_1} = \mathbf{0}$ and that each $\mathbf{B_i}, 1 < i \leq t$, is a linear combination of $\mathbf{A'_1}, \ldots, \mathbf{A'_{\lambda_{i-1}}}$, then we say that A satisfies the *columns condition.*

The columns condition is tremendously complex, which is why we do not prove the following. For the full elucidation, see [GRS].

Theorem 3.18 A system of equations $A\mathbf{x} = \mathbf{0}$ is partition regular if and only if A satisfies the columns condition.

3.5 Additional Results

We conclude our discussion of finite Ramsey Theory by presenting (without proof) some of the other classic results. First we present the Hales-Jewett Theorem, which is often seen as a further generalization of the van der Waerden theorems. Whereas the van der Waerden theorems consider monochromatic progressions in $\mathbb{N}$, a 1-dimensional discrete space, Hales-Jewett extends the idea of a monochromatic progression to n-dimensional discrete space.

Define the *n-cube* over l elements to be the set $C_l^n = [l]^n$. A *line* in C_l^n is a set of points $\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_l}\}$ with

$$\mathbf{x_i} = (x_{i1}, x_{i2}, \ldots, x_{in}) \in C_l^n$$

such that for each "coordinate" $1 \leq j \leq n$ we have either $x_{1j} = x_{2j} = \ldots =$

x_{lj} or $x_{ij} = i$ for each j. For example, if $n = 4$ and $l = 5$ and

$$\begin{aligned} \mathbf{x_1} &= (1, 4, 2, 1, 3) \\ \mathbf{x_2} &= (2, 4, 2, 2, 3) \\ \mathbf{x_3} &= (3, 4, 2, 3, 3) \\ \mathbf{x_4} &= (4, 4, 2, 4, 3) \\ \mathbf{x_5} &= (5, 4, 2, 5, 3) \end{aligned}$$

then $\{\mathbf{x_1}, \ldots, \mathbf{x_5}\}$ is a line in C_5^4.

Theorem 3.19 (Hales-Jewett Theorem) For any $l, r \in \mathbb{N}$ there exists $N = N(l, r)$ such that for any $n \geq N$, every coloring $\chi : C_l^n \to [r]$ produces a monochromatic line.

The other main result in finite Ramsey Theory has to do with colorings of vector spaces. Fix a finite field F; we consider colorings of the subspaces of F^n.

Theorem 3.20 (Graham-Leeb-Rothschild Theorem) Let $r, t, k \in \mathbb{N}$. There exists $N = N(r, t, k)$ such that if the t-dimensional subspaces of F^N are r-colored, then there exists a subspace $X \subseteq F^N$ with $\dim(X) = k$ such that the t-dimensional subspaces of X are monochromatic.

Most of the theorems we have encountered in this section can be generalized further, and computations can be carried out to calculate the exact values of N whose existence these theorems herald (for small cases). Each is a book of its own, and as we desire to plunge forward into infinite Ramsey Theory, we halt the discussion here.

4 Infinite Ramsey Theory

4.1 Preliminaries

Up to this point we have looked at colorings of both finite and infinite sets in search of finite substructures. We will extend this idea presently. It is very believable, given the finite version of Ramsey's Theorem, that if we color all of $\mathbb{N}$ then there is a k-homogeneous set of arbitrarily large (finite) cardinality. The next step will be to color $\mathbb{N}$ finitely and look for *infinite* homogeneous sets. Then we can discuss infinite colorings, etc. We assume the Axiom of Choice throughout.

Our earlier definitions were sufficient for finite Ramsey Theory, but we will need some more general ones now. Recall that a cardinal number is defined to be a "minimal" ordinal; that is, an ordinal number λ such that for all $\lambda' \prec \lambda$ (where $\prec$ well-orders the ordinals) there is no one-to-one function from λ into λ'.

Definition 4.1 Let X be a set. Then $[X]_\kappa$ is defined to be the set of all subsets of X with cardinality κ. That is, $[X]_k \equiv \{Y \subseteq X : |Y| = \kappa\}$.

Definition 4.2 Let X be a set and let λ be a cardinal. A λ-*coloring* of X is a function

$$f : X \to \lambda.$$

Definition 4.3 Let X be a set and let $f : [X]_\kappa \to \lambda$ be a coloring. If $Y \subseteq X$ and all the κ-subsets of Y are the same color, we say that $[Y]_\kappa$ is *monochromatic*, or that Y is *homogeneous*.

We will use the same "arrow" notation as we employed in Section 2. When we write

$$\nu \longrightarrow (\mu)^\lambda_\kappa$$

we mean that if the κ-subsets of a set of cardinality ν are λ-colored, then there exists a homogeneous subset $B \subseteq A$ with $|B| = \mu$. The other uses of the arrow notation carry over as well, so when λ is finite we may write $\nu \longrightarrow (\mu_1, \ldots, \mu_\lambda)_\kappa$ to mean that for some $1 \leq i \leq \lambda$ there is an i-homogeneous subset of cardinality μ_i.

The results of Lemma 2.4 hold; we will restate the relevant ones but not prove them again.

Lemma 4.4 Suppose that $\nu \longrightarrow (\mu)^\lambda_\kappa$. The following hold:

a) If $\mu' \preceq \mu$, then $\nu \longrightarrow (\mu')^\lambda_\kappa$.

b) If $\nu' \succeq \nu$, then $\nu' \longrightarrow (\mu)^\lambda_\kappa$.

What follows is a weaker version of the finite Ramsey Theorem from

Section 2.

Corollary 4.5 Let a, l, k be finite cardinals. Then $\aleph_0 \longrightarrow (a)^l_k$.

Proof The result follows from applying Lemma 4.4(b) to Theorem 2.9. $\square$

4.2 Some Graph Theory and the Infinite Ramsey Theorems

Just as we proved a preliminary version of the finite Ramsey theorem for $k = 2$, we prove a preliminary version of the infinite Ramsey theorem for $k = l = 2$. We show that if the complete graph on a countably infinite number of vertices ($K_{\aleph_0}$) is 2-colored, then there is a monochromatic induced $K_{\aleph_0}$ subgraph.

Before proving this, we will take a quick detour into graph theory.

Definition 4.6 In Section 2, we discussed graphs and defined them to consist of a vertex set V and an edge set E. Extend those definitions to include the possibility of infinite sets. If $v \in V$, the *degree* of v is

$$|\{w \in V : \{v, w\} \in E\}|.$$

In other words, the degree of v is the "number" (in the finite case) of vertices with which v shares an edge. A graph $G = (V, E)$ is *locally finite* if each vertex has finite degree.

Definition 4.7 Let $G = (V, E)$ be a graph. An *infinite path from* v_0 in G is a sequence of edges

$$(\{v_i, v_{i+1}\})_{i=0}^{\infty} \subseteq E.$$

A *path* from v_0 to v_n is a finite subset of an infinite path

$$(\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{n-1}, v_n\}).$$

G is *connected* if there is a path from v to v' for any vertices $v, v' \in V$.

Theorem 4.8 (König's Lemma) Let G be an infinite graph that is connected and locally finite. Then there is an infinite path from v for each vertex v.

Proof Fix $v \in V$ and let w be any other vertex. There is a path from v_0 to w, so there are infinitely many paths in G that "start" at v_0. But v

is of finite degree, so by the Pigeonhole Principle it must be that infinitely many of these paths "go through" one of these edges, say $\{v_0, v_1\}$. Follow the same procedure on $V - \{v_0\}$, starting from v_1, to find a vertex v_2 such that infinitely many of the paths from v_0 go through the edge $\{v_1, v_2\}$, and continue ad infinitum to obtain the infinite path. $\square$

Theorem 4.9 $\aleph_0 \longrightarrow (\aleph_0)^2_2$.

Proof Let A be a set with $|A| = \aleph_0$ and fix a 2-coloring $f$of A (with, say, colors red and blue). For each finite binary sequence $\sigma \in \{0,1\}^n$ we define a subset $C_\sigma \subseteq A$. We will consider each sequence σ to be a set of ordered pairs $\{(1, \sigma(1)), \ldots, (n, \sigma(n))\}$ and define $l(\sigma)$ to be the length of the sequence, so $l(\sigma) \equiv |\sigma|$. Also, given a finite binary sequence σ of length n, define σ^- to be the sequence achieved by extending the length of σ by one term and letting $\sigma(n+1) = 0$. Similarly, σ^+ is the one-term extension of σ with $\sigma(n+1) = 1$. So $\sigma^- \equiv \sigma \cup \{(l(\sigma)+1, 0)\}$, and $\sigma^+ \equiv \sigma \cup \{(l(\sigma)+1, 1)\}$. We will think of the finite binary sequences as forming a graph, where a sequence σ shares an edge with $\sigma|_{n-1}$ (the restriction of σ to its first $n-1$ terms), σ^- and σ^+. To each vertex (i.e., sequence) is associated a subset of A.

Let $C_\emptyset = A$ and define the other subsets C_σ inductively as follows:

- Having defined C_σ, pick $x_\sigma \in C_\sigma$ arbitrarily (if $C_\sigma \neq \emptyset$). If $C_\sigma = \emptyset$, let $C_{\sigma^-} = C_{\sigma^+} = \emptyset$ and let x_σ remain undefined. Otherwise, continue.
- Let $C_{\sigma^-} = \{x \in C_\sigma - \{x_\sigma\} : f(\{x, x_\sigma\}) = \text{red}\}$.
- Let $C_{\sigma^+} = \{x \in C_\sigma - \{x_\sigma\} : f(\{x, x_\sigma\}) = \text{blue}\}$.

We note that given σ,

$$\{\{x_\sigma\}, C_{\sigma^-}, C_{\sigma^+}\}$$

is a partition of C_σ. Therefore, if $\sigma \subsetneqq \sigma'$ then we have:

i) $C_{\sigma'} \subsetneqq C_\sigma$,

ii) $x_\sigma \notin C_{\sigma'}$, and

iii) $x_\sigma \neq x_{\sigma'}$.

Also, if $\sigma^* = \sigma \cup \{(l(\sigma)+1, \sigma'(l(\sigma)+1))\}$ (i.e., σ extended by length one and maintaining its agreement with σ'), then $\sigma^* \subseteq \sigma'$, so

$$x_{\sigma'} \in C_{x_{\sigma'}} \subseteq C_{\sigma^*}.$$

Hence, if $\sigma'(l(\sigma)+1) = 0$, then $f(\{x'_\sigma, x_{\sigma'}\})$ is red, and if $\sigma'(l(\sigma)+1) = 1$, then $f(\{x'_\sigma, x_{\sigma'}\})$ is blue.

We define $\mathcal{T} \equiv \{\sigma \in \{0,1\}^n$ for some finite $n : C_\sigma \neq \emptyset\}$. Thus $\mathcal{T}$ is a subgraph of our original graph of all finite sequences. For any $n \in \mathbb{N}$ we can express our vertex set A as

$$A = \{x_\sigma : \sigma \in \mathcal{T} \text{ and } l(\sigma) < n\} \cup \bigcup \{C_\sigma : l(\sigma) = n\}.$$

The first of these terms is finite so the second must be infinite; that is, for every n we have a nonempty C_σ with $l(\sigma) = n$.

Now $\mathcal{T}$ is an infinite, connected graph that is locally finite (each vertex has degree at most 3) and therefore, by König's Lemma, there is an infinite path P. (Since $|A| = \aleph_0$, it follows that $|P| = \aleph_0$.) Each vertex in P is a finite binary sequence; define

$$p \equiv \bigcup P \in \{0,1\}^{\aleph_0},$$

an infinite binary sequence. Either 0 or 1 occurs infinitely often in the sequence p; without loss of generality say it is 0.

For each $n \in \mathbb{N}$, the restriction of p to n is

$$p|_n = \{(1, p(1)), \ldots, (n, p(n))\}.$$

Let $\mathcal{B} = \{x_{p|_{n-1}} : p(n) = 0\}$. By assumption, $|\mathcal{B}| = \aleph_0$. If $a, b \in \mathcal{B}$, then $a = x_{p|_m}, b = x_{p|_n}$, and $p(m+1) = p(n+1) = 0$. Assume that $m < n$, so $p|_n \subseteq p|_m$ and thus, by our observations above, we have that $f(\{a,b\})$ is red. So $\mathcal{B}$ is the desired homogeneous set. □

Lemma 4.10 Let μ be a cardinal. If $\mu \longrightarrow (\mu)^2_\kappa$, then $\mu \longrightarrow (\mu)^r_\kappa$ for every $r \in \mathbb{N}$.

Proof We proceed by induction on r with trivial base case $r = 1$. The case $r = 2$ follows as well by hypothesis. For the induction step, let $r \geq 3$ and assume $\mu \longrightarrow (\mu)^\rho_\kappa$ for every $2 \leq \rho < r$.

Let M be a set with $|M| = \mu$ and let $f : [M]_\kappa \to [r]$ be a coloring. Define $g : [r] \to [2]$ by

$$g(i) = \begin{cases} 1 & \text{if } 1 \leq i \leq r-1 \\ 2 & \text{if } i = r. \end{cases}$$

So $g \circ f$ is a 2-coloring of $[M]_\kappa$ and thus, by hypothesis, there is a subset $A \subseteq M$ with $|A| = \mu$ and such that A is κ-homogeneous. There are two possibilities: either

i) $(g \circ f)(E) = 1$ for each $E \subseteq A$ with $|E| = \kappa$, or

ii) $(g \circ f)(E) = 2$ for each $E \subseteq A$ with $|E| = \kappa$.

If ii) holds, then $f(E) = r$ for each κ-subset E of A, so A is the desired homogeneous set.

If i) holds, then for each κ-subset $E \subseteq A$, $f(E) \in \{1, \ldots, r-1\}$ so f restricted to A is an $(r-1)$-coloring. Since $|A| = \mu$, by inductive hypothesis there is a homogeneous subset $A' \subseteq A$ with $|A'| = \mu$. □

Corollary 4.11 $\aleph_0 \longrightarrow (\aleph_0)^r_2$ for every $r \in \mathbb{N}$.

Proof Apply Lemma 4.10 to Theorem 4.9. □

At this point, we have that if the 2-subsets of a countably infinite set are r-colored (for any finite r), there is an infinite homogeneous set. The original Ramsey Theorem, the one that Ramsey himself published in 1930, shows that we can extend this result to k-subsets (for $k > 2$). Before proving this, we will need to borrow once again some terminology from graph theory.

Definition 4.12 A connected graph in which every vertex has degree 2 is a *cycle*. A *forest* is a graph of which no subgraph is a cycle. A *tree* is a forest that is connected.

Definition 4.13 An *ordered tree* is a pair $(T, <)$ where T is a tree and $<$ is a partial ordering such that there is a path from a vertex v to a vertex w iff $v < w$ or $w < v$. We will assume that ordered trees have a "root" vertex v_0 such that $v_0 < v$ for every $v \in V - \{v_0\}$.

Theorem 4.14 (Ramsey's Theorem) Let $r, k \in \mathbb{N}$. Then $\aleph_0 \longrightarrow (\aleph_0)^r_k$.

We present a version of Ramsey's original proof and a more modern argument. The second proof provides for an easier extension of the discussion to larger cardinals, as we will see.

Proof A By Lemma 4.10 it suffices to show that $\aleph_0 \longrightarrow (\aleph_0)^2_k$ for any k. We use induction on k; the base case $k = 1$ is a simple extension of the Pigeonhole

Principle–the union of finitely many color classes cannot be infinite if each class is finite. Assume $\aleph_0 \longrightarrow (\aleph_0)^2_k$; we show $\aleph_0 \longrightarrow (\aleph_0)^2_{k+1}$.

Let A be a set with $|A| = \aleph_0$ and let $f : [A]_{k+1} \to [2]$ be a coloring (we will call the color classes "red" and "blue," respectively). It is possible that there is some $x_1 \in A = A_0$ and a subset $A_1 \subseteq A$ such that $|A_1| = \aleph_0$, $x_1 \notin A_1$, and $f(\{x_1, y_1, \ldots, y_k\})$ is red for any $y_1, \ldots, y_k \in A_1$. It is then possible that there is some $x_2 \in A_1$ and a subset $A_2 \subseteq A_1$ with $|A_2| = \aleph_0, x_2 \notin A_2$, and $f(\{x_2, y_1, \ldots, y_k\})$ is red for any $y_1, \ldots, y_k \in A_2$. Then there may be some $x_3 \in A_2$ for which the same hold, and so on.

This process may continue infinitely, in which case we obtain sequences $(x_i)_{i=1}^{\infty}$ and $(A_i)_{i=1}^{\infty}$ such that for each i:

- $|A_i| = \aleph_0$,
- $x_i \in A_{i-1}$ but $x_i \notin A_i$, and
- $f(\{x_i, y_1, \ldots, y_k\})$ is red for any $y_1, \ldots, y_k \in A_i$.

Otherwise, the process fails at some stage (say at stage n) because, having selected A_{n-1}, there is no A_n with $|A_n| = \aleph_0$ and that contains some $x_n \in A_{n-1} - A_n$ such that $f(\{x_n, y_1, \ldots, y_k\})$ is red for all $y_1, \ldots, y_k \in A_n$.

In the former case (where we do obtain infinite sequences), let $\Delta = \{x_1, x_2, \ldots\}$. It is clear by construction that the members x_i are distinct and thus that $|\Delta| = \aleph_0$. Let $\{x_{i_1}, \ldots, x_{i_{k+1}}\} \subseteq \Delta$ with $i_1 < \ldots < i_{k+1}$. We notice that $x_{i_2}, \ldots, x_{i_k}, x_{i_{k+1}} \in A_i$ so $f(\{x_{i_1}, \ldots, x_{i_{k+1}}\})$ is red and thus Δ is our desired homogeneous set.

Next we consider the case in which our process terminates finitely at stage n as described above. There is no A_n that fits the criteria we desire so let A_n be any infinite subset of A_{n-2} and let $y_1 \in A_{n-1}$ be arbitrary. Construct a coloring $g : [A_{n-1} - \{y_1\}]_k \to [2]$ where $g(\{z_1, \ldots, z_k\}) = f(\{y_1, z_1, \ldots, z_k\})$ for $z_1, \ldots, z_k \in A_{n-1} - \{y_1\}$. By inductive hypothesis g yields an infinite 2-homogeneous subset B_1. If $[B_1]_2$ is monochromatically red under g, then having taken $x_n = y_1$ and $A_n = B_1$ would have allowed our previous process to continue past stage n. So it must be that $[B_1]_2$ is monochromatically blue.

Similarly, we then take B_2 to be an infinite subset of B_1 and pick $y_2 \in B_1 - B_2$. We then construct a coloring $h : [B_1 - \{y_2\}]_k \to [2]$ where $h(\{z_1, \ldots, z_k\}) = f(\{y_2, z_1, \ldots, z_k\})$ for any $z_1, \ldots, z_k \in B_1 - \{y_2\}$. By hypothesis we get a homogeneous subset, which must be blue or else it would have allowed our above process to continue at stage n. Repeat this process ad infinitum to obtain a sequence $(y_1)_{n=1}^{\infty}$, which is blue-homogeneous. $\square$

Proof B The proof is by induction on k. The base case $k = 1$ holds for the same reason as in Proof A. We show that if $\aleph_0 \longrightarrow (\aleph_0)_k^r$ for all r then $\aleph_0 \longrightarrow (\aleph_0)_{k+1}^r$ for all r.

Let A be a set with $|A| = \aleph_0$ and let $f : [A]_{k+1} \to [r]$ be a coloring. We will define an ordered tree $(T, \subsetneq)$, where the ordering is proper inclusion, by "levels." That is, we describe all of the vertices at level (depth from the "root") m for each $m \in \mathbb{N}$. Each vertex of our graph will be a nonempty subset of A. Let T_m be the set of all vertices at level m, and set $T_0 = \{A\}$ (so A, the root, is the lone vertex at level 0).

We want to maintain the following invariants for each T_m:

(1) T_m is finite.

(2) $\bigcup T_m = A - \{\min E : E \in \bigcup_{l=0}^{m-1} T_l\}$.

(3) If $E_1, E_2 \in T_m$ with $E_1 \neq E_2$, then $E_1 \cap E_2 = \emptyset$.

(4) For each $0 \leq l < m$ and for each $E \in T_m$, there exists a unique $E' \in T_l$ such that $E \subsetneq E'$.

It is clear that (1)-(4) hold for T_0. We inductively define T_{m+1} and prove that these invariants hold. Assume that we have defined T_l for each $0 \leq l \leq m$ and that conditions $(1) - (4)$ hold for each such T_l.

Let $E \in T_m$. By (4) there exists a unique $E_l \in T_l$ for each $l < m$ such that $E \subsetneq E_l$. Set $y_l = \min E_l$ for $0 \leq l < m$ and set $y_m = \min E$. For each $u \in E - \{y_m\}$ define a function

$$g_u : [\{y_0, y_1, \ldots, y_m\}]_k \to [r]$$

by $g_u(\{z_1, \ldots, z_k\}) = f(\{z_1, \ldots, z_k, u\})$ for any $\{z_1, \ldots, z_k\} \subseteq \{y_0, \ldots, y_m\}$. We define a relation R_E on $E - \{y_m\}$ such that uR_Ev iff $g_u = g_v$. There are only finitely many possible functions (to be precise, there are $r^{\binom{m+1}{k}}$) so there are finitely many equivalence classes. Let Γ_E be the set of all equivalence classes under R_E, so Γ_E is a partition of $E - \{y_m\}$. We define $T_{m+1} \equiv \bigcup_{E \in T_m} \Gamma_E$.

We need to show that (1)-(4) hold for T_{m+1}.

(1) Γ_E is finite for each $E \in T_m$ and T_m is finite, so T_{m+1} is finite.

(2) We calculate:

$$\begin{aligned}\bigcup T_{m+1} &= \bigcup \bigcup_{E \in T_m} \Gamma_E = \bigcup_{E \in T_m} \bigcup \Gamma_E \\ &= \bigcup_{E \in T_m} (E - \{y_m\}) = \bigcup_{E \in T_m} (E - \{\min E\}) \\ &= \bigcup T_m - \{\min E : E \in T_m\} \text{ (since } T_m \text{ is pairwise disjoint)} \\ &= (A - \{\min E : E \in \bigcup_{l=0}^{m-1} T_l\}) - \{\min E : E \in T_m\} \text{ (by (2))} \\ &= A - \{\min E : E \in \bigcup_{l=0}^{m} T_l\},\end{aligned}$$

so (2) holds.

(3) Each element of T_{m+1} is a cell of a partition of a set $E \in T_m$, and the set T_m is pairwise disjoint (by hypothesis), so T_{m+1} is pairwise disjoint as well.

(4) If $E \in T_{m+1}$, it is clear that $E \subsetneqq E'$ for exactly one $E' \in T_m$, and by hypothesis we have $E \subsetneqq E' \subsetneqq E''$ for a unique $E'' \in T_l$ for each $0 \leq l < m$.

We have that $(T, \subsetneqq)$ is an infinite graph, and since each level T_m is finite, the graph is locally finite. Therefore, by König's Lemma, T has an infinite path. (In fact T is a tree with an infinite "branch.") Let this infinite branch be the set

$$\{E_l : l = 0, 1, 2, \ldots\}$$

where $E_l \in T_l$ for each l. For each l, set $x_l = \min E_l$ and let $X = \{x_l : l = 0, 1, 2, \ldots\}$.

Consider $\{x_{i_1}, \ldots, x_{i_k}\} \in [X]_k$ with $i_1 < \ldots < i_k$. If $l > i_k$ then $x_l \in E_l \subseteq E_{l-1} \subseteq \ldots \subseteq E_{i_k+1}$. Since E_{i_k+1} is an equivalence class of the relation $R_{E_{i_k}}$ on E_{i_k}, it follows that

$$f(\{x_{i_1}, \ldots, x_{i_k}, u\}) = g_u(\{x_{i_1}, \ldots, x_{i_k}\})$$

for any $u \in E_{i_k+1}$. So for any $x_{i_1}, \ldots, x_{i_k}, x_{i_{k+1}} \in X$, we have that

$$f(\{x_{i_1}, \ldots, x_{i_k}, x_{i_{k+1}}\})$$

does not depend on i_{k+1}, and hence we can write without loss of generality that

$$f(\{x_{i_1}, \dots, x_{i_k}, x_{i_{k+1}}\}) = g(\{x_{i_1}, \dots, x_{i_k}\}),$$

where $g : [X]_k \to [r]$ is an "induced" coloring on just the k-subsets of X. By our induction hypothesis, therefore, there is a subset $B \subseteq X$ with $|B| = \aleph_0$ and such that g is constant on $[B]_k$. Say g is constantly color j on $[B]_k$; we have, then that $f(\{x_{i_1}, \dots, x_{i_k}, x_{i_{k+1}}\}) = j$ for any $x_{i_1}, \dots, x_{i_k}, x_{i_{k+1}} \in B$. So B is our desired homogeneous set. □

4.3 The Compactness Principle

We take some time for a short digression. In Section 3 we often switched back and forth between a) results stating the existence of a positive integer n sufficiently large enough to produce (under a finite coloring) a large subset with some desirable property and b) results in which we needed to color all of $\mathbb{N}$ in order to achieve the desired substructure.

Often, these two approaches are equivalent. Obviously, if there is an $N \in \mathbb{N}$ such that any coloring of $[n]_k$ for $n \geq N$ produces a monochromatic substructure, then a coloring of $[\mathbb{N}]_k$ will produce the same monochromatic substructure. The Compactness Principle is a crucial result in Ramsey Theory which allows us to go the other way as well: if a coloring of $[\mathbb{N}]$ yields our desired substructure, then there is, in fact, a finite n sufficiently large to generate the substructure we want under any coloring of $[n]_k$.

In particular, we will show that the equivalence holds inasmuch as the infinite version of Ramsey's Theorem (Theorem 4.14) implies the finite version (Theorem 2.9). Those who have studied mathematical logic will note the similarity between this theorem and the Compactness theorems on well-formed first-order logical formulae, which show that if every finite subset of an infinite set has some property, then the entire set has the property.

Proof (of Theorem 2.9) Suppose there are $k, a_1, \dots, a_r \in \mathbb{N}$ such that $R_k(a_1, \dots, a_r)$ does not exist. Then certainly $R_k(a; r)$ does not exist; this means for every $n \in \mathbb{N}$ we have that $n \nrightarrow (a)^r_k$.

For each $n \in \mathbb{N}$ define the set Φ_n to consist of all r-colorings of $[n]_k$ under which there is no homogeneous subset $A \subseteq [n]$ with $|A| = a$. Also let $\Phi_0 = \emptyset$. For $n \geq 1$, each Φ_n is finite, and, by assumption, each $\Phi_n \neq \emptyset$. We form an ordered tree $\mathcal{T} = (T, <)$ where the vertices of our tree are all these finite colorings, or $T = \bigcup_{n=0}^{\infty} \Phi_n$. The order $<$ of our tree is inclusion; that is, if $\chi_n \in \Phi_n$ and $\chi_m \in \Phi_m$ for $m < n$, then there is a path from χ_m to χ_n if $\chi_n|_{[m]} = \chi_m$.

It is clear that our tree $\mathbb{N}$ is infinite and connected. Plus, for a coloring $\chi_n \in \Phi_n$ there are only finitely many (at most r) extensions to colorings in Φ_{n+1}. So $\mathcal{T}$ is locally finite, and we obtain, by König's Lemma, an infinite path $\{\chi_n : n \geq 0\}$ from the "root" vertex, which is the empty coloring in Φ_0.

Let $\chi = \bigcup\{\chi_n : n \geq 0\}$. Then χ is an r-coloring of $\mathbb{N}$ with no homogeneous subset of cardinality a. Certainly, then, there is no homogeneous set of cardinality $\aleph_0$, contradicting the fact that $\aleph_0 \longrightarrow (\aleph_0)^r_k$. $\square$

4.4 Larger Cardinals

When we were dealing with finite cardinals, if we wanted a homogeneous set of size n, we needed to start with a set of cardinality much larger than n, but the finite version of Ramsey's Theorem told us there were sufficiently big *finite* starting sets. But we needed to work with a coloring of a set larger than the homogeneous one we sought–finite cardinals are "too small" to find equinumerous homogeneous subsets under colorings.

In this section, we have seen that $\aleph_0$ is a "perfect" size in that a set of cardinality $\aleph_0$ is big enough so as to find a homogeneous subset of the same cardinality under a finite coloring. That is, $\aleph_0$ is both "big enough" and "not too big." We will now see that the same results do not generalize to larger cardinals, as in this case the size of the desired homogeneous subset is "too big."

Theorem 4.15 (Sierpiński's Theorem) $\aleph_1 \nrightarrow (\aleph_1)^2_2$.

Proof We prove a (possibly) stronger fact: that $2^{\aleph_0} \nrightarrow (\aleph_1)^2_2$. (Remember that $2^{\aleph_0} \succeq \aleph_1$, with the claim of equality being the continuum hypothesis.) To show this, we need to construct a 2-coloring of a set of cardinality $2^{\aleph_0}$ ($\mathbb{R}$ would seem to be a prime choice) such that every subset of cardinality $\aleph_1$ is not 2-homogeneous.

Let $<$ be the natural ordering of $\mathbb{R}$ and let $<^*$ be a well-ordering of $\mathbb{R}$ (which exists by the Axiom of Choice). Define a 2-coloring $f : [\mathbb{R}]_2 \to [2]$ as follows:

$$f(\{x,y\}) = \begin{cases} 1 & \text{if } x < y \text{ iff } x <^* y \\ 2 & \text{if } x < y <^* x \text{ or } x <^* y < x. \end{cases}$$

That is, f assigns $\{x, y\}$ color 1 if $<, <^*$ agree on the relative order of x and y, and color 2 otherwise. Let $A \subseteq \mathbb{R}$ be homogeneous; say $f(\{x, y\}) = 1$ for all $x, y \in A$. So $<$ is a well-ordering on A. We define a function $\psi : A \to \mathbb{Q}$ where $\psi(a)$ is an arbitrary rational between a and the successor of a under $<$. It is clear that ψ is one-to-one; if $a \neq b$ then (without loss of generality) $a < b$ so $\psi(a) < \psi(b)$. So

$$|A| \preceq |\mathbb{Q}| = \aleph_0 \prec \aleph_1.$$

If A is homogeneous for color 2, then the relation $x <' y$ iff $x > y$ is a well-ordering on A, so we follow a similar procedure.

Thus there is no homogeneous subset of cardinality $\aleph_1$. $\square$

The next step is to extend this result and show that for any infinite cardinal $\aleph_\alpha$ we have $2^{\aleph_\alpha} \nrightarrow (\aleph_\alpha^+)^2_2$, where κ^+ denotes the successor of a cardinal κ. We can use a similar argument to the one used in Theorem 4.15 but we first need a lemma. Recall that an ordinal λ is the set consisting of all ordinals less than λ. Also note that since we are assuming the axiom of choice, the alephs are exactly all the infinite cardinals. If the reader is in need of a more thorough refresher in set theory, see [HH]. Otherwise we give just the requisite definitions.

Definition 4.16 Let $(A, <)$ be a well-ordered structure and let $B \subseteq A$. We say B is *cofinal* in A if for every $x \in A$ there is some $y \in B$ such that $x \leq y$.

Definition 4.17 Let $(A, <)$ be a well-ordered structure. The *order type* of $(A, <)$, denoted $\mathrm{ot}((A, <))$ or just $\mathrm{ot}(A)$, is the unique ordinal isomorphic to $(A, <)$. Consider the set

$$\Xi \equiv \{\mathrm{ot}(B) : B \subseteq A \text{ is well-ordered and cofinal in } A\}.$$

Ξ is a nonempty (A is cofinal in A) set of ordinals and therefore has a least element ξ. The order type $\mathrm{ot}(\xi)$ is the *cofinality* of $(A, <)$, denoted $\mathrm{cf}((A, <))$.

Note that the notion of order type is more general and can be used to describe ordered structures that are not necessarily well-ordered. But we will deal only with well-ordered structures.

Definition 4.18 Let ξ be an ordinal. We say ξ is *regular* if $\xi \succ 1$ and $\mathrm{cf}(\xi) = \xi$. We say ξ is *singular* if $1 \prec \mathrm{cf}(\xi) \prec \xi$.

For infinite cardinals there is an important characterization of cofinality: if κ is infinite, then $\mathrm{cf}(\kappa)$ is the least ordinal α for which there is a "sequence" of cardinals

$$\{\kappa_\beta \prec \kappa : \beta \prec \alpha\}$$

such that

$$\sum_{\beta \prec \alpha} \kappa_\beta = \kappa.$$

We will use this fact later; for a full proof see [HH].

The strategy in proving Theorem 4.15 was to use a canonical ordering on $\mathbb{R}$ that we know failed to be a well-ordering on $\mathbb{Q}$. We will do the same for larger cardinals than $\aleph_1$ and $\aleph_0$, but we will need the existence of a canonical ordering on a set of cardinality 2^κ that does not well-order any of its subsets of cardinality κ^+.

We will assume knowledge of the classic set-theoretic facts that every regular ordinal is a cardinal but not every not every cardinal is regular. (The classic example is $\aleph_\omega$). However, if κ is a cardinal, then κ^+ is regular.

Definition 4.19 Let κ be a cardinal and let $x, y \in 2^\kappa$. Then x, y are functions from the cardinal κ into the set $2 = \{0, 1\}$. If $\lambda \preceq \kappa$ is a cardinal, then $x|_\lambda$ is the function x restricted to λ, as usual. The *left-lexicographic ordering* $<$ on 2^κ is such that $x < y$ iff there exists a $\lambda \prec \kappa$ with

$$x|_\lambda = y|_\lambda \quad \text{and} \quad x(\lambda) = 0, y(\lambda) = 1.$$

That is, x is less than y in 2^κ if x and y agree for every $\lambda \prec \kappa$ and $x(\lambda) < y(\lambda)$.

Lemma 4.20 Let κ be an infinite cardinal and let $<$ be the left-lexicographic ordering on 2^κ. If $A \subseteq 2^\kappa$ with $\mathrm{ot}(A) = \kappa^+$, the neither $(A, <)$ or $(A, >)$ (in which the left-lexicographic ordering is reversed) are well-ordered structures with the orderings $>, <$ restricted from 2^κ to A.

Proof Let $A \subseteq 2^\kappa$ with $\mathrm{ot}(A) = \kappa^+$. Assume that the left-lexicographic ordering $<$ well-orders A, so $(A, <)$ is isomorphic to κ^+. Let φ be such an isomorphism. We will show by transfinite induction on μ that

$(*)$ for every $\mu \preceq \kappa$, there is some $\alpha_\mu \prec \kappa^+$ such that for every $\beta \succeq \alpha_\mu$ we have $\varphi(\beta)|_\mu = \varphi(\alpha_\mu)|_\mu$.

If we can show that $(*)$ holds we will let $\mu = \kappa$ and get an $\alpha_\kappa \preceq \kappa$ such that for every $\beta \succeq \alpha_\kappa$ it holds that

$$\varphi(\beta) = \varphi(\beta)|_\kappa = \varphi(\alpha_\kappa)|_\kappa = \varphi(\alpha_\kappa),$$

where $\varphi(\beta) = \varphi(\beta)|_\kappa$ and $\varphi(\alpha_\kappa)|_\kappa = \varphi(\alpha_\kappa)$ hold trivially because $\varphi(\beta), \varphi(\alpha_\kappa)$ are themselves functions on κ. This is a contradiction for any $\beta \succ \alpha_\kappa$, since φ is an isomorphism (and is therefore one-to-one).

The base case for the transfinite induction is $\mu = 0$; in this case we choose $\alpha_0 = 0$ and it is clear that for any $\beta \succeq 0$ we get $\varphi(\beta)|_0 = \varphi(0)|_0$.

Next we show that if $(*)$ holds for an ordinal μ, then it holds for $\mu+1$. We consider two cases. First assume that for each $\beta \succeq \alpha_\mu$ we have $\varphi(\beta)(\mu) = 0$. In this case we choose $\alpha_{\mu+1} = \alpha_\mu$, and let $\beta \succeq \alpha_{\mu+1}$. By hypothesis we have $\varphi(\beta)|_\mu = \varphi(\alpha_{\mu+1})|_\mu = \varphi(\alpha_\mu)|_\mu$. We also have that $\varphi(\beta)(\mu) = 0$ and $\varphi(\alpha_{\mu+1})(\mu) = \varphi(\alpha_{\mu+1})(\mu) = 0$, so $\varphi(\beta)(\mu) = \varphi(\alpha_{\mu+1})(\mu)$. Then $\varphi(\beta), \varphi(\alpha_{\mu+1})$ agree up to μ and at μ, so

$$\varphi(\beta)|_{\mu+1} = \varphi(\alpha_{\mu+1})|_{\mu+1}.$$

Otherwise there is some $\beta' \succeq \alpha_\mu$ with $\varphi(\beta')(\mu) = 1$; in this case, define $\alpha_{\mu+1}$ to be the smallest such β', so that $\varphi(\alpha_{\mu+1})(\mu) = 1$. Let $\beta \succeq \alpha_{\mu+1}$. So $\varphi(\beta)|_\mu = \varphi(\alpha_\mu)|_\mu = \varphi(\alpha_{\mu+1})|_\mu$ (since $\alpha_{\mu+1} \succeq \alpha_\mu$). Since φ is an isomorphism we have $\varphi(\beta) \geq \varphi(\alpha_{\mu+1})$ (note the usage of $\geq$, the lexicographic ordering as opposed to $\succeq$, the regular ordering of the ordinals). And since the two agree up to μ, we have that $\varphi(\beta)(\mu) \geq \varphi(\alpha_{\mu+1})(\mu) = 1$ and hence $\varphi(\beta)(\mu) = 1$ which gets us our desired result that

$$\varphi(\beta)|_{\mu+1} = \varphi(\alpha_{\mu+1})|_{\mu+1}.$$

Finally we consider the case where μ is a limit ordinal. Here we set $\alpha_\mu = \bigcup_{\lambda \prec \mu} \alpha_\lambda = \sup_{\lambda \prec \mu} \alpha_\lambda$. Since we are only working with $\mu \preceq \kappa$, so for every $\lambda \prec \mu$ we have that $\alpha_\lambda \prec \kappa^+$ and hence $\alpha_\mu \preceq \kappa^+$. But since κ^+ is regular, we can eliminate the possibility of equality and get $\alpha_\mu \prec \kappa^+$.

Let $\beta \succeq \alpha_\mu$ and let $\lambda \prec \mu$. Since μ is a limit ordinal we have $\lambda + 1 \prec \mu$ and hence $\alpha_{\lambda+1} \preceq \alpha_\mu$. Applying the inductive hypothesis (that $(*)$ holds for all $\lambda \prec \mu$) gives us that

$$\varphi(\beta)(\lambda) = \varphi(\alpha_{\lambda+1})(\lambda) = \varphi(\alpha_\mu)(\lambda)$$

and hence $\varphi(\beta)|_\mu = \varphi(\alpha_\mu)|_\mu$, as needed.

So we apply (∗) with $\mu = \kappa$ to get a contradiction and therefore we have that the ordering $<$ does not well-order A. The argument that the ordering $>$ does not well-order A is entirely symmetric, we can simply swap every instance of 0 and 1 in the above. □

Theorem 4.21 For any infinite cardinal κ, $\kappa^+ \nrightarrow (\kappa^+)^2_2$.

Proof As in Theorem 4.15, we can prove the stronger result that $2^\kappa \nrightarrow (\kappa^+)^2_2$. The argument is the same as that in Theorem 4.15, except here we work with (i.e., 2-color) the cardinal 2^κ itself by whether its well-ordering $<^*$ agrees with its left-lexicographic ordering $<$ on a pair $\{x, y\}$. By Theorem 4.20 we cannot hope to find a subset A with $|A| = \kappa^+$ on which $<, <^*$ agree. So there is no homogeneous subset of cardinality κ^+. □

Corollary 4.22 For any infinite cardinal κ and finite $r, k \in \mathbb{N}$, $\kappa^+ \nrightarrow (\kappa^+)^r_k$.

Proof If this result were positive it would be in clear contradiction with Theorem 4.21. □

Although we cannot, in general, find an equinumerous homogeneous subset under colorings of uncountable sets, the following result guarantees that if there is not an *equinumerous* homogeneous subset then there is at least an *infinite* homogeneous subset. This is somewhat of a relief. However, the proof is very difficult and we omit it here.

Theorem 4.23 Let μ be any infinite cardinal. Then $\mu \longrightarrow (\mu, \aleph_0)_2$.

The truth is that we have not shown that the result $\kappa \nrightarrow (\kappa)^2_2$ holds for *every* uncountable cardinal κ. We claimed that $\aleph_0$ was "nice" inasmuch as that under finite colorings of (finite) subsets of a set A of cardinality $\aleph_0$ we were able to retrieve homogeneous sets equinumerous to A. We now show that $\aleph_0$ is not necessarily the lone infinite cardinal for which this holds. Theorem 4.18 showed that if an uncountable cardinal is the successor of another cardinal, we cannot retrieve equinumerous, homogeneous sets under arbitrary finite colorings. But the crucial characteristics of $\aleph_0$ that we utilized in the proof of Ramsey's Theorem were that it was a limit cardinal and that we could apply König's Lemma. If we have another infinite cardinal with these properties, we may be able to get "positive" results.

Definition 4.24 An infinite cardinal κ is a *weak limit cardinal* if it is of the form $\aleph_\lambda$ where $\lambda = 0$ or λ is a limit ordinal (that is, for all $\mu \prec \lambda$ also $\mu^+ \prec \lambda$). An infinite cardinal κ is a *strong limit cardinal* if $\kappa \succ 0$ and for all $\mu \prec \kappa$ also $2^\mu \prec \kappa$.

Note that if κ is a strong limit cardinal then it is also a weak limit cardinal, since if $\mu \prec \kappa$ then $\mu^+ \preceq 2^\mu \prec \kappa$. It should be clear that $\aleph_0$ is a strong limit cardinal. Uncountable weak limit cardinals are also said to be *weakly inaccessible*, and uncountable strong limit cardinals are also said to be *strongly inaccessible*, or simply *inaccessible*.

Definition 4.25 Let ξ be a cardinal. Then ξ is *weakly compact* if it is a strong limit cardinal and has the additional property that any ordered tree $\mathcal{T}$ with $|\mathcal{T}| = \xi$ has either a branch of length ξ or a "level" T_m with $|T_m| = \xi$.

Those who have studied axiomatic set theory may know that with the ZFC axioms the existence of uncountable strong limit cardinals (i.e., inaccessible cardinals) cannot be proven or disproven. All the more so can the existence of weakly compact cardinals (other than $\aleph_0$) not be proven. It turns out, though, that these are exactly the properties of $\aleph_0$ that allow us to obtain $\aleph_0 \longrightarrow (\aleph_0)^r_k$ for any finite r, k. If we have a weakly compact cardinal ξ we get the following similar result.

In the case of $\aleph_0$, we have the existence of homogeneous equinumerous sets under r-colorings of $[\aleph_0]_k$ for any *cardinals* $r, k \prec \aleph_0$, which amounts to finite r, k. In the theorem that follows, we allow the k to be any cardinal less than our weakly compact cardinal, but we keep r finite.

Theorem 4.26 Let ξ be a weakly compact cardinal, let $\kappa \prec \xi$ and let $r \in \mathbb{N}$. Then $\xi \longrightarrow (\xi)^r_\kappa$.

We will not prove Theorem 4.26 here, as it is highly technical and the idea is much the same as Proof B of Theorem 4.14 above. For the full account, see [LEV].

The final frontier is to consider infinite colorings.

4.5 Infinite Colorings

By an infinite coloring of a set A we mean one that uses infinitely many colors. The first result deals with the most primitive type of coloring, a coloring of a set itself (rather than its subsets). In Section 2 we called our result on colorings of finite sets the Pigeonhole Principle. That motivates the name we give to the following.

Lemma 4.27 (Infinite Pigeonhole Principle) Let κ be an infinite cardinal and let $\xi \prec \mathrm{cf}(\kappa)$. Then $\kappa \longrightarrow (\kappa)^{\xi}_{1}$.

Proof We mentioned above that the cofinality of an infinite cardinal is the smallest ordinal α such that if we let κ_β be a cardinal less than κ for each $\beta \prec \alpha$ we can have the sum $\sum_{\beta \prec \alpha} \kappa_\beta = \kappa$. Therefore, is $\xi \prec \mathrm{cf}(\kappa)$ colors are used and each color class $A_\alpha, \alpha \prec \xi$ satisfies $|A_\alpha| \prec \kappa$, then

$$\kappa = \sum_{\alpha \prec \xi} |A_\alpha| \prec \kappa,$$

which is ridiculous. So it must be that for some $\alpha \prec \xi$, $|A_\alpha| = \kappa$, which is precisely what we need. $\square$

It is obvious from the discussion within the proof that if $\xi \succeq \mathrm{cf}(\kappa)$ then there is a way of coloring κ with ξ colors without a monochromatic subset of cardinality κ. So $\kappa \not\rightarrow (\kappa)^{\xi}_{1}$. We have the following.

Corollary 4.28 If κ is an infinite cardinal, then $\kappa \longrightarrow (\kappa)^{\xi}_{1}$ if and only if $\xi \prec \mathrm{cf}(\kappa)$.

Throughout our discussion we have been interested in coloring the subsets of a set rather than the elements themselves. With respect to infinite colorings a relevant question to ask is under what circumstances we can find an infinite homogeneous set under an infinite-coloring of, say, the 2-subsets of an infinite set. Our first hope might be that for an infinite cardinal κ the cardinal 2^κ is sufficiently large so as to find at least a countably infinite homogeneous subset under a κ-coloring of the 2-sets of a set of cardinality 2^κ.

Our hopes are dashed, as shown in the following result. The proof will demonstrate that even homogeneous finite sets (with cardinality larger than 2) cannot, in general, be found when infinitely many colors are used.

Theorem 4.29 For any cardinal κ we have $2^\kappa \not\rightarrow (\aleph_0)^\kappa_2$.

Proof When κ is finite the theorem is trivial, so assume κ to be infinite. As usual, we need to construct a coloring of the 2-subsets of a set of cardinality 2^κ under which there is no infinite homogeneous subset. In fact, the coloring we construct does not even have a homogeneous subset of size 3.

Let $A = \{0,1\}^\kappa$ be the set of all functions from κ into $\{0,1\}$. Define a coloring $\varphi : [A]_2 \to \kappa$ as follows: for two functions $g, h \in A$, let

$$\varphi(\{g,h\}) = \min\{\alpha \prec \kappa : g(\alpha) \neq h(\alpha)\}.$$

That is, φ assigns to the pair $\{g,h\}$ the least ordinal on which they disagree. Take any set $\{f,g,h\}$ of distinct functions. If $\varphi(\{f,g\}) = \varphi(\{f,h\}) = \beta$ for some $\beta \prec \kappa$, then we have both that $f(\beta) \neq g(\beta)$ and that $f(\beta) \neq h(\beta)$. But there are only two possible values a function in A can take on at β, so it must be that $g(\beta) = h(\beta)$. Obviously this implies $\varphi(\{g,h\}) \neq \beta$ and hence $\{f,g,h\}$ is not homogeneous under φ. $\square$

We can, however, achieve some positive results with infinite colorings as long as we restrict ourselves to the use of only $\aleph_0$ colors and do not get too greedy with the desired size of our homogeneous sets. Recall that the *beth numbers* are defined by setting $\beth_0 = \aleph_0$ and then by letting $\beth_{\alpha+1} = 2^{\beth_\alpha}$. (Forget, for our purposes, defining $\beth_\alpha$ for other limit ordinals α.)

Theorem 4.30 For any $n \prec \aleph_0$, $\beth_n^+ \longrightarrow (\aleph_1)^{\aleph_0}_{n+1}$.

So if we are coloring finite subsets with countably infinite colors, we have the existence of a sufficiently large cardinal for which we can be guaranteed an infinite (and uncountable) homogeneous subset. We will omit the proof, as it would inevitably lead to us plunging further into the possible and impossible results for colorings of and with infinite cardinals. By this point, however, we have hopefully given the reader a taste of infinite Ramsey Theory, so rather than continue down this road, we will leave this discussion here and continue to some applications of the Ramsey theorems to $\mathbb{R}^n$ and to general metric spaces. For more on this topic, we refer the reader to any and all of [JEC],

[HH], [LEV].

5 Continuous Ramsey Theory

5.1 Two Examples in $\mathbb{R}^n$

The extensions of the Ramsey results to continuous spaces are normally presented as dichotomies. Ramsey's Theorem allows us to color infinite structures (finitely) and find large (infinite) homogeneous substructures. If we take a space and color it in a clever fashion, by which we mean where the coloring is defined based on some meaningful property, the guarantee of an infinite, homogeneous subset can be re-interpreted in an insightful sense. We start with a simple geometric example.

Definition 5.1 Let $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Define the *line from x to y*, $L(x, y)$, to consist of the points $t\mathbf{x} + (1-t)\mathbf{y}$ for any $t \in [0, 1]$. More formally, we let $m_i = \min\{x_i, y_i\}$ and $M_i = \max\{x_i, y_i\}$, and then $L(x, y)$ is the set

$$\{z = (z_1, \ldots, z_n) \in \mathbb{R}^n : z_i = m_i + t(M_i - m_i) \text{ for some } t \in [0, 1]\}.$$

Definition 5.2 Let $E \subseteq \mathbb{R}^n$. If $x, y \in E$, then we say *x is visible from y in E* if the line $L(x, y) \subseteq E$. If $A \subseteq E$, then A is *visually independent in E* if for any $x, y \in A$, x is not visible from y in E. E is *locally star-shaped* at x if there exists an open set O with $x \in O$ such that every point in $E \cap O$ is visible from x in S.

Theorem 5.3 Let $E \subseteq \mathbb{R}^n$ be closed. Then either there is an infinite subset $A \subseteq E$ such that A is visually independent in E or E is locally star-shaped at x for each $x \in E$.

Proof Assume that E is not locally star-shaped at some point $x \in E$. Thus for each $k \in \mathbb{N}$, the ball $B(x, \frac{1}{2^k})$ (centered at x with radius $\frac{1}{2^k}$) contains a point that is not visible from x in E. Call this point x_k; it is clear that the sequence $X = (x_k)_{k=1}^{\infty}$ converges to x. We 2-color the 2-subsets of X, where a pair $\{y, z\} \subseteq X$ is colored red if z is visible from y in E (or vice versa) and blue otherwise. By Theorem 4.9 there is a subset $A \subseteq X$ with $|A| = \aleph_0$ such that each pair in $[A]_2$ is red or such that each pair in $[A]_2$ is blue.

Assume first that $[A]_2$ is monochromatically red. We can write $A = \{x_{i_1}, x_{i_2}, \ldots\}$ where $i_1 < i_2 < \ldots$. For each $n \in \mathbb{N}$ we have that $L(x_{i_1}, x_{i_n}) \subseteq E$. The sequence $(x_{i_n})_{n=1}^{\infty}$ is a subsequence of X and therefore converges to x. Because E is closed, it follows that $L(x_{i_1}, x) \subseteq E$ as well. But this contradicts the fact that x_{i_1} was chosen as a point not visible from x in E. So it must be that $[A]_2$ is monochromatically blue, and therefore A is an infinite, visually independent subset. □

Another example of how we can apply Ramsey's Theorem to sequences to obtain a useful result comes from a lemma to the Bolzano-Weierstrass Theorem in analysis.

Theorem 5.4 (Bolzano-Weierstrass Theorem) Every bounded sequence $(x_n)_{n=1}^{\infty}$ in $\mathbb{R}^n$ has a convergent subsequence.

This theorem is easily proven once the following lemma is in hand.

Lemma 5.5 Every sequence $(x_n)_{n=1}^{\infty}$ in $\mathbb{R}$ has a monotone subsequence. That is, there is a subsequence $(x_{k_i})_{i=1}^{\infty}$ $(k_1 < k_2 < \ldots)$ for which $i < j$ implies $x_{k_i} \leq x_{k_j}$ for all i, j or for which $i < j$ implies $x_{k_i} \geq x_{k_j}$ for all i, j. (These conditions are that the subsequence is weakly increasing or weakly decreasing, respectively.)

In introductory analysis, Lemma 5.5 is typically proven by defining the notion of a "peak" of a sequence. But Ramsey's Theorem gives us an easier method.

Proof (of Lemma 5.5) We define a 2-coloring of $[(x_n)_{n=1}^{\infty}]_2$. If $i < j$, color the pair $\{x_i, x_j\}$ red if $x_i < x_j$ and blue if $x_i \geq x_j$. By Theorem 4.9 there is an infinite, homogeneous subset $A \subseteq \mathbb{N}$. If A is red-homogeneous, then it forms an infinite subsequence of $(x_n)_{n=1}^{\infty}$ which is (strictly) increasing. Similarly, if A is blue-homogeneous then it is a (weakly) decreasing subsequence. □

Theorem 5.4 can then be proven by finding a monotone subsequence in each coordinate and applying the fact that a bounded monotone subsequence converges. We have no need for this result, so we will leave it to the reader to fill in the details.

5.2 The Happy Ending Problem

Another classic result in Ramsey Theory actually relies only on the finite version of Ramsey's Theorem but is a powerful geometric fact about $\mathbb{R}^2$. We are interested in placing points down in the plane (in general position, such that no three points are collinear) in order to form convex n-gons. The following theorem tells us that there is a minimal number of points which guarantees the existence of a convex n-gon.

Theorem 5.6 (Erdős-Szekeres Theorem) Let $n \in \mathbb{N}$. There exists a minimal number $ES(n)$ such that if $N \geq ES(n)$ and $x_1, \ldots, x_N \in \mathbb{R}^2$ are distinct with no three collinear, then there exists a subset $\{x_{i_1}, \ldots, x_{i_n}\}$, $i_1 < \ldots < i_n$, that forms a convex n-gon in $\mathbb{R}^2$.

The proof is simple if we can first solve for a small value of $ES(n)$ and then prove a lemma relating the existence of convex N-gons to the existence of convex n-gons.

Lemma 5.7 $ES(4) = 5$.

Proof Any three non-collinear points in $\mathbb{R}^2$ form a triangle; put a fourth point in the interior of the triangle and we have four non-collinear points that do not form a convex 4-gon. Thus $ES(4) \geq 5$.

Next, take any set of 5 non-collinear points in the plane. Look at 3 of the 5 (call them A, B, C; they form a triangle. The fourth point D is either inside the triangle or outside; if it is outside, then A, B, C, D form a convex 4-gon, so assume it is inside. Now look at the final point E. If E is inside the triangle, then D, E form a convex 4-gon with 2 of A, B, C. If E is outside, enlarge the triangle so that E becomes a vertex; we now have 2 points enclosed in the triangle. But we just pointed out that in this case the two enclosed points form a convex 4-gon with 2 of the exterior points.

We have, therefore, that $ES(4) = 5$. $\square$

Lemma 5.8 Let $n \geq 3$ and let $x_1, \ldots, x_n \in \mathbb{R}^2$ be in general position. If every 4-subset $\Delta \in [\{x_1, \ldots, x_n\}]_4$ forms a convex 4-gon, then $\{x_1, \ldots, x_n\}$ forms a convex n-gon.

Proof The proof is by strong induction on n. The base cases of $n = 3$ and $n = 4$ are trivial.

For the induction step, assume we have $n+1$ points such that any 4 form a convex 4-gon. We have then that if we pick any n points of our $n+1$ total points those n form an n-gon. If adding the last point did *not* produce an $(n+1)$-gon, then the last point x_{n+1} is in the interior of the n-gon formed by $\{x_1, \ldots, x_n\}$. Fix a vertex x_i of the n-gon. This vertex defines a triangle along with $(n-2)$ of the edges (not counting the edges adjacent to x_i). These triangles span the entire interior of the n-gon, so x_{n+1} is in one of these triangles. This, however, violates our assumption that any 4 points form a convex 4-gon. □

Proof (of Theorem 5.6) It suffices, of course, just to find any positive integer that satisfies the condition; we claim that the Ramsey number $R_4(5, n)$ does the trick.

Let $N = R_4(5, n)$ points be placed in the plane. For every foursome of points $\delta = \{x, y, z, w\} \in [N]_4$, color δ red if it forms a convex 4-gon and blue if it does not. By our choice of N there is either a red-homogeneous set A with $|A| = 5$ or a blue-homogeneous set B with $|B| = n$. The former cannot happen by Lemma 5.5.

Thus there is a set B of n points, any 4 of which form a convex 4-gon, so by Lemma 5.6, B forms a convex n-gon. □

This problem is traditionally referred to as the "Happy Ending Problem" because the mathematicians Esther Klein and George Szekeres became close and eventually married after Klein proposed the problem to Szekeres in the 1930s. The original work they did was in the developing of Lemma 5.7, which is the impetus for the full Erdös-Szekeres Theorem, much like our $R(3, 3) = 6$ example from Section 1 is the impetus for the full finite version of Ramsey's Theorem. In fact, this problem is how Erdös and Szekeres became interested in Ramsey Theory and led to their development of the field.

5.3 Application to Topology

Not all Ramsey results directly utilize a version of Ramsey's Theorem. The *idea* of a "Ramsey" result is that if we take a large space and color (or partition) the space (or subsets of the space), the coloring produces a monochromatic or homogeneous subspace with a desirable property. The following discussion is topological rather than combinatorial or geometric, so we present it with less detail than we have to this point. It requires much buildup, but is a good example of the modern extensions of Ramsey Theory into other fields–here, general topology.

Definition 5.9 A metric space (X, d) is *separable* if there is a countable subset $D \subseteq X$ such that D is dense in X. (That is, $\mathrm{cl}_X(D) = X$.) (X, d) is *Polish* if it is separable and complete.

Definition 5.10 Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. A is *nowhere dense* in X if its closure has empty interior; i.e., if $\mathrm{int}_X(\mathrm{cl}_X(A)) = \emptyset$. A is *meager* or *first category* in X if

$$A = \bigcup_{k=1}^{\infty} A_k$$

where each A_k is nowhere dense in X. In other words, A is first category in X if it is the countable union of nowhere dense sets. If A is not first category in X, then A is *second category* in X.

Definition 5.11 Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. A point $x \in A$ is an *isolated point* in A if there is an open set O such that $O \cap A = \{x\}$. A nonempty set $A \subseteq X$ is *perfect* if it is closed and has no isolated points.

Definition 5.12 Let $(X, \mathcal{T})$ be a topological space. A subset $A \subseteq X$ is *Baire* in X if there exists an open set O such that the set

$$(A - O) \cup (O - A)$$

is first category in X.

Definition 5.13 Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$. The function f is *Baire measurable* if for every open set $O \subseteq Y$, the preimage $f^{-1}[O]$ is Baire in X.

If (X, d) is a metric space, we can put a topology on $[X]_k$ as follows. For any pairwise disjoint collection $\mathcal{U} = U_1, \ldots, U_k$ of open sets in X, let

$$B_{\mathcal{U}} = \{\{x_1, \ldots, x_k\} \in [X]_k : x_i \in U_i,\ 1 \leq i \leq k\}.$$

Define the base for the topology of $[X]_k$ as

$$\mathcal{B} = \bigcup \{B_{\mathcal{U}} : \mathcal{U} = U_1, \ldots, U_k \text{ is a pairwise disjoint collection of open sets}\}.$$

Theorem 5.14 (Galvin's Theorem) Let (X, d) be a Polish metric space with no isolated points and let $r \in \mathbb{N}$. Then if $f : [X]_2 \to [r]$ is a Baire measurable coloring, there is a perfect subset $A \subseteq X$ such that A is homogeneous under f.

We do not present this result for its importance in topology but rather as an example of what a continuous Ramsey theorem "looks like." In the next section we will give a more detailed exposition of a famous (but recent) development in continuous Ramsey Theory and its application to analysis.

6 Ramsey Theory in Banach Spaces

6.1 Banach, Hilbert Spaces and Banach's Question

In this section we will present a result that is a paradigm for the application of Ramsey-theoretic techniques to branches of mathematics other than combinatorics and set theory. It is entirely based on a recent paper of W. Timothy Gowers who presented [GOW] a result of a "Ramsey-theoretic nature," which he combined with previous results to give an answer to an open question of Banach.

Definition 6.1 Let V be a vector space with a norm $\|\cdot\|$. Define ρ to be a metric on V, where for any $x, y \in V$ we set $\rho(x, y) = \|x - y\|$. If the metric space (V, ρ) is complete, then we say (V, ρ) (or just V) is a *Banach space.*

Definition 6.2 Let V be a vector space over a field F and define a map $\langle \cdot, \cdot \rangle : V \times V \to F$ such that for any $x, y, z \in V$ and any $c \in F$:

i) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ iff $x = 0$.

ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, where $\overline{\langle \cdot, \cdot \rangle}$ denotes complex conjugation.

iii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

iv) $\langle cx, y \rangle = c\langle x, y \rangle$.

This map is called an *inner product* and the pair $(V, \langle \cdot, \cdot \rangle)$ (or just V) is an *inner product space.*

We can easily define a norm on an inner product space by letting $\|x\| = (\langle x, x \rangle)^{1/2}$. Thus an inner product induces a norm which induces a metric.

Definition 6.3 Let V be an inner product space. if V is a Banach space with respect to the induced norm, then V is a *Hilbert space*. That is, a Hilbert space is an inner product space that is complete with respect to the induced metric.

The quintessential examples of Hilbert spaces are the space ℓ_2 of square-summable sequences and the space L^2, in which the norm of a function f is given by the square root of its integral (over some measure space Ω) and with inner product

$$\langle f, g\rangle \equiv \int_\Omega f(x)\overline{g(x)}d\mu.$$

It is a known fact in analysis that a separable, infinite-dimensional Hilbert space is isomorphic to every infinite-dimensional closed subspace of itself. In 1932, Banach raised the question of whether such spaces represent the only class of Banach spaces that satisfy this condition. The answer was not known until Gowers answered it in 2002. Part of what made this feat so impressive was that he was able to reduce the problem to one which, in his words, has a "much more combinatorial flavor."

6.2 Connection to Ramsey Theory

In the following presentation we will omit the proofs of theorems that are standard results in analysis. Our goal is to understand how Ramsey Theory unifies these results, rather than to become experts in Banach space analysis.

Definition 6.4 If X is an infinite-dimensional Banach space, a *Schauder basis*, or simply a *basis*, for X is a sequence $(x_n)_{n=1}^\infty$ such that if $a \in X$ then there is a unique expression

$$a = \sum_{n=1}^{\infty} a_n x_n$$

for some sequence $(a_n)_{n=1}^\infty$ of scalars (and such that this sum converges with respect to the norm). If $\|x_n\| = 1$ for each n, then the basis $(x_n)_{n=1}^\infty$ is *normalized* and we will write the basis as $(e_n)_{n=1}^\infty$.

The difference between a Schauder basis and a typical basis for an infinite-dimensional vector space is that vectors in a space with a Schauder basis can be *infinite* sums with respect to the basis. The space of polynomials over $\mathbb{R}$, for example, has an infinite basis, but each polynomial in the space is a finite sum of scalar multiples of elements of the basis.

Definition 6.5 If X is a Banach space and $a = \sum_{n=1}^{\infty} a_n e_n \in X$, the *support* of a, denoted supp(a), is the set

$$\text{supp}(a) \equiv \{n \in \mathbb{N} : a_n \neq 0\}.$$

If $x, y \in X$ with $\max(\text{supp}(x)) < \min(\text{supp}(y))$ we write $x < y$. If $x_1, x_2, \ldots, x_n$ are vectors such that

$$x_1 < x_2 < \ldots < x_n < \ldots$$

we say $(x_n)_{n=1}^{\infty}$ is a *block basis* of X. The subspace spanned by a block basis is a *block subspace.*

It should be clear that the elements of a block basis are linearly independent. In fact, the condition $x_1 < \ldots < x_n$ is a much stronger one than linear independence.

We now have the terminology to begin the Ramsey-theoretic part of our discussion. Fix a Banach space X and define $\Sigma(X)$ to be the set of all finite sequences $(x_1, \ldots, x_n)$ with $x_1 < \ldots < x_n$ such that $\|x_i\| \leq 1$ for each i. We construct a "game" between Players A and B, played as follows:

- An arbitrary subset $\sigma \subseteq \Sigma(X)$ is chosen.
- During round n of the game, Player A selects an infinite-dimensional block subspace $X_n \subseteq X$.
- Player B then selects a point $x_n \in X_n$.
- The goal of Player B is, after some stage n, to have constructed a sequence $(x_1, \ldots, x_n) \in \sigma$. Player A desires that after no round n should the sequence of points $(x_1, \ldots, x_n)$ so far selected by Player B be in σ.

If $\Delta = (\delta_n)_{n=1}^{\infty}$ is a sequence of positive scalars, we define the Δ-*expansion* of a subset $\sigma \subseteq \Sigma(X)$, denoted σ_Δ, to be the set of sequences in $\Sigma(X)$ that are within δ_i of a sequence in σ for each i. More formally,

$$\sigma_\Delta \equiv \{(x_1, \ldots, x_n) \in \Sigma(X) : \text{there is some } (y_1, \ldots, y_n) \in \sigma \text{ such that } \|y_i - x_i\| \leq \delta_i \text{ for each } i\}.$$

We similarly define the Δ-*contraction* of σ, denoted $\sigma_{-\Delta}$, to be the set of sequences in $\Sigma(X)$ that are within at least (i.e., no closer than) δ_i of a sequence in $\Sigma(X) - \sigma$ for each i. That is,

$$\sigma_{-\Delta} \equiv \{(x_1, \ldots, x_n) \in \Sigma(X) : \text{every } (y_1, \ldots, y_n) \text{ satisfying} \\ \|y_i - x_i\| \leq \delta_i \text{ for each } i \text{ is in } \sigma\}.$$

More succinctly, $\sigma_{-\Delta} = \Sigma(X) - (\Sigma(X) - \sigma)_\Delta)$.

Lemma 6.6 For any $\sigma \subseteq \Sigma(X)$ and any sequence Δ of positive scalars, we have

$$(\sigma_{-\Delta})_\Delta \subseteq \sigma \subseteq (\sigma_\Delta)_{-\Delta}.$$

Proof The proof is just an exercise in understanding the definitions.

If $\alpha = (\alpha_1, \ldots, \alpha_n) \in (\sigma_{-\Delta})_\Delta$, then there is some $\beta = (\beta_1, \ldots, \beta_n) \in \sigma_{-\Delta}$ such that $\|\beta_i - \alpha_i\| \leq \delta_i$ for each i. But then by definition of $\sigma_{-\Delta}$ we must have $\alpha \in \sigma$.

Similarly, if $\alpha = (\alpha_1, \ldots, \alpha_n) \in \sigma$ and $\beta = (\beta_1, \ldots, \beta_n)$ satisfies $\|\beta_i - \alpha_i\| \leq \delta_i$ for each i, then $\beta \in \sigma_\Delta$ by definition so $\alpha \in (\sigma_\Delta)_{-\Delta}$. $\square$

Some more notation: if $A = (a_1, \ldots, a_m)$ is a sequence and $Y \subseteq X$ is a subspace we write $[A|Y]$ to denote the set of finite sequences $(y_n)_{n=1}^N$ such that $y_i = a_i$ for $1 \leq i \leq m$ and $y_i \in Y$ for $m < i \leq N$. In the case where $A = \emptyset$ we write $[Y]$ instead of $[\emptyset|Y]$. If $\sigma \subseteq \Sigma(X)$, we write $\sigma[A|Y]$ as the set of sequences $(a_1, \ldots, a_m, x_1, \ldots, x_N)$ such that $(a_1, \ldots, a_m, x_1, \ldots, x_N) \in \sigma \cap [A|Y]$.

The following is what makes this discussion relevant in a survey of Ramsey Theory.

Theorem 6.7 Let X be a Banach space with $\Sigma(X)$ as defined above, and let Δ be a sequence of positive scalars. Let $f : \Sigma(X) \to [2]$ be an arbitrary coloring (with, say, colors red and blue). Then there is a subspace $Y \subseteq X$ such that every finite block sequence in Y is red or such that if the game described above is played in Y with σ taken to be the Δ-expansion of all the blue sequences, then Player B can always win the game.

With the notation defined above we can phrase this differently (and perhaps more clearly). The coloring f is really just a selection of which finite block sequences to put into σ. When we refer to a player having a *winning strategy for the game* $\sigma[A|Y]$, we mean that the player can always win when $\sigma \subseteq \Sigma(X)$ is chosen and the subspaces picked by Player A all must be subspaces of Y. So the theorem guarantees, for any choice of $\sigma \subseteq \Sigma(X)$, the existence of a subspace Y such that either $\sigma[Y] = \emptyset$ or such that Player B has a winning strategy for the game $\sigma_\Delta[Y]$.

With respect to coloring the sequences in σ red and those not in σ blue, the theorem gives a subspace in which every finite block sequence is blue or that contains enough sequences that are "close enough" (with respect to Δ-expansion) to red sequences as to provide Player B with a sure-fire win. Indeed, since the Banach spaces we usually deal with are uncountable, this result is in contrast to some of our above failures in Section 4.4 and Section 4.5.

Proving this theorem will take a great deal of work but, theorem in hand, the answer to the question of Banach follows (relatively) quickly.

Definition 6.8 Let X be a Banach space with $A = (x_1, \ldots, x_n) \in \Sigma(X)$ and $Z \subseteq X$ an infinite-dimensional block subspace. The pair (A, Z) is a **-pair* (or just a *pair*) if $x_i < z$ for every $x_i \in A$ and $z \in Z$. If (A, Z) is a *-pair we will write $A < Z$.

More notation: If $\Delta = (\delta_i)_{i=1}^\infty$ and $A = (\alpha_1, \ldots, \alpha_n)$, $B = (\beta_1, \ldots, \beta_n)$ are block bases, we write $d(A, B) \leq \Delta$ to mean that $\|\beta_i - \alpha_i\| \leq \delta_i$ for each $1 \leq i \leq n$. If Π is a set of *-pairs, we will write Π_Δ to refer to the set of pairs (A, Z) such that $d(A, B) \leq \Delta$ for some block sequence B with $(B, Z) \in \Pi$.

If $x_1 < \ldots < x_n$ then $\langle x_1, \ldots, x_n \rangle$ will refer to the block subspace generated by $x_1, \ldots, x_n$. If we have a sequence $(x_1, \ldots, x_n) \in \Sigma(X)$, we write $\Sigma(X; x_1, \ldots, x_n)$ as the set of finite block sequences $(y_1, \ldots, y_m)$ for which each $y_i \in \langle x_1, \ldots, x_n \rangle$.

Definition 6.9 Let Υ be a set of block bases and let $\Delta = (\delta_i)_{i=1}^\infty$ be a sequence of positive scalars. A Δ-*net* of Υ is a set Γ of block bases such that for each $\upsilon \in \Upsilon$ there is some $\gamma \in \Gamma$ with $d(\upsilon, \gamma) \leq \Delta$.

Lemma 6.10 Let X be a Banach space and let $(\Delta_n)_{n=1}^\infty$, $(\Pi)_{n=1}^\infty$ be sequences such that each Δ_i is a sequence of positive scalars and each Π_i is a sequence of $*$-pairs. Assume further that the following conditions are satisfied:

(1) For any pair (A, Z) and any $n \in \mathbb{N}$ there is some $Z' \subseteq Z$ with the pair $(A, Z') \in \Pi_n$.

(2) If $(A, Z) \in \Pi_n$ and $Z' \subseteq Z$ then $(A, Z') \in \Pi_n$.

Then there is a subspace $Y \subseteq X$ such that for any block sequence $A \in Y^t$ with $t \geq n$ (that is, for any block sequence in Y of length n or greater) and any $Z \subseteq Y$, the pair $(A, Z) \in (\Pi_n)_{\Delta_n}$.

Proof We inductively pick vectors $y_1, y_2, \ldots$ and block subspaces $Y_0 \supseteq Y_1 \supseteq Y_2 \supseteq \ldots$ as follows.

- Set $Y_0 = X$.
- Once $y_1, \ldots, y_{n-1}$ and $Y_1, \ldots, Y_{n-1}$ have been selected, let $y_n \in Y_{n-1}$ be such that the sequence $(y_1, \ldots, y_n) \in \Sigma(X)$.
- Let $\{A_1, \ldots, A_m\}$ form a Δ_n-net of the set $\Sigma(X; y_1, \ldots, y_n)$. We can, by (1), form a sequence of subspaces

  $$Y_{n-1} \supseteq V_{11} \supseteq \ldots \supseteq V_{1n} \supseteq V_{21} \supseteq \ldots \supseteq V_{2n} \supseteq \ldots \supseteq V_{m1} \supseteq \ldots \supseteq V_{mn}$$

  such that $(A_i, V_{ij}) \in \Pi_j$ for each $1 \leq i \leq m$ and $1 \leq j \leq n$. We set $Y_n = V_{mn}$.

By (2) we get that for any i, j and for any $Z \subseteq Y_n$ it holds that $(A_i, Z) \in \Pi_j$. Because $\{A_1, \ldots, A_m\}$ is a Δ_n-net we get that for any $A \in \Sigma(X; y_1, \ldots, y_n)$ and any $Z \subseteq Y_n$, $(A, Z) \in (\Pi_j)_{\Delta_n}$ for each $1 \leq j \leq n$.

We claim that if we set $Y = \langle y_1, y_2, \ldots \rangle$, then Y is our desired subspace. To see this, let $Z \subseteq Y$ and $A \in Y^t$ with $t \geq n$. If (A, Z) is a $*$-pair then $A < Z$ so $A \in \langle y_1, \ldots, y_k \rangle^t$ and $Z \subseteq \langle y_{k+1}, y_{k+2}, \ldots \rangle \subseteq Y_k$ for some $k \geq m \geq n$. It follows from our construction of Y that $(A, Z) \in (\Pi_j)_{\Delta_n}$ for any $j \leq k$, and, in particular, $(A, Z) \in (\Pi_n)_{\Delta_n}$. $\square$

If the A in a pair (A, Z) is a singleton, then the notation of Π_Δ above extends to the case where $\Delta = \delta$ is a scalar, such that if Π is a set of $*$-pairs $(\{y\}, Z)$ then Π_δ is the set of pairs $(\{x\}, Z)$ such that $\|x - y\| \leq \delta$ for some y with $(\{y\}, Z) \in \Pi$.

Corollary 6.11 Let X be a Banach space, let δ be a positive scalar and let Π be a set of pairs $(\{y\}, Z)$ such that:

(1) For any pair $(\{y\}, Z)$ there is some $Z' \subseteq Z$ with the pair $(\{y\}, Z') \in \Pi$.

(2) If $(\{y\}, Z) \in \Pi$ and $Z' \subseteq Z$ then $(\{y\}, Z') \in \Pi$.

Then there is a subspace $Y \subseteq X$ such that for any $y \in Y$ and any $Z \subseteq Y$, the pair $(\{y\}, Z) \in \Pi_\delta$.

Proof We need only appeal to Lemma 6.9 in the appropriate fashion. Let each Δ_i be a sequence whose first term is δ. Let each Π_i be the set of all $*$-pairs except for Π_1 which includes a pair of the form $(\{y\}, Z)$ iff $(\{y\}, Z) \in \Pi$. □

Corollary 6.12 Under the hypotheses of Lemma 6.8 and with the additional assumption that

(3) If $(A, Z) \in \Pi_n$ then A has length n.

Then there is a subspace $Y \subseteq X$ such that for any $A \in Y^n$ (so A is a sequence of length *exactly* n) and any $Z \subseteq Y$, the pair $(A, Z) \in (\Pi_n)_{\Delta_n}$.

Proof Again, we simply apply Lemma 6.9. We replace each Π_n with Π_n' which is the set of all $*$-pairs except the pairs (A, Z) where A is of length n but $(A, Z) \notin \Pi_n$. □

Definition 6.13 Let X be a Banach space and let $Y \subseteq X$ be a block subspace. If $\sigma \subseteq \Sigma(X)$, we say σ is *large for* Y if every block subspace of Y contains a sequence in σ. We say σ is *strategically large for* Y if Player B has a winning strategy in the game $\sigma[Y]$.

Lemma 6.14 Let X be an infinite-dimensional Banach space with basis $(x_n)_{n=1}^\infty$. Then there exists a constant C such that for every $N \in \mathbb{N}$ and any $M > N$ we have

$$\Big\| \sum_{n=1}^{N} a_n x_n \Big\| \leq C \Big\| \sum_{n=1}^{M} a_n x_n \Big\|$$

for any sequence $(a_n)_{n=1}^\infty$ of scalars. See [DIL], [DAW], [LT].

Definition 6.15 The smallest such constant C, the existence of which is guaranteed by Lemma 6.14, is called the *basis constant* for the basis. In the case that $C = 1$ we say the basis is *monotone*.

It is an exercise in Banach space analysis to show that if X has a monotone basis $(x_n)_{n=1}^{\infty}$ then any finite sum $\sum_{i=1}^{n} \left\| a_n x_n \right\|$ satisfies

$$\sum_{i=1}^{n} \left\| a_n x_n \right\| \geq \frac{1}{2} \max_{1 \leq i \leq n} |a_i|.$$

The following theorem will help us prove Theorem 6.7 and bring us closer to answering Banach's question.

Theorem 6.16 Let X be a Banach space with a monotone basis $(x_n)_{n=1}^{\infty}$, let $\sigma \subseteq \Sigma(X)$, and let $\Gamma = (\gamma_n)_{n=1}^{\infty}$, $\Delta = (\delta_n)_{n=1}^{\infty}$ be sequences of positive scalars satisfying

$$2 \sum_{i=N}^{\infty} \delta_i \leq \gamma_N$$

for any N. If $\sigma_{-\Gamma}$ is large for X there is a block subspace $Y \subseteq X$ such that $\sigma_{2\Delta}$ is strategically large for Y.

Proof Suppose there is a subset $\sigma \subseteq \Sigma(X)$ for which the theorem did not hold. Then $\sigma_{-\Gamma}$ is large for X and $\sigma_{2\Delta}$ is not strategically large for any block subspace $Y \subseteq X$. Note that since $\sigma_{-\Gamma} \subseteq \sigma$ we have also that σ is large for X.

Define ζ to be the set of all sequences $(x_1, \ldots, x_n) \in \sigma$ such that if $y_1 < \ldots < y_k$ and $\langle y_1, \ldots, y_k \rangle \subsetneqq \langle x_1, \ldots, x_n \rangle$ then $(y_1, \ldots, y_k) \notin \sigma$. (So ζ is "minimal" under inclusion with respect to the subspaces generated by its sequences.) It is clear that ζ remains large for X and that $\zeta_{2\Delta}$ remains not strategically large for any block subspace of X. For $n \geq 0$ we set $\Delta_n = (\delta_1, \ldots, \delta_n, 0, 0, \ldots)$ and set $\Theta_n = 2\Delta - \Delta_n = (\delta_1, \ldots, \delta_n, 2\delta_{n+1}, 2\delta_{n+2}, \ldots)$.

Our next step is to define inductively a sequence of vectors $(x_n)_{n=1}^{\infty}$ and a sequence of subspaces $X = X_0 \supseteq X_1 \supseteq X_2 \supseteq \ldots$ with $x_n \in X_{n-1}$ and such that the following invariants are maintained:

(1) ζ_{Δ_n} is large for $[(x_1, \ldots, x_n)|X_n]$.

(2) For any $Z \subseteq X_n$, ζ_{Θ_n} is not strategically large for $[(x_1, \ldots, x_n)|Z]$.

We have $\zeta_{\Delta_0} = \zeta$ large for $X_0 = X$ and $\zeta_{\Theta_n} = \zeta_{2\Delta}$ not strategically large for any subspace $Z \subseteq X_0 = X$, so the induction holds at the base case.

To show that the process can continue infinitely, assume otherwise. That is, suppose that after $x_1, \ldots, x_n$ and $X_1, \ldots, X_n$ have been selected there are no available choices for x_{n+1} and X_{n+1}. This would mean that for every $x \in X_n$ and for every subspace $Y \subseteq X_n$ there is a subspace $Z \subseteq Y$ such that either (i) $\zeta_{\Delta_{n+1}}$ is not large for $[(x_1, \ldots, x_n, x)|X_n]$ (i.e., $\zeta_{\Delta_{n+1}} \cap [(x_1, \ldots, x_n, x)|Z] = \emptyset$) or (ii) $\zeta_{\Theta_{n+1}}$ is strategically large for $[(x_1, \ldots, x_n, x)|Z]$.

Define Π to be the set of $*$-pairs $(\{x\}, Z)$ with $x \in X_n$, $Z \subseteq X_n$ and such that one of (i), (ii) holds. We apply Corollary 6.11 with $\delta = \delta_{n+1}$; it is clear that conditions (1), (2) of the corollary both hold. That gives us a subspace $Y \subseteq X_n$ such that for any $y \in Y$ and $Z \subseteq Y$ we have $(\{x\}, Z) \in \Pi$ for any x satisfying $\|x - y\| \leq \delta_{n+1}$ and thus either (i) or (ii) holds for the pair $(\{x\}, Z)$. Because x and y are "close," if (i) holds it must be that $\zeta_{\Delta_n} \cap [(x_1, \ldots, x_n, y)|Z] = \emptyset$, and if (ii) holds it must be that ζ_{Θ_n} is strategically large for $[(x_1, \ldots, x_n, y)|Z]$. Taking $Z = Y$, we have either that

(i) $\zeta_{\Delta_n} \cap [(x_1, \ldots, x_n, y)|Y] = \emptyset$ or

(ii) ζ_{Θ_n} is strategically large for $[(x_1, \ldots, x_n, y)|Y]$.

Let $\Xi \equiv \{y \in X_n : \zeta_{\Delta_n} \cap [(x_1, \ldots, x_n, y)|Y] = \emptyset\}$. If there is a subspace Z of Y contained in Ξ, then it must be that $\zeta_{\Delta_n} \cap [(x_1, \ldots, x_n)|Z] = \emptyset$, which contradicts invariant (1) above. Therefore it must be the case that the set $\Xi' \equiv \{y \in X_n : \zeta_{\Theta_n}$ is strategically large for $[(x_1, \ldots, x_n, y)|Y]\}$ is large for Y. But this contradicts invariant (2). Hence we have shown that process of selecting x_i and X_i does not fail at any stage.

Let $X' = \langle x_1, x_2, \ldots \rangle$ and let $(y_1, \ldots, y_k)$ be a finite block sequence in the subspace $\langle x_1, \ldots, x_n \rangle \subseteq X'$. By (1) above we can extend $(x_1, \ldots, x_n)$ to $(x_1, \ldots, x_n, \ldots, x_m) \in \zeta_\Delta$. Thus there is a sequence $(x'_1, \ldots, x'_m) \in \zeta$ with $\|x'_i - x_i\| \leq \delta_i$ for each i. Let $(y'_1, \ldots, y'_k)$ be such that $y'_i - y_i = x'_i - x_i$ for each i.

By definition of ζ we have that $(y'_1, \ldots, y'_k) \notin \sigma$ (ζ is minimal). But we also have that $\|y'_i - y_i\| \leq \Gamma_i$ for each $i \leq k$, proving that $(y_1, \ldots, y_k) \notin \sigma_{-\Gamma}$. Therefore $X' \cap \sigma_{-\Gamma} = \emptyset$, contradicting the fact that $\sigma_{-\Gamma}$ is large for X. □

Theorem 6.7 quickly follows.

Proof (of Theorem 6.7) It suffices to show that for any $\sigma \subseteq \Sigma(X)$ there is a subspace $Y \subseteq X$ such that $\sigma[Y] = \emptyset$ or such that Player B has a winning strategy in the game $\sigma_\Delta[Y]$.

Define $\theta = \sigma_{\Delta/2}$. By Lemma 6.6 we have that $\sigma \subseteq \theta_{-\Delta/2}$; it follows hence that $\theta_{-\Delta/2}$ is large for X. We invoke Theorem 6.16 with $\Gamma = \Delta/2$ to get a subspace $Y \subseteq X$ such that θ_Γ is strategically large for Y. Since $\theta_\Gamma \subseteq \sigma_\Delta$ we have our desired result. □

We introduce some more definitions and some more important results from Banach space analysis.

Definition 6.17 Let $(x_n)_{n=1}^\infty$ be a basis for a Banach space X. If $a = \sum_{n=1}^{\infty} a_n x_n \in X$, we say that this sum *converges unconditionally* if for any permutation $\tau : \mathbb{N} \to \mathbb{N}$ the sum $\sum_{n=1}^{\infty} a_{\tau(n)} x_{\tau(n)}$ converges as well. If the sum for each $a \in X$ converges unconditionally whenever it converges then the basis $(x_n)_{n=1}^\infty$ is said to be *unconditional*.

6.3 Application to Banach's Conjecture

Definition 6.18 We say a Banach space is *homogeneous* if it is isomorphic to all of its subspaces.

Theorem 6.19 Let X be a homogeneous Banach space. Then either X is isomorphic to the space ℓ_2 or X has no unconditional basis.

Definition 6.20 A Banach space X is *decomposable* if it can be written as a direct sum $X = Y + Z$ with continuous projections $\pi_Y : X \to Y$ and $\pi_Z : X \to Z$. If no such expression exists, then X is *indecomposable*. X is *hereditarily indecomposable* if every subspace of X is indecomposable.

Theorem 6.21 If a Banach space X is hereditarily indecomposable then X is not isomorphic to any proper subspace $Y \subsetneq X$.

The preceding theorem implies that if X is hereditarily indecomposable it is not homogeneous. The main theorem of Gowers, and the one which answers Banach's question, is the following. The dichotomy it presents is very much Ramseyian.

Theorem 6.22 Let X be a Banach space. There is a subspace $Y \subseteq X$ such that Y has an unconditional basis or such that Y is hereditarily indecomposable.

There is still some work to be done in proving this theorem, but we are close. Once we have the proof we will show how this leads us to answer Banach's question.

Lemma 6.23 Let X be a Banach space. Then X is hereditarily indecomposable if and only if for any two block subspaces $Y, Z \subseteq X$ and every scalar $C \in \mathbb{R}$ there is a sequence

$$y_1 < z_1 < y_2 < z_2 < \ldots < y_n < z_n,$$

with $y_i \in Y$ and $z_i \in Z$ for each i, such that

$$\Big\| \sum_{i=1}^{n} (y_i + z_i) \Big\| > C \Big\| \sum_{i=1}^{n} (y_i - z_i) \Big\|.$$

Lemma 6.24 Let X be a Banach space. Then the following are equivalent:

(a) No subspace $Y \subseteq X$ has an unconditional basis.

(b) If $Y \subseteq X$ is a block subspace and $C \in \mathbb{R}$ then there is a sequence $y_1 < \ldots < y_n$ with each $y_i \in Y$ such that

$$\Big\| \sum_{i=1}^{n} y_i \Big\| > C \Big\| \sum_{i=1}^{n} (-1)^i y_i \Big\|.$$

Definition 6.25 We say a sequence $x_1 < \ldots < x_n$ in a Banach space X is *C-conditional* block sequence if

$$\Big\| \sum_{i=1}^{n} x_i \Big\| > C \Big\| \sum_{i=1}^{n} (-1)^i x_i \Big\|.$$

An infinite block basis $(x_i)_{i=1}^{\infty}$ is *C-unconditional* if it generates a subspace containing no C-conditional sequence. X is *C-hereditarily indecomposable* if for any two block subspaces $Y, Z \subseteq X$ there exist $y \in Y, z \in Z$ such that

$$\|y + z\| > C\|y - z\|.$$

Lemma 6.26 Let X be a Banach space with a monotone basis and fix $C \in \mathbb{R}$. Then X either has a C-unconditional block basis or else, for every $\epsilon > 0$, X has a $(C - \epsilon)$-hereditarily indecomposable block subspace.

Proof Assume that X has no C-unconditional block basis; that is, every subspace of X has a C-conditional sequence, and let $\epsilon > 0$. Define $\sigma \subseteq \Sigma(X)$ to contain all sequences $(x_1, \ldots, x_n)$ that are C-conditional and that contain at least one vector x_i with $\|x_i\| = 1$. If Y is any subspace of X then Y has a C-conditional sequence $(y_1, \ldots, y_k)$. Let $\alpha = \max\{\|y_i\| : 1 \leq i \leq k\}$; then the sequence $\frac{1}{\alpha}(y_1, \ldots, y_k) \in \sigma$. So σ is large for X.

Therefore, by Theorem 6.7, for any sequence $\Delta = (\delta_n)_{n=1}^{\infty}$ of positive scalars there is a block subspace $W \subseteq X$ such that σ_Δ is strategically large for W. Choose Δ such that $\sum_{i=1}^{\infty} \delta_i = \xi$ for some $\xi > 0$ satisfying

$$\frac{C - 2\xi}{1 + 2\xi} \geq C - \epsilon.$$

Let $Y, Z \subseteq W$ be disjoint block subspaces. Since σ_Δ is strategically large for W, Player B can win against any strategy played by Player A, including the one in which Player A alternates selecting the subspaces Y and Z. Player B wins by selecting a sequence $(y_1, z_1, y_2, z_2, \ldots, y_n, z_n) \in \sigma_\Delta$ such that for each i, $y_i \in Y$ and $z_i \in Z$. (Ignore the technical issue that the game may end after an odd number of turns.)

This means that there is a sequence $(y_1', z_1', \ldots, y_n', z_n') \in \sigma$ with $\|y_i' - y_i\| \leq \delta_{2i-1}$ and $\|z_i' - z_i\| \leq \delta_{2i}$ for each $1 \leq i \leq n$. Because the basis for X is monotone we have both that

$$\Big\|\sum_{i=1}^{n}(y_i' + z_i')\Big\|, \Big\|\sum_{i=1}^{n}(y_i' - z_i')\Big\| \geq \frac{1}{2}.$$

This implies that

$$\Big\|\sum_{i=1}^{n}(y_i' + z_i')\Big\| > C\Big\|\sum_{i=1}^{n}(y_i' - z_i')\Big\| \geq \frac{C}{2}.$$

By our choice of Δ we have

$$\Big\|\sum_{i=1}^{n}(y_i + z_i)\Big\| \geq \Big\|\sum_{i=1}^{n}(y_i' + z_i')\Big\| - \xi$$

and

$$\Big\|\sum_{i=1}^{n}(y_i - z_i)\Big\| \geq \Big\|\sum_{i=1}^{n}(y'_i - z'_i)\Big\| + \xi.$$

We put everything together to obtain

$$\Big\|\sum_{i=1}^{n}(y_i + z_i)\Big\| > (C - \epsilon)\Big\|\sum_{i=1}^{n}(y_i - z_i)\Big\|.$$

This proves that W is $(C - \epsilon)$-hereditarily indecomposable. □

Proof (of Theorem 6.22) Assume X has no subspace with an unconditional basis. Then by Lemma 6.26 we know that every block subspace of X has a C-unconditional block sequence for any $C \in \mathbb{R}$. By the previous lemma we can find a sequence

$$W_1 \supseteq W2 \supseteq \dots$$

of block subspaces of X where each subspace W_k is k-hereditarily indecomposable.

Choose a block basis $(w_n)_{n=1}^{\infty}$ of X with $w_n \in W_n$ and let $W = \langle w_1, w_2, \dots \rangle$. Let Y and Z be block subspaces of W and let $C \in \mathbb{R}$. Let $n \geq C$ be a positive integer; both subspaces $Y \cap W_n$ and $Z \cap W_n$ are infinite-dimensional. Also, since W_n is n-hereditarily indecomposable we can find a sequence

$$y_1 < z_1 < \dots < y_n < z_n$$

that satisfies the condition in Lemma 6.23, so W is hereditarily indecomposable. □

The two possibilities in Theorem 6.22 are mutually exclusive, since if Y is a subspace with an unconditional basis it follows that Y itself is indecomposable. Putting Theorem 6.22 together with Theorem 6.21, we get that if X is homogeneous then it has an unconditional basis. For if:

X has no unconditional basis, then

⇒ No subspace of X has an unconditional basis, since X is homogeneous

⇒ There is a subspace of X that is hereditarily indecomposable, by Theorem 6.22

$\Rightarrow X$ is hereditarily indecomposable, since X is homogeneous

$\Rightarrow X$ is *not* homogeneous, by Theorem 6.21.

By Theorem 6.19, then, since X has an unconditional basis it is isomorphic to ℓ_2, a Hilbert space. This allows us to conclude at last with the answer to Banach's long-standing open problem.

Theorem 6.27 Up to isomorphism, a separable Hilbert space is the only infinite-dimensional Banach space that is isomorphic to every infinite-dimensional closed subspace of itself.

This lengthy discussion is quite fascinating in that it relies heavily on theorems of a combinatorial nature that have, as Gowers puts it, "little to do with the problem." This should speak highly of the import of Ramsey Theory and Ramseyian techniques, if they have been extended even to this realm. Again, ultimately the Ramsey results here obtained (vis-a-vis the coloring of finite block sequences and how that speaks to the outcome of our "game") may not be interesting in their own right, but when combined with other known non-Ramsey results and utilized in the right way they form the basis for a result on the cutting edge of mathematics.

7 Odds and Ends

7.1 Semigroups

The point we have been trying to make throughout our presentation of various topics in Ramsey Theory is that there are applications of Ramsey Theory to be examined in almost every branch of mathematics. While the primitives of Ramsey Theory stem from combinatorial set theory, we have extended the ideas to prove results in number theory, linear algebra, graph theory, geometry, topology and analysis. We conclude our discussion with two additional applications of Ramsey Theory to the fields of abstract algebra and mathematical logic.

Definition 7.1 A *semigroup* is a pair $(G, \cdot)$ where G is a set and $\cdot : G \times G \to G$ such that for any $a, b, c \in G$,

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

That is, the binary operation $\cdot$ is associative.

Of course, if we demand the existence of an identity element e and a unique inverse for each element $a \in G$ we get a group, but we will confine our discussion to semigroups. As usual, we will write $a \cdot b$ as ab and the semigroup $(G, \cdot)$ as G.

Definition 7.2 Let G be a semigroup. An *idempotent* is an element $x \in G$ such that $x^2 = xx = x$.

It is easy to show that there is exactly one idempotent in a group–namely, the identity element e. There is no a priori reason to believe that there necessarily should be an idempotent in a semigroup, but Ramsey Theory allows us to prove that there is.

Theorem 7.3 A finite semigroup G has an idempotent.

Proof Assume $|G| = n$. Set $t = R(3; n)$ and take an arbitrary sequence $(x_1, x_2, \ldots, x_t)$ where each $x_i \in G$. We define a coloring $\chi : [t] \to [n]$ where the set $[t]$ can be viewed as the elements of our sequence and the set $[n]$ can be viewed as the elements of G. For $1 \leq i < j \leq t$ we define

$$\chi(\{i,j\}) = \chi(\{x_i, x_j\}) = x_{i+1} \cdot x_{i+2} \cdot \ldots \cdot x_j = \prod_{\alpha=i+1}^{j} x_\alpha.$$

By our choice of t we have a monochromatic "triangle;" that is, there are $1 \leq i < j < k \leq t$ with $\chi(\{x_i, x_j\}) = \chi(\{x_j, x_k\}) = \chi(\{x_i, x_k\})$. That means

$$\prod_{\alpha=i+1}^{j} x_\alpha = \prod_{\alpha=j+1}^{k} x_\alpha = \prod_{\alpha=i+1}^{k} x_\alpha = x$$

for some $x \in G$. But $\left(\prod_{\alpha=i+1}^{j} x_\alpha\right)\left(\prod_{\alpha=j+1}^{k} x_\alpha\right) = \prod_{\alpha=i+1}^{k} x_\alpha$, so we have $xx = x^2 = x$. □

7.2 The Paris-Harrington Theorem

For our final application, we present a short example of how Ramsey-theoretic propositions can be on the edge of provability.

Definition 7.4 A subset $A \subseteq \mathbb{N}$ is *large* if $|S| \geq \min S$.

So $\{2, 87, 10^{10^{10}}\}$ and $\{3, 4, 5\}$ are large but $\{5, 6\}$ is not. Note that this notion of a subset being large is completely independent of the one we discussed in Section 6 regarding finite sequences in infinite-dimensional Banach spaces.

We extend our arrow notation from Section 2 and write

$$m \rightrightarrows (n)^r_k$$

to mean that any r-coloring of the set $[\{n, n+1, \ldots, m\}]_k$ yields a monochromatic large set.

Theorem 7.5 (Paris-Harrington Theorem) For all $n, r, k \in \mathbb{N}$ there exists $m \in \mathbb{N}$ such that $m \rightrightarrows (n)^r_k$.

Proof Let $n, r, k \in \mathbb{N}$ be given. Define $\mathcal{A}$ to be the collection of finite large sets $A \subseteq \{n, n+1, \ldots\}$. That is,

$$\mathcal{A} \equiv \{A \subseteq \mathbb{N} : n \leq \min A \leq |A| < \infty\}.$$

Let $f : [\{n, n+1, \ldots\}]_k \to [r]$ be a coloring; by Theorem 4.14 there is an infinite homogeneous set $B \subseteq \{n, n+1, \ldots\}$. Let $\alpha = \min B$ and let Δ be the set consisting of the first α terms of B. Then $\Delta \in \mathcal{A}$ is large and monochromatic.

The fact that there is actually a finite m for which this holds on colorings of $\{n, n+1, \ldots, m\}$ follows from the Compactness Principle (see Section 4.3). $\square$

The technical theory needed to prove the following result is beyond the scope of this paper, but we refer the interested reader to [GRS] for a full exposition.

Theorem 7.6 Theorem 7.5 is unprovable with the axioms of Peano arithmetic.

It should be clear by now that whatever field of mathematics one is interested in studying, there is some interesting and nontrivial mathematics at the intersection of that field and Ramsey Theory.

ACKNOWLEDGEMENTS

I would like to thank Russ Woodroofe of the Washington University in St. Louis Department of Mathematics for his guidance in this project and his constant push for me not only to absorb the results I encountered but also to be always thinking creatively toward how known results can be extended to new mathematics.

I would also like to thank John Shareshian for being a part of the honors committee and Blake Thornton and Ron Freiwald for spurring me throughout my undergraduate career to become a better mathematician.

Finally, I thank the entire Department of Mathematics of Washington University in St. Louis for providing me with a foundation for my fledgling career as a mathematician.

Jonathan J. Marshall
Saint Louis, Missouri
March 2010

REFERENCES

[BAK] Baek, Jongmin. *Introduction to Infinite Ramsey Theory.* Available on personal website, Dept. of Computer Science, Stanford Univ. Accessed October 2010.

[BAN] Banach, Stefan. *Théorie des Opérations Linéaires*, Éditions Jacques Gabay, Sceaux, 1993 (Reprint of 1932 original).

[BON] Bóna, Miklós. *A Walk Through Combinatorics.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition, 2006. An introduction to enumeration and graph theory. With a foreword by Richard Stanley.

[DAW] Daws, Matt. *Introduction to Bases in Banach Spaces.* Available on personal website, Mathematics Dept., Univ. of Leeds. Accessed March 2011.

[DIL] Dilworth, S.J., Odell, E., Schlumprecht, Th., and Zsák, András. On the convergence of greedy algorithms for initial segments of the Haar basis, *Math. Proc. Cambridge Philos. Soc,* 148:519-529, 2010.

[END] Enderton, Herbert B. *Elements of Set Theory.* Academic Press [Harcourt Brace Jovanovich Publishers], New York 1977.

[FRE] Freiwald, Ronald. *An Introduction to Set Theory and Topology.* Unpublished, 2009.

[GSC] Geschke, Stefan. Convex decompositions in the plane and continuous pair colorings of the irrationals, *Israel J. Math.*, 131:285-317, 2002.

[GOW] Gowers, An infinite Ramsey theorem and some Banach-space dichotomies. *Ann. of Math. (2)*, 156(3):797-833, 2002.

[GHM] Graham, Ronald L. *Rudiments of Ramsey Theory*, volume 45 of *CBMS Regional Conference Series in Mathematics.* American Mathematical Society, Providence, R.I., 1981.

[GR] Graham, R.L. and Rothschild, B.L. A short proof of van der Waerden's theorem on arithmetic progressions. *Proc. Amer. Math. Soc.*, 42:385-386, 1974.

[GRS] Graham, Ronald L., Rothschild, Bruce L., and Spencer, Joel H. *Ramsey Theory.* John Wiley & Sons Inc., New York, 1980. Wiley-Interscience Series in Discrete Mathematics, A Wiley-Interscience Publication.

[GG] Greenwood, R.E. and Gleason, A.M. Combinatorial relations and chromatic graphs. *Canad. J. Math*, 7:1-7, 1955.

[HH] Hajnal, András and Hamburger, Peter. *Set Theory*, volume 48 of *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge, 1999. Translated from the 1983 Hungarian original by Attila Máté.

[HOF] Hoffman, Paul. *The Man Who Only Loved Numbers.* Hyperion, New York. 1998.

[JEC] Jech, Thomas. *Set Theory*, Academic Press, New York, 1978.

[LEV] Levy, Azriel. *Basic Set Theory.* Springer-Verlag, Berlin, 1979.

[LT] Lindenstrauss, Joram and Tzafriri, Lior. *Classical Banach Spaces I: Sequence Spaces.* Springer-Verlag, Berlin, 1977.

[RAD] Radziszowski, Stanisław P. Small Ramsey numbers, *Electron. J. Combin.*, vol. 1 (1994), Dynamic Survey 1, 30 pp. (electronic).

[RAM] Ramsey, Frank P. On a problem of formal logic, *Proc. London Math. Soc.* series 2, vol. 30 (1930), pp.264-286.

[RUD] Rudin, Walter. *Real and Complex Analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[WIL] Wilson, Robin J. *Introduction to Graph Theory.* Longman, New York, third edition, 1985.

[WD1] Woodroofe, Russ. Lecture notes, 9 April 2010.

[WD2] Woodroofe, Russ. Personal correspondence, 15 March 2011.