

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Adventures in Graph Ramsey Theory /

Permalink

<https://escholarship.org/uc/item/6p94c26m>

Author

Parrish, Andrew T.

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Adventures in Graph Ramsey Theory

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Andrew T Parrish

Committee in charge:

Professor Ronald Graham, Chair
Professor Fan Chung Graham, Co-Chair
Professor Alon Orlicsky
Professor Jeffrey Remmel
Professor Jacques Verstraete

2013

Copyright
Andrew T Parrish, 2013
All rights reserved.

The dissertation of Andrew T Parrish is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2013

EPIGRAPH

*Some numbers are red
The others are blue
But patterns emerge
Whatever you do.*

TABLE OF CONTENTS

	Signature Page	iii
	Epigraph	iv
	Table of Contents	v
	List of Figures	vii
	Acknowledgements	viii
	Vita	ix
	Abstract of the Dissertation	x
Chapter 1	Introduction	1
	1.1 Additive Ramsey theory	3
	1.1.1 The columns condition	7
	1.2 Between additive- and graph-Ramsey theory	10
	1.3 The compactness principle	12
	1.4 Outline of results	15
Chapter 2	Toward Rado’s theorem on graphs	16
	2.1 Introduction	16
	2.2 Negative results	17
	2.2.1 Arithmetic progressions	17
	2.2.2 Schur’s equation and generalizations	17
	2.2.3 Three variable equations with six colors	18
	2.3 Necessary conditions for graph-regularity	21
	2.4 The graph columns condition	27
	2.4.1 Definitions	27
	2.4.2 The weak graph columns condition	29
	2.5 First step: two colors and two dimensions	33
	2.6 Coloring trees	36
	2.6.1 Trees and embeddings	36
	2.6.2 Monochromatic tree embeddings	38
	2.7 The full result	41
	2.7.1 Any colors, two dimensions	41
	2.7.2 Upper bounds	44
	2.7.3 Any colors, any dimensions	45
	2.8 Sufficient conditions for graph-regularity	48
	2.8.1 A $Grid_n$ respects embeddings	50
	2.8.2 Always a monochromatic $Grid_n$	53

	2.8.3	A $Grid_n$ is enough	55
	2.9	Hypergraph-regular equations	58
	2.10	Further work	60
Chapter 3		On adapting a result from graphs to arithmetic progressions .	62
	3.1	Large Ramsey	62
	3.2	Large van der Waerden	63
Bibliography		67

LIST OF FIGURES

Figure 2.1:	A large red grid, and the corresponding sets A and B	35
Figure 2.2:	This 3-ary tree contains a monochromatic embedded T_1 — the nodes λ , 13, 21, and 31 are all red.	37
Figure 2.3:	The sequence of subgrids.	42

ACKNOWLEDGEMENTS

I would like to thank the friends who helped keep graduate school fun, especially Franklin, Jake, Jimmy, Katie, and Scott from my year, and the many others who helped to keep a great department community.

I would also like to thank Bill Gasarch for getting me started in his “van der Waerden gang” at the University of Maryland, and for offering his continued guidance after I left for San Diego.

Chapter 2 consists primarily of previously published material. The majority of content in Sections 2.2, 2.5, 2.6, and 2.7 were first published in the Journal of Combinatorics in volume 2, no. 4 (2011), published by International Press in the article “An additive version of Ramsey’s theorem” by Parrish, A. The majority of content in Sections 2.3, 2.4, 2.8, and 2.9 were first published in the Electronic Journal of Combinatorics, volume 20, no. 1 (2013) in the article “Toward a graph version of Rado’s theorem” by Parrish, A. Both papers were authored solely by the author of this dissertation.

VITA

2007	B. S. in Mathematics, University of Maryland, College Park
2007	B. S. in Computer Science, University of Maryland, College Park
2008–2012	Graduate Teaching Assistant, Department of Mathematics, University of California, San Diego
2011	M. A. in Mathematics, University of California, San Diego
2011	Associate Instructor, Department of Mathematics, University of California, San Diego
2012–2013	Senior Teaching Assistant, Department of Mathematics, University of California, San Diego
2013	Ph. D. in Mathematics, University of California, San Diego

PUBLICATIONS

- A. Parrish, “An additive version of Ramsey’s theorem”, *J. Comb.* 2 (2011), no. 4, 593–613.
- A. Parrish, “A graph version of Rado’s theorem”, *Electron. J. Combin.* 20 (2013) no. 1.

ABSTRACT OF THE DISSERTATION

Adventures in Graph Ramsey Theory

by

Andrew T Parrish

Doctor of Philosophy in Mathematics

University of California, San Diego, 2013

Professor Ronald Graham, Chair
Professor Fan Chung Graham, Co-Chair

We define what it means for an equation to be *graph-regular*, extending the idea of partition-regular equations to a graph setting. An equation is graph-regular if it always has monochromatic solutions under edge-colorings of $K_{\mathbb{N}}$. We find an infinite family of graph-regular equations, and present two Rado-like conditions which are respectively necessary and sufficient for an equation to be graph-regular. In the process, we prove a Ramsey-like theorem for binary and k -ary trees which may be of independent interest.

We also look at a stronger version of Ramsey's theorem from Paris and Harrington, and show a counterexample to the analogous version of van der Waerden's theorem.

Chapter 1

Introduction

Ramsey theory, broadly, is the study of order within colorings. The main questions asked have the form, “Whenever the components of a structure are colored, is there always a substructure whose components are all the same color?”

The namesake of Ramsey theory is the logician Frank P. Ramsey, who in 1928 proved an important result about partitions of graphs [14][4].

Theorem 1.0.1 (Ramsey). *For any size n and any number of colors r , there is a value $R = R(n, r)$ so that, for every r -coloring of the edges of the complete graph K_R , there is some K_n subgraph whose edges are all the same color.*

Formally, an r -coloring of a set A is simply a map $\chi : A \rightarrow [c]$, where the values in $[c] = \{1, 2, \dots, c\}$ are said to be the colors. Two objects are the same color, then, if they have the same value under χ . An object is monochromatic if all of its components — in this case edges — are the same color. Of course the numerical values of these colors have no meaning; any finite set will do. Colloquially, we often use more colorful language, replacing the set of colors by, perhaps, $\{\text{red, blue, green, } \dots\}$. We will use either convention when it suits our needs.

While Theorem 1.0.1 is perhaps the most common version of Ramsey’s theorem, he also proved an infinite version (Theorem 1.0.2) and a hypergraph version (Theorem 1.0.3) — as well as an infinite hypergraph version which we trust the reader can guess.

Theorem 1.0.2. *For any number of colors r and for every r -coloring of the edges of an infinite complete graph, there is a monochromatic complete infinite subgraph.*

Theorem 1.0.3. *For any size n , any uniformity t , and any number of colors r , there is a number $R = R_t(n, r)$ so that for every r -coloring of the edges of the complete t -uniform hypergraph $K_n^{(t)}$, there is a monochromatic complete subgraph on n vertices.*

1.1 Additive Ramsey theory

Ramsey's results are important, and the first to attract wide interest. However, at the same time, a host of similar results were popping up forming the base of Additive Ramsey Theory — results where the objects being colored are natural numbers, and the monochromatic results are based on arithmetic structure, rather than graph structure in Ramsey's theorem.

Perhaps the first such result comes from David Hilbert in 1892 [7].

Given natural numbers a, d_1, \dots, d_n , define

$$H(a; d_1, \dots, d_n) = \left\{ a + \sum_{i \in I} d_i \mid I \subseteq [n] \right\}.$$

We call such a set $H(a; d_1, \dots, d_n)$ a Hilbert cube of dimension n .

The key structure of a Hilbert cube is its inductive nature:

$$H(a; d_1, \dots, d_{n+1}) = H(a; d_1, \dots, d_n) \cup H(a + d_{n+1}; d_1, \dots, d_n).$$

Lemma 1.1.1. *Given a dimension n and number of colors r , there is a number $H = H(r, n)$ so that, for any r -coloring of $[H]$, there is a monochromatic Hilbert cube of dimension n .*

Alex Soifer [16] has credited this theorem as the first result in Ramsey Theory, but in truth Hilbert proved this obscure lemma and subsequently forgot about it. It is notable to us solely because the same structure of Hilbert cubes turns out to be central to one of our results.

Hilbert's proof is a simple repeated application of the pigeonhole principle. We present an argument similar to Hilbert, though his estimate on the value $H(r, n)$ is significantly better.

Proof. Suppose $n = 1$. The Hilbert cube $H(a; d)$ is simply the numbers $\{a, a + d\}$. By the pigeonhole principle tell us that, if we color $r + 1$ numbers with r colors, then two of them must be the same color. Calling the first a and the second $a + d$, we get our cube. This tells us $H(r, 1) = r + 1$.

Now suppose we already know, for a fixed n , that every r -coloring of $[H(r, n)]$ contains a monochromatic n -dimensional Hilbert cube. We want to guarantee an $(n + 1)$ -dimensional cube.

Consider $[(r^H + 1)H]$, divided into $r^H + 1$ blocks of integers, each of size $H = H(r, n)$ — call them $B_1, B_2, \dots, B_{r^H+1}$. An r -coloring, χ , of $[(r^H+1)H]$ can be seen as an r^H -coloring of the blocks. Specifically we may define $\chi' : [r^H + 1] \rightarrow [r]^H$ by

$$\chi'(k) = (\chi(kH + 1), \chi(kH + 2), \dots, \chi(kH + H)).$$

Since χ' assigns one of r^H colors to each of $r^H + 1$ objects, once again the pigeonhole principle tells that two objects must receive the same color — $\chi'(c) = \chi'(c + d)$. By the definition of χ' , equating the corresponding components we get: for every x in $[H]$,

$$\chi(cH + x) = \chi((c + d)H + x) = \chi(cH + dH + x). \quad (1.1)$$

Now consider the block $B_c = \{cH + 1, cH + 2, \dots, cH + H\}$. By definition of $H = H(n, r)$, there is an n -dimensional Hilbert cube $H(a; d_1, d_2, \dots, d_n)$, given by the values

$$\left\{ a + \sum_{i \in I} d_i \mid I \subseteq [n] \right\},$$

all contained in B_c , and all given the same color by χ . By Equation 1.1, we see that $\chi(a + \sum d_i) = \chi(a + \sum d_i + dH)$.

This gives our $(n + 1)$ -dimensional Hilbert cube: $H(a; d_1, \dots, d_n, dH)$. \square

Later, in 1916, while Issai Schur was working on an approach to Fermat's Last Theorem, he discovered a similar lemma [15].

Lemma 1.1.2. *For any number of colors r , there is a number $I = I(r)$ so that, for any r -coloring of $[I]$, there are monochromatic numbers x, y, z so that $x + y = z$.*

Though Schur was ultimately unable to use this to prove Fermat's Last Theorem, this lemma stayed on his mind, as he eventually conjectured a stronger result. In 1927, it was proven by Bartel Leendert van der Waerden [17].

Theorem 1.1.3. *For any length k , and any number of colors r , there is a number $W = W(k, r)$ so that, for any r -coloring of $[W]$, there is a monochromatic k -term arithmetic progression.*

A k -term arithmetic progression (also called a k -AP) is a sequence of k numbers $\{a, a+d, a+2d, \dots, a+(k-1)d\}$. We say that this k -AP is “anchored” at a , and has “common difference” d , since the difference between any two consecutive terms is d .

Van der Waerden’s theorem has a notable variant in higher dimensions, proven independently by Tibor Gallai in the 1930s (unpublished) and by Ernst Witt in 1952 [18].

Theorem 1.1.4 (Gallai-Witt). *For all r, k , there exists $GW = GW(r, p)$ so that, for every r -coloring of $[GW] \times [GW]$, there are numbers x, y, d so that the square grid*

$$\{(x + id, y + jd) \mid i, j = 0, \dots, p - 1\}$$

is all one color.

This result will be central to the results in Chapter 2. Gallai and Witt also proved their generalization to higher dimensions: any r -coloring of \mathbb{N}^d must have arbitrarily large monochromatic square subgrids.

We will also make use of a slight modification to the Gallai-Witt theorem, which follows trivially from the original by translation and re-scaling.

Corollary 1.1.5. *For all r, p, q , there exists $GW = GW(r, p, q)$ so that, for every r -coloring of $[GW] \times [GW]$, there are numbers x, y, d with the following property. There is a monochromatic square grid given by points $(x + id, y + jd)$ over all $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i|, |j| \leq p$.*

Back in one dimension, in the early 1930s Richard Rado, a student of Schur, discovered the holy grail of these efforts [12][13]. To state Rado’s theorem, it will help to introduce some terminology.

Definition 1.1.6. We say an equation $f(\mathbf{x}) = 0$ (or system of equations) is r -regular if there is an N so that for every r -coloring of $[N]$, there is a solution vector $\mathbf{x} = (x_1, \dots, x_n)$ with $\chi(x_1) = \dots = \chi(x_n)$.

If $f(\mathbf{x}) = 0$ is r -regular for every r , then we say it is *regular*.*

Schur's lemma now simply states that $x + y = z$ is regular. Likewise, van der Waerden's theorem states that, for every k ,

$$x_2 - x_1 = x_3 - x_2 = \dots = x_k - x_{k-1}$$

is regular.

Hilbert's cube lemma may also be phrased in this fashion, though it is more cumbersome than informative. The 2-dimensional case translates to the regularity of $w - x = y - z$.

While each of these results gives a family of regular equations, Rado's theorem characterizes all regular linear equations. We state his result for homogeneous linear equations.

Theorem 1.1.7. *The system of equations $A\mathbf{x} = \mathbf{0}$ is regular if and only if the matrix A satisfies the columns condition.*

We will define the columns condition shortly, but first we state the (much simpler) result for single equations.

Corollary 1.1.8. *For $a_1, \dots, a_n \in \mathbb{Z}_{\neq 0}$ fixed, the equation $a_1x_1 + \dots + a_nx_n = 0$ is regular if and only if some nonempty subset of the a_i s sums to zero.*

Although Rado's theorem seems to be the end of the line, Neil Hindman in 1974 [8] proved a conjecture of Ron Graham and Bruce Rothschild. Hindman's result is the first result in additive Ramsey theory which gives an infinite monochromatic structure, extending Folkman's theorem, a finite result already known.

Given a set X , define

$$FS(X) = \left\{ \sum_{y \in Y} y \mid Y \text{ is a finite subset of } X \right\},$$

the set of all finite sums of distinct elements of X .

Hindman then proved:

*Such an equation is sometimes called *partition-regular*, to be distinguished from *density-regular* equations. We will not refer to the latter.

Theorem 1.1.9. *For all r , for every r -coloring of the natural numbers, there is an infinite set X so that $FS(X)$ is monochromatic.*

1.1.1 The columns condition

Definition 1.1.10. Fix a matrix A , and label its columns $\mathbf{c}_1, \dots, \mathbf{c}_n$. We say that A satisfies the columns condition if there is a partition $[n] = I_1 \cup \dots \cup I_T$ so that $\sum_{j \in I_t} \mathbf{c}_j$ is in the rational span of $\{\mathbf{c}_i \mid i \in I_1 \cup \dots \cup I_{t-1}\}$. In particular, $\sum_{j \in I_1} \mathbf{c}_j$ is supposed to be in the span of the empty set, so this requires $\sum_{j \in I_1} \mathbf{c}_j = \mathbf{0}$.

Example 1.1.11. Consider the following matrix:

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

Label the columns $\mathbf{c}_1, \dots, \mathbf{c}_6$.

One possible partition which shows the columns condition is

$$I_1 = \{1, 2, 3\}, I_2 = \{4, 5\}, I_3 = \{6\}.$$

We have one condition to check for each of the three sets I_j :

$$\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{c}_4 + \mathbf{c}_5 = \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix} = 2\mathbf{c}_2 + \mathbf{c}_3$$

$$\mathbf{c}_6 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3}(\mathbf{c}_4 - \mathbf{c}_5).$$

The above is the common definition of the columns condition, but it will also be helpful to look at a new, equivalent reformulation:

Definition 1.1.12. A matrix A with n columns satisfies the columns condition if there is a sequence of vectors $\mathbf{u}_1, \dots, \mathbf{u}_T$ in the nullspace of A and decreasing sequence of sets $R_1 \supseteq \dots \supseteq R_T$ so that:

1. If $i \in R_t$, then $u_t(i) = 0$.
2. If $i \notin R_t$, then there is an $s \leq t$ with $u_s(i) = 1$.
3. $R_T = \emptyset$.

Example 1.1.13. Using the matrix A from the previous example, we may take

$$\begin{aligned} R_1 &= \{4, 5, 6\} & \mathbf{u}_1 &= (1, 1, 1, 0, 0, 0)^T \\ R_2 &= \{6\} & \mathbf{u}_2 &= (0, -2, -1, 1, 1, 0)^T \\ R_3 &= \emptyset & \mathbf{u}_3 &= (0, 0, 0, -\frac{1}{3}, \frac{1}{3}, 1)^T. \end{aligned}$$

Lemma 1.1.14. *Definitions 1.1.10 and 1.1.12 are equivalent.*

Proof. We must prove that A satisfies Definition 1.1.10 \iff A satisfies Definition 1.1.12.

(Definition 1.1.10 \implies Definition 1.1.12):

For $t \in [T]$, define $R_t = I_{t+1} \cup \dots \cup I_T$. For each t , we are given that $\sum_{j \in I_t} \mathbf{c}_j$ is contained in the span of

$$\{\mathbf{c}_i \mid i \in I_1 \cup \dots \cup I_{t-1}\} = \{\mathbf{c}_i \mid i \notin R_{t-1}\}.$$

In other words, there are coefficients λ_i so that

$$\sum_{j \in I_t} \mathbf{c}_j = \sum_{i \notin R_{t-1}} \lambda_i \mathbf{c}_i.$$

Rearranging, this gives us

$$\sum_{j \in I_t} \mathbf{c}_j - \sum_{i \notin R_{t-1}} \lambda_i \mathbf{c}_i = \mathbf{0}. \tag{1.2}$$

We define the vector $\mathbf{u}_t = (u_t(1), \dots, u_t(n))$ by:

$$u_t(i) = \begin{cases} 1 & \text{if } i \in I_t \\ -\lambda_i & \text{if } i \in R_{t-1} \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

Equation 1.2 may now concisely be stated as $A\mathbf{u}_t = \mathbf{0}$ — that is, \mathbf{u}_t is in the nullspace of A , as needed.

By Equation 1.3, the sequence of vectors \mathbf{u}_t satisfies condition (1) of Definition 1.1.12. Since $u_t(i) = 1$ when $i \in I_t$, we also satisfy condition (2). Condition (3) is satisfied since $[n] = \bigcup I_t$.

(Definition 1.1.12 \Rightarrow Definition 1.1.10):

We will essentially simply reverse the previous argument. Define $I_1 = [n] \setminus R_1$, and $I_t = R_{t-1} \setminus R_t$. From conditions (1) and (2) of Definition 1.1.12, we know that, if $i \in I_t$, then $u_t(i) = 1$. Since $A\mathbf{u}_t = \mathbf{0}$, we have

$$\begin{aligned} \sum_{i \in [n]} u_t(i) \mathbf{c}_i &= \mathbf{0} \\ \sum_{i \in I_t} \mathbf{c}_i + \sum_{\substack{j \in I_s \\ s < t}} u_t(j) \mathbf{c}_j &= \mathbf{0}. \end{aligned}$$

This tells us that indeed $\sum_{i \in I_t} \mathbf{c}_i$ is in the span of $\{\mathbf{c}_j \mid j \in I_1 \cup \dots \cup I_{t-1}\}$, as required. \square

1.2 Between additive- and graph-Ramsey theory

These results in Section 1.1 are philosophically related to Ramsey's theorem, but the graph theoretic and additive sides of Ramsey theory are largely distinct fields.

In 1995, however, András Hajnal asked Paul Erdős a question linking the two subfields.

Question 1.2.1. Consider a 2-coloring of the complete graph $K_{\mathbb{N}}$. Must there be either an infinite set X so that the complete graph on $FS(X)$ is red, or a blue triangle?

Here $FS(X)$ is the set of all finite sums of elements of X , as seen in Hindman's theorem (Theorem 1.1.9).

Hajnal's question spurred a number of hybrid results between additive- and graph-Ramsey theory [2][5][6][10]. These beautiful results are fundamentally asymmetric — there is either a red clique with prescribed additive structure, or a blue subgraph with prescribed graph structure.

As it turns out, the first consequence of Hajnal's question is too strong, but the second consequence may be weakened, to get this powerful general result [6].

Theorem 1.2.2. *For any partition-regular equation $A\mathbf{x} = \mathbf{0}$ and any m , for every 2-coloring of $K_{\mathbb{N}}$, there is either a solution vector $\mathbf{x} = (x_1, \dots, x_n)$ so that all edges among $\{x_1, \dots, x_n\}$ are red, or a blue K_m .*

The proof of this (and earlier results) are elementary, coming from a nice use of the Hales-Jewett theorem, another central result in Ramsey Theory.

Various authors have considered the simpler seeming version of the question:

Question 1.2.3. For a fixed equation $f(\mathbf{x}) = 0$, and fixed r , is it true that for every r -coloring of the edges of $K_{\mathbb{N}}$, there is a vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ with $f(\mathbf{x}) = 0$ so that the x_i s are distinct, and the complete subgraph on $\{x_i\}_{i=1}^n$ is monochromatic?

Chapter 2 is devoted to this question. Deceptively, the simplest forms of arithmetic structure — arithmetic progressions, solutions to $x + y = z$ — are very

simple to avoid when coloring edges. For this reason, the question seems initially hopeless, and was long neglected. Our main results show that there are, in fact, examples of linear equations $A\mathbf{x} = \mathbf{b}$ which do have this property, which we call *graph-regular* equations.

1.3 The compactness principle

Often when proving a result in Ramsey theory, it becomes burdensome to keep track of the proposed *Ramsey number* — the size of the structure which is being proved “large enough” to guarantee the desired monochromatic substructure. For example, in van der Waerden’s theorem, just how many integers are we claiming that one must 7-color before creating a monochromatic 10-term arithmetic progression? It can be much simpler (and, given the astronomical bounds which arise, perhaps not must less precise) to prove that coloring *all* of the integers gives the same.

A very nice result by Paul Erdős and Nicolaas Govert de Bruijn tells us that the finite and infinite results are actually equivalent [1]. As stated, the de Bruijn-Erdős theorem is about the chromatic number of a graph.

Definition 1.3.1. For a graph (or hypergraph) $H = (V, E)$, we say that $\chi : V \rightarrow [k]$ is a *proper coloring* if no edge in E is monochromatic — every edge contains vertices of at least two different colors.

The *chromatic number* of H , $\chi(H)$, is the minimum k so that such a proper coloring exists — or ∞ if no such k exists.

We state their result, known as the Compactness Principle, in the general case of hypergraphs.

Theorem 1.3.2. *For any (infinite) hypergraph H whose edges are finite,*

$$\chi(H) = \sup\{\chi(F) \mid F \subset H \text{ is a finite sub-hypergraph}\}.$$

Proof. We give a proof of the Compactness Principle for countable hypergraphs.

It is clear that $\chi(H) \geq \chi(F)$ for any subgraph F (finite or not). Moreover, if $\chi(F)$ grows unbounded, then $\chi(H)$ must be infinite. It remains to show that, if $\chi(F)$ is bounded, then there is some proper k -coloring of H . Define

$$k = \max\{\chi(F) \mid F \subset H \text{ is a finite sub-hypergraph}\}.$$

Without loss of generality, take $V(H) = \mathbb{N}$. We build a partial k -coloring, χ , of \mathbb{N} , one vertex at a time, in such a way that χ maintains this key property:

$$\text{Every finite subgraph of } H \text{ has a proper coloring consistent with } \chi. \quad (1.4)$$

The initial “empty” coloring satisfies this by assumption.

By induction, suppose we have defined $\chi(j)$ for all $j < \ell$, so that χ satisfies Property 1.4. We must now choose a color for ℓ .

For each $n \geq \ell$, let $H[n]$ denote the induced sub-hypergraph of H on the vertex set $[n]$.

For each $i \in [k]$, let A_i be the set of $n \geq \ell$ so that $H[n]$ has a proper coloring consistent with χ on $[\ell - 1]$, and vertex ℓ is given color i .

Due to Property 1.4, χ does extend to a proper coloring of each $H[n]$, each giving ℓ some color. This tells us $\bigcup A_i = \{\ell, \ell + 1, \dots\}$. In particular, since there are only k choices for i , we may pick some i with A_i infinite. Define $\chi(\ell) = i$.

We claim that χ still satisfies Property 1.4. Indeed, every finite subgraph $F \subset H$ contains some maximum vertex N . Since A_i is infinite, there is some $n \geq N$ in A_i , so $F \subset H[n]$ does have a proper coloring extending χ .

Continuing this procedure for all ℓ , we construct $\chi : \mathbb{N} \rightarrow [k]$. Since each edge of H is contained in some $H[n]$, we know it is properly colored by χ . That is, χ is indeed a proper k -coloring of H , as desired. \square

The uncountable case works in exactly the same way, but requires the axiom of choice. Although set theory is outside of the scope of this introduction, we will describe the minor adjustments needed, for the benefit of anyone already familiar with the terms.

Rather than \mathbb{N} , we should assume that $V(H)$ is well-ordered. At each step, we should choose i so that A_i is *cofinal* in $V(H)$ — that is, $\sup A_i = \lambda$ — or alternatively so that $|A_i| = |V(H)|$. Applying transfinite induction to define χ on all vertices, we get the same result.

That’s all well and good, but how is the Compactness Principle relevant to Ramsey theory? In fact, many results in Ramsey theory can be stated in terms of the chromatic number of certain hypergraphs.

Here is a corollary (or perhaps a restatement) of the Compactness Principle which applies to results in Ramsey theory over \mathbb{N} .

Corollary 1.3.3. *Let A be a family of finite subsets of \mathbb{N} .[†] Then the following are equivalent.*

1. *For every r -coloring of \mathbb{N} , there is some monochromatic $a \in A$.*
2. *There is some N so that, for every r -coloring of $[N]$, there is some monochromatic $a \in A$ with $a \subseteq [N]$.*

This can be easily generalized to results in graph Ramsey theory. For any set X , and $n \in \mathbb{N}$, let $\binom{X}{n}$ refer to the family of all n -term subsets of X . Then we have:

Corollary 1.3.4. *Fix $k \in \mathbb{N}$, and let A be a family of finite subsets of $\binom{\mathbb{N}}{k}$.[‡] Then the following are equivalent.*

1. *For every r -coloring of $\binom{\mathbb{N}}{k}$, there is some monochromatic $a \in A$.*
2. *There is some N so that, for every r -coloring of $\binom{[N]}{k}$, there is some monochromatic $a \in A$ with $a \subseteq \binom{[N]}{k}$.*

[†]For example, A could be all k -term arithmetic progressions, or all solutions to $x + y = z$.

[‡]For example, if $k = 2$, A could be the collection of all edge sets of complete graphs on n vertices.

1.4 Outline of results

In this dissertation, we explore several problems in Ramsey theory.

In Chapter 2, we investigate Question 1.2.3 for linear equations: are there equations $A\mathbf{x} = \mathbf{b}$ so that every finite coloring of pairs of natural numbers gives a monochromatic clique whose vertices solve the equation? In short, is $A\mathbf{x} = \mathbf{b}$ a *graph-regular* equation? Taking inspiration from work leading up to Rado’s theorem, we first look at [systems of] equations known early on to be partition-regular in Section 2.2. When this fails horribly, we identify some necessary conditions in Section 2.3. In Section 2.4, we give two extensions of Rado’s “columns condition” to the graph setting — the *weak* and *strong* graph columns conditions — and show that the weak version is necessary for an equation to be graph-regular. In Section 2.5 we give an initial positive result: there is an n so that, any 2-coloring of the edges of the complete graph on $[n]$ gives a monochromatic 2-dimensional Hilbert cube. In Section 2.6, we prove a lemma about coloring k -ary trees which may be interesting in its own right. In Section 2.7 we extend our initial result to any number of colors and to Hilbert cubes of any size. In Section 2.8, we prove that the strong graph columns condition is sufficient for an equation to be graph-regular. In Section 2.9, we show that the notions of partition-regular and graph-regular equations do not have a satisfying extension to hypergraph-regular equations. Finally, in Section 2.10, we suggest possible research directions for this type of problem.

In Chapter 3, we look at a stronger version of Ramsey’s theorem from Paris and Harrington, and show a counterexample to the analogous result for van der Waerden’s theorem.

Chapter 2

Toward Rado's theorem on graphs

2.1 Introduction

Rado's theorem (Theorem 1.1.7) gives a [relatively] simple characterization of regular linear equations. That is, given a linear equation $A\mathbf{x} = \mathbf{b}$, Rado's columns condition (Definition 1.1.10) determines whether every finite-coloring of \mathbb{N} yields a vector $\mathbf{x} = (x_1, \dots, x_n)$ so that $\{x_1, \dots, x_n\}$ are all the same color.

In Section 1.2, we saw a partial generalization of regular equations to graphs. We complete the generalization:

Definition 2.1.1. We say $f(\mathbf{x}) = 0$ is *graph-regular* if, for all r , there is a number $N(r)$ so that, for all $N > N(r)$, every r -coloring of the edges of the complete graph on $[N]$ has a solution $\mathbf{x} = (x(1), \dots, x(k))$ so that (1) the edges $\{x(i), x(j)\}$ are all the same color, and (2) the values $\{x(i)\}$ are distinct.

We require a solution by distinct values due to degeneracy issues which do not appear in the case of coloring points. Further, for non-triviality, we require the equation to contain at least three variables.

2.2 Negative results

We begin the search for graph-regular equations by looking at natural first-guesses — analogs of well-known theorems for partition-regular equations. While none of these will give us any examples of graph-regularity, they begin to help shape our understanding of the difficulties.

2.2.1 Arithmetic progressions

Arithmetic progressions may be the “most popular” structures to look for in additive Ramsey theory. Van der Waerden’s theorem [17] tells us that any finite-coloring of the naturals have arbitrarily long monochromatic arithmetic progressions. What can we say when coloring pairs of naturals? An arithmetic progression of length 3 is given by $a, a + d, a + 2d$. We notice that the triple contains two differences: d and $2d$. This allows us to 2-color the complete graph on the naturals without a monochromatic 3-AP.

The coloring is straightforward. For a pair $\{x, y\}$, write $|x - y| = 2^p q$ where p, q are integers and q is odd. If p is even, color $\{x, y\}$ red. Otherwise, color it blue.

Now let $a, a + d, a + 2d$ be a 3-AP. Write $d = 2^p q$. Then we see $2d = 2^{p+1} q$, so the edges $\{a, a + d\}$ and $\{a, a + 2d\}$ have different colors.

This coloring avoids 3-APs, so we certainly cannot hope for anything longer.

2.2.2 Schur’s equation and generalizations

Schur’s theorem [15] states that any finite-coloring of the naturals has a mono-chromatic solution to $x + y = z$.

If $x + y = z$ then either x or y is smaller than their average, $\frac{1}{2}z$, and the other must be larger than their average. Thus, given a pair $\{u, v\}$ with $u < v$, we color it red if $u \leq \frac{1}{2}v$, and blue if $u > \frac{1}{2}v$. Now we see that whenever $x + y = z$, the largest of the three numbers must be z . The lesser of x and y is smaller than $\frac{1}{2}z$, and the other is larger, so the pairs $\{x, z\}$ and $\{y, z\}$ have different colors. (Recall that we are only interested in solutions by distinct numbers).

A similar approach works for equations of the form

$$a_1x_1 + \dots + a_kx_k = bz, \tag{2.1}$$

with constants $a_1, \dots, a_k \geq b > 0$, and variables x_1, \dots, x_k, z .

Using two colors, we can ensure that every graph induced by a solution to an equation in the form of Equation 2.1 contains both colors.

We see that, for each i , $a_ix_i \leq bz$. Since $a_i \geq b > 0$, we get $x_i \leq z$. Let $M = a_1 + \dots + a_k$. Divide both sides of the equation by M to get

$$\frac{a_1}{M}x_1 + \dots + \frac{a_k}{M}x_k = \frac{b}{M}z.$$

This says that the weighted average of the x_i 's is $\frac{b}{M}z$. Again, one of the x_i 's must be smaller than their average, and another must be larger. Thus, when $u < v$, we should color $\{u, v\}$ red if $u \leq \frac{b}{M}v$, and blue otherwise. We immediately see that one of the pairs $\{x_i, z\}$ must be red and another must be blue.

Remark 2.2.1. The argument given above is really a greedy coloring. At step t , color the pairs $\{1, t\}, \dots, \{t-1, t\}$ in a way that handles those solutions to Equation 2.1 with largest element t . Since we can manage all these solutions at once, we avoid all monochromatic solutions. The incredible thing to notice here is that this coloring is much stronger than needed. If x_1, \dots, x_k, z satisfy Equation 2.1, then the star connecting z to all of the x_i 's is not even monochromatic. Forget about the clique! The strength of this technique suggests that we may be able to handle a larger family of equations.

2.2.3 Three variable equations with six colors

As with many problems in Ramsey theory, we may consider our conjecture as a hypergraph coloring problem. The vertex set is all pairs in \mathbb{N} . For each solution (x_1, \dots, x_k) to $b_1x_1 + \dots + b_kx_k = 0$, there is a hyperedge containing all pairs among the x_i 's. If we properly color this $\binom{k}{2}$ -uniform hypergraph — that is, if we avoid monochromatic hyperedges — then there are no monochromatic solutions to the equation. Thus we may apply theorems about hypergraph coloring.

For an equation in three variables, this hypergraph is *simple* — any two pairs A and B are either disjoint (and have no hyperedges in common), or have the form $A = \{x, y\}, B = \{x, z\}$, leaving only $\{y, z\}$ to complete the hyperedge.

Fix a, b, c, d , and consider the hypergraph formed as above by the equation

$$ax + by + cz = d. \quad (2.2)$$

Consider a pair $\{u, v\}$. How many hyperedges can it be contained in? Well, there are 6 different ways of assigning the values u and v to the variables in Equation 2.2, each determining the third value:

$$\begin{aligned} au + bv + cz = d &\implies z = \frac{d-au-bv}{c} \\ av + bu + cz = d &\implies z = \frac{d-av-bu}{c} \\ au + by + cv = d &\implies y = \frac{d-au-cv}{b} \\ av + by + cu = d &\implies y = \frac{d-av-cu}{b} \\ ax + bu + cv = d &\implies x = \frac{d-bu-cv}{a} \\ ax + bv + cu = d &\implies x = \frac{d-bv-cu}{a}. \end{aligned}$$

Thus, each pair $\{u, v\}$ is contained in at most 6 hyperedges. This gives us a simple hypergraph with maximum degree $\Delta \leq 6$. The hypergraph version of Brooks' theorem [9] applies.

Theorem 2.2.2. *If H is a hypergraph with maximum degree Δ , then $\chi(H) \leq \Delta$ except in these cases:*

1. $\Delta = 1$,
2. $\Delta = 2$ and H contains an odd cycle (an ordinary graph),
3. H contains a K_Δ (an ordinary graph).

Corollary 2.2.3. *For any choice of $a, b, c \in \mathbb{N}$, the equation $ax + by + cz = d$ is not graph-regular.*

Proof. Construct a hypergraph from this equation as described above. All of these cases of Brooks' theorem are irrelevant; ours is a 3-uniform hypergraph, and we don't have any illusions that we can 1-color it. Thus we can properly 6-color

our hypergraph. By construction, this avoids monochromatic solutions to the equation. \square

Moreover, if for example $a = b$, then the six solutions reduce to three distinguishable ones, meaning 3 colors is enough.

2.3 Necessary conditions for graph-regularity

Having ruled out any candidate equations in only three variables, we now turn our focus to a more systematic attempt to find a graph-regular equation.

We define a family of colorings, f_n , of $\binom{\mathbb{N}}{2}$ by

$$f_n(an + x, bn + y) = \begin{cases} \text{blue} & \text{if } x = y \\ \min\{x, y\} & \text{if } x \neq y, \end{cases}$$

where $x, y \in \{0, 1, \dots, n-1\}$, and where we storm past the usual boundary between *actual* colors and mathematician's colors. We claim that all monochromatic triangles under f_n are blue.

Consider a triangle $\{x, y, z\}$ with no blue edge, where $x = an + i, y = bn + j, z = cn + k$, and $i, j, k \in \{0, 1, \dots, n-1\}$. Then the numbers i, j, k must be distinct. Reordering so that $i < j < k$, we see that $f_n(x, y) = f_n(x, z) = i$, while $f_n(y, z) = j \neq i$. Thus a triangle without a blue edge cannot be monochromatic. Turning this around, any monochromatic triangle must be blue.

Going back to the definition of f_n , this means that any monochromatic triangle — and hence any monochromatic clique — must represent only one congruence class mod n .

Lemma 2.3.1. *If $\sum a_i x_i = b$ is graph-regular with $b, a_i \in \mathbb{Z}$, then b is a multiple of $\sum a_i$.*

Proof. Write $\sum a_i = M$. If $M \neq 0$, then consider the coloring f_M . Let $\{x_i\}$ be monochromatic under f_M , so that $x_i = b_i M + c$.

Then we have

$$\begin{aligned} \sum a_i x_i &= \sum a_i (b_i M + c) \\ &= (\sum a_i b_i M) + \sum a_i c \\ &= (\sum a_i b_i) M + cM. \end{aligned}$$

We see that $\sum a_i x_i$ is a multiple of M . If $\sum a_i x_i = b$, then we see that b is a multiple of M as well.

On the other hand, if $M = 0$, then we may repeat the above argument using any f_n . Since the cM term goes away, we learn that b is a multiple of n for every n we choose, forcing $b = 0$ as well. \square

Fix n , and define the coloring g_n of $\binom{\mathbb{N}}{2}$ by

$$g_n(n^j a, n^k b) = \begin{cases} \text{red} & \text{if } j \neq k \\ f_n(a, b) & \text{if } j = k, \end{cases}$$

where a and b are not divisible by n .

Note 2.3.2. Using the same argument as for f_n , we see that any triangle which is monochromatic under g_n must be red or blue. Writing $x_i = b_i n^{r_i}$, this means either all r_i values are distinct (yielding a red clique), or all r_i values are equal and all b_i values are congruent modulo n (yielding a blue clique).

Lemma 2.3.3. *If $\sum a_i x_i = b$ is graph-regular with $b, a_i \in \mathbb{Z}$, then $\sum a_i = b = 0$.*

Proof. Suppose $\sum a_i x_i = b$ is graph-regular, with $M = \sum a_i \neq 0$. By Lemma 2.3.1, we may write $b = kM$. We assume each a_i is non-zero, as removing superfluous variables will preserve graph-regularity.

We apply a new coloring, which should be thought of as a hybrid between the colorings f_n and g_n . There is a prime p which does not divide M nor any of the a_i values, since none of these values is 0. For any x , we may uniquely write $x = cp + d + k$, where $d \in \{0, 1, \dots, p-1\}$ and $k = \frac{b}{M}$ was defined above.

Using this form, we define

$$\chi(cp + d + k, c'p + d' + k) = \begin{cases} \min\{d, d'\} & \text{if } d \neq d', \\ g_p(c, c') & \text{if } d = d'. \end{cases}$$

Note that we treat the colors from the two pieces of this function as distinct — all pairs in a monochromatic clique must have either all used the first piece, or all used the second. In fact, we have already seen from our analysis of f_n that no monochromatic clique can arise from the first piece of the definition of χ , so any solution must come from g_p .

Let $\{x_i\}$ be a monochromatic solution. Since all edges must have been colored by the second piece of χ , we can write $x_i = \beta_i p + d + k$, where $d \in$

$\{0, 1, \dots, p-1\}$ is common for each x_i .* This gives us

$$\begin{aligned} \sum a_i x_i &= b \\ \sum a_i (\beta_i p + d + k) &= kM \\ \sum a_i \beta_i p + dM + kM &= kM \\ (\sum a_i \beta_i) p + dM &= 0. \end{aligned}$$

Since p does not divide M , and d is less than p , we must have $d = 0$. Dividing by p , we are left with

$$\sum a_i \beta_i = 0.$$

Write $\beta_i = (b_i p + c_i) p^{r_i}$, with $c_i \in \{1, 2, \dots, p-1\}$. From Note 2.3.2, we have either a red clique with each r_i distinct, or we have a blue clique with $r_i = r$ and $c_i = c$ common across all i .

Case 1. The clique is blue, so $\beta_i = (b_i p + c) p^r$. We see that

$$\begin{aligned} \sum a_i (b_i p + c) p^r &= 0 \\ \sum a_i (b_i p + c) &= 0 \\ \sum a_i c &\equiv 0 \pmod{p} \\ cM &\equiv 0 \pmod{p}. \end{aligned}$$

Since p divides neither c nor M , this is impossible.

Case 2. The clique is red, so $\beta_i = (b_i p + c_i) p^{r_i}$, with each r_i distinct.

Let r_j be the unique smallest exponent. We find

$$\begin{aligned} \sum a_i (b_i p + c_i) p^{r_i} &= 0 \\ \sum a_i (b_i p + c_i) p^{r_i - r_j} &= 0 \\ a_j (b_j p + c_j) + \sum_{i \neq j} a_i (b_i p + c_i) p^{r_i - r_j} &= 0 \\ a_j c_j &\equiv 0 \pmod{p}. \end{aligned}$$

Again, since p divides neither c nor a_j , this is impossible.

Together, we have seen that $M \neq 0$ is impossible, so $\sum a_i = M = 0$. Since $b = kM$, we also get $b = 0$ for free. \square

*For small values of x_i , the resulting β_i may be zero or negative. A little technical care is required to handle $\beta_i = 0$, but we will ignore it here.

Note that Lemma 2.3.3 extends to systems of linear equations — if $A\mathbf{x} = \mathbf{b}$ is graph-regular, then $\mathbf{b} = \mathbf{0}$ and the columns of A sum to $\mathbf{0}$. This is easily seen since each equation from the system $A\mathbf{x} = \mathbf{b}$ must also be graph-regular.

Consider such an equation, $\mathbf{a}_1x_1 + \dots + \mathbf{a}_kx_k = 0$, where the coefficients sum to 0. We may rewrite this as, for instance,

$$\mathbf{a}_1(x_1 - x_k) + \dots + \mathbf{a}_{k-1}(x_{k-1} - x_k) = 0,$$

now an equation relating differences. This suggests that we should consider colorings based on these differences — colorings of the form $\chi(x < y) = f(y - x)$. We may now take guidance from Rado’s theorem to get a better handle on things.

For a prime p , and $x = p^r(bp + s)$, let $\psi_p(x) = s \in [p - 1]$ be the “super mod p ” coloring, from Rado’s theorem. Rado’s theorem suggests to us that, when looking at colorings based only on differences between endpoints, we need only consider the colorings ψ_p . We will show that Rado’s theorem does apply here, but we begin with a simple consequence to give a feel for how it works.

Theorem 2.3.4. *Let $\sum_{i=1}^k a_i x_i = 0$ be graph-regular with $a_i \in \mathbb{Z}$. Then there is a nonempty set $I \subsetneq [k]$ so that*

$$\sum_{i \in I} a_i = \sum_{j \notin I} a_j = 0.$$

To prove this, we introduce the graph version of ψ_p .

Define $\varphi_p : \binom{\mathbb{N}}{2} \rightarrow [p - 1]$ by

$$\varphi_p(x < y) = \psi_p(y - x).$$

Proof. Fix a prime p and color $\binom{\mathbb{N}}{2}$ by φ_p . Suppose x_1, \dots, x_k are distinct values satisfying $a_1x_1 + \dots + a_kx_k = 0$, with the edges among them all color c . Let x_j be the smallest of these values. As noted earlier, we see that

$$\sum_{i \neq j} a_i(x_i - x_j) = 0.$$

By choice of x_j , each of the terms $x_i - x_j$ is positive. Thus we may write $x_i - x_j = p^{r_i}(b_i p + c)$, since $\varphi_p(x_j < x_i) = \psi_p(x_i - x_j) = c$. Let r be the smallest exponent

among these $k-1$ terms, and let $I = \{i \in [k] \setminus \{j\} \mid r_i = r\}$. Note that $\emptyset \subsetneq I \subsetneq [k]$.

We see that

$$\begin{aligned}
0 &= \sum_{i \neq j} a_i p^{r_i} (b_i p + c) \\
&= \sum_{i \neq j} a_i p^{r_i - r} (b_i p + c) \\
&= \sum_{i \in I} a_i (b_i p + c) + p \left(\sum_{i \notin I \cup \{j\}} a_i p^{r_i - r - 1} (b_i p + c) \right) \\
&\equiv c \left(\sum_{i \in I} a_i \right) \pmod{p}.
\end{aligned}$$

Since c is in $[p-1]$, we see that p divides $\sum_{i \in I} a_i$. If we take $p > \sum_{i=1}^k |a_i|$, then the only way this can happen is if $\sum_{i \in I} a_i = 0$. Since we already know that $\sum_{i=1}^k a_i = 0$, we learn that $\sum_{j \notin I} a_j = 0$ as well. \square

Corollary 2.3.5. *No nondegenerate homogeneous linear equation of three variables is graph-regular.*

We have already proven this as Corollary 2.2.3, but we now have a self-contained proof.

Proof. Let $k = 3$, and let $I \subsetneq \{1, 2, 3\}$ be the nonempty set guaranteed by Theorem 2.3.4. Either I or its complement has a single element. The corresponding coefficient must be 0, meaning the equation depends on at most two variables and is trivial. \square

Corollary 2.3.6. *The only candidate homogeneous linear equation in four variables which might be graph-regular is $w - x + y - z = 0$ (up to re-scaling).[†]*

Proof. Let $aw + bx + cy + dz = 0$ be graph-regular, with $a, b, c, d \in \mathbb{Z}_{\neq 0}$. By Theorem 2.3.4, we know that two complementary subsets of the coefficients must add to zero. Up to permutation, this leaves us with $aw - ax + cy - cz = 0$, or rather $a(w - x) = c(z - y)$. We may assume both a and c are positive by switching w and x , or y and z . We claim $a = c$.

[†]Theorem 2.5.1 will show that this equation is indeed graph-regular.

Suppose not. After canceling common factors, we may assume we may assume c is divisible by some prime p which does not divide a . Pick r so that p^r divides c , but p^{r+1} does not. Consider the 2-coloring given by

$$\chi(x, y) \equiv \left\lfloor \frac{f(x, y)}{r} \right\rfloor \pmod{2},$$

where $f(x, y)$ gives the highest exponent of p which divides $x - y$.

Now suppose $a(w - x) = c(z - y)$, with w, x, y, z distinct. Let $f(w, x) = k$. This means that $a(w - x)$ represents a power of p^k on the left hand side. Dividing by c , we learn that $(z - y)$ represents a power of p^{k-r} , so $f(y, z) = k - r$. Looking at the corresponding χ values of $\{w, x\}$ and $\{y, z\}$, we see that they are different, so the edges among $\{w, x, y, z\}$ are not monochromatic. \square

2.4 The graph columns condition

2.4.1 Definitions

As introduced in Section 1.1, Rado's theorem (Theorem 1.1.7) characterizes the regular equations by way of the columns condition. We restate a stronger version of Rado's theorem here.

Theorem 2.4.1. *The equation $A\mathbf{x} = \mathbf{0}$ has a monochromatic solution under any finite coloring of \mathbb{N} whenever A satisfies the columns condition. If A does not satisfy the columns condition then there is some $p_0 = p_0(A)$ so that, for every prime $p > p_0$, a monochromatic solution is avoided by the coloring ψ_p (which will be introduced in Section 2.3).*

There are several ways to extend the columns condition to apply to edge-colorings. We state two.

Definition 2.4.2. We say a matrix A with n columns satisfies the *weak graph columns condition* (WGCC) if there is a sequence of vectors $\mathbf{1} = \mathbf{u}_0, \dots, \mathbf{u}_T$ in the nullspace of A , and a decreasing sequence of graphs $R_0 \supseteq \dots \supseteq R_T$ with common vertex set $[n]$ so that

1. If $\{i, j\} \in R_t$, then $u_t(i) = u_t(j)$.
2. If $\{i, j\} \notin R_t$, then there is an $s \leq t$ with $|u_s(j) - u_s(i)| = 1$.
3. R_T is empty.

Further, we say A satisfies the *strong graph columns condition* (SGCC) if we may replace (1) and (2) by (1*) and (2*):

- 1*. If $\{i, j\} \in R_t$, $u_t(i) = u_t(j) \in \{0, 1\}$.
- 2*. If $\{i, j\} \notin R_t$, then there is an $s \leq t$ with $u_s(i) = 0$ and $u_s(j) = 1$ (or vice versa).

In words, the \mathbf{u}_t 's are restricted so that, for each i, j pair, as t increases, the values $u_t(i), u_t(j)$ are initially equal, remain equal until they differ by exactly 1, and are unrestricted after that. An edge between i and j in graph R_t means that pair remains restricted through time t . If $\{i, j\} \in R_t$, then we say the pair is *restricted* at time t , otherwise it is *unrestricted*.

The strong graph columns conditions requires conditions on the values $u_t(i)$ in addition to the differences across edges.

Example 2.4.3. Let

$$A = \begin{pmatrix} 2 & -2 & 5 & -3 & 4 & -6 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{pmatrix}.$$

Here is one sequence of vectors showing A satisfies the graph columns condition (weak and strong):

$$\mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

The corresponding restriction graphs may be described simply:

- R_1 : All edges among $\{1, 2\}$ and $\{3, 4, 5, 6\}$ remain restricted.
- R_2 : Edges $\{3, 4\}$ and $\{5, 6\}$ remain restricted.
- R_3 is empty — no edges are restricted.

Note that, at each step, R_t is a union of disjoint cliques. This always happens.

Example 2.4.4. Let

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & -1 & 0 & -1 & 0 & -1 & -r & r+1 \end{pmatrix}.$$

Here is one sequence of vectors showing A satisfies the weak graph columns condition:

$$\mathbf{u}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ r+1 \\ r \end{pmatrix}.$$

Notice that \mathbf{u}_3 relaxes the restriction on the 7th and 8th columns by using values r and $r+1$, rather than 0 and 1 as required by the strong graph columns condition.

We now state our main result:

Theorem 2.4.5. *Fix a matrix A . If $A\mathbf{x} = \mathbf{b}$ is graph-regular, then A satisfies the weak graph columns condition and $\mathbf{b} = \mathbf{0}$. If A satisfies the strong graph columns condition, then $A\mathbf{x} = \mathbf{b}$ is graph-regular.*

We will prove WGCC is necessary in Section 2.4.2, using Theorem 2.4.1. In Section 2.8, we will show SGCC is sufficient.

2.4.2 The weak graph columns condition

Lemma 2.4.6. *Let A be a matrix whose columns sum to $\mathbf{0}$. If the equation $A\mathbf{x} = \mathbf{0}$ has a monochromatic solution under the edge-coloring φ_p for every prime p , then A satisfies the weak columns condition.*

Proof. Let $A\mathbf{x} = \mathbf{0}$ have a monochromatic solution under φ_p for every prime p . Denote the columns of A by $\{\mathbf{a}_i\}_{i=1}^n$.

From A , we will make a larger matrix C with columns indexed by $\binom{[n]}{2} = \{(i, j) \mid 1 \leq i < j \leq n\}$. The columns of C come from gluing the columns of A (or

the zero vector) to new vectors which bind relationships between the columns of A .

$$\mathbf{c}_{1j} = \begin{pmatrix} \mathbf{a}_j \\ - \\ \mathbf{b}_{1j} \end{pmatrix}, \text{ and } \mathbf{c}_{ij} = \begin{pmatrix} \mathbf{0} \\ - \\ \mathbf{b}_{ij} \end{pmatrix} \text{ if } i > 1,$$

where

$$b_{ij}(k, \ell) = \begin{cases} 1 & \text{if } (k, \ell) = (1, j) \\ -1 & \text{if } (k, \ell) = (1, i) \text{ or } (i, j) \\ 0 & \text{otherwise.} \end{cases}$$

Note that the matrix C does not explicitly contain the column \mathbf{a}_1 . However, since $\sum \mathbf{a}_i = \mathbf{0}$, that information is not lost.

Suppose $C\mathbf{y} = \mathbf{0}$, with $y(1, j) = x(j) - x(1)$. The vectors $\{b_{ij}\}$ are designed so that $y(i, j) = x(j) - x(i)$.

Turned around, when $A\mathbf{x} = \mathbf{0}$, and \mathbf{y} is defined above, we get $C\mathbf{y} = \mathbf{0}$. Likewise, if $C\mathbf{y} = \mathbf{0}$ then, for any value $x(1)$, the values $x(i)$ are uniquely defined from \mathbf{y} , and they satisfy $A\mathbf{x} = \mathbf{0}$.

We would like to say that, when $A\mathbf{x} = \mathbf{0}$ is a monochromatic solution under the edge-coloring φ_p , the corresponding solution to $C\mathbf{y} = \mathbf{0}$ is monochromatic under the vertex-coloring ψ_p . However, this is not quite true. The definition says $\varphi_p(x, y) = \psi_p(y - x)$ only when $x < y$. For a monochromatic solution under ψ_p , we would need $x(1) < x(2) < \dots < x(n)$. Instead, for each permutation $\sigma \in S_n$, we must define the matrix $C(\sigma)$ which will “work” when $x(\sigma_1) < x(\sigma_2) < \dots < x(\sigma_n)$. We omit the definition of $C(\sigma)$, but it is essentially the same as C , defined in such a way that $y(i, j)$ is always a positive number when \mathbf{x} is ordered by σ .

If \mathbf{x} is a solution to $A\mathbf{x} = \mathbf{0}$, with $x(\sigma_1) < x(\sigma_2) < \dots < x(\sigma_n)$, then there is a corresponding solution to $C(\sigma)\mathbf{y} = \mathbf{0}$ by positive numbers, where $y(i, j) = x(\sigma(j)) - x(\sigma(i))$. When \mathbf{x} is monochromatic under φ_p , \mathbf{y} is monochromatic under ψ_p .

Claim: some $C(\sigma)$ satisfies the columns condition.

Proof: If not, then Theorem 2.4.1 says each σ , gives a value $p_0(C(\sigma))$ so

that, for $p > p_0(C(\sigma))$ prime, $C(\sigma)$ has no monochromatic solutions under ψ_p .

Let $p_0 = \max_{\sigma \in S_n} \{p_0(C(\sigma))\}$. Take a prime $p > p_0$. Since $A\mathbf{x} = \mathbf{0}$ has a monochromatic solution under φ_p , there must be some $\sigma \in S_n$ so that $C(\sigma)$ has a monochromatic solution under ψ_p . As $p > p_0(C(\sigma))$, this is a contradiction.

Fix σ so that $C(\sigma)$ satisfies the columns condition. This means there are vectors $\mathbf{u}_1, \dots, \mathbf{u}_T$ in the nullspace of $C(\sigma)$, and decreasing graphs $R_1 \supseteq \dots \supseteq R_T = \emptyset$ satisfying:

1. If $i \in R_t$, then $u_t(i) = 0$.
2. If $i \notin R_t$, then there is an $s \leq t$ with $u_s(i) = 1$.
3. $R_T = \emptyset$.

For simplicity, we reorder the columns of A so that σ is the identity, and $C(\sigma)$ is the matrix C described originally.

This means there are vectors $\mathbf{w}_1, \dots, \mathbf{w}_T$ indexed by $\binom{[n]}{2}$, and sets $R_1 \supseteq \dots \supseteq R_T = \emptyset$ with vertex set $\binom{[n]}{2}$ satisfying conditions (1)-(3) of Definition 1.1.12.

Define a sequence of vectors $\mathbf{u}_1, \dots, \mathbf{u}_T$ on $[n]$ by $u_t(1) = 0$, and $u_t(i) = w_t(1, i)$ for $i > 1$. Additionally define $\mathbf{u}_0 = \mathbf{1}$ and $R_0 = [n]$. We just need $\{\mathbf{u}_t\}, \{R_t\}$ to satisfy requirements (1)-(4) of the graph columns condition.

It will be helpful to know that, for $k < \ell$,

$$u_t(\ell) - u_t(k) = w_t(k, \ell). \quad (2.3)$$

To see this, consider the $\{k, \ell\}$ row of the vectors \mathbf{b}_{ij} within C . Since $C\mathbf{w} = \mathbf{0}$, inspecting this row tells us that

$$w_t(1, \ell) - w_t(1, k) = w_t(k, \ell).$$

By definition of \mathbf{u}_t , we see that

$$u_t(\ell) - u_t(k) = w_t(k, \ell),$$

as desired.

Using this equation, properties (1)-(3) are immediate. Property (4) comes from the assumption that the columns of A sum to $\mathbf{0}$. \square

Corollary 2.4.7. *If the equation $A\mathbf{x} = \mathbf{b}$ is graph-regular, then A satisfies the weak graph columns condition, and $\mathbf{b} = \mathbf{0}$.*

Proof. From Lemma 2.3.3 and the note following it, we know that $\mathbf{b} = \mathbf{0}$ and the columns of A sum to $\mathbf{0}$. Since $A\mathbf{x} = \mathbf{0}$ is graph-regular, it must have a monochromatic solution under φ_p for every prime p . By Lemma 2.4.6, A satisfies the weak graph columns condition. \square

We end this section by considering the sufficiency of the WGCC.

Corollary 2.4.8. *If a matrix A satisfies the weak graph columns condition but is not graph-regular, then the offending coloring is not of the form $\chi(x < y) = f(y - x)$.*

To see this in action, consider the coloring φ_p and the matrix A from Example 2.4.4. Suppose that $1 < r < p^k - 1$. Consider the vector \mathbf{x} in the nullspace of A given by

$$\mathbf{x} = (p^{k+2} + 1)\mathbf{u}_1 + (p^{k+1} + p)\mathbf{u}_2 + p^2\mathbf{u}_3,$$

where $\{\mathbf{u}_i\}$ come from the analysis of this example earlier. It is easy to check that $x(1) > x(2) > \dots > x(8)$, and that φ_p colors all edges by “1”. Indeed, Lemma 2.4.6 actually shows that all colorings based on the difference of the endpoints will yield a monochromatic solution. Therefore, if the equation $A\mathbf{x} = \mathbf{0}$ is not graph-regular, it must be from some other type of coloring.

2.5 First step: two colors and two dimensions

The following result will be our goal through Section 2.7.

Theorem 2.5.1. *For all r, n , there is a number $N = N(r, n)$ so that any r -coloring of the edges of the complete graph on $[N]$ gives a Hilbert cube $H = H(a; d_1, \dots, d_n)$ so that all edges in H are the same color, and the 2^n elements of H are distinct.*

We first prove the theorem for $r = n = 2$. Note that a 2-dimensional Hilbert cube is four numbers of the form $a, a + b, a + c, a + b + c$. We will then extend those ideas to any number of colors, and then to Hilbert cubes of any dimension.

The proof will rely on the Gallai-Witt theorem [18] (Theorem 1.1.4), with help from a consequence of Rado's theorem [12] (Theorem 1.1.7):

Theorem 2.5.2 (Corollary to Rado). *There is a number T so that any 2-coloring of $[T]$ gives distinct numbers $i, j, i + j, j - i$, all the same color.*

Note: Rado's theorem gives conditions for a system of linear equations to have monochromatic solutions by distinct numbers. It is a simple exercise to check that the above satisfies them.

We now prove Theorem 2.5.1 for the case of $r = k = 2$.

Proof. Define $S = GW(T + 1, 2)$, where T comes from Theorem 2.5.2. We will show that $N = 2S$ suffices.

Fix an 2-coloring $\chi : \binom{[N]}{2} \rightarrow [2]$. We would like to find a solution to $w + x = y + z$ which forms a monochromatic clique. We view χ as a coloring of the upper half of the lattice $[N] \times [N]$ — for $x < y$, the color of (x, y) is $\chi(\{x, y\})$.

Consider the top left quadrant of our grid: $\{1, \dots, S\} \times \{S + 1, \dots, 2S\}$. Define $\chi' : [S] \times [S] \rightarrow [2]$ by

$$\chi'(a, b) = \chi(a, S + b).$$

Since $S = GW(T + 1, 2)$, and χ' is a 2-coloring of $[S] \times [S]$, we may apply Gallai-Witt to find x, y, d so that all points of the form

$$\{(x + id, y + jd) \mid i, j = 0, \dots, T\}$$

are the same color, say red, under χ . We will consider each subsquare of this large grid.

For now, consider a red square given by the points

$$(a, b) \quad (a + h, b) \quad (a, b + h) \quad (a + h, b + h).$$

We may rewrite the underlying numbers as $a, a + h, a + (b - a), a + h + (b - a)$ to see they form a Hilbert cube of dimension 2.

There are six edges in the graph on these four numbers, and we know that four of them are red. Thus, we only need to consider the edges $\{a, a + h\}$ and $\{b, b + h\}$. If these are both red (and the four values are distinct), then we have the desired monochromatic 4-clique. Thus, either we have our goal, or every red square gives us two points which cannot both be red.

Well, we have a great many red squares. Each has corner $(x + id, y + jd)$ and side-length ℓd , for every choice of i, j, ℓ with $i, j, i + \ell, j + \ell$ all in $\{0, \dots, S\}$. The four underlying numbers are all distinct by the choice of our initial grid $\{1, \dots, S\} \times \{S + 1, \dots, 2S\}$. The other two edges for this square are $\{x + id, x + (i + \ell)d\}$ and $\{y + jd, y + (j + \ell)d\}$, so these two cannot both be red without reaching our goal.

All of our red squares will give us many interacting conditions, which we record in a graph. Let $G = (A, B, E)$ be a bipartite graph, where $A = B = \binom{\{0, \dots, T\}}{2}$. We say $\{a, a'\} \sim \{b, b'\}$ if $\{x + ad, x + a'd\}$ and $\{y + bd, y + b'd\}$ are the last two edges for some red square. There is an induced 2-coloring of both A and B — namely

$$\chi_A(\{i, j\}) = \chi(x + id, x + jd),$$

$$\chi_B(\{i, j\}) = \chi(y + id, y + jd).$$

We see immediately that $\{i, i + \ell\} \sim \{j, j + \ell\}$ so long as those numbers are all in $\{0, \dots, T\}$. This means that each pair in A with difference ℓ is connected to every pair in B with that difference. This means that if one pair in A is red, all pairs in B with that difference must be blue (and vice versa). In fact, this is the entire structure of G .

Write $A = A_1 \cup A_2 \cup \dots \cup A_T$, where A_ℓ contains all pairs in A of the form $\{i, i + \ell\}$. We now 2-color $[T]$, the index set of the A_ℓ 's. Say $\phi(\ell) = \text{red}$ if *any* pair

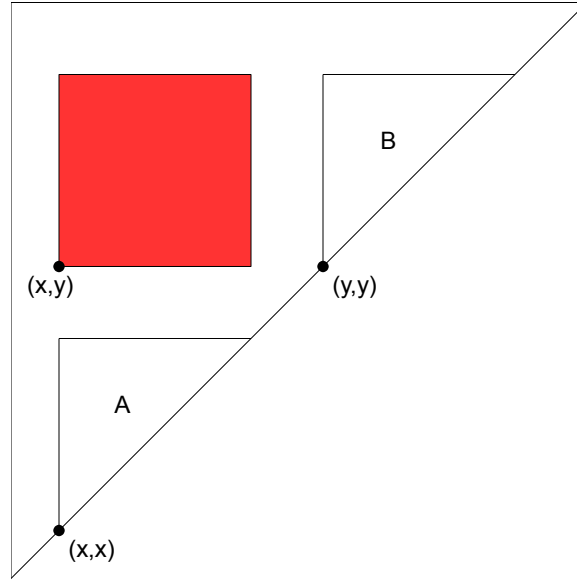


Figure 2.1: A large red grid, and the corresponding sets A and B .

in A_ℓ is red. Otherwise, $\phi(\ell) = \text{blue}$, meaning that A_ℓ is entirely blue. Since ϕ is a 2-coloring of $[T]$, Theorem 2.5.2 tells us there are distinct numbers $i, j, i + j, j - i$ which are monochromatic.

Case 1: The numbers are red. This means each set $A_i, A_j, A_{i+j}, A_{j-i}$ contains a red pair. Therefore the corresponding sets in B , what we should call $B_i, B_j, B_{i+j}, B_{j-i}$, are all entirely blue. The proof continues as in case 2 below, but with all A 's changed to B 's, and all x 's changed to y 's.

Case 2: The numbers are blue, so all pairs in $A_i, A_j, A_{i+j}, A_{j-i}$ are blue. We list the relevant blue pairs:

$$\begin{aligned} \text{In } A_i &: \{0, i\}, \{j, i + j\} \\ \text{In } A_j &: \{0, j\}, \{i, i + j\} \\ \text{In } A_{i+j} &: \{0, i + j\} \\ \text{In } A_{j-i} &: \{i, i + (j - i)\} = \{i, j\}. \end{aligned}$$

Taken together, we see that $0, i, j, i + j$ form a blue K_4 under χ_A . Recalling the relationship between χ and χ_A , this gives us a blue K_4 under χ with vertices $x, x + id, x + jd, x + (i + j)d$. This is the desired 2-dimensional Hilbert cube. \square

2.6 Coloring trees

In order to achieve Theorem 2.5.1 for any number of colors, we need to find more order in the type of hierarchical coloring system that we saw in the previous section. To achieve this order, we prove a Ramsey-type theorem for k -ary trees.

2.6.1 Trees and embeddings

Notation 2.6.1. We use $[k]^*$ denote all finite sequences (strings) of elements of $[k] = \{1, \dots, k\}$. If $s, t \in [k]^*$, we use $s \cdot t$ to denote concatenation — all characters of s followed by all characters of t . The length of a string $s \in [k]^n$ is $|s| = n$. Its characters may be written as $s = s_0 \cdot s_1 \cdots s_{n-1}$.

Definition 2.6.2. For $k \geq 2$, a perfect k -ary tree, $T_n^{(k)}$, of height n is the collection of nodes

$$T_n^{(k)} = \{s \in \{1, \dots, k\}^j \mid 0 \leq j \leq n\}.$$

We say λ , the empty string, is the root of the tree. The j^{th} level of $T_n^{(k)}$ consists of all those strings of length exactly j . The initial segments of s are called its *ancestors*.

Since we are only interested in perfect k -ary trees in this paper, we may occasionally refer to them simply as “ k -ary trees.” Since k will generally be fixed, we will often write $T_n^{(k)}$ as simply T_n .

Next, we define what it means to embed one k -ary tree into another.

Definition 2.6.3. Let T, R be two k -ary trees. A map $\varphi : T \rightarrow R$ is an *embedding* if it satisfies

$$\varphi(s \cdot i) = \varphi(s) \cdot f_\varphi(i, |s|),$$

where $f_\varphi(i, n) \in [k]^*$ is a string beginning with i , and $|f_\varphi(i, n)|$ is independent of i .

Lemma 2.6.4. *If $\varphi : T \rightarrow R$ is an embedding, then there is a function g_φ so that $\varphi(s \cdot t) = \varphi(s) \cdot g_\varphi(t, |s|)$, where t and $g_\varphi(t, n)$ agree on their first character, and $|t| = |t'| \Rightarrow |g_\varphi(t, n)| = |g_\varphi(t', n)|$.*

In particular, if φ is an embedding, and $|s| = |t|$, then $|\varphi(s)| = |\varphi(t)|$.

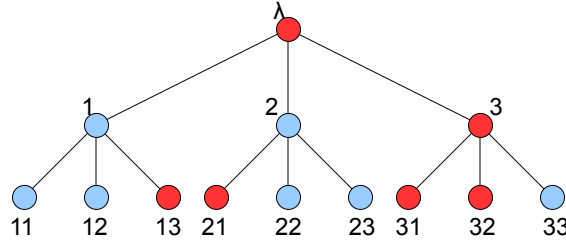


Figure 2.2: This 3-ary tree contains a monochromatic embedded T_1 — the nodes λ , 13, 21, and 31 are all red.

Proof. Let $s, t \in [k]^*$, with $t = t(1) \cdots t(m)$, and $|s| = n$.

$$\begin{aligned}
 \varphi(s \cdot t) &= \varphi(s \cdot t(1) \cdots t(m)) \\
 &= \varphi(s \cdot t(1) \cdots t(m-1)) \cdot f_\varphi(t(m), n+m-1) \\
 &\quad \vdots \\
 &= \varphi(s) \cdot f_\varphi(t(1), n) \cdots f_\varphi(t(m), n+m-1).
 \end{aligned}$$

We define $g_\varphi(t, n) = f_\varphi(t(1), n) \cdots f_\varphi(t(m), n+m-1)$. Since $g_\varphi(t, n)$ begins with $f_\varphi(t(1), n)$, its first character is $t(1)$, as needed. The length of $g_\varphi(t, n)$ is given by

$$\begin{aligned}
 |g_\varphi(t, n)| &= |f_\varphi(t(1), n)| + \dots + |f_\varphi(t(m), n+m-1)| \\
 &= \ell_n + \dots + \ell_{n+m-1},
 \end{aligned}$$

where ℓ_j represents the common lengths of $f_\varphi(i, j)$ over all $i \in [k]$. Thus we see that $|g_\varphi(t, n)|$ depends only on $|t|$ and n , not on the contents of t . \square

Example 2.6.5. A map $\varphi : T_1^{(k)} \rightarrow [k]^*$ is an embedding if there is some node $r \in [k]^*$, and strings $s_1, \dots, s_k \in \{1, \dots, k\}^j$ for some j , so that

$$\begin{aligned}
 \varphi(\lambda) &= r \\
 \varphi(i) &= r \cdot i \cdot s_i.
 \end{aligned}$$

An important property of embeddings is that they are composable.

Lemma 2.6.6. *If $\varphi : A \rightarrow B$ and $\psi : B \rightarrow C$ are embeddings, then the composition $\psi \circ \varphi : A \rightarrow C$ is also an embedding.*

Proof. Notice that

$$\begin{aligned}\psi(\varphi(s \cdot i)) &= \psi(\varphi(s) \cdot f_\varphi(i, |s|)) \\ &= \psi(\varphi(s)) \cdot g_\psi(f_\varphi(i, |s|), |\varphi(s)|).\end{aligned}$$

Thus, we write

$$f_{\psi \circ \varphi}(i, n) = g_\psi(f_\varphi(i, n), |\varphi(1^n)|).$$

Here 1^n is a stand-in for any string of length n , since $|\varphi(1^n)|$ only depends on n . Notice that, by Lemma 2.6.4, the first character of $g_\psi(f_\varphi(i, n), |\varphi(1^n)|)$ is the first character of $f_\varphi(i, n)$, which by definition begins with i . Moreover, its length is $|g_\psi(f_\varphi(i, n), |\varphi(1^n)|)|$ depends only on n and $|f_\varphi(i, n)|$, which in turn depends only on n .

Thus $\psi \circ \varphi$ meets the definition of of an embedding, as needed. \square

2.6.2 Monochromatic tree embeddings

Our main goal is to show that we may always find monochromatic tree embeddings.

We start by getting just the first (non-trivial) level:

Lemma 2.6.7. *If $r < \binom{k}{k-1}^n$, then every r -coloring of $[k]^n$ contains k points of the form $a \cdot 1 \cdot b_1, \dots, a \cdot k \cdot b_k$ which are all the same color.*

For $k = 2$, the proof amounts to the pigeonhole principle — as soon as there are more entries than colors, two must have the same color, and the strings a, b_1, b_2 may be determined. For general k , we must be more careful.

Proof. We work by induction on r . For $r = 1$, our bound only forces $n > 0$. Indeed, taking $n = 1$, the points $1, 2, \dots, k$ give us our goal (with $a = b_1 = \dots = b_k = \lambda$).

Now fix $r > 1$, and assume the result holds for any smaller number of colors. Also fix n so that $r < \binom{k}{k-1}^n$. We write $[k]^n = 1 \cdot [k]^{n-1} \cup \dots \cup k \cdot [k]^{n-1}$. If some color is contained in each of these subsets, then we may set $a = \lambda$ and find appropriate strings b_1, \dots, b_k to reach our goal. Otherwise, we may assume each color appears in at most $k - 1$ of these subsets. This means the average subset

contains at most a $\left(\frac{k-1}{k}\right)$ fraction of the colors. Thus we may find some specific i so that $i \cdot [k]^{n-1}$ uses r' colors, where $r' \leq \left(\frac{k-1}{k}\right) r$.

We observe that

$$r' \leq \left(\frac{k-1}{k}\right) r \leq \left(\frac{k-1}{k}\right) \left(\frac{k}{k-1}\right)^n \leq \left(\frac{k}{k-1}\right)^{n-1}.$$

Thus, by induction, there are strings a', b_1, \dots, b_k so that the points $i \cdot a' \cdot 1 \cdot b_1, \dots, i \cdot a' \cdot k \cdot b_k$ are all the same color. Taking $a = i \cdot a'$, we are done. \square

We now use this lemma to guarantee a “rainbow” tree embedding — an embedding so that each level of the embedded tree uses only one color.

Lemma 2.6.8. *For all k, r, n , there is an $N = N(k, r, n)$ so that for every r -coloring of the k -ary tree T_N , there is an embedding $\varphi : T_n \rightarrow T_N$ so that the image of each level $[k]^\ell$ uses only one color (though each level’s color may be different).*

Proof. We work by induction on n . For $n = 0$, the lemma is trivial.

Fix $n > 1$, and assume the result for $n - 1$. Fix an r -coloring χ of $[k]^*$. Let $M = N(k, r, n - 1)$. For each $s \in [k]^*$, we define a function $\tilde{\chi}(s) : T_M \rightarrow [r]$, given by $\tilde{\chi}(s)(t) = \chi(s \cdot t)$. Let $Y = |T_M| = k^M + k^{M-1} + \dots + k + 1$. We note that there are r^Y possible functions $\tilde{\chi}(s)$, so $\tilde{\chi}$ acts as an r^Y -coloring of $[k]^*$.

By Lemma 2.6.2, there is an N' and strings a, b_1, \dots, b_k so that $a \cdot 1 \cdot b_1, \dots, a \cdot k \cdot b_k \in [k]^{N'}$ are monochromatic under $\tilde{\chi}$. Call their common “color” χ' , so that for any i and any string $t \in T_M$, $\chi(a \cdot i \cdot b_i \cdot t) = \chi'(t)$.

By choice of M , there is an embedding $\psi : T_{n-1} \rightarrow T_M$ so that each level of the embedded tree contains only one color under χ' .

We may now define our desired embedding $\varphi : T_n \rightarrow [k]^*$ by $\varphi(\lambda) = a$, and $\varphi(i \cdot t) = a \cdot i \cdot b_i \cdot \psi(t)$ for $i \in [k]$.

We would like to show that each level of the image of φ is monochromatic. The zeroth level, a , consists of a single point, and so is trivially monochromatic. For $s = i \cdot t$ on level ℓ , we observe that

$$\chi(i \cdot t) = \chi(a \cdot i \cdot b_i \cdot \psi(t)) = \chi'(\psi(t))$$

is the color of level $\ell - 1$ of the image of ψ , as desired. \square

Now we have a rainbow tree. Of course, what we really want is a tree with all its points on every level the same color. Fortunately, we may now get that essentially for free.

Theorem 2.6.9. *For every k, r, n , there is a number $E = E(k, r, n)$ so that every r -coloring of the k -ary tree T_E yields a monochromatic embedding of T_n .*

We say that a coloring is n -balanced if the conclusion holds.

Proof. We take $E = E(k, r, n) = N(k, r, rn)$. Fix a coloring $\chi : T_E \rightarrow [r]$. By Lemma 2.6.8, there is an embedding $\psi : T_{rn} \rightarrow T_E$ so that, on each level of the image, all the points have the same color. That is, there are strings a_0 and $a_{i,\ell}$ for $i \in [k], \ell \leq rn$ so that

$$\psi(\sigma) = a_0 \cdot \sigma(1) \cdot a_{\sigma(1),1} \cdots \sigma(rn) \cdot a_{\sigma(rn),rn},$$

and, for each ℓ , $\chi(\psi(\sigma))$ is constant over all $\sigma \in [k]^\ell$.

Counting the zeroth level, there are $rn + 1$ levels in our embedded tree. Since there are only r colors, $n + 1$ of these levels must use the same color. Call these levels $\ell_0, \ell_1, \dots, \ell_n$, and call their color red.

Define an embedding $\varphi : T_n \rightarrow T_{rn}$ by

$$\varphi(\sigma) = 1^{\ell_0} \cdot \sigma(1)^{\ell_1 - \ell_0} \cdots \sigma(j)^{\ell_j - \ell_{j-1}},$$

where $j = |\sigma|$. To state it precisely in the language of embeddings, φ is defined by $a_0 = 1^{\ell_0}$, and $a_{i,j} = i^{\ell_j - \ell_{j-1} - 1}$. We see that $|a_{i,j}| = \ell_j - \ell_{j-1} - 1$ is independent of i .

We claim that $\psi \circ \varphi$ is the desired embedding. To check this, we observe that, if $|\sigma| = j$, then

$$\begin{aligned} |\varphi(\sigma)| &= |1^{\ell_0} \cdot \sigma(1)^{\ell_1 - \ell_0} \cdots \sigma(j)^{\ell_j - \ell_{j-1}}| \\ &= \ell_0 + (\ell_1 - \ell_0) + \cdots + (\ell_j - \ell_{j-1}) \\ &= \ell_j. \end{aligned}$$

Thus we see that $\varphi(\sigma) \in [k]^{\ell_j}$, so $(\psi \circ \varphi)(\sigma) = \psi(\varphi(\sigma))$ is red. Since this holds for all $j \leq n$ and all $\sigma \in [k]^{\ell_j}$, we see that our tree is entirely red. \square

2.7 The full result

In this section, we give the full proof of Theorem 2.5.1, first for any number of colors, but $n = 2$, and then for any n . As before, we view pairs of integers as ordered pairs (x, y) with $x < y$. When we have a grid $\{(x + id, y + jd)\}$ for a range of values i and j , we will say the grid is in position (x, y) with scale d .

2.7.1 Any colors, two dimensions

We now prove Theorem 2.5.1 for the case $n = 2$. In fact, this case captures all of the difficulty of the full theorem. We begin with this case in order to present the idea without being overburdened by notation.

Proof. We first give the arguments ignoring the numbers involved, and in the next section we determine a bound on $N(r, 2)$.

Begin with an r -coloring $\chi_0 = \chi$ of a large initial grid, G_λ . By Gallai-Witt, find a large monochromatic subgrid of color c_λ in position (x_0, y_0) with scale d_0 .

As in the proof with two colors, this yields two grids, G_1 and G_2 of equal size, in positions (x_0, x_0) and (y_0, y_0) respectively, both with scale d_0 . Note that these grids contain points on, above, and below the line $x = y$ — we only consider those points above the line. As in the proof in Section 2.5, if two points in these grids of the form $(x_0 + id, x_0 + jd)$ and $(y_0 + id, y_0 + jd)$ are both the same color as the grid G_λ , then we get our monochromatic Hilbert cube of dimension 2. The colorings of G_1 and G_2 correspond to χ_A and χ_B from the initial proof. We consider the coloring of a new grid, where the point (i, j) is colored by

$$\chi_1(i, j) = (\chi_0(x_0 + id, x_0 + jd), \chi_0(y_0 + id, y_0 + jd)).$$

We now use Gallai-Witt with r^2 colors, to find a large subgrid under χ_1 with color (c_1, c_2) in position (x_1, y_1) with scale d_1 . This grid actually corresponds to two grids: one of color c_1 in position $(x_0 + x_1d_0, x_0 + y_1d_0)$, and the other of color c_2 in position $(y_0 + x_1d_0, y_0 + y_1d_0)$. Both grids have scale d_0d_1 , and they are entirely contained in grids G_1 and G_2 respectively. If the grids G_λ, G_1 , and G_2 were all the same color, then we would have all of the points we need.

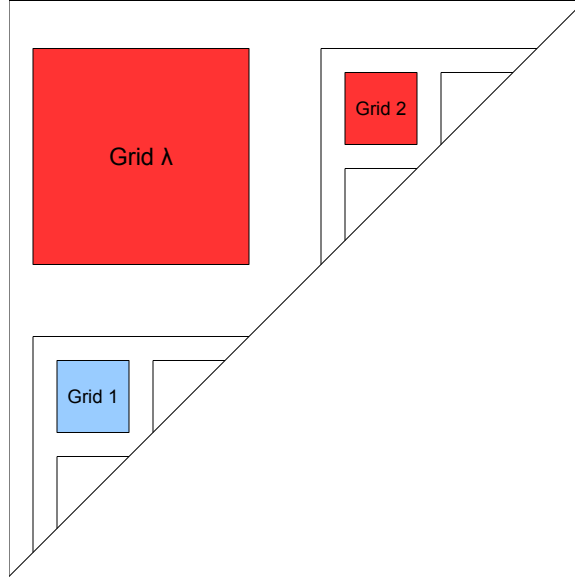


Figure 2.3: The sequence of subgrids.

Instead, we again pass to subgrids. The grid in G_1 yields two subgrids G_{11} and G_{12} , in positions $(x_0 + x_1d_0, x_0 + x_1d_0)$ and $(x_0 + y_1d_0, x_0 + y_1d_0)$ respectively, both with scale d_0d_1 . Likewise G_2 give us two subgrids, G_{21} , and G_{22} . Now we have more ways to win: the colorings of G_{11} and G_{12} restrict each other, in the way that A and B did in the proof for two colors. Similarly, G_{21} and G_{22} restrict each other, and both of G_{11} , G_{12} restrict both of G_{21} , G_{22} . Note that, whether the position of the grid involves x_0 or y_0 is determined by the first part of the subscript, and whether it involves x_1 or y_1 is dependent on the next part.

The next step, which we briefly state, is to define a grid-coloring χ_2 with r^4 colors corresponding to each of the four grids $G_{11}, G_{12}, G_{21}, G_{22}$. We find a subgrid of color $(c_{11}, c_{12}, c_{21}, c_{22})$ under this coloring, which corresponds to four grids, which further restrict one another.

Continue this process until the final grids are indexed by strings of length $E = E(2, r, 1)$, from Theorem 2.6.9. At the final step, the “large” monochromatic grid we find under χ_E need only be a 2×2 grid, giving G_s a single point for all s of length E .[‡] The color of this point is c_s .

[‡]If the 2×2 grid is in position (x, y) with scale d , the point corresponds to either $(x, x + d)$ or $(y, y + d)$.

We now recognize the map $s \mapsto c_s$ as an r -coloring of $T_E^{(2)}$. By Theorem 2.6.9, this coloring must contain a monochromatic embedded $T_1^{(2)}$. That is, there is an embedding φ so that the following points are the same color:

$$s := \varphi(\lambda) \quad u := \varphi(1) = s \cdot f_\varphi(1, k) \quad v := \varphi(2) = s \cdot f_\varphi(2, k),$$

where $s \in [2]^k$, $u, v \in [2]^\ell$ for some ℓ , and $f_\varphi(i, k)$ begins with i . Call their common color red.

Write $s = s_0 \cdot s_1 \cdots s_{k-1}$. Since s is red, the monochromatic grid found in grid G_s is red. Let

$$z_i(s) = \begin{cases} x_i & \text{if } s_i = 1 \\ y_i & \text{if } s_i = 2. \end{cases}$$

Then the grid G_s is in position $(X(s), Y(s))$, where

$$\begin{aligned} X(s) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{k-1}(s) + d_{k-1}x_k)\dots)) \\ Y(s) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{k-1}(s) + d_{k-1}y_k)\dots)). \end{aligned}$$

and has scale $D = d_0d_1 \cdots d_k$. Note that the only difference between X and Y is the x_k and y_k respectively in the inner-most term.

Now we look at the grids G_u and G_v , using only a single point from each. Define z_i , X , and Y in the same way as above for u and v . Note that

$$\begin{aligned} u_0 = s_0, \quad u_1 = s_1, \quad \dots \quad u_{k-1} = s_{k-1}, \quad u_k = 1 \\ v_0 = s_0, \quad v_1 = s_1, \quad \dots \quad v_{k-1} = s_{k-1}, \quad v_k = 2. \end{aligned}$$

Thus, we see that G_u is in position $(X(u), Y(u))$ with

$$\begin{aligned} X(u) &= X(s) + D(x_k + d_k(\dots(z_{k+\ell-1}(u) + d_{k+\ell-1}x_{k+\ell})\dots)) \\ Y(u) &= X(s) + D(x_k + d_k(\dots(z_{k+\ell-1}(u) + d_{k+\ell-1}y_{k+\ell})\dots)), \end{aligned}$$

and similarly G_t is in position $(X(t), Y(t))$ with

$$\begin{aligned} X(v) &= Y(s) + D(y_k + d_k(\dots(z_{k+\ell-1}(v) + d_{k+\ell-1}x_{k+\ell})\dots)) \\ Y(v) &= Y(s) + D(y_k + d_k(\dots(z_{k+\ell-1}(v) + d_{k+\ell-1}y_{k+\ell})\dots)). \end{aligned}$$

We claim that $X(u), X(v), Y(u), Y(v)$ form our Hilbert cube. Indeed, writing $a = X(u)$, $b = X(v) - X(u) = Y(v) - Y(u)$, and

$$c = Dd_k \cdots d_{k+\ell}(y_{k+\ell+1} - x_{k+\ell+1}),$$

we see that they have the form $a, a + b, a + c, a + b + c$ respectively.

Now consider the colors of the six points among these values (still only considering points with $x < y$). Since the points $(X(u), Y(u))$ and $(X(v), Y(v))$ are in G_u and G_v respectively, we know that both points are red.

Now we recognize that these values are given by

$$\begin{aligned} X(u) &= X(s) + iD, \\ Y(u) &= X(s) + jD, \\ X(v) &= Y(s) + iD, \\ Y(v) &= Y(s) + jD, \end{aligned}$$

so the four points we need look like

$$\begin{aligned} (X(u), X(v)) &= (X(s) + iD, Y(s) + iD) \\ (X(u), Y(v)) &= (X(s) + iD, Y(s) + jD) \\ (Y(u), X(v)) &= (X(s) + jD, Y(s) + iD) \\ (Y(u), Y(v)) &= (X(s) + jD, Y(s) + jD). \end{aligned}$$

By design, these fall nicely into the grid G_s , so these points are red as well. \square

2.7.2 Upper bounds

The process repeats to a depth of $E = E(2, r, 1)$, at which point we have 2^E grids, meaning r^{2^E} colors. At this level, we are looking for a square, so these grids must have size

$$S_E = 2.$$

At the prior level, our 2^{E-1} grids must have monochromatic subgrids of size S_E , and the joint coloring has $r^{2^{E-1}}$ colors. Thus

$$S_{E-1} = 2GW(S_E, r^{2^{E-1}}),$$

where the factor of 2 allows us to take the top-left quadrant of the grid. As before, this ensures distinct values in the x and y components. Repeating this reasoning, we find that

$$S_k = 2GW(S_{k+1}, r^{2^k}),$$

which leaves us with this bound for the size of the initial grid:

$$N(r, 2) \leq S_0 = 2GW(S_1, r).$$

2.7.3 Any colors, any dimensions

We now restate Theorem 2.5.1 and extend the previous proof to prove its full claim — finding a Hilbert cube of any dimension.

Theorem 2.5.1. *For all r, n , there is a number $N = N(r, n)$ so that for any r -coloring of the edges of the complete graph on $[N]$, there is a Hilbert cube $H = H(a; b_1, \dots, b_n)$ so that all edges within H are monochromatic.*

Proof. Let χ be an r -coloring of a large grid. Repeat the process from the proof in Section 2.7.1, only now continuing until we have a tree of height $E = E(2, r, n - 1)$.

As before, we label the grids as G_s for $s \in T_E^{(2)}$. Let the vectors $\mathbf{x}, \mathbf{y}, \mathbf{d}$ record the grid position information as before. That is, for $s = s_0 \cdot s_1 \cdots s_j$, and

$$z_i(s) = \begin{cases} x_i & \text{if } s_i = 1 \\ y_i & \text{if } s_i = 2, \end{cases}$$

we have

$$\begin{aligned} X(s) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{j-1}(s) + d_{j-1}x_j)\dots)) \\ Y(s) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{j-1}(s) + d_{j-1}y_j)\dots)). \end{aligned}$$

With this notation, Grid G_s is in position $(X(s), Y(s))$, and its scale is $d_0d_1 \cdots d_j$.

By Theorem 2.6.9, there is an embedding, φ , of $T_{n-1}^{(2)}$ which is entirely, say, red under χ . The nodes are labeled $\varphi(s)$ for $s \in [2]^j$ for $0 \leq j < n$.

For each $s \in [2]^{n-1}$, consider the red point $(X(\varphi(s)), Y(\varphi(s))) \in G_{\varphi(s)}$. We claim that the 2^n values

$$\{X(\varphi(s)) \mid s \in [2]^{n-1}\} \cup \{Y(\varphi(s)) \mid s \in [2]^{n-1}\}$$

are distinct, have the form $a + \sum_{i \in I} b_i$, and comprise an entirely red clique.

The first is easy. The values $X(s)$ and $Y(s)$ are all distinct by design. Each G_s uses disjoint x - and y -values, and as such each subgrid is disjoint from its parent grid and from those on its same level.

Next, we show that these values form a Hilbert cube. It follows immediately from the formulas for X and Y that

$$Y(s) - X(s) = d_0 d_1 \cdots d_{|s|-1} (y_{|s|} - x_{|s|}). \quad (2.4)$$

In particular, for $|s| = n - 1$, we define this value to be b_n .

Next, consider the value of $X(\varphi(s \cdot 2 \cdot t)) - X(\varphi(s \cdot 1 \cdot t))$. Their difference may be written as a sum of $|\varphi(s \cdot 1 \cdot t)| + 1 = |\varphi(s \cdot 2 \cdot t)| + 1$ differences — the i^{th} being

$$q_i = d_0 \cdots d_{i-1} (z_i(\varphi(s \cdot 2 \cdot t)) - z_i(\varphi(s \cdot 1 \cdot t))).$$

Let $j = |\varphi(s)|$ and $\ell = |\varphi(s \cdot 1)| = |\varphi(s \cdot 2)|$. Recall that, for $i \in \{1, 2\}$,

$$\varphi(s \cdot i \cdot t) = \varphi(s \cdot i) \cdot g_\varphi(t, j + 1) = \varphi(s) \cdot f_\varphi(i, j) \cdot g_\varphi(t, j + 1).$$

We see that $\varphi(s \cdot 1 \cdot t)$ and $\varphi(s \cdot 2 \cdot t)$ agree on their first j entries, as well as the entries from ℓ onward. Thus $q_0 = \dots = q_{j-1} = 0$, and $q_\ell = 0 = q_{\ell+1} = \dots = 0$ as well. What we are left with is

$$X(\varphi(s \cdot 2 \cdot t)) - X(\varphi(s \cdot 1 \cdot t)) = q_j + \dots + q_{\ell-1}.$$

For simplicity, write $u = f_\varphi(1, j)$ and $v = f_\varphi(2, j)$. Now, the remaining values q_{j+i} are determined by the values of u_i and v_i . If $u_i = v_i$, then $q_{j+i} = 0$. Otherwise, the value is $\pm d_0 \cdots d_{j+i-1} (y_{j+i} - x_{j+i})$ — “+” if $u_i = 1, v_i = 2$, and “−” if they are reversed. What q_{j+i} does *not* depend on, though, are s or t , other than $j = |s|$. Thus we may safely define b_1, \dots, b_{n-1} by

$$b_{j+1} = X(\varphi(s \cdot 2 \cdot t)) - X(\varphi(s \cdot 1 \cdot t)) \text{ for } s \in [2]^j, t \in [2]^*.$$

In fact, the same exact analysis for Y gives us the same differences:

$$b_{j+1} = Y(\varphi(s \cdot 2 \cdot t)) - Y(\varphi(s \cdot 1 \cdot t)) \text{ for } s \in [2]^j, t \in [2]^*.$$

It is now easily seen that

$$X(s) = X(1^{|s|}) + \sum_{\{i|s_i=2\}} b_i$$

and

$$Y(s) = Y(1^{|s|}) + \sum_{\{i|s_i=2\}} b_i = X(1^{|s|}) + b_k + \sum_{\{i|s_i=2\}} b_i.$$

Finally, writing $a = X(\varphi(1^{n-1}))$, our chosen values have the structure of a Hilbert cube, as needed. It remains to check that all the edges among these values are red.

Let $s \in [2]^{n-1}$. By virtue of $(X(\varphi(s)), Y(\varphi(s)))$ being a point in the grid $G_{\varphi(s)}$, we know that this edge is red. Now pick any $t \in [2]^{n-1}$ with $s < t$ lexicographically. Let σ be the longest initial substring that s and t agree on — their closest common ancestor. Since $s < t$ and $|s| = |t|$, we must have that $s = \sigma \cdot 1 \cdot u$ and $t = \sigma \cdot 2 \cdot v$ for some u and v of the same length.

We may now write

$$\begin{aligned} \varphi(s) &= \varphi(\sigma) \cdot f_\varphi(1, |\sigma|) \cdot g_\varphi(u, |\sigma| + 1) = \varphi(\sigma) \cdot 1 \cdot u' \\ \varphi(t) &= \varphi(\sigma) \cdot f_\varphi(2, |\sigma|) \cdot g_\varphi(v, |\sigma| + 1) = \varphi(\sigma) \cdot 2 \cdot v', \end{aligned}$$

where we use the crucial property that $f_\varphi(i, j)$ begins with i .

As in the previous proof, because $G_{\varphi(\sigma)}$ is red, we immediately learn that the four points

$$\begin{aligned} &(X(\varphi(s)), X(\varphi(t))) \\ &(X(\varphi(s)), Y(\varphi(t))) \\ &(Y(\varphi(s)), X(\varphi(t))) \\ &(Y(\varphi(s)), Y(\varphi(t))) \end{aligned}$$

are all red.

By considering all possible $s, t \in [2]^{n-1}$, this argument says that all edges among these values are red, so we have reached our goal. \square

Along the same lines as Section 2.7.2, we may define the recurrence

$$\begin{aligned} T_{E(2,r,n)} &= 2, \text{ and} \\ T_n &= 2GW(T_{n+1}, r^{2^n}), \end{aligned}$$

to get an upper bound of

$$N(r, n) \leq T_0 = 2GW(T_1, r).$$

2.8 Sufficient conditions for graph-regularity

We know that all graph-regular equations satisfy the weak graph columns condition. We now prove the strong graph columns condition guarantees graph-regularity. In order to do this, we first define a large, hierarchical parametrized grid.

Definition 2.8.1. Fix $p, q \in \mathbb{N}$, and $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^{n+1}$, $\mathbf{d} \in \mathbb{N}^{n+1}$. As usual, we write

$$z_i(s) = \begin{cases} x_i & \text{if } s_i = 1 \\ y_i & \text{if } s_i = 2. \end{cases}$$

For $s \in T_n$ and $i \in \mathbb{Q}$, we define

$$\begin{aligned} X(s, i) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{|s|-1}(s) + d_{|s|-1}(x_{|s|} + id_{|s|})) \dots)) \\ Y(s, i) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{|s|-1}(s) + d_{|s|-1}(y_{|s|} + id_{|s|})) \dots)) \\ Z(s, i) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{|s|-2}(s) + d_{|s|-2}(z_{|s|-1}(s) + id_{|s|-1})) \dots)). \end{aligned}$$

For convenience, we will often write

$$\begin{aligned} X(s) &:= X(s, 0) \\ Y(s) &:= Y(s, 0) \\ Z(s) &:= Z(s, 0). \end{aligned}$$

Note that $X(s, i) = Z(s \cdot 1, i)$, and $Y(s, i) = Z(s \cdot 2, i)$.

We say the hierarchical grid of depth n with parameters $p, q, \mathbf{x}, \mathbf{y}, \mathbf{d}$ is the collection of points

$$Grid_n(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d}) = \{(X(s, i), Y(s, j)) \mid s \in T_n, i, j \in \frac{1}{q}\mathbb{Z}, |i|, |j| \leq b(|s|)\},$$

where $b(k)$ is given by

$$b(k) = p \max\{|x_{k+1}|, |y_{k+1}|\} + b(k+1)d_{k+1} \quad (2.5)$$

(and $b(n+1) = x_{n+1} = y_{n+1} = 0$ for convenience).

We say that a set G is a $Grid_n(p, q)$, or just a $Grid_n$ if p and q are understood, if there are vectors $\mathbf{x}, \mathbf{y}, \mathbf{d}$ so that $G = Grid_n(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$.

We say that a $Grid_n$ is “proper” if none of its points $(X(s, i), Y(s, j))$ have $X(s, i) = Y(s, j)$. Equivalently, it is proper if, for all $k \leq n$ and $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i|, |j| \leq b(k)$, we have $x_k + id_k \neq y_k + jd_k$.

For convenience of notation, we will treat a proper $Grid_n$ as a graph, since it uniquely stores pairs $\{x, y\}$.

We make some observations about the structure of a $Grid_n$.

Lemma 2.8.2. *For $i \in \frac{1}{q}\mathbb{Z}$, then there are values*

$$\begin{aligned} Z(s, x_{|s|} + id_{|s|}) &= Z(s \cdot 1, i) = X(s, i) \\ Z(s, y_{|s|} + id_{|s|}) &= Z(s \cdot 2, i) = Y(s, i). \end{aligned}$$

Moreover, if $|i| \leq b(|s|)$, then $x_{|s|} + id_{|s|} \leq b(|s| - 1)$.

Proof. We show the first is true — the other is similar. Fix $s \in [2]^k$.

$$\begin{aligned} Z(s, x_k + id_k) &= z_0(s) + d_0(z_1(s) + d_1(\dots(z_{k-1}(s) + (x_k + id_k)d_{k-1})\dots)) \\ &= X(s, i) \\ &= Z(s \cdot 1, i), \end{aligned}$$

where

$$\begin{aligned} |x_k + id_k| &\leq |x_k| + |i|d_k \\ &\leq \max\{|x_k|, |y_k|\} + b(k)d_k \\ &= b(k - 1). \end{aligned}$$

□

Lemma 2.8.3. *Fix $G = Grid_n(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$, and let $s, t \in T_n$ be such that $s < t$, and there is some k so that $s_k \neq t_k$. Let $A = Z(s, i)$ and $B = Z(t, j)$, where $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i| \leq b(|s|), |j| \leq b(|t|)$. Then $(A, B) \in Grid_n(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$.*

Proof. Let σ be the closest common ancestor of s and t . Since $s \leq t$ and they disagree somewhere, we may write $s = \sigma \cdot 1 \cdot u, t = \sigma \cdot 2 \cdot v$. By repeated applications of the previous lemma, we may write

$$\begin{aligned} A &= Z(\sigma \cdot 1, i') = X(\sigma, i') \\ B &= Z(\sigma \cdot 2, j') = Y(\sigma, j'), \end{aligned}$$

where $|i'|, |j'| \leq b(|\sigma|)$. Thus $(A, B) = (X(\sigma, i'), Y(\sigma, j'))$ has exactly the needed form, so it is contained in G . □

2.8.1 A $Grid_n$ respects embeddings

We now show an essential fact — any tree embedding of T_n into T_m gives us a new $Grid_n$ within a $Grid_m$.

Lemma 2.8.4. *If G is a $Grid_m(p, q)$, and φ is an embedding of T_n into T_m , then*

$$G \circ \varphi := \{(X(\varphi(s), i), Y(\varphi(s), j)) \mid s \in T_n, i, j \in \frac{1}{q}\mathbb{Z}, |i|, |j| \leq b(|\varphi(s)|)\}$$

contains a $Grid_n(p, q)$. If G is proper, then so is the $Grid_n$.

Although this lemma is not very deep, we devote this section to carefully working through the technical details.

Proof. Let $G = Grid_m(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$, and φ be an embedding.

We are going to define vectors $\mathbf{x}', \mathbf{y}' \in \mathbb{Z}^{n+1}$, $\mathbf{d}' \in \mathbb{N}^{n+1}$ so that

$$Grid_n(p, q, \mathbf{x}', \mathbf{y}', \mathbf{d}') \subseteq G \circ \varphi.$$

In anticipation of this, we define $z'_k(s) \in \{x'_k, y'_k\}$ in the usual way, and the label the points in our $Grid_n$ as $(X'(s, i), Y'(s, j))$. We will bound the extent of its subgrids by $b'(k) = p \max\{|x'_{k+1}|, |y'_{k+1}|\} + b'(k+1)d'_{k+1}$.

Define $\ell(k) = |\varphi(s)|$ for $0 \leq k \leq n$ and $s \in [2]^k$.[§]

To start, define

$$\begin{aligned} x'_0 &= z_0(\varphi(1)) + d_0(\dots(z_{\ell(1)-1}(\varphi(1)))\dots) \\ y'_0 &= z_0(\varphi(2)) + d_0(\dots(z_{\ell(1)-1}(\varphi(2)))\dots) \\ d'_0 &= d_0 \cdots d_{\ell(1)-1} \\ x'_n &= x_{\ell(n)} \\ y'_n &= y_{\ell(n)} \\ d'_n &= d_{\ell(n)}. \end{aligned}$$

In between, for $0 < k < n$ and any $s \in [2]^k$, we define

$$\begin{aligned} x'_k &= z_{\ell(k)}(\varphi(s \cdot 1)) + d_{\ell(k)}(\dots(z_{\ell(k+1)-1}(\varphi(s \cdot 1)))\dots) \\ y'_k &= z_{\ell(k)}(\varphi(s \cdot 2)) + d_{\ell(k)}(\dots(z_{\ell(k+1)-1}(\varphi(s \cdot 2)))\dots) \\ d'_k &= d_{\ell(k)}d_{\ell(k)+1} \cdots d_{\ell(k+1)-1}. \end{aligned}$$

[§] $\ell(k)$ is well-defined by Lemma 2.6.4.

More simply, for $0 < k < n$, we may write

$$z'_k(s) = z_{\ell(k)}(\varphi(s)) + d_{\ell(k)}(\dots(z_{\ell(k+1)-1}(\varphi(s)))\dots).$$

Claim 2.8.5. *The definitions of x'_k, y'_k do not depend on choice of s .*

Proof. The question only arises for $0 < k < n$. the definition depends on $z_j(\varphi(s \cdot i))$ for $|\varphi(s)| \leq j < |\varphi(s \cdot i)|$ and $i = 1$ or 2 . Writing $\varphi(s \cdot i) = \varphi(s) \cdot f_\varphi(i, k)$, we see that this only refers to characters in $f_\varphi(i, k)$, so these values x'_k and y'_k do not depend on the choice of s . \square

Claim 2.8.6. *For all $s \in T_n$ there are values $i, j \in \mathbb{Z}$ so that $X'(s) = X(\varphi(s), i_s)$, and $Y'(s) = Y(\varphi(s), j_s)$.*

Proof. For $s \in [2]^n$, it is easy to see that $X'(s) = X(s), Y'(s) = Y(s)$.

Otherwise, we have

$$X'(s) = z_0(\varphi(s \cdot 1)) + d_0(\dots(z_{\ell(k)}(\varphi(s \cdot 1)) + d_{\ell(k)}(\dots(z_{\ell(k+1)-1}(\varphi(s \cdot 1)))\dots))\dots).$$

Notice that $z_{\ell(k)}(\varphi(s \cdot 1))$ looks at the first character of $f_\varphi(1, |s|)$, which is guaranteed to be a 1. Thus, this term becomes $x_{\ell(k)}$, giving us

$$X'(s) = X(\varphi(s)) + d_{\ell(k)}(z_{\ell(k)+1}(\varphi(s \cdot 1)) + \dots(z_{\ell(k+1)-1}(\varphi(s \cdot 1)))\dots).$$

Thus, writing

$$i_s = (z_{\ell(k)+1}(\varphi(s \cdot 1)) + \dots(z_{\ell(k+1)-1}(\varphi(s \cdot 1)))\dots),$$

we find $X'(s) = X(\varphi(s), i_s)$.

Similarly, if we let

$$j_s = (z_{\ell(k)+1}(\varphi(s \cdot 2)) + \dots(z_{\ell(k+1)-1}(\varphi(s \cdot 2)))\dots),$$

then we find $Y'(s) = Y(\varphi(s), j_s)$. \square

Claim 2.8.7. *For all $s \in T_n$,*

$$\{(X'(s, i), Y'(s, j)) \mid i, j \in \frac{1}{q}\mathbb{Z}, |i|, |j| \leq b'(|s|)\}$$

is completely contained in

$$\{(X(\varphi(s), i), Y(\varphi(s), j)) \mid i, j \in \frac{1}{q}\mathbb{Z}, |i|, |j| \leq b(\ell(|s|))\}.$$

Proof. For $|s| = n$, we have $b'(n) = 0 = b(\ell(n))$, and the two singleton sets are equal, so we will only look at $|s| = k < n$.

From the proof of Claim 2.8.6, we only need to verify that

$$b'(k) \leq b(\ell(k)) - |i_s|,$$

and similar for j_s . Because the work is identical, we will only work with i_s . It will suffice to show that for any $s \in [2]^k$.

Unraveling the definition of $b(\ell(k))$, and writing $w_k = \max\{|x_k|, |y_k|\}$, we get

$$b(\ell(k)) = pw_{\ell(k)+1} + d_{\ell(k)+1}(\dots(pw_{\ell(k+1)-1} + d_{\ell(k+1)-1}b(\ell(k+1) - 1))\dots).$$

Meanwhile, we have

$$\begin{aligned} |i_s| &= |(z_{\ell(k)+1}(\varphi(s \cdot 1)) + \dots(z_{\ell(k+1)-1}(\varphi(s \cdot 1)))\dots)| \\ &\leq w_{\ell(k)+1} + d_{\ell(k)+1}(\dots(w_{\ell(k+1)-1})\dots). \end{aligned}$$

Subtracting these, we see that $b(\ell(k)) - |i_s|$ is at least

$$(p-1)w_{\ell(k)+1} + d_{\ell(k)+1}(\dots((p-1)w_{\ell(k+1)-1} + d_{\ell(k+1)-1}b(\ell(k+1) - 1))\dots).$$

Ignoring the non-negative terms $(p-1)w_j$, and recognizing $d_j \geq 1$, we see

$$b(\ell(k)) - |i_s| \geq b(\ell(k+1) - 1).$$

It remains to show that, for $0 \leq k < n$, $b'(k) \leq b(\ell(k+1) - 1)$. Instead, for more reasonable indices, we will show that

$$b'(k-1) \leq b(\ell(k) - 1) \text{ for } 0 < k \leq n.$$

We will show this by reverse induction, starting at $k = n$ and working downward.

At $k = n$, we have equality:

$$\begin{aligned} b'(n-1) &= p \max\{|x'_n|, |y'_n|\} + b'(n)d'_n \\ &= p \max\{|x_{\ell(n)}|, |y_{\ell(n)}|\} + 0 \\ &= p \max\{|x_{\ell(n)}|, |y_{\ell(n)}|\} + b(\ell(n))d_{\ell(n)} \\ &= b(\ell(n) - 1). \end{aligned}$$

For smaller k , we have

$$b'(k-1) = pw'_k + b'(k)d'_k,$$

where $w'_k = \max\{|x'_k|, |y'_k|\}$ is bounded by

$$C = w_{\ell(k)} + d_{\ell(k)}(\dots(w_{\ell(k+1)-2} + d_{\ell(k+1)-1}w_{\ell(k+1)-1})\dots).$$

We may unravel $b(\ell(k) - 1)$ as

$$\begin{aligned} b(\ell(k) - 1) &\leq pw_{\ell(k)} + d_{\ell(k)}(\dots(pw_{\ell(k+1)-2} + d_{\ell(k+1)-1}b(\ell(k+1) - 1))\dots) \\ &= pC + b(\ell(k+1) - 1)d_{\ell(k)} \cdots d_{\ell(k+1)-1} = C + b(\ell(k+1) - 1)d'_k. \end{aligned}$$

Subtracting, we see that

$$\begin{aligned} (\ell(k) - 1) - b'(k-1) &\geq (pC + b(\ell(k+1) - 1)d'_k) - (pC + b'(k)d'_k) \\ &= d'_k(b(\ell(k+1) - 1) - b'(k)). \end{aligned}$$

By induction, this last expression is positive, as needed. \square

With this last claim, we have seen that each piece of $Grid_n(p, q, \mathbf{x}', \mathbf{y}', \mathbf{d}')$ is contained within $G \circ \phi$, which was the goal. Moreover, if G is proper, this last claim tells us that the $Grid_n$ must be, too.

2.8.2 Always a monochromatic $Grid_n$

We will modify the proof of Theorem 2.5.1 to get the following lemma.

Lemma 2.8.8. *Fix $p, q \in \mathbb{N}$. There is a number $Q = Q(r, n, p, q)$ so that, for every r -coloring of $[Q] \times [Q]$, there is a monochromatic proper $Grid_n(p, q)$.*

Proof. This proof borrows heavily from the several generations of proofs of Theorem 2.5.1. However, because it differs in some details, and this is the most general of these proofs, we present it in full detail.

Let $E = E(2, r, n)$ from Theorem 2.6.9. Let Q_0, \dots, Q_E and N_0, \dots, N_E be constants, given by this recursion:

$$\begin{aligned} N_k &= Q_{k+1}(p + N_{k+1}) \\ Q_k &= GW(r^{2^k}, N_k, q), \end{aligned}$$

terminating at $N_E = 0$. Note that we use the version of the Gallai-Witt theorem stated in Corollary 1.1.5.

As before, we will use strings $s \in [2]^*$ to index monochromatic square grids G_s .

We show that $Q = 2Q_0$ is a sufficient bound for $Q(r, n, p, q)$. Let χ be an r -coloring of $[Q] \times [Q]$. Consider the top-left quadrant of this lattice in which all x -values are less than all y -values. Q was chosen so that this region has size $Q_0 = GW(r, N_0, q)$. By the Gallai-Witt theorem, and using our usual notation, there are values x_0, y_0, d_0 so that the square grid of points $(X(\lambda, i), Y(\lambda, j))$ all have color $c(\lambda)$, for $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i|, |j| \leq N_0$.

Now, we are given k (initially 1) and values $\{x_j\}, \{y_j\}, \{d_j\}$ for $0 \leq j < k$. We know that, for $s \in T_{k-1}$, and for each choice of $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i|, |j| \leq N_{|s|}$, the point $(X(s, i), Y(s, j))$ has color $c(s)$. We will show how to find x_k, y_k, d_k so that the same holds for all $s \in [2]^k$ — and hence for all $s \in T_k = T_{k-1} \cup [2]^k$.

We define a r^{2^k} coloring χ_k of $(-[Q_k]) \times [Q_k]$.[¶] Our colors are functions $\chi_k(i, j) : [2]^k \rightarrow [r]$ given by

$$\chi_k(i, j)(s) = \chi(Z(s, i), Z(s, j)).$$

Since $Q_k = GW(r^{2^k}, N_k, q)$, we apply Gallai-Witt to find some x_k, y_k, d_k so that the grid $\{(x_k + id_k, y_k + jd_k)\}$ is monochromatic under χ_k for $i, j \in \frac{1}{q}\mathbb{Z}$ with $|i|, |j| \leq N_k$. Each of these grid points assigns the same color to each $s \in [2]^k$ — thus we may say that all points $(Z(s, x_k + id_k), Z(s, y_k + jd_k))$ have color $c(s)$. Using Lemma 2.8.2, we may concisely write this by saying

$$(X(s, i), Y(s, j)) \text{ has color } c(s), \text{ for } s \in [2]^k, i, j \in \frac{1}{q}\mathbb{Z} \text{ with } |i|, |j| \leq N_k.$$

Noting that all values $x_k + id_k$ are negative while all $y_k + jd_k$ are positive, we see that $X(s, i) \neq Y(s, j)$ for such i, j .

After continuing this procedure for E steps, we now have many grids, indexed over $s \in T_E$, each of whose points is monochromatic, with color $c(s)$. We

[¶]We use $-[N]$ to represent the set $\{-1, -2, \dots, -N\}$.

claim these points contain $Grid_E(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$. If so, it must be proper, since we saw $X(s, i) \neq Y(s, j)$. Because the needed structure is already present, we must only check that $N_k \geq b(k)$. We will show this by a reverse induction, starting at $N_E = 0 = b(E)$ and working downward.

Suppose that $N_k \geq b(k)$. Recall that $N_k = Q_{k+1}(p + N_{k+1})$ and $b(k) = p \max\{|x_{k+1}|, |y_{k+1}|\} + b(k+1)d_{k+1}$. By construction, for $k > 0$, we have $|x_k|, |y_k|, d_k$ all bounded by Q_k , since these describe the position and scale of a grid contained in $(-[Q_k]) \times [Q_k]$. Thus we may write

$$\begin{aligned} b(k-1) &= p \max\{|x_k|, |y_k|\} + b(k)d_k \\ &\leq pQ_k + b(k)Q_k \\ &\leq pQ_k + N_k Q_k \\ &= N_{k-1}, \end{aligned}$$

as desired.

Thus our points do contain $G = Grid_E(p, q, \mathbf{x}, \mathbf{y}, \mathbf{d})$, which is proper. Now, viewing c as an r -coloring of T_E , we apply Theorem 2.6.9 to find an embedded T_n whose image is entirely, say, red. Call the embedding φ , so the nodes are labeled $\varphi(s)$ for $s \in T_n$. By Lemma 2.8.4, $G \circ \varphi$ contains our monochromatic proper $Grid_n(p, q)$. \square

2.8.3 A $Grid_n$ is enough

Since we know every finite-coloring of $[Q] \times [Q]$ contains a large monochromatic proper $Grid_n$ (for Q sufficiently large), we only need to show the following.

Lemma 2.8.9. *Let A satisfy the strong graph columns condition. Then there is some $n, p, q \in \mathbb{N}$ so that the following holds. If G is a proper $Grid_n(p, q)$, then there is a solution to $A\mathbf{w} = \mathbf{0}$ so that, for all i, j , $w(i) \neq w(j)$ and the edge $\{w(i), w(j)\}$ is in G .*

In particular, if A satisfies the strong graph columns condition in T steps with vectors $\mathbf{u}_0, \dots, \mathbf{u}_T$ with entries in $\frac{1}{q}\mathbb{Z}$, then we may take $n = T - 1, p = 1 + 2 \max |u_t(i)|$, and use the given q .

Proof. Let A satisfy the strong graph columns condition, and n, p, q be as suggested. Let $R_0 \supseteq \dots \supseteq R_{n+1} = \emptyset$ and vectors $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ demonstrate the SGCC.

Fix $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^{n+1}, \mathbf{d} \in \mathbb{N}^{n+1}$. For $0 \leq k \leq n$, define the vector $\{\mathbf{v}_k\}$ by

$$\mathbf{v}_k = x_k \mathbf{1} + (y_k - x_k) \mathbf{u}_{k+1},$$

which takes the value x_k when $u_{k+1}(i) = 0$, and y_k when $u_{k+1}(i) = 1$.

Additionally, define

$$\mathbf{w} = \mathbf{v}_0 + d_0(\mathbf{v}_1 + d_1(\dots(\mathbf{v}_{n-1} + d_{n-1}\mathbf{v}_n)\dots)).$$

Note that the vectors \mathbf{v}_k and \mathbf{k} are all in the nullspace of A , as linear combinations of vectors \mathbf{u}_k (including $\mathbf{u}_0 = \mathbf{1}$). We will refer to the entries of \mathbf{w} as our vertices, since we will show that the complete graph on these values is contained within our $Grid_n$.

For a fixed a , consider the sequence $u_0(a), u_1(a), \dots, u_{n+1}(a)$. The SGCC requires that this sequence begins with a sequence of 0s and 1s, and is then allowed to take any value. This means the sequence $v_0(a), v_1(a), \dots, v_n(a)$ will initially have $v_k(a) \in \{x_k, y_k\}$ through some k^* (depending on a). Let

$$i = v_{k^*+1}(a) + d_{k^*+1}(v_{k^*+2}(a) + d_{k^*+2}(\dots(v_{n-1}(a) + d_{n-1}v_n(a))\dots)).$$

Looking at \mathbf{w} , there is some $s \in [2]^{k^*+1}$ so that

$$w(a) = Z(s, i).^{\parallel}$$

Note that, as a \mathbb{Z} -linear combination of values $z_k(a)$, we have $i \in \frac{1}{q}\mathbb{Z}$. Also observe that

$$\begin{aligned} |v_k(a)| &= |x_k + (y_k - x_k)u_{k+1}(a)| \\ &\leq |x_k| + (|y_k| + |x_k|)|u_{k+1}(a)| \\ &\leq \max\{|x_k|, |y_k|\}(1 + 2 \max\{|u_{k+1}(a)|\}) \\ &= p \max\{|x_k|, |y_k|\}. \end{aligned}$$

[∥]In fact, s is given by $s_k = u_{k+1}(a) + 1$.

Applying this bound recursively, it is immediate that

$$|i| \leq b(k^*) = p \max\{|x_k|, |y_k|\} + b(k^* + 1)d_{k^*+1}.$$

Finally, we would like to show that, for any two of our vertices $w(a)$ and $w(b)$, the edge $\{w(a), w(b)\}$ is in our $Grid_n$. Write these vertices as $w(a) = Z(s, i)$ and $w(b) = Z(t, j)$. The SGCC tells us that, between any two of our vertices, their edge at some time changes from restricted to unrestricted. This gives us a first time k so that, without loss of generality, $u_k(a) = 0$ and $u_k(b) = 1$ (or vice versa). Moreover, for each $k' < k$, we know that either $u_{k'}(a) = u_{k'}(b) = 0$ or $u_{k'}(a) = u_{k'}(b) = 1$.

This means that the strings s and t must disagree at some position. We may thus apply Lemma 2.8.3 to confirm that the edge is in our $Grid_n$. Moreover, since G is proper, it also shows that $w(a) \neq w(b)$ — the vertices are distinct. Since this observation holds for every edge, we now have our complete graph on distinct vertices, so we are done. \square

Corollary 2.8.10. *Let A satisfy the strong graph columns condition. Then $A\mathbf{x} = \mathbf{0}$ is graph-regular.*

Proof. Let n, p, q be the values given from Lemma 2.8.9, and let $r \in \mathbb{N}$ be the number of colors. We claim that, if $Q \geq Q(r, n, p, q)$ from Lemma 2.8.8, then any r -coloring of $\binom{[Q]}{2}$ will contain a solution to $A\mathbf{x} = \mathbf{0}$ so that the values $\{x(i)\}$ are distinct, and the edges $\{x(i), x(j)\}$ are monochromatic.

Indeed, by Lemma 2.8.8, viewing χ as an r -coloring of $[Q] \times [Q]$, we find a monochromatic proper $Grid_n(p, q)$. By Lemma 2.8.9, this $Grid_n$ contains a solution to $A\mathbf{x} = \mathbf{0}$ with distinct values x_k as desired. \square

2.9 Hypergraph-regular equations

There is a natural extension of graph-regularity to the hypergraph Ramsey theorem.

Unfortunately, this extension is not fruitful. Say a homogeneous linear equation is “ r -graph-regular” if, for every coloring of the r -sets of \mathbb{N} , it has a monochromatic solution by distinct numbers. As with graphs, when considering an r -uniform hypergraph, we require the equations to have at least $r + 1$ variables, or else every solution will be trivially monochromatic.

Theorem 2.9.1. *For $r \geq 3$, no homogeneous linear equation of at least $r + 1$ variables is r -graph-regular for r -uniform hypergraphs.*

Proof. We show the result for $r = 3$, and suggest the appropriate modifications for higher r .

Fix an equation $\sum a_i x_i = b$ in at least $r + 1$ variables $\{x_i\}$. Assume each a_i is nonzero, since discarding trivial variables only makes it easier to be graph-regular.

We first show that $\sum a_i = 0 = b$.

For any n , define an $(n + 1)$ -coloring $f_n^{(3)}$ of $\binom{\mathbb{N}}{r}$ by

$$f_n^{(3)}(an + x, bn + y, cn + z) = \begin{cases} \text{blue} & \text{if } x = y = z \\ \min\{x, y, z\} & \text{if one of } x, y, z \text{ is smallest} \\ \max\{x, y, z\} & \text{otherwise,} \end{cases}$$

where $x, y, z \in \{0, 1, \dots, n - 1\}$. Similar to before, any set of four elements which is monochromatic under this coloring must be blue.

Now define $g_n^{(3)}$ on $\binom{\mathbb{N}}{r}$ by

$$g_n^{(3)}(n^i a, n^j b, n^k c) = \begin{cases} f_{n-1}^{(3)}(a, b, c) & \text{if } i = j = k \\ \text{red} & \text{if one of } i, j, k \text{ is smallest} \\ \text{green} & \text{otherwise,} \end{cases}$$

where a, b, c are not divisible by n . Again, similar to before, any monochromatic clique under this coloring on at least four points must be red or blue. The proof of Lemma 2.3.3 may be slightly reworked to show that the coefficients of a hypergraph-regular equation must sum to zero, and $b = 0$.

Therefore, we only consider $\sum_{i=1}^k a_i x_i = 0$ where $\sum a_i = 0$.

Define a new coloring, $h_p(x, y, z) = g_p(y - x, z - x)$, where $x < y < z$, and g_p is the graph-coloring used in Section 2.3.

Suppose x_1, \dots, x_k are distinct values satisfying $\sum a_i x_i = 0$, with the hyperedges among them monochromatic — either red or blue. Let x_j be the smallest of these values. Since $a_j = -\sum_{i \neq j} a_i$, we see that

$$\sum_{i \neq j} a_i (x_i - x_j) = 0.$$

By choice of x_j , we see that $\{x_i - x_j\}_{i \neq j}$ is monochromatic under g_p . As before, a red clique means some a_i is 0. If the clique is blue, then $\sum_{i \neq j} a_i = 0$, meaning $a_j = 0$. Since none of the coefficients are 0, we have reached a contradiction. Thus no homogeneous linear equation in at least 4 variables is hypergraph-regular under colorings of 3-sets.

For a general r -uniform hypergraph with $r > 3$, one can easily modify the definition of $g_p^{(3)}$ to find a suitable $g_p^{(r)}$, which will force coefficients to add to zero. Likewise, one may define a coloring similar to h_p which is built upon $g_p^{(r-1)}$, which will force one of the coefficients to be zero. These two colorings together will avoid solutions to any equation in at least $r + 1$ variables. \square

Evidently, the ability to color 3-sets (or higher) of integers is too strong to guarantee monochromatic solutions to linear equations.

2.10 Further work

We have only begun the research into graph-regular equations — there are many interesting problems still to consider. We list several that appeal to us.

First and foremost, we would like to see the gap closed in our characterization of graph-regular equations. A good start would be to determine whether the equation separating the strong and weak graph columns conditions is graph-regular.

Open Problem 2.10.1. Determine whether the equation from Example 2.4.4 is graph-regular.

What happens when we relax the search for a monochromatic complete graph? For a given graph G , consider equations $A\mathbf{x} = \mathbf{b}$ where each entry of \mathbf{x} corresponds to a vertex of G .

Open Problem 2.10.2. Does every finite coloring of pairs of natural numbers contain a monochromatic copy of G so that the corresponding vector \mathbf{x} solves the equation?

Although this problem seems quite complex to fully answer, we do have some insight into it. Suppose A partially satisfies the strong graph columns condition, in that there is a sequence of vectors $\mathbf{1} = \mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_T$ and restriction graphs $K_n = R_0, R_1, \dots, R_T$ satisfying conditions 1* and 2*, but R_T is not empty. The proof of Lemma 2.8.9 can easily be modified to guarantee the complement of R_T is contained in every sufficiently large proper $Grid_{T-1}$, whose vertices solve $A\mathbf{x} = \mathbf{0}$.

Inspired by results like Hindman's theorem (Theorem 1.1.9), we wonder:

Open Problem 2.10.3. Is there some non-trivial infinite system of equations which is graph-regular?

Our initial attempt — an monochromatic infinite-dimensional Hilbert cube — can be avoided by a simple coloring: for $x < y$,

$$\chi(x, y) = \begin{cases} \text{red} & y < 2x \\ \text{blue} & y \geq 2x. \end{cases}$$

It is easy to see that (1) any infinite monochromatic clique must be blue, and (2) the vertices of such a clique must grow exponentially, which is impossible in a Hilbert cube.

Chapter 2 consists primarily of previously published material. The majority of content in Sections 2.2, 2.5, 2.6, and 2.7 were first published in the Journal of Combinatorics in volume 2, no. 4 (2011), published by International Press in the article “An additive version of Ramsey’s theorem” by Parrish, A. The majority of content in Sections 2.3, 2.4, 2.8, and 2.9 were first published in the Electronic Journal of Combinatorics, volume 20, no. 1 (2013) in the article “Toward a graph version of Rado’s theorem” by Parrish, A. Both papers were authored solely by the author of this dissertation.

Chapter 3

On adapting a result from graphs to arithmetic progressions

The logicians Jeff Paris and Leo Harrington, while searching for a natural result outside of Peano arithmetic, found a strengthening of Ramsey's theorem [11]. Their main result was that this combinatorial result necessarily requires the use of the infinite version of Ramsey's theorem. Since van der Waerden's theorem has no infinite analog, Bill Gasarch asked* whether it could allow the same strengthening. We show that the direct analog is false, though a similar extension could still be possible.

3.1 Large Ramsey

For any set X , let K_X denote the complete graph on the vertex set X . Consider a set $A \subseteq \mathbb{N}$ to be *large* if $|A|$ is at least as large as the smallest element of A . Paris and Harrington proved what sometimes called the Large Ramsey Theorem:

Theorem 3.1.1. *For all $m \in \mathbb{N}$, there is some $N = N(m)$ so that for every 2-coloring of the edges of $K_{\{m, \dots, N\}}$, there is a large set $A \subseteq \mathbb{N}$ so that the edges of K_A are monochromatic.*

*Personal communication.

Proof. Ramsey’s theorem tells us that any 2-coloring of the edges of $K_{\{m,m+1,\dots\}}$ yields an infinite complete monochromatic subgraph on the vertices $A \subseteq \mathbb{N}$. As an infinite set, A is clearly large.

By the compactness principle (Corollary 1.3.4), there must also be some finite bound $N = N(m)$ so that we only need to use vertices up to N . \square

The key here is that the finite case follows from the infinite. This is no accident: the more interesting result of Paris and Harrington — known as the Paris-Harrington theorem — states that, in fact, the Large Ramsey theorem is independent from Peano arithmetic, so only an infinitary proof like the above will suffice.

3.2 Large van der Waerden

Van der Waerden’s theorem tells us that, for every $k \in \mathbb{N}$, there is an N so that any 2-coloring of $[N]$ yields a monochromatic k -term arithmetic progression (a k -AP). Unlike Ramsey’s theorem, this fails in the infinite case — it is easy to 2-color \mathbb{N} without an infinite arithmetic progression. Does a “large” form of van der Waerden’s theorem hold? The Paris-Harrington result seems to suggest it will not. Indeed, the immediate analogue of the Large Ramsey theorem fails.

Theorem 3.2.1. *The numbers $\{3, 4, 5, \dots\}$ may be 2-colored without a large monochromatic arithmetic progression.*

Note: “Large” van der Waerden holds trivially for $m = 1$ or 2 , so this shows the first nontrivial case fails.

Proof. First we present the coloring. We color by solid blocks, of length 2, 2, 4, 8, 16, etc, to look like this (using 0 and 1 as our colors):

Color:	11001111000000001111111111111110...
Position:	345 7 11 19 35.

Explicitly, we assign the color 1 to 3 and 4. From then on, for each $n \geq 0$, we have a block of length 2^{2n+1} of color 0 beginning at $3 + 2^{2n+1}$, and a block of length 2^{2n+2} of color 1 beginning at $3 + 2^{2n+2}$. This coloring has one key property:

$$\text{The length of each block is the total number of points preceding it.} \quad (3.1)$$

Now that we have a coloring, suppose that $A = \{a, a + d, \dots, a + (k - 1)d\}$ is a monochromatic k -AP. That this AP is not large will be seen from the following facts.

Fact 1. If $a = 3$, then $k \leq 2$.

Proof. The first three terms would be $3, 3 + d, 3 + 2d$, all colored 1. For $3 + d$ to be in a block of 1s, we must have some n such that $3 + 2^{2n} \leq 3 + d < 3 + 2^{2n+1}$. This lets us bound $3 + 2d$ by doubling all parts and subtracting 3. This says $3 + 2 \cdot 2^{2n+1} \leq 3 + 2d < 3 + 2^{2n+2}$, which says the third entry is in a block of 0s, so it is not a monochromatic AP.

Fact 2. If $k \geq a$ (that is, if A is large), then A is not contained in a single block.

Proof. The block beginning at 3 is too short by 2. All the rest are too short by 3.

Fact 3. $\{a + d, \dots, a + (k - 1)d\}$ are all in the same block.

Proof. Pick any two terms a and $a + d$ of the same color to begin the AP. This fixes the common difference as d . By property (1), d is smaller than the length of the next block, which has a different color from a and $a + d$. This means that the next term, and indeed all remaining terms must be in the same block as $a + d$, as there is no way to “jump over” the next block.

Fact 4. If a and $a + d$ are in different blocks, then $k \leq 3$.

Proof. The length of the block before $a + d$ is half the length of the one holding $a + d$, so d is at least this large. This means at most one more term will fit in this block. Fact 3 shows that this ends the AP.

Fact 1 says a large monochromatic AP could not begin at 3. Facts 2 and 3 tell us any such AP must have a in one block, and $a + d, \dots, a + (k - 1)d$ all in

one other. Fact 4 tells us a monochromatic AP of this form must be of length 3, so it is not large after all. \square

Although this analogue of the Large Ramsey theorem fails, there may still be something salvageable. Consider this generalization of what it means to be “large”.

Definition 3.2.2. For an increasing function f , call a set A f -large if $|A|$ is at least as large as $f(a)$, where $a \in A$ is the smallest element.

However, a slight weakening will not do.

Theorem 3.2.3. Fix k , and let $f(x) = \frac{x}{k}$. There is a number $m = m(k)$ and a 2-coloring of $\{m, m + 1, m + 2, \dots\}$ so that there are no f -large monochromatic arithmetic progressions.

Proof. We only provide a sketch, with a coloring based on the proof of Theorem 3.2.1. We make no claims that our choice of m is the best possible.

For simplicity, let $k = 2^r$. Let $m = 3 + 2^r$. The coloring from Theorem 3.2.1 from this point consists of blocks of alternating colors, with lengths $2^r, 2^{r+1}, \dots$. Cut each block of length 2^n into 2^r blocks each of length 2^{n-r} , alternating colors. A similar analysis to before will show that this colorings does the trick. \square

Is there an unbounded function f so that, for all m , for every 2-coloring of $\{m, m + 1, m + 2, \dots\}$, there is a monochromatic arithmetic progression which is f -large? This is not the right question, either. The answer is trivially yes: we may select f based on the inverse van der Waerden function. Specifically, let $g(n)$ be the length of the longest arithmetic progression guaranteed under all 2-colorings of $[n]$.

Lemma 3.2.4. Let $f(n) = g(\lfloor n/2 \rfloor)$. Then, for all m , for every 2-coloring of $\{m, m + 1, m + 2, \dots\}$, there is a monochromatic f -large arithmetic progression.

Proof. Fix a 2-coloring χ , and restrict it to $\{m, m + 1, \dots, 2m\}$. By definition of g , there is a monochromatic arithmetic progression A of length $g(m + 1)$, beginning at some $a \leq 2m$. Observing that

$$f(a) \leq f(2m) = g(m) \leq g(m + 1) = |A|,$$

we see that A is f -large. □

This fails to shine any new light onto the nature of monochromatic arithmetic progressions — we are simply converting information about where the progression must end into a bound on where it must start. We are now left with an ill-posed question: is there any “surprising” function f so that 2-colorings always yield f -large arithmetic progressions?

Bibliography

- [1] N.G. de Bruijn, P. Erdős, A colour problem for infinite graphs and a problem in the theory of relations, *Indagationes Math.* 13 (1951), 369-373.
- [2] W. Deuber, D. S. Gunderson, N. Hindman, D. Strauss, Independent finite sums for K_m -free graphs, *J. Combin. Theory Ser. A* 78 (1997), no. 2, 171-198.
- [3] W. Gasarch, C. Kruskal, A. Parrish, *Van der Waerden's theorem and its variants*, to appear.
- [4] R.L. Graham, B.L. Rothschild, J. Spencer, *Ramsey Theory*, John Wiley & Sons Inc., New York, second edition 1990.
- [5] D. S. Gunderson, I. Leader, H. J. Prömel, and V. Rödl, Independent arithmetic progressions in clique-free graphs on the natural numbers, *J. Combin. Th. Ser. A* 93 (2001), 1-17.
- [6] D. S. Gunderson, I. Leader, H. J. Prömel, and V. Rödl, Independent Deuber sets in graphs on the natural numbers, *J. Combin. Th. Ser. A* 103 (2003), 305-322.
- [7] D. Hilbert, Über die Irreduzibilität ganzer rationaler Funktionen mit Ganzzahligen Koeffizienten, *J. Reine Angew. Math.* 110 (1892), 104-129.
- [8] N. Hindman, Finite Sums from Sequences within Cells of a Partition of \mathbb{N} , *J. Combin. Theory Ser. A* 17 (1974), 1-11.
- [9] A. V. Kostochka, M. Stiebitz, B. Wirth, The colour theorems of Brooks and Gallai extended, *Discrete Mathematics*, Volume 162, Issues 1-3, 25 December 1996, 299-303.
- [10] T. Łuczak, V. Rödl, T. Schoen, Independent finite sums in graphs defined on the natural numbers, *Discrete Math.* 181 (1998), 289-294.
- [11] J. Paris and L. Harrington, A Mathematical Incompleteness in Peano Arithmetic, *Handbook of Mathematical Logic*, v. 90 (1977) 1133-1142.
- [12] R. Rado, Studien zur Kombinatorik, *Math. Zeit.* 36 (1933), 424-480.

- [13] R. Rado, Note on Combinatorial Analysis, Proc. Lond. Math. Soc. 48(2) (1943) 122-160.
- [14] F. P. Ramsey, On a problem in formal logic, Proc. London Math. Soc. (2), 30 (1930), 264-286.
- [15] I. Schur, Über die Kongruenz $x^m + y^m = z^m \pmod{p}$, Jber. Deutsch. Math. Verein 25 (1916), 114-116.
- [16] A. Soifer, Ramsey Theory before Ramsey, Prehistory and Early History, in *Ramsey Theory: Yesterday, Today and Tomorrow*, edited by Alexander Soifer, Progress in Mathematics, vol. 285, Springer-Birkhuser, 2010, 1-26.
- [17] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, Nieuw Arch. Wiskunde 15 (1927), 212-216.
- [18] E. Witt, Ein kombinatorischer Satz der Elementargeometrie, Math. Nachrichten 6 (1952), 261-262.