

# An Introduction to Ramsey Theory on Graphs

James Odziemiec Dickson

Thesis submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Master of Science  
in  
Mathematics

Ezra A. Brown, Chair  
Martin Klaus  
Nicholas A. Loehr  
Mark M. Shimosono

May 2, 2011  
Blacksburg, Virginia

Keywords: Ramsey Theory, Graph Theory, Combinatorics

Copyright 2011, James O. Dickson

# An Introduction to Ramsey Theory on Graphs

James O. Dickson

(ABSTRACT)

Ramsey theory deals with finding order amongst apparent chaos. Given a mathematical structure of interest and a setting where it may appear, Ramsey theory strives to identify conditions on this setting under which our mathematical structure of interest *must* appear. On a clear night one can look into the sky and observe patterns amongst the stars. Suppose we are interested in finding a constellation of 10 stars forming a cup shape. Are we guaranteed to find this constellation if we can see 100 stars, 1000 stars, or infinitely many stars? The mathematical generalization of this question has been named the Happy End problem after its proposer and solver, Esther Kline and George Szekeres who married shortly after solving it (the role of the proposer being reversed!). At the time Esther and George's mathematics circle included Paul Erdos who became the most prolific mathematical author ever and the leading exponent of Ramsey theory. When Erdos lectured about Ramsey theory on graphs he drew in his audience with two problems. The first problem has been named the Party problem. Given 6 people who have been invited to a party can we always find a subset of 3 people all of whom know each other or all of who do not know each other? The problem is equivalent to asking if every coloring of the edges of the complete graph on 6 vertices in the colors maroon and orange contains a subgraph of 3 vertices for which the edges running between these vertices are either all maroon or all orange. The least number of vertices on which the complete graph on these vertices guarantees such a set of 3 vertices is denoted  $R(3,3)$ . Ramsey type problems typically include some form of partitioning. In the above example we partitioned the pairs of invitees into 2 sets, those pairs who knew each other and those pairs that did not. Then we asked if we could find 3 pairs in either of the partitions with the property that they formed a triangle of 3 people. In a typical Ramsey problem we not only insist that our object of interest appear as a substructure of some superstructure, we ask: how large must our superstructure be so that no matter how we partition it into a given

number of parts, one of the partitions contains the desired substructure? Erdos's second problem asks us to pretend we have encountered aliens who will destroy us unless we can tell them  $R(5,5)$ , what should we do, what if they asked for  $R(6,6)$ ? For  $R(5,5)$  Erdos says that all mathematicians and computers should work together to find the solution. For  $R(6,6)$  Erdos recommends trying to figure out a way to destroy the aliens before they destroy us! It has been shown that  $R(6,6)$  is at least 102. Before considering symmetries there are  $2^{102}$  or more than  $10^{30}$  graphs on 102 vertices. The number of atoms in the observable universe is estimated as less than  $10^{80}$ . The computational power required to obtain  $R(6,6)$  by brute force may never be available, the upper bound of 165 requires checking nearly  $10^{50}$  graphs. It is fascinating that we even have upper bounds given our computational shortcomings. But even more fascinating is that Ramsey theory gives us upper bounds for  $R(m,n)$  for any natural numbers  $m$  and  $n$ . It is these techniques that drew me to this topic and which I hope to relate. All sorts of mathematical weaponry have been brought to bear on Ramsey theory: constructive methods, computer algorithms, random graphs and the probabilistic method. Despite the difficulty of classical Ramsey theory the beauty of the philosophy behind it has led mathematicians to other elegant areas: Euclidean Ramsey theory, the problem of the chromatic number of the plane, Schur's theorem, van Der Waerden's theorem, the Hales-Jewett theorem, and other results in extremal graph theory are critical parts in the growing Ramsey theory.

# Dedication

This work is firstly dedicated to my little brother Will. Your potential is unbounded. To my twin sister Sara, I would not be where I am today if I had not grown up along side you. To my mother Camille and father John, your dedication to family is inspiring. My grandmother, Nana, your positive influence has been ever present. I love you guys.

# Acknowledgments

I wish to acknowledge Dr. Ezra Brown for his patience, guidance, and support without which this thesis could not have been written. Bud's mentoring and sage advice is invaluable. Thank you B<sup>2</sup>!

I also wish to acknowledge the useful corrections and encouragement I received from my committee members: Martin Klaus, Nick Loehr, and Mark Shimosono.

Thanks to the Math Department of Virginia Tech for the support, financial and otherwise.

Andrew Wills and Maria Beane your compatriotism, assistance, and advice have made this process not only bearable but fun.

Michael Stewart and Jennifer Pavlak your friendship has been awesome during my time at Virginia Tech.

Laurie Heyer, Sarah Mason, Tim Chartier, Irl Bivens, Rich Niedinger, John Swallow, and Robert Whitton from Davidson College's math department: your passion for math inspired me to pursue graduate school.

# Contents

- 1 Introduction** **1**
- 1.1 Introduction . . . . . 1
- 1.2 Preliminaries . . . . . 3
- 1.3 Frank P. Ramsey . . . . . 4
- 1.4 Ramsey’s Theorem . . . . . 7
- 1.5 Generalized and Infinite versions of Ramsey’s Theorem . . . . . 11
  - 1.5.1 More Colors . . . . . 11
  - 1.5.2 Arbitrary Graphs . . . . . 12
  - 1.5.3 Hypergraphs . . . . . 13
  - 1.5.4 Infinite version of Ramsey’s theorem . . . . . 13
  - 1.5.5 Compactness . . . . . 15
  - 1.5.6 Density . . . . . 15
  - 1.5.7 Compactness Continued . . . . . 18
- 1.6 Paul Erdős . . . . . 20
  
- 2 Ramsey Numbers** **24**

2.1	Exact Values . . . . .	24
2.2	Erdős' lower bound . . . . .	34
2.3	Lovasz Local Lemma . . . . .	35
2.4	Diagonal Ramsey numbers . . . . .	37
2.5	A recursive upper bound for diagonal Ramsey numbers . . . . .	39
2.6	Comparing near diagonal Ramsey numbers . . . . .	41
2.7	$R(3, k)$ . . . . .	43
2.8	Combinatorial Games and Ramsey Numbers . . . . .	46
<b>3</b>	<b>Chromatic number of the plane</b>	<b>52</b>
3.1	Vertex Coloring . . . . .	52
3.2	Coloring the plane . . . . .	54
3.3	Generalizations of the chromatic number of the plane . . . . .	60
<b>4</b>	<b>Ramsey-type theorems for the Integers</b>	<b>64</b>
4.1	Schur's theorem . . . . .	64
4.2	Van der Waerden's theorem . . . . .	69
4.3	Rado's theorem . . . . .	72
4.4	Hilbert's Cube Lemma . . . . .	74
4.5	The happy end problem . . . . .	75
4.6	Conclusion . . . . .	76
	<b>Bibliography</b>	<b>77</b>

# List of Figures

1.1	$R(3, 3) > 5$ . . . . .	7
2.1	$R(3, 4) > 8$ . . . . .	25
2.2	$R(3, 5) > 13$ . . . . .	26
2.3	$R(4, 4) > 17$ . . . . .	27
2.4	$R(3, 3, 3) > 16$ . . . . .	51
3.1	Moser's Spindle . . . . .	56
3.2	Golomb graph . . . . .	57
3.3	Square tiling . . . . .	57
3.4	Hexagon tiling . . . . .	58
3.5	Stechkin's example . . . . .	61



# Chapter 1

## Introduction

### 1.1 Introduction

Ramsey theory deals with finding order amongst apparent chaos. Given a mathematical structure of interest and a setting where it may appear, Ramsey theory strives to identify conditions on this setting under which our mathematical structure of interest *must* appear. On a clear night one can look into the sky and observe patterns amongst the stars. Suppose we are interested in finding a constellation of 10 stars forming a cup shape. Are we guaranteed to find this constellation if we can see 100 stars, 1000 stars, or infinitely many stars? The mathematical generalization of this question has been named the Happy End problem after its proposer and solver, Esther Kline and George Szekeres who married shortly after solving it (the role of the proposer being reversed!). At the time Esther and George's mathematics circle included Paul Erdős, who became the most prolific mathematical author ever and the leading proponent of Ramsey theory.

When Erdős lectured about Ramsey theory on graphs he drew in his audience with two problems. The first problem has been named the Party problem. Given 6 people who have been invited to a party, can we always find a subset of 3 people all of whom know each other

or all of whom do not know each other? The problem is equivalent to asking if every coloring of the edges of the complete graph on 6 vertices in the colors maroon and orange contains a subgraph of 3 vertices for which the edges running between these vertices are either all maroon or all orange. The least number of vertices on which the complete graph on these vertices guarantees such a set of 3 vertices is denoted  $R(3,3)$ .

Ramsey-type problems typically include some form of partitioning. In the above example we partitioned the pairs of invitees into 2 sets, those pairs who knew each other and those pairs that did not. Then we asked if we could find 3 pairs in either of the partitions with the property that they formed a triangle of 3 people. In a typical Ramsey problem we not only insist that our object of interest appear as a substructure of some superstructure, we ask: how large must our superstructure be so that no matter how we partition it into a given number of parts, one of the partitions contains the desired substructure?

Erdős's second problem asks us to pretend we have encountered aliens who will destroy us in six months unless we can tell them the exact value of  $R(5,5)$ , what should we do, what if they asked for  $R(6,6)$ ? For  $R(5,5)$  Erdős says that all mathematicians and computers should work together to find the solution. For  $R(6,6)$  Erdős recommends trying to figure out a way to destroy the aliens before they destroy us! It has been shown that  $R(6,6)$  is at least 102. Before considering symmetries there are  $2^{102}$  or more than  $10^{30}$  graphs on 102 vertices. The number of atoms in the observable universe is estimated as less than  $10^{80}$ . The computational power required to obtain  $R(6,6)$  by brute force may never be available, as the upper bound of 165 requires checking nearly  $10^{50}$  graphs.

It is fascinating that we even have upper bounds given our computational shortcomings. But even more fascinating is that Ramsey theory gives us upper bounds for  $R(m, n)$  for any natural numbers  $m$  and  $n$ . It is these techniques that drew me to this topic and which I hope to relate. All sorts of mathematical weaponry have been brought to bear on Ramsey theory: constructive methods, computer algorithms, random graphs and the probabilistic method. Despite the difficulty of classical Ramsey theory the beauty of the philosophy behind it has

led mathematicians to other elegant areas: Euclidean Ramsey theory, the problem of the chromatic number of the plane, Schur's theorem, van der Waerden's theorem, the Hales-Jewett theorem, and other results in extremal graph theory are critical parts in the growing Ramsey theory.

This thesis is broken into four chapters. The first chapter discusses Ramsey's theorem, its proof and its generalizations. The second chapter covers Ramsey numbers, the known exact values and the known upper and lower bounds. The third chapter is devoted to the chromatic number of the plane problem. The final chapter summarizes known results for Ramsey-type theorems dealing with the integers.

## 1.2 Preliminaries

A *graph*  $G = (V(G), E(G))$  is a pair of sets, a vertex set and an edge set. A *vertex*  $v$  is drawn as a point and an *edge*  $e = uv$  is drawn as an arc connecting the vertices  $u$  and  $v$ .  $G$  will always represent a graph.

Vertices  $u$  and  $v$  are *adjacent* in  $G$  if  $uv \in E(G)$ . An *independent set*  $S$  is a set of pairwise nonadjacent vertices in  $G$ . If  $k$  vertices form an independent set this is called a *k-IS*. Vertex  $v$  is *incident* with edge  $e$  if one of  $e$ 's endpoints is  $v$ .

Let  $[n] = \{1, 2, \dots, n\}$  denote the set containing the first  $n$  natural numbers. The *complete* graph on  $n$  vertices  $K_n = ([n], \{ij | i, j \in [n] \text{ and } i < j\})$  is the graph drawn by placing  $n$  points and connecting each pair of points with an arc.

The *cycle* graph  $C_n = ([n], \{ij | i + 1 = j \text{ for } i \in [n - 1]\} \cup \{n1\})$  may be drawn as  $n$  distinct points on a circle.

$G$  is called *simple* provided  $G$  has no loops or multiple edges. A *loop* is an edge connecting a vertex to itself. *Multiple* edges occur when two or more identical edges appear in  $E(G)$ . All graphs will be simple unless otherwise stated.

$G$  is called *directed* if any of  $G$ 's edges are oriented from one vertex to another. In this case the arc drawn includes an arrow agreeing with the orientation. All graphs will be undirected in this thesis.

$H$  is a *subgraph* of  $G$  means that  $V(H) \subseteq V(G)$  and that  $E(H) \subseteq E(G)$ .

$G$  is *isomorphic* to  $H$  means there is a bijective function  $f : V(G) \rightarrow V(H)$  where  $uv \in E(G)$  if and only if  $f(u)f(v) \in E(H)$ .

An  $r$ -*coloring* of  $G$  is an assignment of one of  $r$  colors to each edge in  $E(G)$ . More generally an  $r$ -coloring of a set  $S$  is a map  $\chi : S \rightarrow [r]$  where each  $s \in S$  is given a color  $\chi(s)$ . If  $T \subseteq S$  and for some  $r \in [r]$  we have  $\chi(t) = r$  for all  $t \in T$  then  $T$  is *monochromatic*.

Vertex  $v$  has *degree*  $k$  if  $v$  is incident with exactly  $k$  edges. Likewise  $v$  has *maroon-degree*  $k$  if  $v$  is incident with  $k$  maroon colored edges.  $G$  is  $k$ -*regular* if each vertex in  $G$  has degree  $k$ .

An  $r$ -coloring  $\chi$  of  $G$  is  $(G_1, G_2, \dots, G_r)$ -*good* provided for each  $i$  in  $[r]$   $\chi$  colors no subgraph isomorphic to  $G_i$  monochromatically in color  $i$ .

Vertex  $v$ 's *deleted neighborhood*  $n'(v) = \{u \in V \mid uv \in E\}$  is the set of vertices adjacent to  $v$ .  $v$ 's *neighborhood* is  $n(v) = n'(v) \cup \{v\}$ . As above  $v$  may have a *maroon- $n(v)$*  if a coloring is present.

The *complement*  $\bar{G}$  of  $G$  has  $V(\bar{G}) = V(G)$  but  $uv \in E(\bar{G})$  if and only if  $uv \notin E(G)$ .

### 1.3 Frank P. Ramsey

Frank Plumpton Ramsey was born February 22, 1903 in Cambridge. His father Arthur Ramsey was a mathematician who served as the president of Magdalene College. At a young age Frank Ramsey was home schooled. Between 13 and 16 Ramsey attended Winchester, the oldest public school in England, before attending Trinity College of Cambridge University. At Winchester Ramsey's brilliance was undeniable. He learned German in a few weeks to

the point where he could critique German texts. Later he would travel to Austria to consult with these authors about translating their works to English. He approached these authors with such clear logic that they felt compelled to revise their works in German. Ramsey's shrewd and highly logical mind was anticipated at Cambridge. At 16 when Ramsey arrived at Cambridge, Keynes and his followers approached Ramsey for approval of their ideas. Ramsey was interested in many subjects ranging from classics, English literature, and the German language, to current day politics. However with the encouragement of Keynes, Frank Ramsey's scholarly publications came from mathematical economics, mathematics, and philosophy of logic.

In mathematical economics, Ramsey's three main published contributions each came as responses to questions posed to him by Keynes, and the brilliance of each would only be recognized decades after Ramsey's death. John Maynard Keynes was perhaps the most famous economist of the twentieth century. Keynes recognized Ramsey's brilliance freely; he relates "from a very early age, about sixteen I think, his (Ramsey's) precocious mind was intensely interested in economic problems" (Keynes 1933). Also Keynes's referred to one of Ramsey's papers as

"one of the most remarkable contributions to mathematical economics ever made, both in respect of the intrinsic importance and difficulty of its subject, the power and elegance of the technical methods employed, and the clear purity of illumination with which the writer's mind is felt by the reader to play about its subject. The article is terribly difficult reading for an economist, but it is not difficult to appreciate how scientific and aesthetic qualities are combined in it together" (Keynes 1933).

In economics, savings play a critical role in modeling the economy as a whole. For the current year, savings represent a drain on how much the economy may produce and hence limits how much an economy can consume. However, the famous assumption that savings=investment indicates that savings are the key to future prosperity. One of Ramsey's influential papers determined the optimal rate of saving for an economy as measured by the utility received in

the present and future.

Ramsey also published on the optimal rate of taxation on a regulated monopolist. The idea is that the monopolist, say the water company, should neither profit nor lose money and at the same time maximize the consumer surplus (difference in what the consumer is willing to pay and is charged). Ramsey pricing refers to the solution of this problem where the price is set so that price less marginal cost is inversely proportional to the price elasticity of demand.

Finally for Keynes, economic actors participated with full information, therefore their actions reflected the true probabilities of the possible outcomes. Ramsey disagreed with Keynes on this point, insisting that economic actors formed their beliefs about the probability of an event based on incomplete or faulty information, and their actions only represented their personal subjective probabilities. Ramsey further believed that an individual's subjective probability of an event could be identified by offering them sequentially worse odds on the occurrence of that event until they declined to wager that the event would occur.

Ramsey's theorem arose from Ramsey's attempt to come to grips with the ideas of David Hilbert, and specifically the ideas of Bertrand Russell and Alfred Whitehead in *Principia Mathematica*. Russell and Whitehead tried to show that any mathematically true statement could be derived from a set of axioms and a set of logical rules. Hilbert took this a step further by trying to prove that there could be a procedure that did this. Ramsey found these goals desirable and yet found logical holes that he set out to fix. Ramsey showed that a certain class of first-order logic problems were decidable. We now know that some first-order logic statements are undecidable in that their truth or falsehood does not affect the truth of any other mathematically true statements. Therefore Ramsey's aim could not have been fulfilled in full. The theorem that is given his name was a mere lemma in his 1930 paper "On a Problem of Formal Logic".

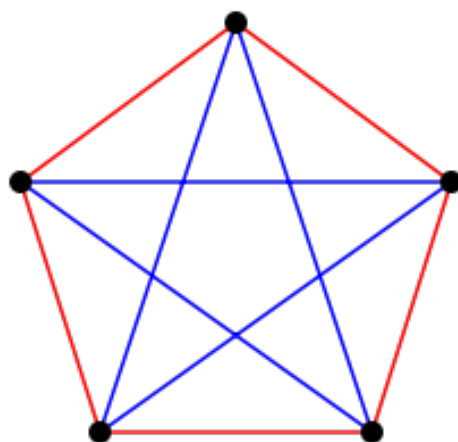
Not only were none of his ideas given their proper recognition at the time of their publication, Ramsey's chronic liver problems led to his untimely death at the age of 26 before Ramsey's theorem was even published as a lemma.

## 1.4 Ramsey's Theorem

The party problem asks: what is the smallest number of people attending a party for which we are guaranteed to find 3 people none of whom know the other two or 3 people each of whom know the other two? Assume that knowing is a symmetric relationship.

Suppose we have 5 people at a math party: Andy, Bud, Camille, Dan, and Emily. If Andy knows Bud and Emily, Bud knows Andy and Camille, Camille knows Bud and Dan, Dan knows Camille and Emily, and Emily knows Dan and Andy, then no three of the partygoers are mutual non-acquaintances or mutual friends. To see this quickly, we encode this information in a graph by taking  $K_5$  and assigning each partygoer a vertex and then coloring the edge between partygoers maroon if they know each other and orange if they do not know each other. If three partygoers all know each other, there will be a maroon  $K_3$  subgraph, and if three partygoers all do not know each other there will be a orange  $K_3$  subgraph. The Party problem appeared in this graph-theoretic manner in the March 1953 W. L. Putnam Mathematical Competition. Figure 1.1 shows this graph and it has no monochromatic  $K_3$ . Say Frank Ramsey joins the party. Now we have 6 attendees. As above,

Figure 1.1:  $R(3, 3) > 5$



form the graph  $K_6$  and consider any 2-coloring of the edges. By the pigeonhole principle, as Frank is incident with 5 edges and we have used 2 colors, at least 3 of the edges incident

with Frank are the same color. Without loss of generality assume the three edges are colored orange and these edges are incident with Andy, Bud, and Camille meaning Frank knows none of these people. If any of the edges between Andy, Bud, and Camille is also orange, Frank and the two people incident with this orange edge all do not know each other and they are part of a monochromatic orange  $K_3$ . If none of the edges between Andy, Bud, and Camille are orange, then they are all maroon so Andy, Bud, and Camille are part of a maroon monochromatic  $K_3$  and they all know each other. So 6 people are enough to guarantee a clique of 3 people all of whom know each other or all of whom are strangers. We have solved the problem, and in view of the following definition have proved that  $R(3,3)=6$ .

**Definition** The *Ramsey number*  $R(m, o)$  is defined to be the smallest  $n$  for which any 2-coloring of  $K_n$  in maroon and orange contains a monochromatic maroon  $K_m$  or a monochromatic orange  $K_o$ .

The 2-coloring of  $K_5$  above is the only 2-coloring avoiding monochromatic  $K_3$ s (up to graph isomorphisms). Indeed in the proof above if there is a vertex with three incident edges of the same color, then a monochromatic  $K_3$  is formed amongst this vertex and these neighbors. So in any coloring of  $K_5$  avoiding monochromatic  $K_3$ s, each vertex is incident to two edges of each color. For finite graphs, if each vertex is degree two or more then there is a cycle, as any maximal path not repeating vertices must enter a vertex of degree one of which there are none. If in one color the smallest cycle is a  $K_3$ , then this coloring fails. If in one color the smallest cycle is a 4-cycle then the other two edges between these 4 vertices are the second color, and to avoid a triangle in the first color the fifth vertex sends two edges of each color to non-consecutive vertices on the 4-cycle, but then there is a triangle in the second color involving the fifth vertex and the. The 2-coloring in the proofs of  $R(3, 5)$  and  $R(4, 4)$ , as well as the critical graph for  $R(3, 9)$  are also unique. However, generally the 2-colorings used to prove Ramsey numbers are not unique. For instance, there are 430,215 non isomorphic critical 2-colorings to choose from when proving  $R(3, 8)$ !

**Proposition 1.4.1** For all  $m \in \mathbb{N}$ ,  $R(m, 2) = m$



**Proof**  $K_2$  is a graph consisting of a single edge. If  $\chi$  is a 2-coloring of  $K_m$  in maroon and orange and  $\chi$  assigns an edge the color orange then we have a monochromatic orange  $K_2$ . But if  $\chi$  assigns no edge the color orange, then all edges are colored maroon and we have a monochromatic maroon  $K_m$ . So  $R(m, 2) \leq m$ . On the other hand consider a monochromatic maroon  $K_{m-1}$ . It has neither a maroon  $K_m$  subgraph nor an orange  $K_2$  subgraph. So  $R(m, 2) \geq m$ . This implies  $R(m, 2) = m$ . ■

**Proposition 1.4.2** *The Ramsey function is symmetric i.e  $R(a, b) = R(b, a)$  for all  $a, b \in N$ .*

**Proof** Let  $a, b \in N$ . Suppose  $R(a, b) = n$  then there is a 2-coloring  $\chi$  of  $K_{n-1}$  avoiding maroon  $K_a$  and avoiding orange  $K_b$ . Define a coloring  $\delta$  on  $G$  by  $\delta(e) = 1$  if  $\chi(e) = 2$  and  $\delta(e) = 2$  if  $\chi(e) = 1$ . Then  $\delta$  and  $\chi$  always swap colors. So  $\delta$  is a coloring of  $K_{n-1}$  that avoids maroon  $K_b$  and orange  $K_a$ . Therefore  $R(b, a) \geq R(a, b)$ . Nothing we did above depends on the order of the inputs. So swapping  $a$  and  $b$  in the above argument gives  $R(a, b) \geq R(b, a)$ , hence we must conclude  $R(a, b) = R(b, a)$ . ■

**Lemma 1.4.3** *The Ramsey numbers satisfy  $R(a, b) \leq R(a, b - 1) + R(a - 1, b)$ .*

**Proof** Let  $n = R(a, b - 1) + R(a - 1, b)$ . Consider a vertex  $v$  of  $K_n$  which has been 2-colored maroon and orange. Since  $v$  is degree  $n - 1$ ,  $\text{maroon-degree}(v) + \text{orange-degree}(v) = n - 1$ . Then one of the following two statements must be true: i)  $\text{maroon-degree}(v) \geq R(a - 1, b)$  or ii)  $\text{orange-degree}(v) \geq R(a, b - 1)$ . For if neither of these is true then  $n - 1 = \text{maroon-degree}(v) + \text{orange-degree}(v) \leq R(a, b - 1) + R(a - 1, b) - 2 = n - 2$ , a contradiction. In case i),  $\text{maroon-}n'(v)$  contains an orange  $K_b$  or a maroon  $K_{a-1}$  which when combined with  $v$  forms a maroon  $K_a$ . In case ii),  $\text{orange-}n'(v)$  contains a maroon  $K_a$  or an orange  $K_{b-1}$  which when combined with  $v$  forms an orange  $K_b$ . Since in both case i), and case ii), we must always have either a maroon  $K_a$  or an orange  $K_b$ , this shows that  $R(a, b) \leq R(a, b - 1) + R(a - 1, b)$ . ■

**Proposition 1.4.4** *The Ramsey numbers are monotone.*

**Proof** Let  $u_1 \geq u_2$  and  $v_1 \geq v_2$  then if  $n$  is large enough to guarantee the existence of either a maroon  $K_{u_1}$  or an orange  $K_{v_1}$  then  $n$  also guarantees the existence of a maroon  $K_{u_2}$  or an orange  $K_{v_2}$ , as  $K_{u_2}$  is a subgraph of  $K_{u_1}$  and  $K_{v_2}$  is a subgraph of  $K_{v_1}$ . Hence  $R(u_1, v_1) \geq R(u_2, v_2)$  ■

**Theorem 1.4.5** (Ramsey's Theorem for 2 colors) *For any  $m, o \geq 2 \in N$   $R(m, o)$  exists, i.e. is finite.*

**Proof** Combining propositions 1.4.1 and 1.4.2 we have  $R(2, 2) = 2$  and  $R(3, 2) = R(2, 3) = 3$ . Proceed by induction on  $o + m = n$ . We have done the base cases  $n = 4$  and  $n = 5$  above. Suppose that for  $n \in N$  we know  $R(m, o)$  exists whenever  $m + o = n$  and  $m, o \geq 2$ . Let  $m + o = n + 1$  and  $m, o \geq 2$ . If either  $m$  or  $o$  are 2, proposition 1.4.1 shows  $R(m, o)$  exists. Otherwise  $m, o \geq 3$  and by Lemma 1.4.3,  $R(m, o) \leq R(m, o - 1) + R(m - 1, o)$  where  $m, o, m - 1, o - 1 \geq 2$  and the inductive hypothesis gives that both  $R(m, o - 1)$  and  $R(m - 1, o)$  exist. So  $R(m, o)$  is bounded above and therefore exists. ■

The recursion in Lemma 1.4.3 as used in this proof provides a way to bound the Ramsey numbers above in a closed form.

**Theorem 1.4.6** *For all  $a, b \geq 2$ ,  $R(a, b) \leq \binom{a+b-2}{a-1}$ .*

**Proof** For  $r = s = 2$ , Proposition 1.4.1 gives  $R(2, 2) = 2 \leq \binom{2+2-2}{2-1} = \binom{2}{1} = 2$ . Again, induct on the sum of the inputs  $a + b$ . If the theorem holds whenever this sum is  $n - 1$  and  $a + b = n$  then  $R(a, b) \leq R(a - 1, b) + R(a, b - 1) \leq \binom{(a-1)+b-2}{(a-1)-1} + \binom{a+(b-1)-2}{a-1} = \binom{a+b-3}{a-2} + \binom{a+b-3}{a-1} = \binom{a+b-2}{a-1}$ , where the last equality is Pascal's identity. ■

This is by no means the only way to prove the existence of the Ramsey numbers. We will later establish upper bounds on  $R(m, m)$ . Proposition 1.4.4 then bounds  $R(s, t)$  above by  $R(\max(s, t), \max(s, t))$  when  $R(s, t)$  exists for  $s, t \in N$ .

## 1.5 Generalized and Infinite versions of Ramsey's Theorem

### 1.5.1 More Colors

First consider allowing colorings with more than 2 colors. Define for natural numbers  $I_1, \dots, I_r$  the *Ramsey number*  $R(I_1, I_2, \dots, I_r)$  to be the least number  $n$  for which any  $r$ -coloring of  $K_n$  contains a monochromatic  $K_{I_t}$  in color  $t$  for at least one index  $t$ .

One might worry that  $R(I_1, I_2, \dots, I_r)$  could be undefined because the extra colors provide too much flexibility to the coloring. However, Ramsey's theorem for 2 colors generalizes quite readily to an arbitrary number of colors.

**Lemma 1.5.1** *The  $r$ -color Ramsey numbers satisfy  $R(I_1, I_2, \dots, I_r) \leq 2 + (\sum_{i=1}^r R(I_1, I_2, \dots, I_i - 1, \dots, I_r) - 1)$*

**Proof** As in Lemma 1.4.3, if  $n \geq 2 + (\sum_{i=1}^r R(I_1, I_2, \dots, I_i - 1, \dots, I_r) - 1)$  then any  $v \in K_n$  has degree  $\geq 1 + (\sum_{i=1}^r R(I_1, I_2, \dots, I_i - 1, \dots, I_r) - 1)$  so that in some color  $j$   $v$  has  $j$ -degree  $\geq R(I_1, I_2, \dots, I_j - 1, \dots, I_r)$  when either for  $i \neq j$   $n'(v)$  contains a monochromatic  $K_{I_i}$  in color  $i$  or  $n'(v)$  contains a monochromatic  $j$ -colored  $K_{I_j-1}$  which when  $v$  is added becomes a monochromatic  $j$ -colored  $K_{I_j}$ . ■

**Proposition 1.5.2** *For all  $I_1, I_2, \dots, I_r \in \mathbb{N}$   $R(I_1, I_2, \dots, I_r) = R(I_1, I_2, \dots, I_r, 2)$ .*

**Proof** As in Proposition 1.4.1; any  $(I_1, I_2, \dots, I_r, 2)$ -good  $r+1$  coloring does not use color  $r+1$ , so this coloring is a  $(I_1, I_2, \dots, I_r)$ -good  $r$ -coloring. Also  $(I_1, I_2, \dots, I_r)$ -good  $r$ -colorings are  $(I_1, I_2, \dots, I_r, 2)$ -good  $r+1$ -colorings. This establishes both  $\geq$  and  $\leq$  giving equality. ■

**Definition** When all  $r$  inputs to the Ramsey function are the same number  $n$ , we may write  $R(n; r)$  to indicate  $R(n, n, \dots, n)$  when in the latter expression we have  $r$   $n$ 's.

**Theorem 1.5.3** (Ramsey's theorem for  $r$  colors) *For any  $I_1, I_2, \dots, I_r \geq 2 \in \mathbb{N}$   $R(I_1, I_2, \dots, I_r)$  exists.*

**Proof** Let  $r$  and  $n \in \mathbb{N}$ . As in Theorem 1.4.5 induct on  $n = \sum_{i=1}^r I_i$ . We know  $R(2, 2) = 2$ . Repeated application of Proposition 1.5.2 gives  $R(2; r) = 2$ , settling our base case of  $n = 2r$ . Assume for  $n \geq 2r$  that whenever  $I_1, I_2, \dots, I_r \geq 2 \in \mathbb{N}$  and  $\sum_{i=1}^r I_i = n$  that  $R(I_1, I_2, \dots, I_r)$  exists. Consider  $I_1, I_2, \dots, I_r \geq 2 \in \mathbb{N}$  and  $\sum_{i=1}^r I_i = n + 1$ , Lemma 1.5.1 gives  $R(I_1, I_2, \dots, I_r) \leq 2 + (\sum_{i=1}^r R(I_1, I_2, \dots, I_i - 1, \dots, I_r) - 1)$  where the inductive hypothesis gives that  $R(I_1, I_2, \dots, I_i - 1, \dots, I_r)$  exists for each  $i$  in the sum, as the inputs sum to  $n + 1 - 1 = n$ . So  $R(I_1, I_2, \dots, I_r)$  is bounded above and finite. ■

The above results give a quick upper bound on  $R(3; r)$ , the multicolor Ramsey numbers for triangles with  $r$  colors.

**Theorem 1.5.4**  $R(3; r) \leq 3r!$

**Proof** Combining Proposition 1.5.2 and Lemma 1.5.1,  $R(3; r) \leq rR(3; r-1)$  since the sum in the lemma has  $r$  terms each of which appear with  $r-1$  3s as  $R(3, \dots, 3, 2, 3, \dots, 3) = R(3; r-1)$ . As  $R(3; 2) = 6 = 3 * 2$  repeatedly applying the above relation gives the result. ■

## 1.5.2 Arbitrary Graphs

In addition to increasing the number of colors allowed by the colorings, the Ramsey function need not be restricted to complete graphs.

**Definition** For graphs  $G_1, G_2, \dots, G_r$ , the Ramsey number  $R(G_1, G_2, \dots, G_r)$  is the least  $n$  for which there are no  $(G_1, G_2, \dots, G_r)$ -good  $r$ -colorings of  $K_n$ . Erdős and company have studied the special cases when the  $G_i$  are cycles or bipartite.

### 1.5.3 Hypergraphs

**Definition** A *hypergraph*  $H = (V(H), E(H))$  has a vertex set  $V(H)$  just as for normal graphs. However, edges  $e$  in  $E(H)$  can be any subset of  $V(H)$ , not just subsets of order 2. If each  $e$  in  $E(H)$  joins  $k$  vertices, then  $H$  is called *k-regular*.

**Theorem 1.5.5** (Ramsey's theorem for Hypergraphs) *For any integers  $I_1, I_2, \dots, I_r, k$  there exists  $n \in \mathbb{N}$  such that if the edges of the complete  $k$ -regular hypergraph on  $n$  vertices is  $r$ -colored then for some  $i \in [r]$  there is a subhypergraph on  $I_i$  vertices that is monochromatic in color  $i$ .*

This theorem could be proved by inducting on  $k$ ; for instance,  $k = 2$  is the base case and the normal  $r$ -color Ramsey theorem. However, to demonstrate a different proof technique and introduce another area of Ramsey theory, first consider an infinite version of Ramsey's theorem that will imply both the  $r$ -color Ramsey theorem and more generally Ramsey's theorem for hypergraphs. This is the approach Ramsey took in his seminal paper.

### 1.5.4 Infinite version of Ramsey's theorem

Ramsey viewed his result as a set-theoretic result. This theorem will be presented from the set-theoretic perspective. Experience dictates that for finite cases, speaking about graphs aids intuition as one can "see" what is going on. On the other hand when dealing with infinite cases, mathematicians develop intuition from working with common infinite sets.

**Theorem 1.5.6** (Infinite version of Ramsey's theorem)

*If  $X$  is a countably infinite set and for any  $n \in \mathbb{N}$  the subsets of size  $n$ ,  $X^n$ , are colored in finitely many colors, then there is an infinite set  $M \subseteq X$  where all subsets of  $M$  of size  $n$  are colored in the same color.*

**Proof** Let  $r$  be the finite number of colors. Let  $X$  be countably infinite and proceed by induction on  $n$ . For  $n = 1$  we are coloring the members of a countably infinite set in finitely many colors, so by the infinite pigeonhole principle, one color, say red, is used on infinitely many elements. The set of red elements certainly has the property that all 1-element subsets are red. Now assume that for  $n \in \mathbb{N}$  we may always find a countably infinite set  $M \subseteq X$  where the  $n$ -element subsets of  $M$  are all colored the same color. Let  $\chi$  be a coloring of the  $n+1$ -element subsets of  $X$ . Take any element  $x_0$  of  $X$  and form  $X_0 = X - \{x_0\}$ .  $\chi$  induces an  $r$ -coloring  $\chi_0$  on the  $n$  element subsets of  $X_0$  by  $\chi_0(a_1, a_2, \dots, a_n) = \chi(x_0, a_1, a_2, \dots, a_n)$ . Now the inductive hypothesis provides a countably infinite set  $M_0 \subseteq X_0$ , where all  $n$ -element subsets of  $M_0$  are colored the same color by  $\chi_0$ . This means that for any  $n+1$  set in  $X$  including  $x_0$  and  $n$  elements from  $M_0$  is colored the same color. Now repeat this process by taking any element  $x_1 \in M_0$ , forming  $X_1 = M_0 - \{x_1\}$  and finding a countably infinite  $M_1 \subseteq M_0$  where all  $n+1$  sets in  $X$  including  $x_1$  and  $n$  elements from  $M_1$  are colored the same color. Repeating this process gives a sequence  $(x_0, x_1, \dots)$  where if  $(x_{i_1}, x_{i_2}, \dots, x_{i_{n+1}})$  is an ordered  $n+1$ -tuple from elements in the above sequence (so  $i_1$  is the smallest index), then  $\chi(x_{i_1}, x_{i_2}, \dots, x_{i_{n+1}})$  depends only on  $x_{i_1}$  since  $(x_{i_1}, x_{i_2}, \dots, x_{i_{n+1}})$  includes  $x_{i_1}$  and  $n$  elements from  $M_{i_1}$ . Color the sequence  $(x_0, x_1, \dots)$  where  $x_i$  gets colored the color  $\chi$  gives to any  $n+1$  tuple from this sequence with smallest index  $i$ . Since we have used finitely many colors and the sequence is countably infinite in length, some color, say blue, is used infinitely many times when coloring the sequence. Then the blue elements form a countably infinite set  $B$  with the property that any  $n+1$  tuple from  $B$  is given the color blue by  $\chi$  as no matter what the smallest index of this  $n+1$  tuple is, it indicates the color blue. This shows the theorem holds for  $n+1$  and by induction the theorem holds for each  $n \in \mathbb{N}$ . ■

The above technique will be used later in the finite setting to achieve a bound on  $R(m, m)$ .

### 1.5.5 Compactness

Here the goal is to prove that the infinite version of Ramsey's theorem implies the finite version. The *chromatic number*  $\chi(H)$  is defined to be the least number of colors needed to color the vertices of  $H$  so that the endpoints of each edge are not all the same color. This definition is particularly tailored to hypergraphs where edges may have more than 2 endpoints, where it is fine for the edge  $uvw$  to have  $\chi(u) = \text{blue}$ ,  $\chi(v) = \text{blue}$ , and  $\chi(w) = \text{red}$ .

There is a particularly useful connection between chromatic numbers and the Ramsey property. Say we are wondering if we can color the edges of  $K_n$  in two colors so that we form no monochromatic  $K_m$ s, e.g. , if  $R(m, m) > n$ . We can reformulate this problem into one of finding the chromatic number of a hypergraph as follows. For each edge in  $K_n$  associate a vertex  $v$  in a new hypergraph  $H$ , for each  $K_m$  subgraph of  $K_n$  place an edge in  $H$  connecting the vertices in  $H$  corresponding to the edges of the  $K_m$ . If  $\chi(H) > 2$  then for every 2-coloring  $\chi_1$  of the vertices in  $H$ , some edge in  $E(H)$  has all  $\binom{m}{2}$  endpoints the same color. But  $\chi_1$  induces a coloring of  $K_n$  where each edge of  $K_n$  is colored the color its corresponding vertex in  $H$  was colored by  $\chi_1$ . Alternatively we could have started with a 2-coloring of the edges of  $K_n$  and similarly induced a 2-coloring of the vertices of  $H$ . The bijection assures us that if  $\chi(H) > 2$  then every 2-coloring of  $K_n$  contains a monochromatic  $K_m$ . So if  $\chi(H) > 2$  then  $R(m, m) > n$ . Compactness discussion will continue after doing some brief preparation.

### 1.5.6 Density

Paul Turán, a member of Erdős' Hungarian circle, asked what is the maximum number of edges a graph on  $n$  vertices may have and yet have no complete subgraphs on more than  $r$  vertices?

Turán showed in 1941 that the graph on  $n$  vertices with the most edges lacking a  $K_{r+1}$  subgraph could be constructed by partitioning the  $n$  vertices into  $r$  sets of almost equal size

and connecting pairs of vertices in different parts of the partition with an edge. These graphs are now called *Turán* graphs. Turán graphs have at most  $((r-1)/r)(n^2/2)$  edges as, at best, of the set of  $n^2$  ordered pairs the portion  $((r-1)/r)$  may be edges (a vertex does not form an edge in its own part of the partition), and each edge counts an ordered pair twice. Parts of a partition are said to be of *almost equal* size if the size of the parts differs by at most 1.

**Theorem 1.5.7** (Turán's Theorem) *The simple graphs on  $n$  vertices containing no  $K_{r+1}$  subgraphs have at most  $((r-1)/r)(n^2/2)$  edges.*

**Proof** Let  $G$  be a graph on  $n$  vertices with no  $K_{r+1}$  subgraphs and the maximum number of edges amongst all such graphs. The first thing to show is that  $G$ 's vertices may be partitioned into sets where pairs of vertices within sets are nonadjacent but pairs of vertices in different sets are adjacent. This is exactly what the following claim asserts.

Claim 1: Amongst any set of three vertices  $u, v, w$  of  $G$  it is not the case that  $uv \in E(G)$  but  $uw, vw \notin E(G)$ .

Suppose we do have the situation in claim 1.

Case 1:  $d(w) < d(u)$  or  $d(w) < d(v)$ . Without loss of generality assume  $d(w) < d(u)$ . Remove the vertex  $w$  and replace it with a copy  $u'$  of  $u$  (so  $u'$  is adjacent to every vertex  $u$  is adjacent to) and call the new graph  $G'$ . Since  $u$  and  $u'$  are not adjacent and the original graph had no  $K_{r+1}$  subgraphs, neither does  $G'$ . However,  $G'$  has more edges than  $G$  since the number of edges in  $G'$  is the number of edges in  $G$  plus  $d(u) - d(w) > 0$ . This contradicts that  $G$  had the maximum number of edges amongst  $K_{r+1}$ -avoiding graphs on  $n$  vertices.

Case 2:  $d(w) \geq d(u)$  and  $d(w) \geq d(v)$ . Remove the vertices  $u$  and  $v$  and replace each with a copy of  $w$ ; call the new graph  $G'$ .  $G'$  has no  $K_{r+1}$  subgraphs, since  $w$  and its two copies are pairwise nonadjacent and the original graph had no  $K_{r+1}$  subgraphs.  $G'$  has more edges than  $G$ , as the number of edges in  $G'$  is the number of edges in  $G$  plus  $2d(w) - (d(u) + d(v) - 1) \geq 1$  since we have added  $w$ 's edges to both copies of  $w$  and removed the edges incident with  $u$  and  $v$ , but as  $uv \in E(G)$   $d(u) + d(v)$ , overcounts the quantity of edges removed by 1. This



contradicts that  $G$  had the maximum number of edges amongst  $K_{r+1}$ -avoiding graphs on  $n$  vertices.

So in fact claim 1 holds and  $G$  is a *complete  $r$ -partite* graph.

Claim 2: The number of edges in a complete  $r$ -partite graph is maximized when the parts differ by at most 1. Suppose  $P_1$  and  $P_2$  are parts of the partition and  $|P_1| - 1 > |P_2|$ . Then move a vertex from  $P_1$  to  $P_2$ , then at least  $|P_1| - 1$  edges are added and  $|P_2|$  edges are removed, so the graph gains at least 1 edge and the graph is still  $r$ -partite and thus contains no  $K_{r+1}$ .

Indeed we have shown that the Turán graph is the only graph (up to a permutation of the vertices) on  $n$  vertices that achieves the maximum number of edges a graph can have and still avoid  $K_{r+1}$ s. ■

**Definition** For a hypergraph  $H = (V(H), E(H))$ , define  $T(H)$  as the *Turán number* of  $H$ .  $T(H)$  is the minimum number such that no matter how this number of vertices are chosen from  $V(H)$  at least one edge in  $E(H)$  may be formed from these vertices. Define  $\tau(H) = T(H)/|V(H)|$ .

At this point as a warm-up consider as in p.12 [GRS2] the statements for hypergraphs and sequences of hypergraphs  $H_n = (V(H_n), E(H_n))$ :

$$A: \chi(H) > r$$

$$B: \tau(H) \leq r^{-1}$$

$$A^*: \lim_{n \rightarrow \infty} \chi(H_n) = \infty$$

$$B^*: \lim_{n \rightarrow \infty} \tau(H_n) = 0$$

[GRS2] comments: “B says that any sufficiently large set of vertices contains a hyperedge...A says that if the vertex set is partitioned into  $r$  classes one class contains a hyperedge”.

**Proposition 1.5.8**  $B \Rightarrow A, B^* \Rightarrow A^*$ .

**Proof** Largely from [GSR2]: Let  $H = (V(H), E(H))$  be a hypergraph. Assume B holds for  $H$  and consider an  $r$ -coloring  $\chi$  of  $V(H)$ . Since  $rr^{-1} = 1$ , at least one color is used on at least  $r^{-1}|V(H)| \geq \tau(H)|V(H)| = T(H)$  vertices. By definition the color with at least  $T(H)$  vertices contains a hyperedge. Let  $H_n = (V(H_n), E(H_n))$  be a sequence of hypergraphs and assume B\* holds for  $H_n$ . Now for any given  $r$  there is  $n \in \mathbb{N}$  such that for  $m \geq n$   $\tau(H_n) \leq r^{-1}$ , but since we just showed  $B \Rightarrow A$  then for each  $m \geq n$   $\chi(H_n) \geq r$  which is exactly what it means for A\* to hold. ■

Now A\* does not imply B\*. To see this use the connection between Ramsey numbers and chromatic numbers above. Say we care about  $R(m; r)$ . The connection above can be described by the following notation from [GRS2]: take  $H_n = (V(H_n), E(H_n))$  where  $V_n = [n]^2$  and  $E_n = \{[S]^2 \mid S \in [n]^m\}$ . Remember from above that  $\chi(H_n) > r \iff R(m; r) > n$ . On the right as  $n$  approaches infinity  $r$  goes to infinity, but this means  $\lim_{n \rightarrow \infty} \chi(H_n) = \infty$  so A\* holds. But a calculation on  $\tau(H_n)$  shows that  $\lim_{n \rightarrow \infty} \tau(H_n) = (m-2)/(m-1) \neq 0$ . So B\* does not hold.

### 1.5.7 Compactness Continued

Here the point is to give conditions for when the infinite version of a Ramsey-type problem implies the finite version. Often it is easier to prove the infinite version as it does not require producing explicitly a sufficiently large  $n$ , but rather proving the statement only for colorings of  $N$  or  $N^k$  which give the largest margin for simplification. This follows [GSR2] P.14-17 closely.

For a hypergraph  $H = (V(H), E(H))$  and a set  $W \subseteq V(H)$  the restriction  $H_W$  has  $W$  as its vertex set and an edge set consisting of those edges in  $H$  all of whose endpoints are in  $W$ . So  $H_W = (W, E_W)$  where  $E_W = \{e \in E(H) \mid e \subseteq W\}$ .

**Theorem 1.5.9** (Compactness Principle [GSR2]) *Let  $H = (V(H), E(H))$  be a hypergraph*

where all edges have finitely many endpoints (possibly infinitely many vertices in  $V(H)$ ). If for any finite set  $W \subseteq V$ ,  $\chi(H_W) \leq r$  then  $\chi(H) \leq r$ .

**Proof** For  $V(H)$  uncountable see [GSR2] proof 2, p.15. For  $V(H)$  countable a proof of Theorem 6 requires the Axiom of Choice and [GSR2] gives a topological argument. The uncountable case is important for the chromatic number of the plane, but in typical Ramsey-type problems  $V(H)$  is  $N$  or  $N^k$ , so  $V(H)$  is countably infinite. Here only the countably infinite case is presented.

Let  $H = (V(H), E(H))$ . Assume  $V(H)$  is countable, it suffices to consider  $V(H) = N$ . This proof is direct. Let  $r \in N$  be given and assume for each  $n \in N$  there is a coloring  $\chi_n : [n] \rightarrow [r]$  where no edge is monochromatic. Since any finite set  $W \subset N$  has a maximum element  $m$ , considering  $\chi_m$  restricted to  $W$  shows  $\chi(H_W) \leq r$ . This assumption considers only some of the finite sets  $W \subset N$  and the theorem allows making the assumption for all finite sets, so this assumption is safe to make.

Now define a coloring  $\chi^* : N \rightarrow [r]$  using an inductive approach. The goal is to show  $\chi^*$  never colors the vertices of an edge of  $H$  monochromatically. Since we only use  $r$  colors, infinitely many of the  $\chi_n$  color 1 with the same color. Define  $\chi^*(1)$  to be this color. Take  $S_1$  to be the set of  $n \in N$  where  $\chi_n(1) = \chi^*(1)$ , and note that  $S_1$  is countably infinite. Likewise infinitely many of the  $n \in S_1$  color 2 in the same color. Define  $\chi^*(2)$  to be this color and define  $S_2$  to be the set of  $n \in S_1$  where  $\chi_n(2) = \chi^*(2)$ . So  $S_2$  is a countably infinite collection of colorings from the  $\chi_n$  that agree on the color of the first 2 vertices. Build  $\chi^*$  in this manner for each  $n \in N$ . Take  $e = (x_1, x_2, \dots, x_m) \in E(H)$  with largest entry  $x_m$ . Now  $S_{x_m}$  is countably infinite, so nonempty, and there is a coloring  $\chi_t \in S_{x_m}$ ,  $t \geq x_m$  where  $\chi_t(i) = \chi^*(i)$  for  $i \leq x_m$ . Now  $\chi_t$  is an  $r$ -coloring of at least  $[x_m]$ , so  $e$  is not monochromatic under  $\chi_t$ . As  $\chi_t$  matches  $\chi^*$  on at least  $[x_m]$ ,  $e$  is not monochromatic under  $\chi^*$  and this is true for all edges. ■

Now consider the contrapositive of the Compactness Principle, if  $\chi(H) > r$  then there is a

finite  $W$  for which  $\chi(H_W) > r$ . Since the infinite version of Ramsey's theorem has already been proven, which was a statement about the hypergraph with vertex set  $X^n$  and edge set  $\{e \in 2^X : |e| = \infty\}$  having no  $r$ -coloring for any  $r$  where  $X$  was countably infinite, the following spacial cases are evident.

**Corollary 1.5.10** (Compactness Principle (version B) [GSR2]) *“Let  $k$  be a fixed positive integer. Let  $\mathcal{A}$  be a family of finite subsets of  $N$ . Suppose that for any  $r$ -coloring of  $[N]^k$ , there is an  $A \in \mathcal{A}$  so that  $[A]^k$  is monochromatic. Then there exists  $n_0$  so that, for  $n \geq n_0$  if  $[n]^k$  is  $r$ -colored there is an  $A \in \mathcal{A}$ ,  $A \subseteq [n]$  so that  $[A]^k$  is monochromatic”.*

Version B is for one arbitrary  $r$ , but if it is true for an arbitrary  $r$  then it is true for all  $r$  which gives version C.

**Corollary 1.5.11** (Compactness Principle (version C) [GSR2]) *“Let  $k$  be a fixed positive integer. Let  $\mathcal{A}$  be a family of finite subsets of  $N$ . Suppose that for any finite-coloring of  $[N]^k$ , there is an  $A \in \mathcal{A}$  so that  $[A]^k$  is monochromatic. Then for all  $r$  there exists  $n_0(r)$  so that, for  $n \geq n_0(r)$  if  $[n]^k$  is  $r$ -colored there is an  $A \in \mathcal{A}$ ,  $A \subseteq [n]$  so that  $[A]^k$  is monochromatic”.*

**Corollary 1.5.12** *The infinite version of Ramsey's theorem implies the finite version.*

**Proof** Compactness Principle (version C) gives that for any  $r$  and  $m$  there is  $n$  so that  $R(m; r) = n$ . The monotonicity of the Ramsey numbers implies the existence of each Ramsey number. ■

## 1.6 Paul Erdős

Paul Erdős was born on March 26, 1913 in Budapest, Hungary. Erdős died in 1996 in Warsaw. Paul is the Westernized version of Pal. I never met Erdős and refer here to comments made in

*My Brain is Open* by Bruce Schechter and *Erdős on Graphs* by Fan Chung and Ron Graham. Schechter never met Erdős but relies on stories told by Ron Graham, Andrew Vazsonyi, Esther Kline, and George Szekeres. I will try to give a glimpse of this unparalleled man.

Erdős parents were both Jewish mathematicians who taught high school math courses in Budapest. Erdős never met his two sisters who died of scarlet fever shortly before he was born. Paul's loving parents focused all of their energy on him and at 4 years old Paul was quite the prodigy. At 4, Paul could calculate the number of seconds someone had lived and was comfortable with the concept of a number being negative.

Early 20th century Budapest was an ideal place to raise a young mathematician. The Hungarian educational system was excellent, and intellectual endeavors were highly valued in the Jewish community to which Erdős belonged. Though Erdős claimed to never have noticed, when pressed on the subject by Vazsonyi, Erdős could not recall a single non-Jewish friend from his childhood. In 1920 the Hungarian Commune was dissolved leading to a wave of anti-Semitism. By the age of 12 Erdős did notice that "eventually I'd have to leave Hungary because I am a Jew." Erdős was initially home schooled and then joined the public school system where he spent his free time solving problems published in KoMal, a journal of math problems for advanced high school students.

Erdős joined the Science University of Budapest and became friends with a group of fellow young Jewish mathematicians whose solutions had likewise been published in KoMal. This group included Esther Kline, George Szekeres, Andrew Vazsonyi, Paul Turán, Tibor Gallai, Marta Sved, and half a dozen others. They met at the Statue of Anonymous in the City Park of Budapest to explore their mutual passion for mathematics. Erdős' first major publication was an elegant re-proof of Bertrand's Theorem that between  $n$  and  $2n$  there always lies a prime. When his advisor Leopold Fejer noticed that Ramanujan had similarly proved the same result, Erdős showed that for  $n > 7$  one could find two primes in the  $n$  to  $2n$  window, one being  $1 \pmod{4}$  and the other being  $3 \pmod{4}$ . At the age of 21 Erdős earned his doctorate. He only spent 2 years as a university student for both undergraduate and dissertation!

The year was 1934 and anti-Semitism had engulfed Hungary. Erdős left Hungary for the West. First Erdős was a lecturer at Manchester University but soon took a scholarship at Princeton. Shortly thereafter he spent time at Purdue and Notre Dame but never more than a year and a half. By this time Erdős had gained a reputation of being incredibly sharp. When encountering other mathematician he would solve their problems or give them promising directions to work in. Moreover Erdős was a master at seeing the natural progression of a topic and quickly posed a new sequence of interesting questions building on what he had just solved. Erdős quickly made friends in the West and carved out a unique lifestyle for himself. For the rest of his life he became a nomad. Erdős moved from university to university visiting friends and forging new friendships. His typical method of deciding where to visit next was to ask the advice of the department he was at. His brilliance guaranteed that wherever he chose to go, and however little advance warning he gave, that accommodations were made for him until he decided to leave. Erdős spent much time at US institutions except between 1952 and 1963 due to McCarthy era mistrust of immigrants from Communist countries.

Erdős' list of collaborators grew rapidly. When Erdős helped someone with a problem it almost always resulted in a paper with Erdős' name on it written by his collaborator. In this way Erdős amassed over 1500 publications, more than any other mathematician. Mathematicians created a convention to determine how close a mathematician was to being a collaborator of Erdős called an *Erdős* number. One's Erdős number is 1 if they published a paper with Erdős, 2 if they published a paper with someone having Erdős number 1, and so on. However, the 1s were too plentiful, so Erdős' coauthors took to defining their Erdős number as  $1/(\text{number of publications authored with Erdős})$ . Nowadays almost all faculty members have Erdős numbers.

Erdős was extraordinarily quirky. First, he referred to things in his own unique way known as *Erdős*es. He referred to children as “epsilons” (because they were small), “The Book” was a collection of the most important theorems and their most elegant proofs accessible only to a supreme being, to “torture” was to give an oral exam, to be “dead” was to have stopped publishing. Erdős never drove. He also was nearly incapable of doing anything domestically

and did not like to be left alone. This often placed a burden on his hosts who had to arrange everything. Erdős was never comfortable with physical contact and could barely manage to brush someone's hand when expected to shake hands. Certainly Paul never showed interest in physical intimacy. He was briefly accompanied by a Dutch physicist named Jo Brining but otherwise did not have relationships with women outside of math. His mother was very protective of him and discouraged such relationships in his childhood.

Erdős lived life simply. All of his possessions fit into two suitcases. Erdős pursued mathematics nearly single-mindedly. Despite his nomadic lifestyle Erdős never faced serious financial problems. The fees he collected when lecturing and the generosity of his friends supported him sufficiently. Erdős often offered prizes for solutions to problems he had not solved but wished to know more about. These prizes ranged from \$1 to \$10,000. Many of these problems remain unsolved and his friends, principally Ron Graham, are willing to pay out these prizes for good solutions. Had all of Erdős' problems been solved he would have been incapable of paying out the prizes, but no one ever managed to solve any worth more than \$1000. Many of these problems dealt with Ramsey theory and have served to draw mathematicians to the topic, not for the financial gain, but to solve one of his problems. No one ever got rich from solving Erdős' problems. When Erdős won prizes that carried substantial financial benefits he often donated these to friends in need.

Erdős was charismatic and treated his friends and collaborators well. He was also exceptionally good with children. Despite his quirks he was almost universally loved by the mathematics community. Mathematicians all have their favorite Erdős story. Through his extensive travels he touched and inspired many.

# Chapter 2

## Ramsey Numbers

### 2.1 Exact Values

Proposition 1.4.1 gave  $R(m, 2) = m$  for  $m \in \mathbb{N}$  and the party problem showed that  $R(3, 3) = 6$ . The next smallest case to consider is that of  $R(4, 3)$ . This requires strengthening Lemma 1.4.3 with Corollary 2.1.1.

**Corollary 2.1.1** *For  $a, b \in \mathbb{N}$  if  $R(a, b - 1)$  and  $R(a - 1, b)$  are both even, then  $R(a, b) \leq R(a - 1, b) + R(a, b - 1) - 1$ .*

**Proof** Let  $n = R(a - 1, b) + R(a, b - 1) - 1$  and suppose there is a  $(K_a, K_b)$ -good 2-coloring of  $K_n$  in maroon and orange. Then as in Lemma 1.4.3 each vertex  $v$  has maroon-degree at most  $R(a - 1, b) - 1$  and orange-degree at most  $R(a, b - 1) - 1$  since  $v$ 's maroon- $n'[v]$  must be  $(K_{a-1}, K_b)$ -good and  $v$ 's orange- $n'[v]$  must be  $(K_a, K_{b-1})$ -good. But  $v$  has degree  $R(a - 1, b) + R(a, b - 1) - 2$  allowing the conclusion that  $\text{maroon-degree}(v) = R(a - 1, b) - 1$  and  $\text{orange-degree}(v) = R(a, b - 1) - 1$  both of which are odd. Since  $n$  is odd, there are an odd number of vertices. So in each color the sum of the color degrees is odd (odd  $\times$  odd = odd). However, each colored edge is incident with 2 vertices so the sum of the color degrees must



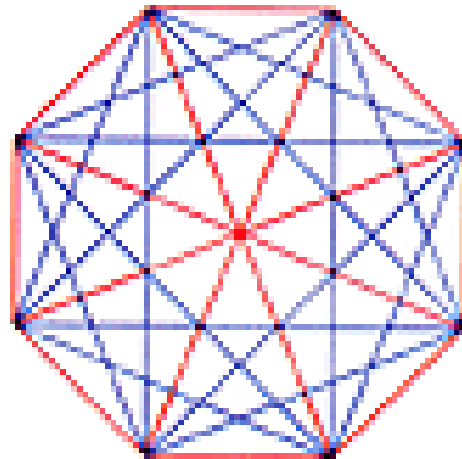
be even. This contradiction implies the claim. ■

$R(3, 4) = 9$ ,  $R(3, 5) = 14$ , and  $R(4, 4) = 18$  were known to Greenwood and Gleason in 1955. These can all be proved quickly from the above results.

**Proposition 2.1.2**  $R(3, 4) = 9$ .

**Proof** Corollary 2 gives  $R(3, 4) \leq R(2, 4) + R(3, 3) - 1 = 4 + 6 - 1 = 9$ . Figure 2.1 shows  $R(3, 4) > 8$ . ■

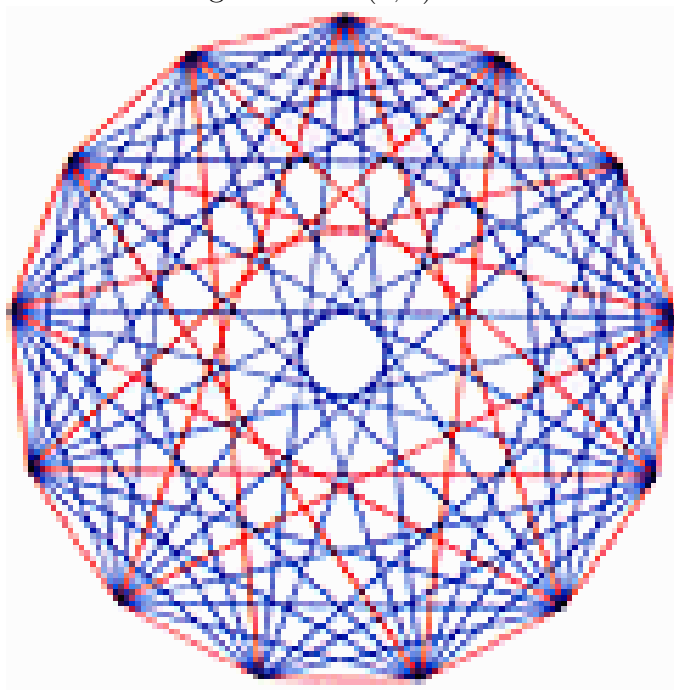
Figure 2.1:  $R(3, 4) > 8$ .



**Proposition 2.1.3**  $R(3, 5) = 14$ .

**Proof** Lemma 1.4.3 gives  $R(3, 5) \leq R(2, 5) + R(3, 4) = 5 + 9 = 14$ . Figure 2.2 shows  $R(3, 5) > 13$ . ■

**Proposition 2.1.4**  $R(4, 4) = 18$

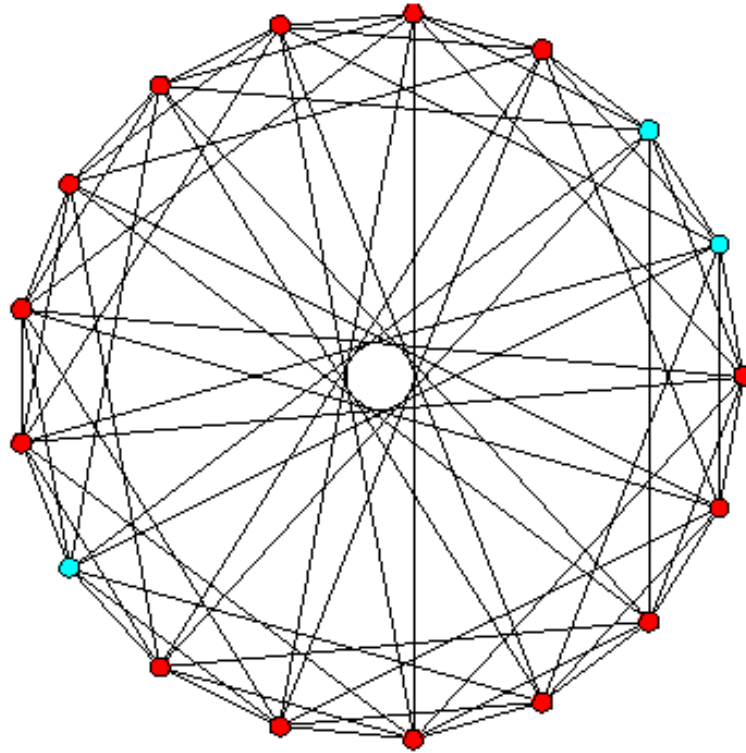
Figure 2.2:  $R(3, 5) > 13$ .

**Proof** Lemma 1.4.3 gives  $R(4, 4) \leq R(3, 4) + R(4, 3) = 9 + 9 = 18$ . Figure 2.3 shows  $R(4, 4) > 17$  by showing one of the colors. ■

**Proposition 2.1.5**  $R(3, 3, 3) = 17$ .

**Proof** Combining Lemma 1.5.1, Propositions 1.4.2, and 1.5.2, and  $R(3, 3) = 6$ ,  $R(3, 3, 3) \leq 2 + ((R(2, 3, 3) - 1) + (R(3, 2, 3) - 1) + (R(3, 3, 2) - 1)) = 2 + 3(R(3, 3, 2) - 1) = 2 + 3(R(3, 3) - 1) = 2 + 3(6 - 1) = 17$ . Figure 2.4 shows  $R(3, 3, 3) > 16$ ; note that some vertices appear twice to improve the layout. Also Figure 2.4 shows two graphs, these are the only two non-isomorphic critical colorings. ■

For a graph  $G$  on  $n$  vertices, there is an associated 2-coloring of  $K_n$  where every edge in  $G$  is colored maroon and every edge in  $G$ 's complement is colored orange. Thus,  $R(a, b) \geq n$  if

Figure 2.3:  $R(4, 4) > 17$ .

there is a graph  $G$  with no  $K_4$  subgraph and no b-IS. This is why Figure 2.3 is a good way of demonstrating  $R(4, 4) > 17$ .

Corollary 2.1.1 gives that  $R(3, 6) \leq R(3, 5) + R(2, 6) - 1 = 14 + 6 - 1 = 19$ , and for the first time in the  $R(3, k)$  sequence this bound is not best possible. The proof is lengthy and shows the effort and ingenuity required to improve the easy bounds above by only 1. The value of  $R(3, 6)$  was first known and proved by Kalbfleisch in his PhD thesis at University of Waterloo, which according to tradition was not immediately published. Indeed Kalbfleisch continued to improve these results before publishing them. At approximately the same time (1964-1965), Kery published a proof of this result. Others have proven the result since and

there seems to be some competition regarding finding a clear and quick proof. Cariolaro's is the most recent in this vein; and I reproduce it here while reminding the reader of the relevant previously proved statements when they are applied.

**Theorem 2.1.6**  $R(3,6)=18$ .

**Proof** Let  $G$  be a triangle-free (no  $K_3$  subgraphs) graph on 18 vertices. Suppose that  $G$  has no 6-IS. Some counting will show that  $G$  is 5-regular. Since  $G$  is triangle-free, the deleted neighborhood of each vertex forms an independent set, so for each  $v \in V(G)$ ,  $|n'(v)| \leq 5$  as there is no 6-IS. If  $\deg(v) = |n'(v)| < 4$  then removing  $n(v)$  from  $G$  yields a graph on at least  $18 - 4 = 14 = R(3, 5)$  vertices, so as  $G$  is triangle-free, this graph on  $R(3, 5)$  or more vertices has a 5-IS none of whose vertices are adjacent to  $v$ , so adding  $v$  to the 5-IS gives a 6-IS. If  $\deg(v) = |n'(v)| = 4$  then again consider  $H$ , the graph formed by removing  $n(v)$  from  $G$ ,  $H = G - n(v)$ ,  $|H| = 18 - 4 - 1 = 13$ . Each vertex  $h \in H$  has  $\deg(h) \geq 4$  in  $H$  as otherwise  $K$ , the graph formed by removing  $n(h)$  from  $H$  has  $|K| = |H - n(h)| \geq 13 - 4 = 9 = R(3, 4)$  and so has a 4-IS which forms a 6-IS when  $v$  and  $h$  are added as nothing in  $H$  is adjacent to  $v$  and nothing in  $K$  is adjacent to  $h$ . Let  $t \in n'(v)$ . we showed above that for all  $v \in G$   $\deg(v) \geq 4$ , since the only vertex in  $n(v)$  adjacent to  $t$  is  $v$ ,  $t$  is adjacent to three vertices  $h_1, h_2, h_3 \in H$ .  $h_1, h_2, h_3$  are independent as any adjacent pair with  $t$  forms a triangle.  $n'(v) - t$  is also a 3-IS. Finally  $h_1, h_2, h_3 \in H$  and  $n'(v) - t$  are independent as the degree in  $G$  of each  $h_i$  is at most 5, but each  $h_i$  is adjacent to at least 4 vertices in  $H$  and to  $t$ , so  $(n'(v) - t), h_1, h_2, h_3$  is a 6-IS. Since this is ruled out by assumption, each  $v \in G$  has degree 5.

It is tempting to end the proof by noting that we have the computational power to check all 5-regular graphs on 18 vertices. However, Cariolaro's proof is elegant, not difficult (which is quite an accomplishment), and readers starting out in Ramsey theory may benefit from additional explanation. The labellings and ideas presented here are largely his. Above is Cariolaro's proof of his first claim that  $G$  would need to be 5-regular. Cariolaro's second claim is as follows:

For any vertex  $v \in G$  there are exactly 4 non-neighbors  $p_1, p_2, p_3, p_4$  of  $v$  with  $|n(p_i) \cap n(v)| = 1$  and exactly 8 non-neighbors  $q_1, \dots, q_8$  of  $v$  with  $|n(q_i) \cap n(v)| = 2$ . Moreover the  $p_i$ s share 4 distinct neighbors with  $v$  and the  $q_i$ s share 8 distinct pairs of neighbors with  $v$ .

First note that 5-regularity implies each vertex has  $18 - 5 - 1 = 12$  non-neighbors. For  $u$  and  $v$  nonadjacent vertices in  $G$   $1 \leq |n(u) \cap n(v)| \leq 2$ . If this intersection were empty, then  $u$  would be independent of  $n'(v)$  and  $u \cap n'(v)$  is a 6-IS since 5-regularity and triangle-free gives  $n'(v)$  as a 5-IS. If there were 3 or more vertices in this intersection then  $|n(u) \cup n(v)| = |n(u)| + |n(v)| - |n(u) \cap n(v)| \leq 6 + 6 - 3 = 9$ . Then  $H = G - (n(u) \cup n(v))$ , the graph with the neighborhoods of  $u$  and  $v$  removed, has at least  $18 - 9 = 9 = R(3, 4)$  vertices, and thus  $H$  has a 4-IS since  $G$  is triangle-free. Combining  $u$  and  $v$  with the 4-IS gives a 6-IS since  $u$  and  $v$  are non-adjacent and their neighborhoods are not in  $H$ . Each vertex in  $n'(v)$  has 4 edges to vertices outside of  $n(v)$ . As  $|n'(v)| = 5$  there are  $4 \cdot 5 = 20$  edges, from non-neighbors of  $v$  into  $n(v)$ . Since  $v$  has 12 non-neighbors and each sends either 1 or 2 edges to  $n(v)$  we must have 8  $q_i$ s sending 2 and 4  $p_i$ s sending 1. If  $p_i$  and  $p_j$  are adjacent to the same vertex  $u \in n(v)$  then  $\{p_i, p_j\} \cup \{n'(v)u\}$  is a 6-IS which is not allowed, so the  $p_i$ s are adjacent to distinct neighbors of  $v$ . If  $q_i$  and  $q_j$  are both adjacent to the same  $x$  and  $y$  in  $n(v)$  then  $|n(x) \cap n(y)| = |\{q_i, q_j, v\}| = 3$  contradicting that for non-adjacent  $u$  and  $v$   $1 \leq |n(u) \cap n(v)| \leq 2$  since  $x$  and  $y$  are non-adjacent members of the independent set  $n'(v)$ . So the  $q_i$ s are adjacent to distinct pairs of  $v$ 's neighbors.

The final claim is that  $p_1, p_2, p_3, p_4$  induce a 4-cycle. This means that the subgraph of  $G$  on these 4 vertices is the 4-cycle. The strategy is to show that there are exactly 4 edges in this subgraph, the fact that  $G$  has no triangles will imply this subgraph is 4-cycle.

Throughout this part, remember that  $G$  is 5-regular. Let  $n'(v) = \{t, s_1, s_2, s_3, s_4\}$  and by the above let  $s_1p_1, s_2p_2, s_3p_3, s_4p_4$  be the edges from the  $p_i$ s to  $n(v)$ . Note that the  $p_i$ s are not adjacent to  $t$ , as each  $p_i$  is only adjacent to one vertex in  $n(v)$  which is denoted  $s_i$ . The neighbors of  $t$  other than  $v$  are outside  $n(v)$  and these four neighbors are not  $p_i$ s, so they are  $q_i$ s. Label these  $q_i$ s, specifically vertices in  $n(t) - v$ , as  $t_1, t_2, t_3, t_4$ . Label the other four

$q_i$ s  $w_1, w_2, w_3, w_4$ . For record-keeping, in  $V(G)$  we have  $v, t$ , and four each of  $s_i$ s,  $p_i$ s,  $t_i$ s, and  $w_i$ s.

By definition each  $s_i$  sends one edge to  $v$ , one edge to the corresponding  $p_i$ , one edge to the  $t_i$ s (the  $t_i$  send edges to  $t$  and one other neighbor of  $v$  and if there were 2 edges from an  $s_i$  to the  $t_i$ s then the neighborhoods of the non-adjacent vertices  $s_i$  and  $t$  have these 2  $t_i$ s and  $v$  in common, but the neighborhoods of two non-adjacent vertices may have at most 2 vertices in common), and thus as each  $s_i$  has degree 5, each  $s_i$  sends 2 edges to the  $w_i$ s. Likewise since the  $s_i$  are adjacent to  $v$  no pair of  $s_i$ s may be adjacent to the same pair of  $w_i$ s since the  $s_i$  are in  $n'(v)$  and are non-adjacent their neighborhoods cannot have  $v$  and 2  $w_i$ s in common. Likewise no  $w_i$  is adjacent to three or more of the  $s_i$  since the non-adjacent  $v$  and this  $w_i$  would have 3 common neighbors. So each  $w_i$  is adjacent to a unique pair of  $s_i$ s and there are 4 such triples, one for each  $w_i$ .

Say  $s_1$  and  $s_2$  are adjacent to  $w_1$ , and note that the vertices in the 3-IS  $t, s_3, s_4$  are not connected to  $s_1, s_2, p_1, p_2$ , or  $w_1$ . So if  $G$  has no 6-IS the subgraph induced by  $s_1, s_2, p_1, p_2, w_1$  has no 3-IS, and must be triangle-free, so this induced subgraph is the unique  $R(3, 3)$  avoiding graph on 5 vertices, a 5-cycle.  $p_1$  is adjacent to  $s_1$ , and thus not adjacent to  $s_2$  as the  $p_i$ s are only adjacent to one  $s_i$  nor is  $p_1$  adjacent to  $w_1$  as then  $s_1, w_1, p_1$  would form a triangle, so  $p_1$  is also adjacent to  $p_2$ ! As this reasoning holds for each of the four triples with a  $w_i$  and two adjacent  $s_i$ s, the subgraph induced by the  $p_i$ s has 4 edges. The only triangle-free graph on four vertices with four edges is a 4-cycle. Thus the  $p_i$  induce a 4-cycle. Without loss of generality, this 4-cycle is  $p_1 p_2 p_3 p_4 p_1$ .

Since  $p_i$  and  $t$  are non-adjacent, as above their neighborhoods contain one or two common vertices. Each pair of  $p_i$ s has no common neighbors except amongst the  $p_i$ s, for if two  $p_i$  are adjacent, a common neighbor forms a triangle, and if two  $p_i$  are not adjacent, since the  $p_i$  lie on a 4-cycle, both of the other  $p_i$ s are common neighbors, and any common neighbor outside the  $p_i$ s would be a third common neighbor which is a case dismissed above. Since  $n(t)$  has  $v$  and the  $t_i$  and the  $p_i$  are not adjacent to  $v$ , each  $p_i$  is adjacent to a unique  $t_i$ , and

by relabeling assume for each  $i$  that  $p_i s_i$  is an edge. Each  $p_i$  is not adjacent to  $v$ , sends one edge to  $s_i$ , two edges to other  $p_i$ s, an edge to  $t_i$  and thus by 5-regularity, sends exactly one edge to the  $w_i$ s. By relabeling the  $w_i$ s and recalling that the  $p_i$  have no common neighbors outside of  $p_i s_i$ , assume  $p_i w_i$  is an edge for each  $i$ . Now since  $p_i$  is adjacent to  $s_i, t_i$ , and  $w_i$  to avoid triangles with  $p_i, s_i w_i, s_i t_i$  and  $t_i w_i$  must not be edges for each  $i$ .

Consider the  $t_i$ s, they are not adjacent to  $v$ , each sends one edge to the  $s_i$ s, one edge to  $t$ , one edge to the  $p_i$ s ( $t_i, p_i$ ) and thus sends two edges to the  $w_i$ s. No  $w_i$  is adjacent to three  $t_i$ s as then  $w_i$  and  $t$  have 3 common neighbors, so each  $w_i$  sends two edges to the  $t_i$ s. This means that  $t$  and  $w_1$  share exactly two neighbors amongst  $t_2, t_3, t_4$  and by definition of  $w_1$  it shares two neighbors with  $v$  from amongst  $s_2, s_3, s_4$ . Then there is an  $i \neq 1$  with  $w_1$  adjacent to both  $s_i$  and  $t_i$ . If  $i = 2$  or  $i = 4$  then  $p_i$  and  $w_1$  have the three common neighbors  $p_1, s_i$  and  $t_i$  so  $i = 3$ .  $w_1$  sends a second edge to the  $s_i$ s, but above we noted not to  $s_1$ , by symmetry we may assume it goes to  $s_2$ . Then  $w_1$  is adjacent to  $t_4$  as  $w_1$  sends two edges to the  $t_i$ s, one of which is to  $t_3$  but the second may not be to  $t_1$  as then  $w_1, p_1, t_1$  forms a triangle and the second may not be to  $t_2$  as then  $t_2, t_3, s_2$  are common neighbors of  $w_1$  and  $p_2$ .

Finally, consider  $s_2$ . It is adjacent to some  $t_i$ , but if  $s_2$  is adjacent to  $t_2$  then there is a triangle  $s_2, t_2, p_2$ , and if  $s_2$  is adjacent to  $t_3$  then there is the triangle  $s_2, w_1, t_3$  and if  $s_2$  is adjacent to  $t_4$  then there is the triangle  $s_2, w_1, t_4$ , so  $s_2$  is adjacent to  $t_1$ . However, then  $s_2$  and  $p_1$  have common neighbors  $p_2, w_1, t_1$ . Again, the case of 3 or more common neighbors was dismissed above. So  $G$  can not be both triangle-free and avoid 6-IS. So  $R(3, 6) \leq 18$ . Carliolaro shows  $R(3, 6) > 17$  with a specific graph completing the proof. ■

We have that  $R(4, 5) \leq R(3, 5) + R(4, 4) = 14 + 18 = 32$ .

**Theorem 2.1.7**  $R(4, 5) = 25$ .

**Proof** Two computer programs written independently by Brendan D. McKay and Stanislaw P. Radziszowski both came to this conclusion in March 1993 though published in 1995.

Suppose there were a  $(K_4, K_5)$ -good 2-coloring of  $K_{25}$ . Removing any vertex gives a  $(K_4, K_5)$ -good 2-coloring of  $K_{24}$ . They make an observation about always being able to find a vertex having certain properties in any  $(K_4, K_5)$ -good 2-coloring of  $K_{25}$ , and decided to always remove this vertex. This imposes constraints on the  $(K_4, K_5)$ -good 2-colorings of  $K_{24}$  that might appear when this vertex is removed. They then build a catalog of all 350,866  $(K_4, K_5)$ -good 2-colorings of  $K_{24}$  meeting this constraint and show that adding a vertex with these properties does not extend any of them to a  $(K_4, K_5)$ -good 2-coloring of  $K_{25}$ . ■

Brendan D. McKay and Stanislaw P. Radziszowski comment that the catalog built “likely contains most but not all  $(K_4, K_5)$ -good 2-coloring of  $K_{24}$ ” which gives the impression that the real benefit gained from their observations was in restricting the number of extensions they needed to examine.

There is another ingredient to their code of great mathematical interest, particularly their use of optimization techniques. The constraints imposed by removing their vertex deal intricately with sums of other relevant vertex degrees, and are linear in nature. Essentially the graphs in the catalog are approximate solutions to a very large linear programming problem. What this means is they have an objective function that rewards being a  $(K_4, K_5)$ -good 2-coloring, decision variables which encode a 2-coloring, and the linear constraints on the decision variables imposed by the observations referenced above. For instance, a standard approach might have decision variables looking like  $x_{i,j} = 1$  if the edge between  $i$  and  $j$  is maroon and  $x_{i,j} = 0$  if the edge between  $i$  and  $j$  is orange. However, finding and enumerating all solutions to this type of discrete optimization problem is NP-hard. The standard approach is to relax the decision variables by allowing them to take any value between 0 and 1. The problem becomes a continuous optimization problem on a polytope (imposed by the constraints) which can be solved very quickly (polynomial time). Then as McKay and Radziszowski did, since now an edge may be .6 maroon and .4 orange, typically to give these solutions meaning, the decision variables are rounded to 0 or 1. The question becomes: are all solutions to the original optimization problem solutions to the relaxed



version? Designing the objective function so that this is the case is a task often requiring insight and skill.

It is hard to do justice to what Brendan D. McKay and Stanislaw P. Radziszowski have done for Ramsey theory, particularly in calculating bounds for Ramsey numbers. Their efforts go far beyond  $R(4, 5)$ , and they have improved too many bounds to list here, the most important of which may be  $R(5, 5) \leq 49$ , though they suspect the true value to be 43. Fortunately Radziszowski maintains the excellent dynamic survey on Small Ramsey Numbers in the *Electronic Journal of Combinatorics*. Radziszowski and McKay are both housed in computer science departments and while they use ingenious mathematics to cut down on computation, the computations they perform are truly enormous. In the case of  $R(4, 5)$ , 3.2 years of CPU time were required by the faster of their two codes.

How do the proofs of  $R(4, 5)$  and  $R(3, 6)$  compare? At first glance the technique of looking at the size of intersections of neighborhoods in the proof of  $R(3, 6)$  seems powerful. This technique is appealing as counting something local and having ramifications for the larger graph is a common theme in determining other small Ramsey numbers.

There is no longer a large mistrust or confusion regarding computer-aided proofs [see ch.22 of Soifer's *Mathematical Coloring Book*], but still a feeling that non-computer proofs equip the reader with more insight. As such a non-computer proof of  $R(4, 5) = 25$  is desirable and I have tried to use techniques from the proof of  $R(3, 6)$  on this problem with no success. The main difficulty seems to be that  $K_4$  has twice the number of edges of  $K_3$ , and building up enough edges to contradict the absence of a  $K_4$ , when starting from a single vertex, without breaking down the problem into tons of cases does not appear to be feasible. This difficulty does seem to be relevant as Cariolaro notes that a non-computer proof of  $R(3, 7) = 23$  exists and was known to Kalbfleisch in 1966. There is a fascinating connection between  $R(3, 8)$  and  $R(3, 9)$  in that  $R(3, 9) = 36$  was known to Grinstead and Roberts in 1982 while  $R(3, 8) = 28$  was first proved by McKay and Min in 1992.

## 2.2 Erdős' lower bound

The first real progress on lower bounds came from Paul Erdős' probabilistic techniques. To introduce the counting aspect I will initially disguise the role of probability.

**Theorem 2.2.1**  $e^{-1}2^{-1/2}n2^{n/2} \leq R(n, n)$ .

**Proof** The number of ways to 2-color a  $K_m$  is  $2^{\binom{m}{2}}$  since there are  $\binom{m}{2}$  edges each of which can be colored in 2 ways. Say we want to insist on having a monochromatic  $K_n$  subgraph. We may pick the  $n$  vertices where the monochromatic  $K_n$  is to occur in any of  $\binom{m}{n}$  ways. We then have 2 choices for the color of the  $K_n$  and this determines the color of the  $\binom{n}{2}$  edges in this subgraph. Having guaranteed a monochromatic  $K_n$  we can then color each of the other edges either color. Every 2-coloring of  $K_m$  with a monochromatic  $K_n$  can be formed in this manner. Therefore the number of 2-coloring of  $K_m$  with a monochromatic  $K_n$  subgraph is at most  $2\binom{m}{n}2^{\binom{m}{2}-\binom{n}{2}}$ . This over-counts the true number since a coloring with  $k$  monochromatic  $K_n$ s is counted  $k$  times. However, if the number of colorings of  $K_m$  exceeds the over-estimate on the number of such colorings containing a monochromatic  $K_n$ , then there must be some coloring without a monochromatic  $K_n$ . This means that if  $2^{\binom{m}{2}} > \binom{m}{n}2^{\binom{m}{2}-\binom{n}{2}+1}$  then  $R(n, n) > m$ . Dividing through by  $2^{\binom{m}{2}}$  on both sides gives if  $1 > \binom{m}{n}2^{1-\binom{n}{2}}$  then  $R(n, n) > m$ .

Now imagine a  $K_m$  is to be 2-colored where the color of each edge is determined by a coin flip, so each edge has probability 0.5 of being colored in each color. What is the probability of finding a monochromatic  $K_n$  in the resulting coloring? Given two events  $A$  and  $B$  in a probability space, the probability of either occurring is  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ , the sum of their individual probabilities minus the probability they both occur. The events we care about are potential monochromatic  $K_n$ s, there are  $\binom{m}{n}$  of these events, and each occurs with probability  $2^{1-\binom{n}{2}}$  since each of these has  $\binom{n}{2}$  edges and once one of them is colored the only way to get a monochromatic  $K_n$  is for all remaining edges to be given this color. From above we have  $P(A \cup B) \geq P(A) + P(B)$  and therefore the probability of such

a 2-coloring having a monochromatic  $K_n$  is at most  $\binom{m}{n}2^{1-\binom{n}{2}}$  which sums the probabilities of the individual events. If this is less than 1 by the definition of a probability there must be some 2-coloring of  $K_m$  not having a monochromatic  $K_n$  and again if  $1 > \binom{m}{n}2^{1-\binom{n}{2}}$  then  $R(n, n) > m$ . If the events were mutually exclusive, then the probability estimate would be exact, but for  $m > n$  coloring all edges orange shows that the events are not mutually exclusive. This method of overestimating a probability is called Boole's inequality.

Take  $n$  to be fixed (we are trying to get a bound on  $R(n, n)$ ). Define  $N$  to be the least  $m$  for which  $1 \leq \binom{m}{n}2^{1-\binom{n}{2}}$ .  $R(n, n) \geq N$  because by the definition of  $N$  and the argument above we have shown that there is a 2-coloring of  $K_m$  on  $N - 1$  vertices with no monochromatic  $K_n$ . Therefore  $R(n, n) \geq N = (N^n)^{1/n} > \left(\binom{N}{n}n!\right)^{1/n} \geq (2^{\binom{n}{2}-1}n!)^{1/n} = 2^{(n/2)-(1/2)-(1/n)}(n!)^{1/n} \approx n2^{n/2}e^{-1}2^{-1/2}(2^{3/(2n)}(\pi)^{1/(2n)}n^{1/(2n)}) > n2^{n/2}e^{-1}2^{-1/2}$ , where Stirling's approximation has been used to replace the  $n!$ . This result is just a factor of 2 short of the best known lower bound for which we will need the Lovasz Local Lemma. ■

## 2.3 Lovasz Local Lemma

If some collection of events  $A_1, A_2, \dots, A_n$  are independent of each other and have probabilities  $p_1, p_2, \dots, p_n$  all of which are  $< 1$ , then the probability that none of them occur is given by  $\prod_{i=1}^n (1 - p_i)$  and this product is greater than 0 so there is some scenario where none of the events occur. However, all too often the assumption of independence does not hold. Roughly speaking, the Lovasz Local Lemma says that if the probabilities are small enough and the events are for the most part independent, then there is some scenario where none of them occur. A *dependency graph* is a graph whose vertices represent events and there is an edge between two vertices if the events they represent are not independent.

**Lemma 2.3.1** (Lovasz Local Lemma [GSR2])

Let  $A_1, A_2, \dots, A_n$  be events with dependency graph  $G$ . Suppose that there exist  $x_1, x_2, \dots, x_n$

with  $0 < x_i < 1$  so that, for all  $i$ ,  $P(A_i) < x_i \prod_{j|(i,j) \in E(G)} (1 - x_j)$ . Then  $P(\cap_i \bar{A}_i) > 0$ .

**Proof** The idea is to prove, for any given set  $S$  of events not containing event  $i$ , that the probability of  $A_i$  given none of the events in  $S$  occur is less than  $x_i$ . This means the probability of none of the events in  $S$  occurring and  $A_i$  not occurring is larger than  $1 - x_i > 0$  as by assumption  $x_i < 1$ . Then taking  $S$  to be all but one of the  $A_i$ s shows that the probability of none of the  $A_i$ s occurring is greater than 0. [GSR2] accomplishes this by induction on the size of the set  $S$ . ■

**Corollary 2.3.2** (Lovasz Local Lemma (Symmetric case))

Let  $A_1, A_2, \dots, A_n$  be events with dependency graph  $G$  of maximal degree  $d$ . Suppose that for all  $i$ ,  $P(A_i) < p$  and  $ep(d+1) < 1$  then  $P(\cap_i \bar{A}_i) > 0$ .

**Proof** Take each  $x_i = 1/(d+1)$  in the Lovasz Local Lemma, which then demands that each  $P(A_i)$  be less than  $d^d/(d+1)^{d+1}$ . By definition of  $p$ , this is true if  $ep(d+1) < 1$ . The only purpose of introducing  $e$  is to make computations easier. ■

Lovasz who was working with Paul Erdős published this lemma in 1975. Yet during the same year it was Joel Spencer who used the lemma to improve the probabilistic work above by a factor of 2 and obtained the best known lower bound on  $R(n, n)$ . Only after the year 2000 was a constructive proof of the Lovasz Local Lemma discovered by Robin Moser and Gabor Tardos. As a warm-up first consider the following example using the symmetric version.

$11n$  points are to be placed on a circle and each of  $n$  colors will be used 11 times to color all of the points. No matter how the coloring is done some set of  $n$  points containing one point of each color has no consecutive pair on the circle.

Imagine that each point of each color has the same chance ( $1/11$ ) of being chosen as the representative of this color. Choosing consecutive points is to be avoided. The probability of choosing a consecutive pair is  $1/121$  if the colors of the points are different and 0 if they

are the same. So  $p = 1/121$  for the purposes of the lemma. Consider any pair of consecutive points  $(a, b)$ . There are 10 other points colored in the same color as  $a$  and 10 other points colored in the same color as  $b$ , each of which are in 2 consecutive pairs of points,  $a$  and  $b$  are each in one other pair other than  $(a, b)$ . Some of these pairs may be counted twice, but this shows that choosing  $(a, b)$  depends on at most  $2 \cdot 10 + 2 \cdot 10 + 1 + 1 = 42$  other pairs allowing  $d = 42$  to be used in the lemma. Since  $e \cdot (1/121) \cdot (42 + 1) < .97$ , there must be a way to select  $n$  points with one of each color and no two consecutive.

**Theorem 2.3.3**  $R(n, n) > e^{-1} 2^{1/2} n 2^{n/2}$ .

**Proof** The events to avoid are monochromatic  $K_n$ s. As in the previous theorem, the probability of a monochromatic  $K_n$  is  $2^{1-\binom{n}{2}}$ . The coloring of two  $K_n$ s is independent if they share no edges in common. Therefore in an  $m$  vertex graph the number of events a single event may depend on is at most  $\binom{n}{2} \binom{m}{n-2}$  which counts, for each edge in the  $K_n$ , the number of other  $K_n$ s this edge belongs to. By the symmetric version of the Lovasz Local Lemma, if  $e 2^{1-\binom{n}{2}} (\binom{n}{2} \binom{m}{n-2} + 1) < 1$  then  $R(n, n) > m$ . Simplifying using Stirling's formula as in the previous theorem yields the result. ■

## 2.4 Diagonal Ramsey numbers

**Definition** The  $n$ -th diagonal Ramsey number is  $R(n, n)$  and sometimes is denoted simply  $R(n)$ .

Since the Ramsey numbers are monotone in each input, to prove Ramsey's theorem for graphs, one only needs to show  $R(n)$  is finite. The next theorem uses a direct proof to prove this theorem as opposed to the inductive argument given in the introduction. The method of proof is fairly typical and one can adapt it to other settings where induction (on  $n$  or for  $R(m, n)$  on  $m + n$ ) does not work as smoothly.

This bound and the previous one both give that  $\lim_{n \rightarrow \infty} R(n, n)^{1/n} \geq 2^{1/2}$ .

**Theorem 2.4.1**  $R(n, n) \leq 2^{2^{(n-1)+1}} - 1$  [GSR2]

**Proof** Consider any 2-coloring of  $K_{2^{2^{(n-1)+1}}}$ . Call the set of vertices  $S_1$  and pick any vertex from  $S_1$  and label it  $x_1$ .  $x_1$  is incident to some number of maroon edges and some number of orange edges and one of these colors is used more frequently (if they are each used the same number of times choose either). Take  $S_2$  to be the set of vertices adjacent to  $x_1$  by an edge of the more frequently used color and color the vertex  $x_1$  this color. Then select any vertex from  $S_2$  and label it  $x_2$ . Amongst the edges from  $x_2$  to its neighbors in  $S_2$  some color is used most frequently. Define  $S_3$  to be the set of vertices in  $S_2$  adjacent to  $x_2$  by an edge of this (possibly new) color and color  $x_2$  this color. Continue in this manner picking  $x_i$  from  $S_i$  arbitrarily and looking at the color of the edges between  $x_i$  and the other vertices in  $S_i$  and then coloring  $x_i$  accordingly. Since originally there are  $2^{2^{(n-1)+1}} - 1$  vertices,  $2(n-1)$  times the vertices in  $S_1$  may be cut in half as above and still have at least 1 vertex to select, so at least  $2(n-1) + 1$  vertices are selected. By the pigeonhole principle  $n$  of these  $2(n-1) + 1$  vertices have the same color. Moreover, for each of pair  $x_i, x_j$  with  $i < j$  of these  $n$  vertices the edge between them is the color of  $x_i$  since  $x_j \in S_i$ , and since each of these  $n$  vertices are given the same color, these vertices and the edges between them form a monochromatic  $K_n$  subgraph. ■

Note the only role of the  $-1$  that is not in the exponent is to allow, on each splitting, the cardinality of  $S_i$  to possibly be one less than a power of 2. This is why, after  $2(n-1)$  splits there is still at least  $2^{2^{(n-1)+1-2(n-1)}} - 1 = 1$  vertex remaining. Generally for larger  $n$  one can subtract more and still eventually get within one of a power of 2.

This bound gives that  $\lim_{n \rightarrow \infty} R(n, n)^{1/n} \leq 4$ .

Using  $a = b = n$  in the upper bound derived from the proof of Ramsey's theorem in the introduction gives  $R(n, n) \leq \binom{2n-2}{n-1}$  which is a better bound than the one just proved. In

fact this bound gives  $R(n, n) \leq c4^n n^{-1/2}$  for some constant  $c$ . To get a feel for how much these bounds differ consider the following identity from the proof of Bertrand's postulate. It uses the binomial theorem.

$4^m = 0.5(1+1)^{2m+1} = 0.5 \sum_{k=0}^{2m+1} \binom{2m+1}{k} > 0.5(\binom{2m+1}{m} + \binom{2m+1}{m+1}) = \binom{2m+1}{m}$  Taking  $m = n-1$  gives  $R(n, n)^{1/n} \leq (4^{n-1})^{1/n}$  so that  $\lim_{n \rightarrow \infty} R(n, n)^{1/n} \leq 4$ .

Of course, the  $n^{-1/2}$  factor was lost when sacrificing terms from the summation. In these types of problems Stirling's formula can often be used to move from exponential functions to binomial coefficient with more efficiency.

**Corollary 2.4.2**  $R(n; r) \leq r^{(n-1)r+1} - 1$  .

**Proof** Use  $r$  colors and replace 2 with  $r$  in the proof of Theorem 2.4.1. ■

**Corollary 2.4.3**  $2^{1/2} \leq \lim_{n \rightarrow \infty} R(n, n)^{1/n} \leq 4$ .

In 1947 Erdős offered 100 dollars for a proof that  $\lim_{n \rightarrow \infty} R(n, n)^{1/n}$  existed and another 250 for its exact value. A dollar in 1947 had about 10 times the purchasing power of a dollar today.

## 2.5 A recursive upper bound for diagonal Ramsey numbers

In 1968 K. Walker published the bound  $R(k, k) \leq 4R(k, k-2) + 2$  by generalizing methods Kalbfleisch used to determine specific small Ramsey numbers. They examine *extremal colorings*, which use the maximum possible number of edges of a certain color and still have the typical properties of a Ramsey-type problem. With a small change in notation Walker's proof is followed and expanded upon below.

**Lemma 2.5.1** *The number of monochromatic triangles,  $\Delta$  in a maroon-orange coloring of  $K_n$  where the maroon-degree of vertex  $i$  is  $m_i$  is  $\Delta = \binom{n}{3} - (1/2) \sum_{i=1}^n m_i(n - m_i - 1)$*

**Proof**  $(1/2) \sum_{i=1}^n m_i(n - m_i - 1)$  counts the number of non-monochromatic triangles. To see this, note that a non-monochromatic triangle has two vertices at which edges of different colors meet. Since  $m_i(n - m_i - 1)$  counts the number of pairs of opposite colored edges meeting at vertex  $i$ ,  $\sum_{i=1}^n m_i(n - m_i - 1)$  double-counts the number of non-monochromatic triangles which multiplying by  $(1/2)$  resolves. Finally  $\binom{n}{3}$  is the total number of potential monochromatic triangles from which the number of non-monochromatic triangles is subtracted to obtain the number of monochromatic triangles. ■

**Theorem 2.5.2**  $R(k, k) \leq 4R(k, k - 2) + 2$ .

**Proof** Let  $n < R(k, l)$  and define  $M(n, k, l)$  to be the maximum number of maroon edges in a  $(K_k, K_l)$ -good coloring of  $K_n$ . Likewise, define  $O(n, k, l)$  to be the maximum number of orange edges in such a coloring. For  $n < R(k - 1, l) + R(l, k - 1)$  suppose that a  $(K_k, K_l)$ -good coloring exists and define  $p_j$  to be the number of vertices in this color incident with  $j$  maroon edges.  $p_j = 0$  unless  $n - R(k, l - 1) \leq j < R(k - 1, l)$  since  $n - j - 1$ , which is the number of orange edges, must be less than  $R(k, l - 1)$  to avoid monochromatic orange  $K_l$  and  $j$  must be smaller than  $R(k - 1, l)$  to avoid monochromatic maroon  $K_k$ . A vertex  $v$  with  $j$  maroon edges forces these neighbors to form a subgraph with a  $(K_k - 1, K_l)$ -good coloring and each maroon edge in this subgraph forms a monochromatic maroon triangle with  $v$ , so  $v$  is in at most  $M(j, k - 1, l)$  monochromatic maroon triangles. The same reasoning shows that  $v$  is in at most  $O(n - j - 1, k, l - 1)$  orange triangles. Of course, triangles have three vertices so that  $\Delta \leq (1/3) \sum_{j=n-R(k,l-1)}^{R(k-1,l)-1} [M(j, k - 1, l) + O(n - j - 1, k, l - 1)]p_j$ . Combining this result with the lemma and multiplying both sides by 6 gives that  $n(n - 1)(n - 2) \leq \sum_{j=n-R(k,l-1)}^{R(k-1,l)-1} [2M(j, k - 1, l) + 2O(n - j - 1, k, l - 1) + 3j(n - j - 1)]p_j$ .

In a  $(K_k, K_l)$ -good 2-coloring each vertex has at most  $R(k - 1, l) - 1$  maroon edges, so  $M(n, k, l) \leq (n/2)(R(k - 1, l) - 1)$ . In particular,  $M(j, k - 1, k) \leq (1/2)j(R(k - 2, k) - 1)$ .



Likewise  $O(n, k, l) \leq (n/2)(R(k, l-1) - 1)$ . In particular  $O(n-j-1, k, k-1) \leq (1/2)(n-j-1)(R(k, k-2) - 1)$ . Now if  $n \geq 4R(k, k-2) + 2$  had a  $(K_k, K_k)$ -good 2-coloring then for each  $j$   $[2M(j, k-1, k) + 2O(n-j-1, k, k-1) + 3j(n-j-1)] \leq [2(1/2)j(R(k-2, k) - 1) + 2(1/2)(n-j-1)(R(k, k-2) - 1) + 3j(n-j-1)] = (n-1)(R(k, k-2) - 1) + 3j(n-j-1) \leq (n-1)(R(k, k-2) - 1) + 3((n-1)/2)^2$ , where clearly  $j = (1/2)(n-1)$  maximizes  $3j(n-j-1)$ . By assumption  $R(k, k-2) - 1 \leq (1/4)(n-2) - 1 = (1/4)(n-6)$  gives  $[2M(j, k-1, k) + 2O(n-j-1, k, k-1) + 3j(n-j-1)] \leq (n-1)(R(k, k-2) - 1) + 3((n-1)/2)^2 \leq (n-1)((1/4)(n-6) + (1/4)3(n-1)) = (n-1)(n-9/4) < (n-1)(n-2)$ . This was true for each  $j$ , and  $\sum_{j=n-R(k, k-1)}^{R(k-1, k)-1} p_j = n$  since the  $p_j$ s count the vertices by their red degree. So  $\sum_{j=n-R(k, k-1)}^{R(k-1, k)-1} [2M(j, k-1, l) + 2O(n-j-1, k, l-1) + 3j(n-j-1)] p_j < n(n-1)(n-2)$ . But above it was shown that if  $n \leq R(k, k-1) + R(k-1, k)$  then the last inequality is  $\geq$  and since  $R(k, k) \leq R(k, k-1) + R(k-1, k)$ , so  $n = 4R(k, k-2) + 2 \geq R(k, k)$ . ■

In this paper Walker also proves that  $R(5, 5) \leq 57$ , a small improvement on  $R(5, 5) \leq 4R(5, 3) + 2 = 4 * 14 + 2 = 58$ . He then goes on to show that  $R(4, 5) \leq 29$ . Each of these results comes from using linear programming to maximize quantities in the inequalities given above.

## 2.6 Comparing near diagonal Ramsey numbers

$R(k, k) \leq 4R(k, k-2) + 2$  and  $R(k, k) \leq R(k, k-1) + R(k-1, k) = 2R(k, k-1) \leq 2R(k, k-2) + 2R(k-1, k-1)$  have now both been proven. The question remains, which is a better upper bound? This reduces to which is smaller,  $R(k, k-2) + 1$  or  $R(k-1, k-1)$ ? Despite feeling dangerously close to proving that  $R(k, k-2) + 1$  was smaller, I never quite managed it. I can prove that for  $p = 0.5$  the expected number of monochromatic  $K_{k-2}$  red subgraphs exceeds the number of expected red or blue monochromatic  $K_{k-1}$  subgraphs for  $2 \leq n < R(k, k)$ , but because there are plenty of graphs on  $n < R(k, k)$  vertices with multiple monochromatic subgraphs, these expectations exceed 1 and are useful only as a

heuristic and not for showing the existence of graphs with specific properties. Of course, I expect that whenever the sums of the inputs are the same, the more central Ramsey number is larger. I have a proof that shows if this is eventually true then it is always true using backwards induction.

In a 1980 problem paper Erdős remarks that “Faudree, Schelp, Rousseau needed recently a lemma stating  $\lim_{n \rightarrow \infty} (R(n+1, n) - R(n, n))/n = \infty$ ; we could prove this without much difficulty.” Only much later, in 1989, does anything resembling this result get used in a paper of Erdős but that paper only shows  $(R(n+1, n) - R(n, n)) > 2n - 3$ . So at this point it is a mystery — Ron Graham did not know — whether the result Erdős quoted but never published is in fact true. I found this most intriguing. One possible strategy is to consider as  $n$  goes to infinity, 2-colorings of  $K_{R(n, n) - 1}$  avoiding monochromatic  $K_n$  and try to extend it by  $(c_n)n$  vertices, where  $c_n$  is a sequence going to infinity, to a 2-coloring avoiding blue  $K_{n+1}$  and red  $K_n$ . Perhaps with high probability the edges amongst these  $c_n$  vertices are colored red and with high probability the edges between the original vertices and the  $c_n$  vertices are colored blue. Then essentially the only way to form a blue  $K_{n+1}$  is to find a blue  $K_{n-1}$  in the original graph and a vertex amongst the added vertices that are totally connected by blue edges. The first two of those things seem to be a bit rare. However, the concept of Ramsey multiplicity, which tries to count the least number of monochromatic  $K_{n-1}$ s in a 2-coloring of  $K_{R(n, n) - 1}$  does not have very well-developed bounds. Essentially the only way to find a red  $K_n$  is to look amongst the added vertices. A word of caution though, even if Erdős’ 1980 claim is correct, to prove it the sequence  $c_n$  ought not to be taken as  $c_n = n$  since Soifer points out it is not known whether  $(R(n+1, n) - R(n, n))/n^2$  is bounded below by any  $c > 0$ . In the 1980 paper Erdős wanted to prove that this difference increased faster than any polynomial but admitted he was unable to. Also of interest would be to show that  $R(n+1, n) \geq (1+c)R(n, n)$  for some  $c > 0$ .

## 2.7 $R(3, k)$

In Soifer's very recent book *Ramsey Theory: Yesterday, Today, and Tomorrow* he presents a collection of writings on the topic by modern day experts. Joel Spencer wrote the excellent chapter on  $R(3, k)$ .  $R(3, k)$  can be described efficiently in graph-theoretic terms, rather than by coloring terminology, and this section will take advantage of that. Recall that  $R(3, k)$  is the least  $n$  so that any triangle-free graph on  $n$  vertices has an independent set of size  $k$ . Spencer likes to give a computer science flavor to his proofs.

**Theorem 2.7.1**  $R(3, k) \leq k^2$ .

**Proof** Let  $G$  be triangle-free and have  $k^2$  vertices. If  $G$  has a vertex  $v$  of degree  $k$  or higher,  $n'(v)$  forms an independent set of size at least  $k$ . Assume all vertices of  $G$  have degree less than  $k$ . Select any vertex  $v$  and remove  $n(v)$  from  $G$ . Since at most  $k - 1 + 1 = k$  vertices are removed,  $G$  has at least  $k^2 - k$  vertices remaining. Continue removing the neighborhoods of vertices. After  $k - 1$  removals at least  $k^2 - (k - 1)k = k$  vertices remain, so at least  $k$  removals can occur. In the original graph the  $k$  vertices whose neighborhoods were removed were independent since if two such vertices were adjacent they would be removed on the same removal and only one of their neighborhoods would get removed. ■

**Theorem 2.7.2** *There is a triangle-free graph on  $n$  vertices with no independent set of size  $cn^{2/3} \ln(n)$ .*

**Proof** Unfortunately Spencer's proof needs a small tweak, he drops a  $2^3$  but it is important to his argument. In the probability space of random graphs  $G(2n, p)$  with  $2n$  vertices and edge probability  $p$  consider  $p = n^{-2/3}$ . By the linearity of expectation the expected number of triangles is  $\binom{2n}{3} n^{-2/3^3} < 4n/3$ . At this point Spencer has  $< n/6$  as if the graph had only  $n$  vertices, however, in Spencer's proof, this expectation needs to be  $< n$ . But if we use  $p = n^{-2-\epsilon/3}$  the expectation is  $4n^{1-3\epsilon}/3$  and the theorem eventually holds for independent

sets of size  $cn^{(2+3\epsilon)/3}\ln(n)$ . The proof philosophy is more important than the details and in any case Spencer's next theorem improves the result. So assume the expected number of triangles is less than  $n$ . By the choice of  $p$  the expected number of independent sets of size  $2.01n^{(2+3\epsilon)/3}\ln(n)$  is less than one. Therefore eventually the expected number of triangles or independent sets of the above size is less than  $n$ , and thus eventually there are always such graphs. Then on a graph with less than  $n$  triangles or independent sets, a vertex from each triangle or independent set may be removed when the resulting graph has at least  $2n - n = n$  vertices and no triangles or independent sets! ■

**Theorem 2.7.3** *There are graphs on  $n$  vertices with no triangles or independent sets of size  $cn^{1/2}\ln(n)$ .*

**Proof** The result is Erdős' from 1961. Firstly note that this is an improvement over the previous result because the permissible size of the independent sets has shrunk, so the graph is more strongly connected, and yet triangles are still avoided.

Consider  $G(n, p)$  with  $p = \epsilon n^{-1/2}$  for a small constant  $\epsilon$  and take  $x = cn^{-1/2}\ln(n)$  where  $c$  is a large constant. Define a set  $I$  of  $x$  vertices as a failure if whenever  $u, v \in I$  are adjacent in  $G$ , then there is a triangle  $u, v, z$  in  $G$  where  $z$  is not in  $I$ . Erdős shows that in  $G(n, p)$  there is positive probability of having no failures. As a heuristic Spencer gives a concise argument for why this would be true if for vertices  $u, v, w$  in  $I$  the probability of extending the edge  $uv$  to a triangle with third vertex outside  $I$  were independent of the probability of extending  $uw$  to a similar triangle. But these events clearly depend on each other for if  $z$  is not in  $I$  the knowledge that  $u, v, z$  is not a triangle in  $G$  increases the probability that  $u, w, z$  is not a triangle in  $G$ . For a graph in  $G(n, p)$  having no failures, find a triangle-free subgraph  $H$  in a greedy manner. To do this order the edges of  $G$  arbitrarily and include an edge in the subgraph if it does not form a triangle with the edges already added to the subgraph.  $H$  is triangle-free by design. If  $I$  is a set of  $x$  vertices then  $I$  is not a failure and  $I$  has an edge  $uv$  that is a member of at most triangles in  $I$ . If  $uv \in H$  then  $I$  is not independent in the subgraph  $H$ . If  $uv$  is not in  $H$  there must be  $w$  in  $I$  forming a triangle  $u, v, w$  in  $G$  with

$uw$  and  $wv$  in  $H$ , but the mere presence of  $uw$  in  $H$  then shows  $I$  is not independent in  $H$ . Therefore  $H$  has no independent sets of  $x = cn^{-1/2} \ln(n)$  vertices. ■

**Theorem 2.7.4** *Let  $G$  be triangle-free on  $n$  vertices and the average degree in  $G$  be at most  $k$ . Then there is an independent set  $I$  of size  $|I| \geq c(n/k) \ln(k)$*

**Proof** The proof requires Szemerdi's regularity lemma which roughly speaking says that for  $n$  large enough, every graph can be partitioned in such a way that the edges between the partitions behave randomly.

The Lovasz Local Lemma can be applied to the case of  $R(3, k)$  as it was for  $R(k, k)$ . The result given by Spencer is that if there are constants  $p, y, z \in [0, 1]$  with  $p^3 \leq y(1-y)^{3n}(1-z)^{\binom{n}{k}}$  and  $(1-p)^{\binom{k}{2}} \leq z(1-y)^{nk^2/2}(1-z)^{\binom{n}{k}}$  then  $R(3, k) > n$ . Spencer himself used this setup to prove Erdős' 1961 result in a different manner.

Finally in 1995 Heong-Jan Kim was able to mostly resolve the asymptotics of  $R(3, k)$ . His solution was an analysis of an algorithm. Recall, above forming  $H$  as a subgraph of  $G$  in a greedy triangle-free manner. Erdős, Winkler, and Suen looked at this algorithm where  $K_n$  plays the role of  $G$  and looked at the independent sets as the number of edges added increased. Kim modified their approach by using a nibble algorithm where at each step rather than choosing one edge randomly, a carefully selected collection of edges is chosen. Through analyzing this algorithm with martingales Kim was able to show the following:

**Theorem 2.7.5**  $c_1 k^2 / \ln^2 k < R(3, k) < c_2 k^2 / \ln^2 k$ .

Finding a  $c$  so that these expressions match as  $k$  gets large is still an open problem. Moreover, in the spirit of the previous section Erdős wanted a proof or disproof that  $R(3, n+1) - R(3, n) > o(n)$ , even the boundedness of  $R(n+1, 3) - R(n, 3)$  is an open question.

## 2.8 Combinatorial Games and Ramsey Numbers

This is an attempt at bringing Ramsey numbers closer to real life. As noted by Slany, combinatorial games “serve as models that simplify the analysis of competitive situations with opposing parties that pursue different interests” and “finding a winning strategy to a combinatorial game can be translated into finding a strategy to cope with many kinds of real world problems such as found in telecommunications, circuit design, scheduling, as well as a large number of other problems of industrial relevance.”

Ramsey graph games start with a complete graph  $K_n$  and the players color an edge on their turn, each player uses a unique color. In an avoidance game each player is given a graph which if she colors a subgraph isomorphic to this graph monochromatically in her color, she loses. In an achievement game she would win by coloring such a subgraph monochromatically. The avoidance version is called Sim after Simmons who introduced it in 1969. Sim has been studied extensively, notably by Frank Harary who is known for his work in graph theory.

Starting with the achievement game, which resembles the popular Tic-Tac-Toe, suppose as in the case of  $R(3, 3) = 6$  that there are 2 players and they are trying to form monochromatic  $K_3$ s. If the graph they are coloring is  $K_6$ , then there will be a winner because we have proven  $R(3, 3) = 6$  which means that any way the two players color their graph, someone will eventually color a  $K_3$ . However, if they are playing on a  $K_5$  then the unique  $R(3, 3)$  critical graph shows that they may tie, meaning neither achieves their goal. If they play on  $K_3$  they must tie as there is only a single  $k_3$  to color and each will color at least one of the edges.

**Theorem 2.8.1** *In the achievement Ramsey games, if all players are trying to achieve the same graph, and there is a winning strategy, then it belongs to the first player.*

**Proof** This is a standard strategy stealing argument. If someone other than the first player, say the  $k$ th-player, had a winning strategy then the first player imagines she is this player.

On each turn she imagines  $k - 1$  additional edges have been colored one each in the colors of the  $k - 1$  players preceding her, and then selects an edge to color dictated by the  $k$ th-player's winning strategy. But in following this strategy, the first player does not lose because each of her moves were part of a winning strategy which means that she could not have lost before her next turn. If there are not  $k$  edges left on player one's turn, this means player one did not follow player  $K$ 's strategy as player  $K$  wins at latest on the turn before. If the first player does not lose, then the  $k$ th-player does not win. So only player one may have a winning strategy.

Continuing the example above. If the players are attempting to form monochromatic  $K_3$ s on  $K_4$ , then as  $K_4$  has only 6 edges, if the first player wins, she must do so on her third turn. However, after two turns she threatens to complete at most one  $K_3$  on her next turn, and if player 2 after seeing player one's second move, is sure to color the missing edge in the  $K_3$  that player one threatens to complete, player 2 achieves at least a draw. So again, with best play on a  $K_4$  the result is a tie because the above theorem indicates player one should not lose. If they play on a  $K_5$  player one wins. Player one need only stop player 2 and avoid forming a 5-cycle as this is the unique  $R(3, 3)$  critical graph. It turns out that this is always possible, the observation that achieving 3 edges of the same color incident to a vertex disallows 5-cycles will help in executing this strategy. Finally, for  $n \geq 6$  if they play on  $K_n$ , Ramsey theory promises that one of the player wins, when the theorem indicates that the first player has a winning strategy. So Ramsey theory is intimately linked to the outcome of this combinatorial game.

**Corollary 2.8.2** *In the achievement Ramsey games there is an  $n$  for which if  $m \geq n$  and two players play on  $K_m$  and strive to achieve the same graph, then the first player wins with best play.*

**Proof** If both players are attempting to achieve graph  $G$ , take  $n = R(G, G)$  then if  $m \geq n$  any 2-coloring of  $K_m$  yields a monochromatic  $G$  subgraph, so some player wins. If player two

had a winning strategy, then player one would steal it. So player two does not have a winning strategy, but player one is the only other player, so there is some strategy for player one in which player two can not force victory. As somebody must win on this largely chosen graph, using this strategy player one must win, and therefore player one has a winning strategy.

■

Note that the corollary does not hold for more than 2 players since I have removed the assumption that there is a winning strategy. With three or more players, a group may team up against the first player when no player has a winning strategy. Many business and social dynamics mirror this situation where each player is aiming at a common goal, but if one player gets too close to achieving it the others will set aside their own ambitions to block this player.

Sim, the avoiding Ramsey game, is less intuitive. Aside from avoiding coloring edges that would complete a triangle for the opponent, which can be colored at any time, at first glance it is hard to see much strategy.

Here the strategy stealing argument used above does not work because unlike in achievement the absence of  $k - 1$  edges to arbitrarily color does not indicate that player  $k$  could have won on the previous turn (and thus that player one could have won on the previous turn). The difference is that in this game being the player to move can be a disadvantage or an advantage, whereas in the achievement game it could never hurt to be the player to move. Even in the 2-player avoidance setting neither player will have a strategy stealing argument.

For standard Sim played on  $K_6$  with two players trying to avoid  $K_3$ , an exhaustive computer search determines a win for the second player. Of course even in the most general avoidance setting with multiple players trying to avoid different graphs, the Ramsey number of these graphs provides an upper bound on  $n$  for which if they play on  $K_n$  at least one player losses.

The natural next game to examine is where one player tries to achieve some monochromatic subgraph in their color and the opponent tries to prevent this (without needing to achieve a



specific monochromatic subgraph). This is a modification of the first game where there are no ties and is called Pekec's game.

**Lemma 2.8.3** *If Pekec's game is played on  $K_m$  and the achiever is trying to create a monochromatic  $H$  where  $m \geq 2R(H, H)$  then the achiever wins with best play.*

**Proof** Divide the edges into three sets, two which form the complete subgraph on  $R(H, H)$  vertices and the other having the remaining edges. The achiever's strategy is to assign an isomorphism between the two  $K_{R(H, H)}$  subgraphs and whenever the opponent colors an edge in one of these subgraphs the achiever makes sure to have the corresponding edge colored her color in the corresponding subgraph. Notice that coloring of edges in the third set only serves to determine which player colors first in the two  $K_{R(H, H)}$  graphs, whenever the opponent colors in the third set, so does the achiever until there are no more edges in the third set. Regardless of which player colors first in the  $K_{R(H, H)}$  graphs the achiever is able to ensure that these graphs swap colors on each edge. By definition of  $R(H, H)$  if in the first of these graphs the achiever's monochromatic  $H$  does not appear then in her color, it appears in her opponents. But this means that in the second of these graphs the monochromatic  $H$  appears in the achiever's color. ■

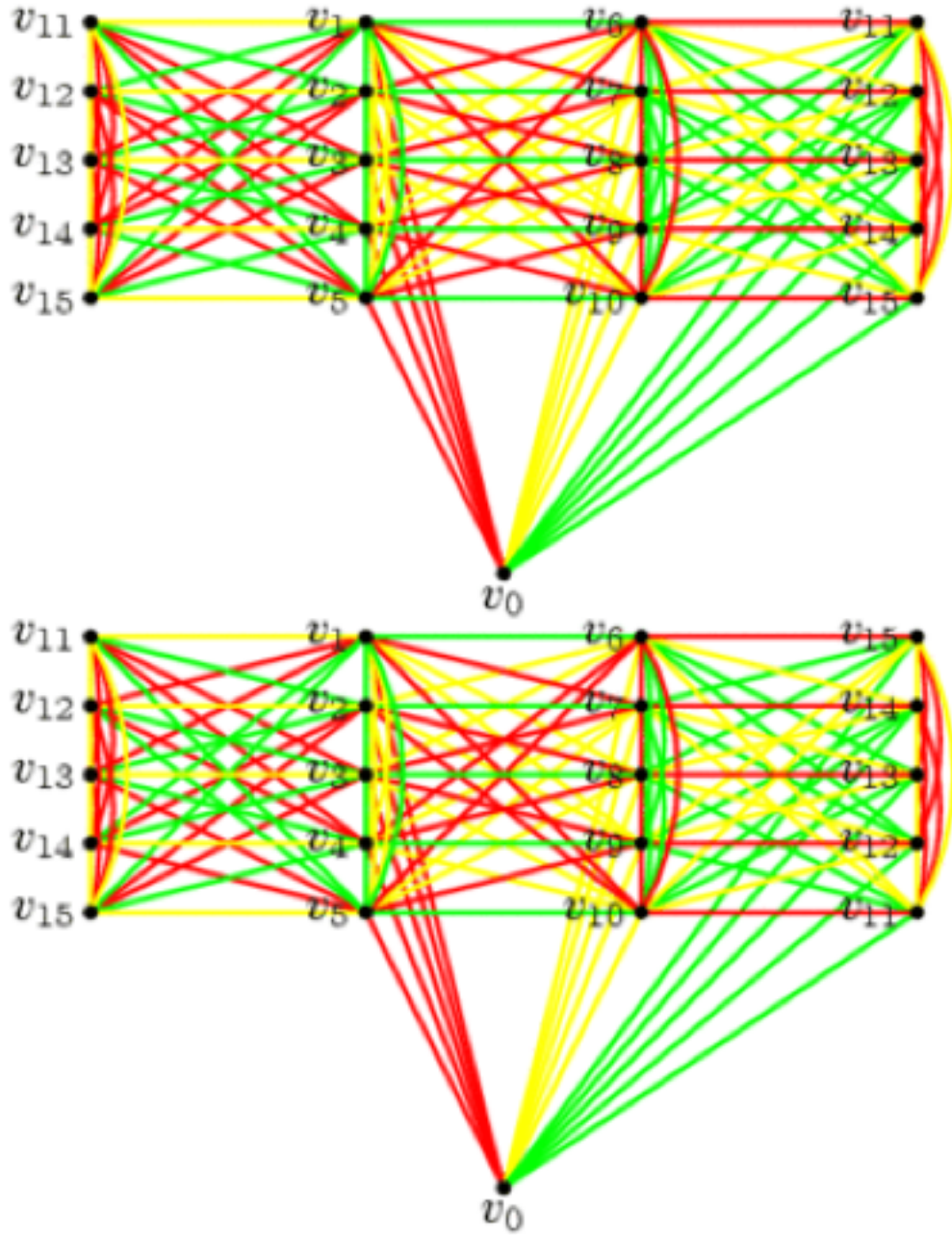
Pekec has other strategies to help the achiever win on smaller graphs, but this is another nice application of graph Ramsey theory.

**Definition** The *size Ramsey numbers*,  $\bar{R}(G, G)$  is the minimum number of edges a graph  $H$  must have so that when 2-colored a monochromatic  $G$  subgraph must appear. For clarity this means we find the graph  $H$  with the least number of edges with the property that any 2-coloring has a monochromatic  $G$  and then we count the edges in  $H$ . These have been studied quite a bit, originally motivated as a tool to help compute Ramsey numbers.

In this final combinatorial game, called an *Online-Ramsey* game, the size Ramsey numbers play a role. In this two player game the Builder places edges on a graph one at a time and

the Painter must choose one of two colors for this edge. The Painter is given a graph to avoid painting any monochromatic copies of. The Builder may be constrained in the sense that at each step the added edge must create a graph in a family of graphs known to both players. The connection to size Ramsey numbers is that if the family is all graphs, then the size Ramsey number of  $G$  with itself  $\bar{R}(G, G)$  is an upper bound on the number of turns required for Builder to win since Builder builds the corresponding  $H$ . However, Builder can often win in far fewer turns because Painter must commit to colors on edges without knowing which edges come next.

Figure 2.4:  $R(3, 3, 3) > 16$



# Chapter 3

## Chromatic number of the plane

### 3.1 Vertex Coloring

Here consider only simple graphs and no hypergraphs. Recall from the compactness section that a *legal* vertex coloring  $\chi : V(G) \rightarrow [r]$  is an assignment of one of  $r$  colors to each vertex of  $G$  in such a way that no adjacent vertices are assigned the same color. The least number of colors permitting a legal coloring is the chromatic number  $\chi(G)$ .

$n$  colors are required to legally color  $K_n$  since each vertex is adjacent to all other vertices and no adjacent vertices may be assigned the same color. However, coloring is much more tricky. The assertion that  $\chi(G) \geq r$  tells very little about the graph  $G$ . Consider  $C_5$  the cycle on 5 vertices.  $\chi(C_5) = 3$  for if there was a red-blue 2-coloring and vertex 1 were red, then vertices 2 and 5 are blue, but 2 being blue means 3 is red and 5 being blue means 4 is red when the adjacent vertices 3 and 4 share the same color. If we had 3 colors, we could color 4 green and produce a legal 3-coloring. This shows that  $\chi(G) = 3$  does not imply  $G$  has a  $K_3$  or  $C_3$  subgraph. Triangles are components of every  $K_n$  for  $n \geq 3$ . However, it was known to Blanche Descartes in the 1940s that triangle-free graphs could have arbitrarily high chromatic number. These two theorems are given in *Proofs from the Book* by Martin

Aigner, Günter M. Ziegler.

**Theorem 3.1.1** *There is a sequence of triangle-free graphs  $G_3, G_4, \dots$  with  $\chi(G_n) = n$ .*

**Proof** As above take  $G_3 = C_5$  which has  $\chi(G_3) = 3$  and has no triangles. To use induction on  $n$ , assume that there is a graph  $G_n$  with  $\chi(G_n) = n$  with vertex set  $V$ . Define  $G_{n+1}$  as having  $V(G_{n+1}) = V \cup V' \cup x$ . Construct  $G_{n+1}$  from  $G_n$  as follows. For each  $v \in V$  add a copy  $v' \in V'$  with the same neighbors as  $v$  and neighboring  $x$  which is an additional vertex. By construction the  $v'$  are not members of triangles since they are not adjacent to each other and if  $v'$  formed a triangle with vertices  $a, b \in G_n$  then  $v, a, b$  is a triangle in  $G_n$ . Since  $x$  is only adjacent to  $v'$ s and these are not adjacent,  $x$  is not in any triangles. The vertices of  $G_n$  by assumption do not form triangles amongst themselves. So  $G_{n+1}$  is triangle-free. Finally,  $\chi(G_{n+1}) = n + 1$ . To see this first note that because  $\chi(G_n) = n$  if  $G_n$  is legally  $n$ -colored, for each color  $i \leq n$  there is some  $v_i \in V$  with  $\chi(v_i) = i$  and each other color appears in  $n(v_i)$ . If no such vertex existed, by changing the color of all vertices  $v_i$  of color  $i$  to a color not used in  $n(v_i)$  a  $(n - 1)$ -coloring is formed which is ruled out by assumption. Since the neighborhood of  $v'$  matches the neighborhood of  $v$ , if we try to maintain a legal  $n$ -coloring  $\chi(v') = \chi(v)v \in V$ . Thus  $x$ , which is adjacent to each vertex of  $V'$  can not take on any of the  $n$  colors, but an assignment of color  $n + 1$  to  $x$  gives a legal  $(n + 1)$ -coloring. ■

This proof appears to be fairly simple because it applies to all legal  $n$ -colorings the strategy one would naturally use if trying to accomplish the task for a specific  $n$ -coloring of  $G_n$ . However, some authors prove this theorem in substantially different ways.

One observation is that the graph formed in the above proof contains many 4-cycles. Specifically if  $y, z \in n(v)$  then  $y, v, z, v'$  forms a 4-cycle. So, while avoiding 3-cycles, 4-cycles are seemingly everywhere. Define the *girth*  $\gamma(G)$  to be the smallest cycle in  $G$ . The following theorem is attributed to Erdős.

**Theorem 3.1.2** *There exists a graph  $G$  with  $\chi(G) > k$  and  $\gamma(G) > k$  for  $k \geq 2$ .*

**Proof** The complete proof can be found in the *Proofs from the Book* referenced above. The strategy is to look at the probability space  $G(n, p)$  of random graphs on  $n$  vertices with edge probability  $p$  and compute the probability of  $\gamma(G) \leq k$  and separately compute the probability of  $\chi(G) \leq k$ . Then for fixed  $k$  the idea is to find pairs  $(n, p)$  where both of these probabilities are smaller than .5. This then means that the probability of  $\gamma(G) > k$  and  $\chi(G) > k$  both occurring simultaneously is greater than  $1 - .5 - .5 = 0$  showing the existence of the desired type of graph. ■

There are explicit constructions that prove the same result. However, the  $n$  given by the probabilistic argument is far smaller than the number of vertices used by these constructions. This is to be expected since if the probabilistic analysis is performed without estimating the smallest possible  $n$  will be obtained and this problem is easy enough to allow good estimates. Since it is easy to define small graphs of large girth, where these constructions fall short is forcing a high chromatic number on only a few vertices. These themes are at play in the fascinating problem of the chromatic number of the plane.

## 3.2 Coloring the plane

The remaining material in this chapter summarizes results from Alexander Soifer's excellent *Mathematical Coloring Book*. The pretty diagrams are also from this book. *The plane* refers to its normal interpretation,  $R^2$ . The chromatic number of the plane problem can be stated as:

What is the least number of colors required to color the plane in such a way that a color is never assigned to points that are distance one apart?

**Definition** Define the unit distance graph by letting each point in the plane be a vertex and place an edge between every pair of points that are distance one apart.

An equivalent formulation of the chromatic number of the plane problem is to ask for the chromatic number of the unit distance graph. Soifer describes in detail how Edward Nelson, as an 18-year-old at the University of Chicago, invented this problem in 1950 and presents an unfortunate history of other authors attributing the problem's creation to other individuals. Let  $\chi$  be the answer to the chromatic number of the plane problem.

**Proposition 3.2.1**  $\chi \geq 3$

**Proof** Consider any equilateral triangle in the plane with side length one. The vertices of this triangle must be colored uniquely, so at least three colors are required to color the unit distance graph legally. ■

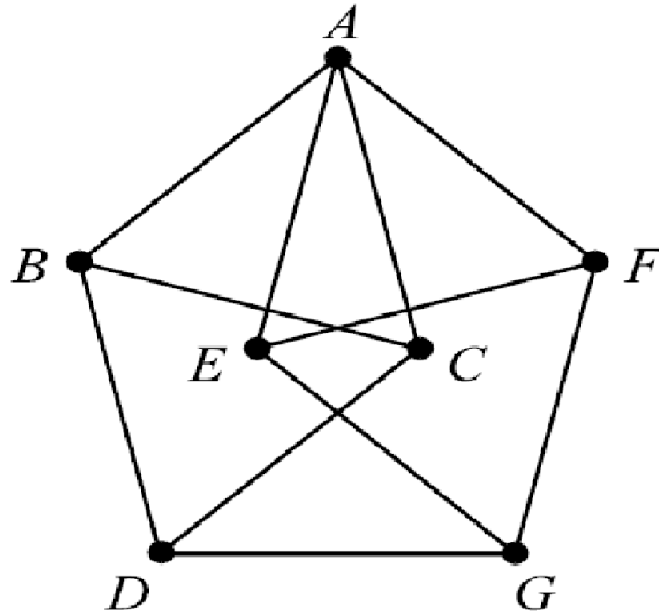
**Theorem 3.2.2**  $\chi \geq 4$

**Proof** Consider Moser's Spindle all of whose edges are length one. Suppose there were a three coloring of Moser's Spindle. The equilateral triangles  $ABC$  and  $BCD$  show that  $A$  and  $D$  are the same color while the equilateral triangles  $AEF$  and  $EFG$  show that  $A$  and  $G$  are the same color. However, this means that the adjacent vertices  $D$  and  $G$  are the same color, a contradiction. Therefore at least four colors are required to color the unit distance graph legally. The Moser brothers showed that their spindle has the smallest number of vertices, seven, that any 4-chromatic unit distance graph can have. ■

Another view of this proof is to just look at  $ABCD$  and swing it in a circle with center  $A$ . Then on the circle traced by  $D$  there is a point  $D'$  a distance one from  $D$ . But three-colorings have every point on this circle the same color as  $A$ , so  $D$  and  $D'$  are the same color, and four colors are required.

A second proof is available that may give a more general understanding. Any three points of Moser's spindle contains two that are distance one apart. Suppose not. Amongst the

Figure 3.1: Moser's Spindle

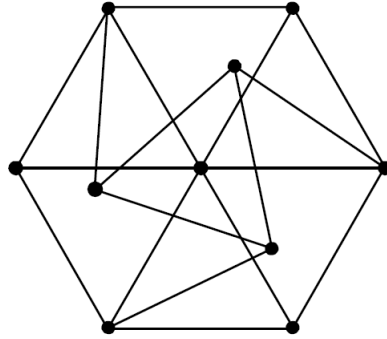


two pairs of equilateral triangles in the first proof, two of any three vertices are in one pair of equilateral triangles, allowing, without loss of generality, the assumption that two of the three points are  $A$  and  $D$ . But  $A$  and  $D$  dominate the graph. This implies that each color is used at most twice. Since Moser's spindle is on seven vertices at least four colors are required. Generally if any  $k$  points contain two that are distance one apart and there are more than  $r(k-1)$  points where  $r$  is the number of colors, then there is not a legal  $r$ -coloring.

The Golomb graph supplies a third proof with a sufficiently important and different theme. Divide a hexagon of unit side into 6 equilateral triangles. If this is to be three-colored the center color is not on the hexagon. So along the hexagon the two non center colors alternate and each color class has three vertices. It is possible to draw an equilateral triangle of side one where each vertex is distance one from a vertex of the same same color class. Since the



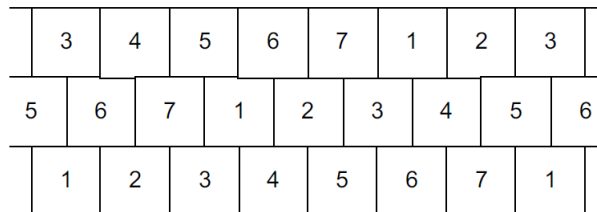
Figure 3.2: Golomb graph



equilateral triangle uses three colors outside of this color class, four colors are required to color the Golomb graph.

**Theorem 3.2.3**  $\chi \leq 7$

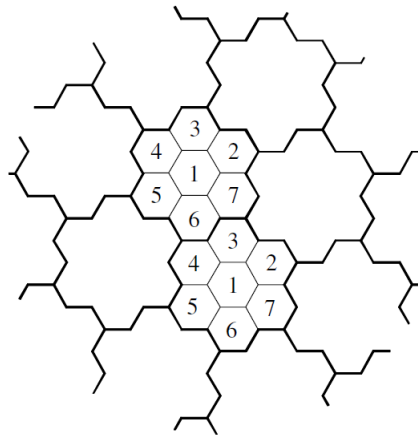
Figure 3.3: Square tiling



**Proof** As in the figure the plane can be tiled in 7 colors with squares of diagonal length one and the next row is shifted two and a half squares right. This proof is due to Lazlo Szekely.

Since the bottom right hand corner of a square is distance one from the upper left hand corner of a square of the same color two rows below, it is necessary to color the borders so that these points are different. Soifer does this by coloring the upper and right edges the color of the square except for the upper left and lower right corners which inherit their color from the squares left and below the square in question. ■

Figure 3.4: Hexagon tiling



An alternate construction uses hexagons in a flower pattern, see figure 3.4. The flowers have 18 sides and each is oriented the same way. Since the flowers fit together, they tile the plane. If the pattern used side length one, there would be no points of the same color a distance  $d$  from each other where  $2 < d < 7^5$ , so using side length  $1/2.1$  gives a suitable 7-coloring of the plane.

One of Soifer's students, Edward Pegg, has worked on using the seventh color as infrequently as possible. Soifer presents a construction where six of the colors use the same house like shape and the seventh color is a square. Despite such structure the squares tile less than one third of one percent of the plane. Therefore  $4 \leq \chi \leq 7$ . Paul Erdős was of the opinion that  $\chi \geq 5$  using such strong language as "sure" and "almost surely". Erdős interrupted a talk of Soifer's in 1994 to explain his position:

“Excuse me for interrupting, I am almost sure that the chromatic number of the plane is greater than four. It is not a proof, but any measurable set without distance one in a very large circle has measure less than one quarter. I also do not think it is seven”.

Ron Graham shares Erdős’ opinion that  $\chi$  is either five or six. Soifer predicts 7. I am willing to believe Erdős’ heuristic, however, having tried to create unit distance graphs of chromatic number 5, creating a unit distance graph of chromatic number 6 seems extraordinary so I will predict  $\chi = 5$ . As proven earlier, every graph of chromatic number  $k$  has a finite subgraph of chromatic number  $k$ , note because the plane is uncountable, when applied to the chromatic number of the plane this result requires the axiom of choice. Producing finite subgraphs seems easier than producing a coloring of an uncountable space. I believe that if progress is to be made on this problem it will be in the form of raising the lower bound. Erdős believed that equilateral triangles were essential to the chromatic color of the plane problem despite his result on the existence of graphs of arbitrary chromatic number and arbitrary girth. As it turns out Wormald in 1979 was able to show the existence of a triangle-free unit distance graph on 6,448 vertices aided by a computer. This number appears much too large and finding the smallest number of vertices for a triangle-free unit distance graph is an open problem. Paul O’Donnell studied the problem of the chromatic number of the plane for his PhD thesis from Rutgers in 1999. He showed the existence of a 23 vertex 4-chromatic unit distance graph and this is the current record, but did not include this in his dissertation. One of many interesting results from that thesis is as follows:

**Theorem 3.2.4** *There are 4-chromatic unit distance graphs of arbitrary girth.*

The following list of results on the chromatic number of the plane can all be found in Soifer’s *Mathematical Coloring Book*.

If there is a 7-chromatic unit distance graph it has at least 6198 vertices. This result is Dan Pritikin’s from 1998. In 1979 Stephen Phillip Townsend showed that the chromatic number of the plane under map type coloring is six or seven. From this result it follows by standard

topological arguments given by Douglas Wodall for closed sets and Nathaniel Brown, Nathan Dunfield, and Greg Perry as undergraduates for open sets, that the chromatic number of the plane under colorings entirely of open sets or entirely of closed sets is also six or seven. In 1981 Kenneth Falconer showed that for colorings of measurable sets that the chromatic number of the plane was at least five.

### 3.3 Generalizations of the chromatic number of the plane

In a coloring of the plane a color is said to realize a distance  $d$  if there are two points distance  $d$  apart of this color. In this terminology the chromatic number of the plane problem asks for the fewest colors required to color the plane none of which realize distance one. Erdős then asked what if instead of each color avoiding distance one, each color was only required to avoid at least one distance. In the above terminology Erdős is asking for the smallest number of colors that can be used to color the plane and no color realizes all distances. Soifer named this number the polychromatic number of the plane and denoted it  $\chi_p$ .

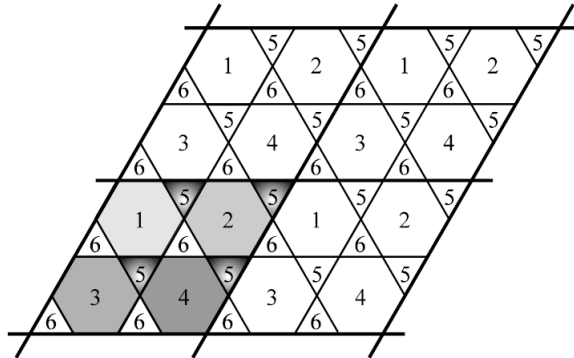
**Proposition 3.3.1**  $\chi_p \leq 7$

**Proof** The chromatic number of the plane is at most seven which is to say that there is a seven-coloring of the plane where each color does not realize distance one, so none of the seven colors realizes all distances. ■

**Theorem 3.3.2**  $\chi_p \leq 6$

In Stechkin's example the border of every hexagon of side one-half is colored the color of the hexagon except the two lower and the rightmost vertices which inherit their color from the surrounding hexagons. Stechkin's example has the triangles avoiding distance one-half

Figure 3.5: Stechkin's example



and the hexagons avoiding distance one. Soifer introduced the term coloring type to capture the distances avoided. Stechkin's example is coloring type  $(1, 1, 1, 1, 1/2, 1/2)$ . To improve the upper bound on  $\chi$  it is required to find a coloring of type  $(1, 1, 1, 1, 1, 1)$  and it is thus desirable to have the least number of non-one entries in the coloring type. Soifer created an example of coloring type  $(1, 1, 1, 1, 1, 1/5^{.5})$  inspired by a commonly used design in Russian toilettes. Soifer also found a  $(1, 1, 1, 1, 1, 1 - 2^{.5})$  6-coloring with the help of then 15 year old Ilya Hoffman, a son of Soifer's cousin Leonid Hoffman. By rotating the small squares down one moves from the  $(1, 1, 1, 1, 1, 1 - 2^{.5})$  to the  $(1, 1, 1, 1, 1, 1/5^{.5})$  colorings which shows, after much analysis, the existence of  $(1, 1, 1, 1, 1, a)$  colorings for  $a \in [2^{.5}, 1/5^{.5}]$ . ■

This further inspired the notation  $\chi_a$  for the almost chromatic number of the plane which is defined to be the least number of colors required to color the plane where all but one color avoids distance one and the last color avoids at least one distance. Determining  $\chi_a$  is also an open problem. In the case of six colors it is interesting to try to get the sixth color to avoid a distance close to one. Of course, the same logic as the above proposition shows  $\chi_p \leq \chi_a \leq \chi$ . So a lower bound on  $\chi_p$  is desirable.

**Theorem 3.3.3**  $\chi_p \geq 4$ .

This theorem was first discovered by Dima Raiskii in 1970 as a tenth-grade student in a Russian high school known for its mathematical prowess. Stechkin's example above is part of Raiskii's paper. Soifer knows Raiskii's history better than almost anyone. Despite his talent Raiskii chose not to pursue mathematics. Raiskii lived a troubling adolescence that severely affected his high school record. Had it not been for the international recognition afforded him by this result, particularly letters sent to his school addressed to professor Raiskii, he would not have been allowed to graduate high school! Soifer presents a newer proof of this result that makes extensive use of Moser's spindle and generalizes the counting argument given in the second proof of  $\chi \geq 4$  above. Remarkably this new proof is also due to a Russian high school student!

Erdős then introduced an interesting problem. For a set  $S$  of  $r$  distances, what is the chromatic number of the plane where adjacency is defined by being a distance in  $S$  apart. Furthermore, define  $\chi_r$  as the maximum chromatic number across all sets of  $r$  distances. Erdős says that it is easy to see that  $\lim_{r \rightarrow \infty} \chi_r / r = \infty$ . It could be that  $S_r = [r]$  gives a sequence of sets that shows this. In the case of  $S_r$  each horizontal line has chromatic number  $r + 1$  as an interval of  $r + 1$  integers forms a  $K_{r+1}$  while on the integer lattice the coloring of a horizontal line not only depends on the line itself but every other horizontal line and we may hope that eventually each horizontal line requires  $r + 1$  colors distinct from the colors of other horizontal lines. For Erdős the question became whether  $\chi_r$  was bounded by a polynomial in  $r$  or not.

For higher dimensional Euclidean spaces the question of the chromatic number of the distance one graph is also interesting. Erdős promoted the problem both in terms of exact values and asymptotics. Oren Nechushtan showed in 2000 that  $\chi(E^3) \geq 6$  and David Coulson submitted in 1998 a paper showing  $\chi(E^3) \leq 15$  which was published in 2002. Kent Cantwell in 1996 showed  $\chi(E^4) \geq 7$  and  $\chi(E^5) \geq 9$  and Josef Cibulka showed  $\chi(E^6) \geq 11$  all of which are still the best known lower bounds. Erdős conjectured that  $\lim_{n \rightarrow \infty} \chi(E^n) = \infty$  with an exponential growth rate. Frankl and Wilson gave an exponential lower bound in 1981 which when combined with Larman and Rogers' exponential upper bound of 1972 proved Erdős

correct. Moreover, both exponential bounds hold for the polychromatic number of the plane which thus has a similar growth rate.

**Theorem 3.3.4** *The chromatic number of the rationals in the plane,  $\chi(Q^2) = 2$*

**Proof** Douglas Woodall proved this in 1973. Soifer sketches an outline and asks the reader for many details as an exercise which I supply. Firstly consider partitioning  $Q^2$  into classes where  $(a, b)$  and  $(c, d)$  have the same class iff  $a - c$  and  $b - d$  have odd denominator when written in lowest terms. If these differences in lowest terms are  $m/n$  and  $u/v$  with the points distance one apart, then  $(m/n)^2 + (u/v)^2 = 1^2 = 1$  so  $m^2v^2 + u^2n^2 = n^2v^2$  and since  $n$  does not divide  $m$  while  $v$  does not divide  $u$ , dividing by  $n$  shows  $n$  divides  $v$  while dividing by  $v$  shows  $v$  divides  $n$  and thus  $v=n$  as we may assume denominators are positive. So  $m^2 + u^2 = n^2$  when if  $n$  where even, both  $m$  and  $u$  are odd and the right hand side is  $0 \pmod 4$  but the left hand side is  $2 \pmod 4$ . So  $n$  and  $v$  are both odd and this implies  $m$  and  $u$  differ in parity. Therefore these classes are disconnected components of the distance one graph that are clearly translates of one another. To show a 2-coloring it suffices to show the 2-coloring of the class of  $(0/1, 0/1)$  integers are considered to have denominator 1. This class has points  $(s_1/r_1, s_2/r_2)$  where both  $r_1$  and  $r_2$  are odd since 1 in  $0/1$  is odd and the above argument says the denominator of each coordinate of the difference between adjacent points is odd. Taken together this means the denominator of any coordinate in this class, which must divides the common denominator of a component of a difference vector of length one plus a component of a point previously shown to be in the class, is a product of odd integers and thus odd. Finally in the class of  $(0, 0)$  color  $(odd/odd, odd/odd)$  and  $(even/odd, even/odd)$  red and other points blue. Recall from above if two points are distance one apart the numerators of the difference vector differ in parity, which each color class avoids. ■

Miro Benda and Micha Perles are responsible for showing  $\chi(Q^3) = 2$  and  $\chi(Q^4) = 4$ . Josef Cibulka has shown  $\chi(Q^5) \geq 8$  and  $\chi(Q^7) \geq 15$  while Matthias Mann has shown  $\chi(Q^6) \geq 10$  and  $\chi(Q^8) \geq 16$ .

# Chapter 4

## Ramsey-type theorems for the Integers

Bruce Landman and Aaron Robertson recently wrote a fantastic book titled *Ramsey Theory on the Integers* which is both thorough and accessible at the advanced undergraduate level and beyond. Here is a small preview of the area. While Erdős lived much of the excitement in Ramsey theory dealt with graph theory. While the following results are all older results, only relatively recently has proving Ramsey-type theorems in other settings surpassed the graph theoretic questions in popularity. Credit for this change of focus may be given to the brilliant work of Ron Graham most of whose work is beyond the scope of this thesis. The following theorems are Ramsey-type results in that they demonstrate some property is invariant under all colorings of the integers or some subset of the integers.

### 4.1 Schur's theorem

Isaai Schur was born in Russia in 1875 to a Jewish family and spent much of his life in Germany including his university education and lecturing at the University of Berlin. Isaai



Schur has many important theorems to his name. He was also an extraordinary lecturer attracting over 500 students to his number theory course in 1930 at the University of Berlin. He advised many PhD students who would generalize his theorems, specifically Brauer and Rado. Schur remained in Germany as long as he could. He was very attached to the Germany of Gauss and Beethoven and considered himself a German. As Schur was a lecturer at Berlin, a civil servant position, during the first world war he was able to continue lecturing until 1935 and advise students from his home until 1939 when he was forced to flee Germany. Schur had been devastated by the turn of events in Germany for a while, but when invitations from western universities offered him safety he declined feeling that he was in less danger than younger Jewish professors. Schur died in 1941 from failing health but not before suffering a heart attack mid lecture and continuing to lecture in his clear and insightful manner.

Schur proved the following theorem in 1916. It was only a lemma that he was using to prove a theorem of Leonard Dickson who was trying to prove Fermat's Last Theorem. Theorems of this type had not been seen before and it was nearly entirely ignored by the mathematics community. For a theorem discovered before Ramsey's theorem the number of connections to Ramsey numbers is surprisingly large. Schur met with Erdős in 1936 and in many ways the baton of Ramsey theory was Schur's to pass to Erdős.

**Theorem 4.1.1** (Schur's theorem) *For any  $r \geq 1$  there exists a smallest number  $s(r)$  such that if the natural numbers between 1 and  $s(r)$  are  $r$ -colored there is a monochromatic solution to  $x + y = z$  where  $x, y$  need not be distinct.*

Note that Schur proved this differently. For a given  $r$  this will show that  $s(r) \leq R(3; r) - 1$ . Consider any coloring  $\chi$  of the integers between 1 and  $R(3; r) - 1$  and for the graph  $K_{R(3; r)}$  label the vertices with these integers and color the edge between vertices  $i$  and  $j$  the color of vertex  $|i - j|$ . By definition of  $R(3; r)$  this graph has a monochromatic triangle with vertices say  $a, b, c$  with  $a < b < c$ . Then by the method of coloring of the edges vertices  $b - a, c - a$ , and  $c - b$  are the same color. Take  $x = b - a, y = c - b$ , and  $z = c - a$  when indeed  $x + y = z$

and  $x, y, z$  are all the same color. Note since  $x, y, z$  are formed by a subtraction of at least 1 from at most  $R(3; r)$  the largest of  $x, y, z$  is at most  $R(3; r) - 1$ . ■

**Corollary 4.1.2**  $s(r) \leq R(3; r) - 1 \leq 3r! - 1$

**Proof** See the proof of Schur's theorem above and the result on the multicolor Ramsey number for triangles from the introduction. ■

It is also possible to prove that the  $x, y, z$  can be taken to be distinct in Schur's theorem. When this is insisted upon the least integer forcing a monochromatic triple is denoted  $s'(r)$  and note that  $s(r) \leq s'(r)s(2r)$  [see theorem 32.2 in A. Soifer: *Mathematical Coloring Book*].

First  $s(1) = 2$ , since if 1 and 2 are given the same color then  $1 + 1 = 2$  in a monochromatic manner. Next  $s(2) = 5$ , since if 1 is maroon then the above suggests 2 is orange when  $2 + 2 = 4$  implies 4 is maroon, then  $1 + 3 = 4$  implies 3 is orange. By this coloring all compositions with two parts  $1 + 1 = 2, 1 + 2 = 3, 1 + 3 = 4, 2 + 2 = 4$  and sums at most 4 are not monochromatic since 1, 4 are maroon and 2, 3 are orange, so  $s(2) > 4$ . Yet since  $1 + 4 = 2 + 3 = 5$  whichever color 5 is assigned has a monochromatic solution to  $x + y = 5$ , so  $s(2) = 5$ . From experience, one can prove  $s(3) = 14$  with a pencil and a couple sheets of paper in less than an hour. To do this employ the methods above. It will be necessary to break the problem into cases but the number of cases is remarkably small if you are sharp about realizing when an assignment of color is forced. When breaking into more cases, try to do so in such a way that all but one of your cases is easily resolved, a little foresight goes a long way. To see  $s(3) > 13$ , color 1, 4, 10, 13 in color 1, 2, 3, 11, 12 in color 2, and 5, 6, 7, 8, 9 in color 3. The last known Schur number is  $s(4) = 45$ . For  $s(4)$  the amount of forcing is significantly less. However, try starting with a random legal 4-coloring of the first 10 natural numbers. It should not take more than a few hours to show that this can't be extended to a legal 4-coloring of the first 45 natural numbers. Since there are only  $\binom{10}{4} = 210$  such starting positions not all of which are legal, one may be able to do this by hand in less than a month. Fortunately it was proven in 1961 by computer by Baumert. Why haven't computers found

$s(5)$ ? The jump in difficulty between  $s(4)$  and  $s(5)$  is probably about that of between  $R(4, 4)$  and  $R(5, 5)$ . The lack of forcing and the frequent inability to restrict the number of cases means that even with a clever depth first search the best one can hope for is  $2^{s(5)}$  cases. The best current lower bound of  $s(5) \leq 161$  shows brute force will never prove  $s(5)$ .

Fredricksen and Sweet searched symmetric partitions in 2000. that is, if they are trying to find a legal 5-coloring of the natural numbers up to 160, they assign the same color to 1 as they do 160 and the same color to 2 as they do 159 etc. In using this process things get forced much quicker. They were able to obtain  $s(6) \geq 537$  and  $s(7) \geq 1681$ . While they were not responsible for the bound  $s(5) \geq 161$  they were able to find a legal symmetric partition of 161. Their algorithm works as suggested above, they use a depth first search next coloring the integer with fewest choices of color remaining. They remark that the color classes are all roughly the same size in the colorings that color the most numbers legally, for  $s(5) \geq 161$  they found partitions as small as 24 and as large as 44. They also make the observation that once about 20 percent of numbers are colored, with high probability the color of the remaining numbers is forced. Their work also determined new lower bounds for  $R(3; 6)$  and  $R(3; 7)$  which follow from the proof of Schur's theorem above. While symmetric colorings make up only a tiny sliver of the colorings allowed by Schur's theorem, I am quite pleased with their approach. The discussion above shows that their algorithm could not have run exhaustively and they mention they would continue to work on the problem. I want to see the results a similar approach could obtain with 10 years newer technology.

Geoffrey Exoo is one of the champions of producing lower bounds for Ramsey numbers. Radziszowski and McKay do this as well but also focus on using technology to get the upper bounds to match. The lower bound  $s(5) \leq 161$  is due to Exoo in 1994. Exoo viewed the problem as a combinatorial optimization problem. His objective function essentially asks to maximize the size of the largest legal restriction in the coloring. He uses simulated annealing and genetic algorithms, at all times he has a coloring of a large number of natural numbers and these algorithms slowly make the first  $n$  sum-free. Again, I wish to see how 15 years of newer technology might improve his results. Exoo's work shows  $R(3; 5) \geq 162$ .

Note that both of these papers define the Schur number to be the largest  $n$  avoiding monochromatic sums, so their Schur numbers are one short of the numbers I use. The numbers above are consistent with the definition I use which is consistent with Landman's. Both papers determine new lower bounds on multicolor Ramsey numbers for triangles. Amazingly these bounds have not been improved in the 10+ years since.

**Lemma 4.1.3**  $s(r + 1) \geq 3s(r) - 1$

**Proof** Let  $\chi$  be a  $r$ -coloring of  $[n]$  with no monochromatic solutions to  $x + y = z$ .  $\chi$  may be extended to a  $(r + 1)$ -coloring of  $[3n - 1]$  having no monochromatic solutions to  $x + y = z$  as follows. Color  $[n + 1, 2n + 1]$  color  $r + 1$  and color  $m \in [2n + 2, 3n + 1]$  the color of  $m - (2n + 1)$ . Since any the sum of two numbers in color  $r + 1$  is at least  $2(n + 1) = 2n + 2$  is not colored in  $r + 1$  there are no  $r + 1$  colored solutions. But there are no solutions of the other colors as well.  $n + n = 2n < 2n + 2$  and the fact that  $\chi$  was a legal  $r$ -coloring of  $[n]$  shows no pair of the first  $[n]$  numbers are in a monochromatic solution. All numbers of a monochromatic solution do not come from the last  $n$  since any such sum is at least  $2(2n + 2) = 4n + 4 > 3n + 1$ . Finally, if  $x \in [1, n]$  and  $y, z \in [2n + 2, 3n + 1]$  with  $x + y = z$  then  $y' = y - (2n + 1), z' = z - (2n + 1)$  gives  $x + y' = z'$  with all three numbers from  $[n]$  which is ruled out by assumption. This has shown  $s(r) \geq n + 1, s(r + 1) \geq 3n + 2$ , solving for  $n$  and substituting gives the claim. The bound is tight for  $s(1), s(2), s(3)$  and close for  $s(4)$ .

■

**Theorem 4.1.4**  $s(r) \geq .5(3^r + 1)$

**Proof**  $s(1)=2$  is a satisfactory base case. Assume the theorem holds for  $r \geq 1$  then using the lemma  $s(r + 1) \geq 3s(r) - 1 \geq .5(3^{r+1} + 3) - 1 = .5(3^{r+1} + 1)$  as desired. ■

One way to generalize Schur's theorem is to associate with each color some number of variables used in the equation. For instance  $S(3,4)$  is the least integer forcing  $x + y = z$  in

red or  $a + b + c = d$  in blue. Again  $S(I_1, I_2, \dots, I_r) \leq R(I_1, I_2, \dots, I_r)$  is not hard to prove and is similar to the proof of Schur's theorem.

## 4.2 Van der Waerden's theorem

This theorem resembles Schur's theorem and was proven in 1927 by Bartel L. van der Waerden. Van der Waerden was born in 1903 in the Netherlands. While demonstrating his prodigious understanding from an early age, van der Waerden became famous for his two volume treatise on algebra for Springer's yellow book series and a series of algebraic geometry articles. He was fortunate to study under Emil Artin, Emily Noether, Richard Dedekind, and David Hilbert.

Van der Waerden was criticized by the mathematics community for remaining in Germany during the Second World War and thus lending his prestige to the Nazi regime. He did speak out against the Nazis and never did join their party. As a result after the war van der Waerden found it difficult to find a teaching appointment in his native Netherlands and instead spent many years at Zurich. Also his achievements were downplayed or not recognized by some western algebraists.

Van der Waerden believed himself to be proving a conjecture of Baudet's another brilliant mathematician from Groningen University who died prematurely at 30. Note that van der Waerden attended Groningen before doing PhD work at Gottingen. The facts are that Schur had conjectured this result when proving the previous theorem but that van der Waerden heard that the conjecture was Baudet's and published the result as a solution to Baudet's conjecture. Van der Waerden had never heard of Schur and Baudet was certainly very interested in this problem. The question that can not be answered definitively is whether Baudet independently conjectured the result or if Baudet heard the conjecture from Frederick Schuh, a colleague of Schur's and an instructor at Groningen who gave problem sessions attended by Baudet. There is also evidence suggesting that Schuh accepted the conjecture

as Baudet's. Schur, ever the professional, never asked for credit for the conjecture but the testimony of his students shows that he was the first to conjecture this result.

**Definition** An arithmetic progression is a sequence of numbers  $a, a + d, a + 2d, \dots$  whose terms have a common distance  $d$  between them. When only the first  $k$  terms are being referred to it is said that the progression has length  $k$ .

The conjecture was: if the natural numbers are 2-colored then there exists monochromatic arithmetic progressions of arbitrarily large length. The proof of van der Waerden establishes compactness and generalizes the result to an arbitrary amount of colors. The story goes that van der Waerden had only began to think on the conjecture the previous day and introduced it to Emil Artin and Otto Schreier over lunch. They returned to Artin's office and as van der Waerden relates Artin and Schreier were having flashes of insight (compactness, multiple colors, strong induction etc) until van der Waerden gave a procedure for two colors and arithmetic progressions of length three. Van der Waerden was then instantly convinced that his procedure could be generalized to progressions of any length while Artin and Schreier were only convinced after insisting that van der Waerden show them the procedure could work for 3 colors and progressions of length 3, and then for 3 colors and length 4. The story is strong evidence that mathematics ought to be done as a team effort, what readily appeared to Artin and Schreier had not quite occurred to van der Waerden and van der Waerden's procedure for building on their insight seemed suspicious to Schreier and Artin. Yet in only a single afternoon they established the result.

**Theorem 4.2.1** (van der Waerden's theorem) *For all positive integers  $k$  and  $r$  there exists  $W(k, r)$  such that if  $[W(k, r)]$  is  $r$ -colored then there is a monochromatic arithmetic progression of length  $k$ .*

**Proof** While  $W(3, 2) = 9$  by enumeration, here is a demonstration from [GSR2] that  $W(3, 2) \leq 325$ . This will demonstrate van der Waerden's procedure although he allowed

the blocks to overlap and got  $W(3, 2) \leq 69$  (see Soifer's *Mathematical Coloring Book* ch.33). After this demonstration [GSR2] comments that "the general proof is now just a double induction on  $k$  and  $r$ ." However, [GSR2] chooses to strengthen the hypotheses and prove a stronger result. Landman and Robertson also give a proof of the general theorem using a lot of elementary machinery.

The strategy is to get arithmetic progressions of length 2 in each color having their third term be the same integer. Break [325] into 65 blocks where block  $B_1 = 1, 2, 3, 4, 5$ ,  $B_2 = 6, 7, 8, 9, 10$  etc. The blocks have 5 elements and so  $2^5 = 32$  choices of color pattern, thus amongst the first 33 blocks, two blocks share the same color pattern. In the first of these two blocks, amongst the first three elements are two of the same color, say  $j$  and  $j + d$  are red, and  $j + 2d$  is also in this block (since  $d = 1$  or  $d = 2$  and the block has 5 element) but to avoid a monochromatic red arithmetic progression,  $j + 2d$  must be blue. Then the other block with this color pattern has  $i$  and  $i + d$  red and  $i + 2d$  blue as well. Now consider the arithmetic progressions  $j, j + (i + d - j), j + 2(i + d - j) = j, i + d, 2i + 2d - j$  and  $j + 2d, j + 2d + (i - j), j + 2d + 2(i - j) = j + 2d, i + 2d, 2i + 2d - j$  the first two terms of the first progression are red and the first two terms of second progression are blue and the third term of both progressions is the the same, so whichever color is assigned to  $2i + 2d - j$  results in a length three arithmetic progression in that color. ■

Also note that if van der Waerden's theorem holds for two colors then it holds for arbitrarily many, on that afternoon this was an observation of Artin's. To see this consider the case of four colors red, yellow, green, and blue and arithmetic progressions of length  $k$ . Then imagine those colored red and yellow are colored orange, and those colored green and blue are colored maroon. Then since the theorem holds for two colors in either maroon or orange there is a progression of the length  $W(k, 2)$  of the form  $a, a + d, a + 2d, \dots, a + (W(k, 2) - 1)d$  and in the original setting these are colored in two colors. If these terms are labeled  $1, 2, 3, W(k, 2)$  there is a monochromatic  $k$  term arithmetic progression, for example 5, 8, 11, and this has the corresponding  $a + 4d, a + 7d, a + 10d$  which is monochromatic. It is clear that empty

color classes are allowed in this argument and that the method of reduction from  $2^k$  to  $2^{k-1}$  colors can be repeated. So for fixed  $k$ , the 2-color case implies the  $r$ -color case.

The best known lower bound for  $w(k, 2)$  occurs when  $k = p$  a prime. This bound is  $w(p + 1, 2) \geq p2^p$  and the proof requires field extensions but can be found in [GSR2 sec 4.3]. For general  $k$  Erdős conjectured that  $\lim_{n \rightarrow \infty} w(k, 2)/2^k = \infty$ . For upper bounds, the upper bound given by van der Waerden's proof is  $w(k, 2) \leq ack(k)$  which uses the Ackermann function which is one of the fastest growing functions ever used in a serious proof. Timothy Gowers earned a Fields Medal in 1998 in large part due to his proof that  $w(k, 2) \leq 2^{2^{2^{2^{k+9}}}}$ . Of course  $w(2, r) = r + 1$  since once a color is used twice it forms an arithmetic progression of length 2. The only known exact values are  $w(3, 2) = 9$ ,  $w(3, 3) = 27$ ,  $w(3, 4) = 76$  and  $w(4, 2) = 35$ . The computational difficulties of van der Waerden numbers are similar to those of Schur numbers. This just scratches the surface, Landman and Robertson's book is an excellent source on the topic.

### 4.3 Rado's theorem

Rado was born in 1906 in Germany and studied under Schur at Berlin. In his PhD thesis Rado proved this theorem which is firstly a generalization of Schur's theorem, but also generalizes van der Waerden's theorem. Rado viewed the equation  $x + y = z$  that Schur studied not as a statement about sum free sets, but rather as a statement about the linear homogeneous equation  $x + y - z = 0$ . All that is meant by homogeneous is that the coefficients are integer valued and non-zero while the right hand side is 0.

**Definition** For a system of equations  $S = S(x_1, x_2, \dots, x_n)$  on a set  $A$ ,  $S$  is said to be  $r$ -regular if any  $r$ -coloring of  $A$  yields a monochromatic solution to  $S$ .  $S$  is called *regular* if for every  $r$   $S$  is  $r$ -regular.



By Schur's theorem the equation  $x + y - z$  is regular on  $N$  and van der Waerden's theorem indicates that the system  $x_1 = x_0 + d, x_2 = x_1 + d, \dots, x_{k-1} = x_{k-2} + d$  which represents finding a  $k$  term arithmetic progression is regular on  $N$ . In the following the set  $A$  will be  $Z^+$ , so all results refer to regular on  $Z^+$ .

**Theorem 4.3.1** (Rado's single equation theorem) *Let  $S$  be the homogeneous linear equation  $\sum_{i=1}^k c_i x_i = 0$  with non-zero integer coefficients then  $S$  is regular iff some nonempty subset of the  $c_i$ s sum to 0.*

Note if  $S$  is replaced by a non homogeneous equation  $\sum_{i=1}^k c_i x_i = b$  then the above theorem is generally not true. However under certain conditions it does hold, see Theorem 9.10 in Landman and Robertson which gives:

**Theorem 4.3.2** *Let  $S$  be the non-homogeneous linear equation  $\sum_{i=1}^k c_i x_i = b$  with non-zero integer coefficients. Then  $S$  is regular iff for  $s = \sum_{i=1}^k c_i$ , either  $b/s \in Z^+$  or  $\sum_{i=1}^k c_i x_i = 0$  is regular and  $b/s \in Z^-$*

The *column condition* for a  $k \times n$  matrix with columns  $c_1, c_2, \dots, c_n$  asks for a reordering of the columns and a sequence of indices  $1 = i_0 < i_1 < \dots < i_t = n$  such that  $s_j = \sum_{i=i_{j-1}+1}^{i_j} c_i$  has  $s_1 = 0$  and for  $2 \leq j \leq t$ ,  $s_j$  may be written as a linear combination of  $c_1, c_2, \dots, c_{i_{j-1}}$ . If this is possible, then the matrix satisfies the column condition.

**Theorem 4.3.3** (Rado's theorem for systems of equations) *Let  $S$  be a system of linear homogeneous equations represented in matrix form by  $Ax = 0$ , then  $S$  is regular if and only if  $A$  satisfies the column condition. Moreover there is a monochromatic solution to  $Ax = 0$  formed from distinct integers if  $A$  satisfies the column condition and there is a not necessarily monochromatic solution of distinct integers to  $Ax = 0$ .*

When calculating Rado numbers, Landman and Robertson come across a useful fact about 2-colorings on the integer lattice. They study for relatively prime  $a$  and  $b$  the integers  $1 + as + tb$

and by looking at the pairs  $(s, t)$  find the need to prove that for  $0 \leq y \leq x \leq 4$ , any 2-coloring of these points admits a monochromatic isosceles right triangle with the same orientation as the space. They suggest that this presents a new Tic-Tac-Toe game where on this board the winner is the player to first complete such an isosceles right triangle in their symbol (X or O) and note that the game can not end in a draw. However, a strategy stealing argument shows that the first player will always win. The connection between Rado's theorem, colorings of the plane, and combinatorial games shows how interwoven these topics are.

## 4.4 Hilbert's Cube Lemma

David Hilbert was born in 1862 near Koenigsberg which in Hilbert's time was part of the German empire but today is Russian territory. Hilbert's influence on modern day mathematics is enormous. By 1900 Hilbert was recognized as one of the world's top mathematicians. At the 1900 International Congress of Mathematics Hilbert presented 23 problems that would guide the course of mathematics through the next century. Hilbert proved this cube lemma in 1892 while studying rational functions with integer coefficients. The importance of this lemma is mostly historical as it is the first known instance of a Ramsey-type result.

Given a set of not necessarily distinct positive integers  $x_0, x_1, \dots, x_n$  the *affine  $n$ -cube* generated by this set is the set of possible values of  $x_0 + \sum_{i \in I} x_i | I \subseteq [n]$ , so the set of sums where  $x_0$  is included with any subset of  $x_1, \dots, x_n$ . For example, the affine cube generated by 1, 3, 9 is 1, 4, 10, 13.

**Theorem 4.4.1** (Hilbert's cube lemma) *For each pair of positive integers  $r, n$  there exists a least integer  $m = H(r, n)$  such that for any  $r$ -coloring of  $[m]$  there is a monochromatic affine  $n$ -cube in some color.*

Hilbert did not prove it this way but consider the arithmetic progression  $a, a+d, a+2d, \dots, a+nd$  this corresponds to the affine cube generated by  $x_0 = a$  and  $x_i = d$  for  $1 \leq i \leq n$ .

Since van der Waerden's theorem states that in any  $r$ -coloring of  $[W(n+1, r)]$  there is a monochromatic  $n+1$  term arithmetic progression which may be viewed as a monochromatic affine cube by the above method, such an  $m$  must exist. ■

Hilbert did not touch Ramsey-type problems again. In fact no one seems to have used this lemma in a Ramsey-type way until Erdős, Sarkosy, and Sos published the following density theorem in 1989.

**Theorem 4.4.2** (Hilbert Cube Lemma Density Version) *For every positive integer  $n$  there is a natural number  $m_0 = H(n)$  such that for any  $m_0 > m$ ,  $B \subseteq [m]$  with  $|B| > 3m^{1-2^{-n}}$  there exists positive integers  $x_0, x_1, \dots, x_n$  that generate an affine cube contained in  $B$ .*

## 4.5 The happy end problem

Returning to Erdős and his mathematical circle, this problem was what allowed the rediscovery of Ramsey theory after Ramsey's untimely death. Esther Kline had been scribbling points in the plane in her notebooks and suddenly a problem occurred to her that seemed of a new and novel nature. She asked: is it true that for all  $n$  there is a least number  $g(n)$  such that any collection of  $g(n)$  points in the plane in general position (i.e, no three points on a line) contain a convex  $n$ -gon? Esther was able to prove  $g(4)=5$ . To see this draw a triangle, the fourth and fifth points are interior to the triangle, otherwise a convex quadrilateral is formed or some other collection of three points traps the fourth and fifth in a triangle. Wherever these fourth and fifth points are placed in the triangle they form a convex four-gon with one two of the vertices of the triangle. Esther presented the problem to the circle of young mathematicians who shared Esther's intuition that they were working on something new and important. Both Erdős and George Szekeres noted that the desire amongst the heavily male-dominated group to be the first to solve this problem also stemmed from the fact that Esther had proposed the problem. George won the race and within a year married Esther. Paul and George wrote the paper "On a Combinatorial Problem in Geometry together".

Call an  $n$ -cup a sequence of  $n$  points with increasing consecutive slopes and an  $n$ -cap a sequence of  $n$  points with decreasing consecutive slopes. Let  $f(n, m)$  be the maximum number of points without a  $n$ -cup or  $m$ -cap. Let  $S$  be a set of points without  $n$ -cups and  $m$ -caps, and  $T$  be the set of right endpoints of  $(n-1)$ -cups. Points in  $T$  are not left endpoints of  $(m-1)$ -caps because the point to the left of  $t \in T$  is more left than  $t$ . Since there are no  $n$ -cups, and  $T$  does not contain a left hand endpoint of a  $(m-1)$ -cap,  $T$  contains neither an  $n$  cup nor an  $m-1$  cap, so  $f(n, m-1) \geq |T|$ . Likewise  $S - T$  has no right endpoints for  $(n-1)$ -cups, so as there are no  $m$ -caps,  $|S - T| \geq f(n-1, m)$ , combining with the first inequality gives  $f(n, m) \geq |S| = |T| + |S - T| \geq f(n, m-1) + f(n-1, m)$ . Now an  $n$ -cup is an example of a convex  $n$ -set, likewise an  $m$ -cap is an example of a convex  $m$ -set. From this it follows that  $g(n) \leq \binom{2n-4}{n-2} + 1$ .

Esther Kline and George Szekeres escaped Budapest for Hong Kong before World War II and settled in Australia where they remained active in mathematics until their deaths which occurred hours apart in 2005.

## 4.6 Conclusion

Ramsey theory is a wide and varied field of combinatorics and graph theory. There is much more that has and will be written on this subject than I could possibly cover here. Some things to explore include Ramsey number for cycles, paths, stars etc as well as problems of colored polygons in the plane. I hope that this has given the reader a thorough introduction to the problems studied, important players, and proof strategies employed in Ramsey theory.

# Bibliography

- [1] Martin Aigner, Gunter M. Ziegler, and Karl H. Hoffman. *Proofs from the Book*. Springer, fourth edition, 2009.
- [2] David Cariolaro. On the Ramsey Number  $R(3,6)$ . *Australasian Journal of Combinatorics*, 2006.
- [3] Fan Chung and Ron Graham. *Erdős on Graphs: His Legacy of Unsolved Problems*. A K Peters, 1998.
- [4] Paul Erdős. Some New Problems and Results in Graph Theory and other Branches of Combinatorial Mathematic. *Lecture Notes in Math 885*, 1981.
- [5] Geoffrey Exoo. A Lower Bound for Schur Numbers and Multicolor Ramsey Numbers of  $K_3$ . *Electronic Journal of Combinatorics 1*, 1994.
- [6] Harold Fredricksen and Melvin M. Sweet. Symmetric Sum Free Partitions and Lower Bounds for Schur Numbers. *Electronic Journal of Combinatorics 7*, 2000.
- [7] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey Theory*. Wiley-Interscience series in discrete mathematics and optimization. John Wiley and Sons, second edition, 1990.
- [8] J. G. Kalbfleisch and R. G. Stanton. On the Maximal Triangle Free Edge Chromatic Graphs in 3 Colors. *Journal of Combinatorial Theory 5*, 1968.

- [9] Bruce M. Landman and Aaron Robertson. *Ramsey Theory on the Integers*. Student Mathematical Library Volume 24. American Mathematical Society, 2003.
- [10] B. D. McKay and S. P. Radziszowski.  $R(4,5)=25$ . *Journal of Graph Theory*, 1995.
- [11] Jarislav Nešetřil and Vojtěch Rödl. *Mathematics of Ramsey Theory*. Algorithms and Combinatorics 5. Springer-Verlag, 1990.
- [12] Frank P. Ramsey. On a Problem of Formal Logic. *Proceedings London Mathematical Society*, 1930.
- [13] Bruce Schechter. *My Brain is Open: The Mathematical Journeys of Paul Erdős*. Simon and Schuster, 1998.
- [14] Alexander Soifer. *The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of its Creators*. Springer, 2008.
- [15] Alexander Soifer. *Ramsey Theory: Yesterday, Today, and Tomorrow*. Progress in Mathematics Volume 285. Springer, 2010.
- [16] K. Walker. Dichromatic Graphs and Ramsey Numbers. *Journal of Combinatorial Theory* 5, 1968.
- [17] Douglas B. West. *Introduction to Graph Theory*. Prentice Hall, second edition, 2001.