The Mathematician's Colouring Book

An Introduction to Ramsey Theory

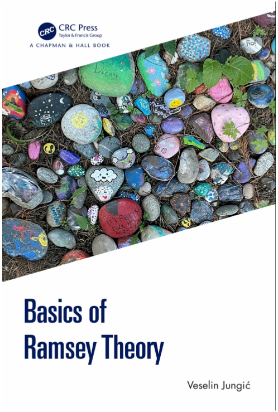
Brandon Papandrea

University of Toronto - Mississauga
Department of Mathematical and Computational Sciences

UTM MCS Reading Course Showcase
December 3, 2025

Acknowledgements

The book used for this reading course was *Basics of Ramsey Theory (2023)* by Veselin Jungić.



What is Ramsey Theory

Ramsey Theory is the study of "finding order within chaos."

Results at the core of this field include:

- Pigeonhole Principle
- Hilbert's Cube Lemma (1892)
- Schur's Theorem (1916)
- van der Waerden's Theorem (1927)
- Ramsey's Theorem (1930)

Figure: Frank Ramsey, Paul Erdős, & Ron Graham

Table of Contents

Definitions

Definition

A **finite colouring** of $\mathbb{N}$ is a way of assigning to each n one of k colours; we sometimes call this a k-colouring.
If $A \subset \mathbb{N}$ is such that each $a \in A$ is the same colour, we say A is **monochromatic**.

Definition

A k-term **arithmetic progression** for a fixed $a, d \in \mathbb{N}$ is any set of the form

$$A = \{a + id :, i \in \{0, \ldots, k-1\}\}$$

Baudet's Conjecture

van der Waerden's Theorem begins with a conjecture of Pierre Baudet:

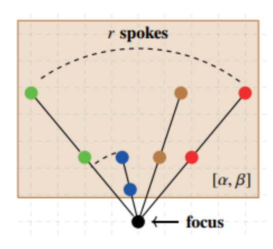
Baudet's Conjecture (1921)

Given any finite colouring of $\mathbb{N}$, there exists a monochromatic 3-term arithmetic progression.

Colour Focused Arithmetic Progressions

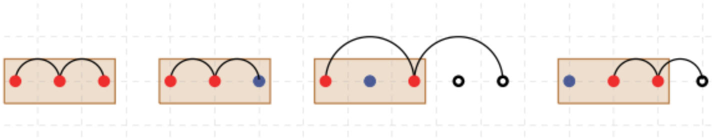
Definition

Given a finite colouring of $\mathbb{N}$ and r k-term arithmetic progressions $A_1, \ldots, A_r$, we say the arithmetic progressions are **colour focused** at a value f if each A_i is monochromatic, different colours from each other, and for all i, $a_i + kd_i = f$. If so, we call each A_i a **spoke** of length k.

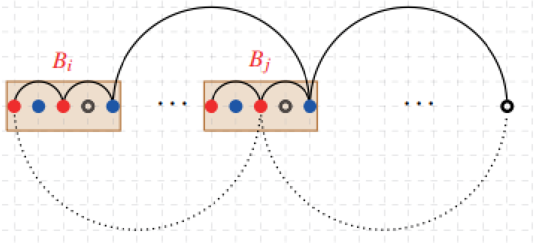


We proceed by strong induction on the number of colours used, ℓ. First suppose that $\ell = 2$:

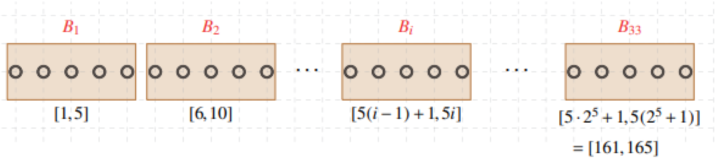
Notice that there are 4 distinct colourings of $\{1, 2, 3\}$ using 2 colours. Each contains either a monochromatic 3-term arithmetic progression, or a single colour focused spoke of length 2 with focus in $\{1, 2, 3, 4, 5\}$.



If two of the same block show up in our colouring of $\mathbb{N}$, we're done:

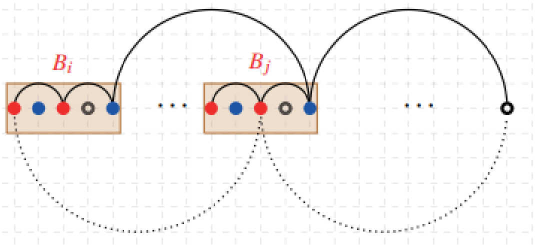


Divide the interval $\{1, \ldots, 165\}$ into 33 blocks of 5 numbers each. By the Pigeonhole Principle (PHP), two of these blocks are coloured the same.



Proving Baudet's Conjecture: Base Case

Assuming these blocks have no monochromatic 3-term arithmetic progression, 2 colour focused spokes of length 2 may be constructed. Their focus must be the same colour as one of them.



Proving Baudet's Conjecture: A Lemma

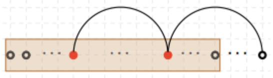
Now let $\ell \geq 2$. We need a lemma:

Lemma

For all $r \in \{1, \ldots, \ell\}$, there is an M such that any ℓ-colouring of $\{1, \ldots, M\}$ contains either a monochromatic 3-term arithmetic progression, or r spokes of length 2.

Proof.

We proceed by induction on r. If $r = 1$, take $M = \ell + 1$. By PHP, two values in $\{1, \ldots, M\}$ are coloured the same, so there is a colour focused spoke of length 2.

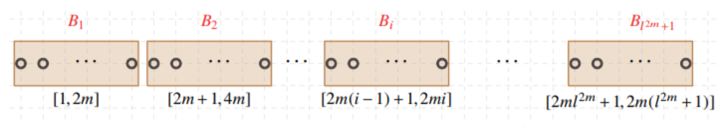


$\square$

Proving Baudet's Conjecture: A Lemma

Proof.

Now suppose the claim holds for $r-1$ using m, and consider r. Divide the interval $\{1,\ldots,2m(\ell^{2m}+1)\}$ into $\ell^{2m}+1$ blocks of $2m$ elements each. Suppose we have an ℓ-colouring of this interval with no monochromatic 3-term arithmetic progression; in particular, no block has one.

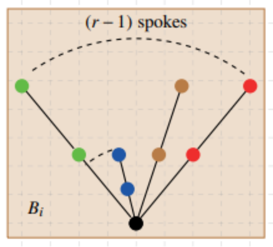


$\square$

Proving Baudet's Conjecture: A Lemma

Proof.

By the induction hypothesis, each block must contain $r - 1$ spokes and a focus. Moreover, as there are ℓ^{2m} ways to colour each block, by PHP, two blocks are coloured the same.

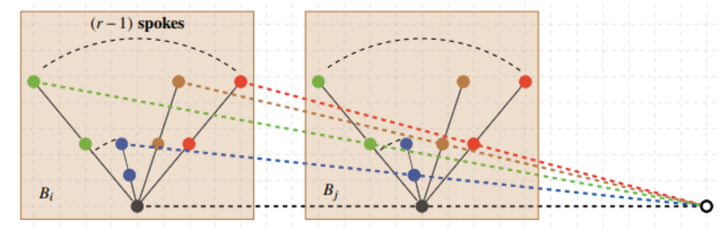


$\square$

Proving Baudet's Conjecture: A Lemma

Proof.

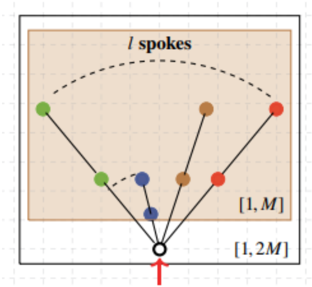
We can then construct r spokes as follows:



Thus, $M = 2m(\ell^{2m} + 1)$. $\square$

Proving Baudet's Conjecture: Inductive Step

Now, with $\ell \geq 2$, consider an ℓ-colouring of $\{1, \ldots, 2M\}$, where M is the integer guaranteed by the Lemma for $r = \ell$. Assuming this ℓ-colouring creates no monochromatic 3-term arithmetic progression in $\{1, \ldots, M\}$, by the Lemma, there are ℓ spokes of length 2 with their focus in $\{1, \ldots, 2M\}$. This focus must be the same colour as one of the spokes, giving us a monochromatic 3-term arithmetic progression.



van der Waerden's Theorem

Bartel van der Waerden proved Baudet's Conjecture in 1926. In doing so, he proved a fundamental result of Ramsey Theory:

Theorem (van der Waerden, 1927)

For all $\ell, k \in \mathbb{N}$, any ℓ-colouring of $\mathbb{N}$ contains a monochromatic k-term arithmetic progression.
In particular, there is a number $W(\ell : k)$ such that any ℓ-colouring of $\{1, \ldots, W(\ell : k)\}$ contains a monochromatic k-term arithmetic progression.

To prove this, we fix ℓ and induct on k. The base cases $k = 1, 2, 3$ have already been dealt with, so let $k \geq 4$ such that $W(\ell : k - 1)$ exists.

Proving vdW: A Lemma

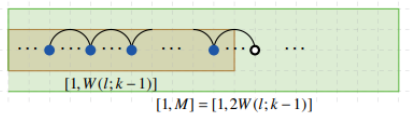
We first prove a lemma:

Lemma

For all $r \in \{1, \ldots, \ell\}$, there is an M such that any ℓ-colouring of $\{1, \ldots, M\}$ contains either a monochromatic k-term arithmetic progression, or r spokes of length $k-1$.

Proof.

We proceed by induction on r. If $r = 1$, take $M = 2W(\ell : k-1)$. There is a monochromatic $(k-1)$-term arithemetic progression in $\{1, \ldots, W(\ell : k-1)\}$, so we have a spoke of length $(k-1)$ in $\{1, \ldots, M\}$.

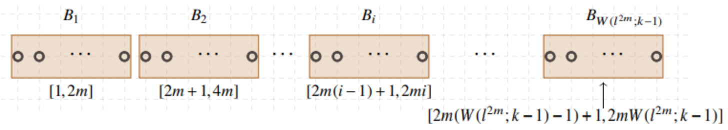


□

Proving vdW: A Lemma

Proof.

Now suppose the claim holds for $r-1$ using m, and consider r. Divide the interval $\{1, \ldots, 2mW(\ell^{2m}; k-1)\}$ into $W(\ell^{2m}; k-1)$ blocks of $2m$ elements each. Suppose we have an ℓ-colouring of this interval with no monochromatic k-term arithmetic progression.

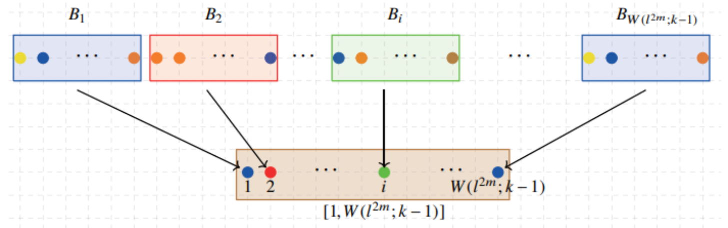


$\square$

Proving vdW: A Lemma

Proof.

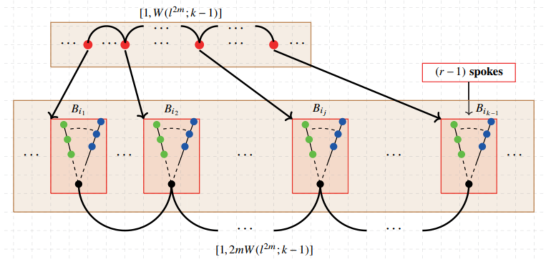
This induces an ℓ^{2m}-colouring of $\{1, \ldots, W(\ell^{2m}; k-1)\}$.



$\square$

Proof.

This new colouring must contain a monochromatic $(k-1)$-term arithmetic progression, meaning there are $k-1$ blocks that are coloured the same and are equally spaced out. Each one contains $r-1$ spokes.

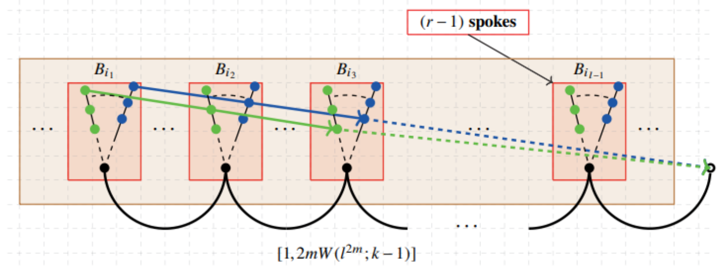


□

Proving vdW: A Lemma

Proof.

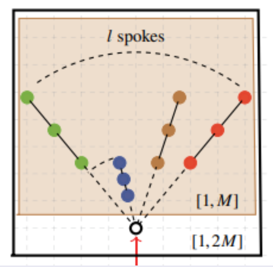
We can now construct r spokes:



Thus, $M = 2mW(\ell^{2m}; k-1)$. □

Proof of van der Waerden's Theorem.

The proof is now identical to that of Baudet: Take an ℓ-colouring of $\{1, \ldots, M\}$, where M comes from the Lemma for $r = \ell$, such that there is not k-term arithmetic progression in $\{1, \ldots, M\}$. Then apply the Lemma:



□

Table of Contents

A Philosophical Question

Question

What allows a monochromatic set of numbers to contain an arbitrarily large arithmetic progression?

Answer

That set is “large” enough.

Density

Definition

Let $A \subset \mathbb{N}$. The **upper/lower density** of A is given by

$$\limsup_{n\to\infty} \frac{|A \cap \{1, \dots, n\}|}{n} \quad \text{and} \quad \liminf_{n\to\infty} \frac{|A \cap \{1, \dots, n\}|}{n}$$

respectively. If these values agree, then we say A has **density** given by this value.

Examples

- The set of multiples of 3 has density $\frac{1}{3}$.
- The set of primes have density 0.
- The set $\left(\bigcup_{k\geq 1}\{2^{4k-1}, \dots, 2^{4k+1}\}\right)$ does not have density.

Density, in a way, measures the "sparseness" of a set.

Szemerédi's Theorem and Ergodic Ramsey Theory

Theorem (Szemerédi, 1975)

Any set with positive upper density contains arbitrarily long arithmetic progressions.

Hillel Furstenberg's 1977 proof of this began the field of **Ergodic Ramsey Theory**, which uses analysis and statstics to solve problems in additive combinatorics.

Theorem (Green & Tao, 2004)

The set of primes contains arbitrarily long arithmetic progressions.

Syndetic Sets

Definition

A set $A = \{a_1, a_2, a_3, \ldots\}$ is **syndetic** if there is an M such that for all i, $a_{i+1} - a_i \leq M$.

Examples

- The set of powers of 2 are not syndetic.
- The set of multiples of 3 and 5 are syndetic.

Theorem

If A is syndetic, it contains arbitrarily long arithmetic progressions.

IP Sets

Definition

Let $A \subset \mathbb{N}$. $\sum_A$ is the set of finite sums of elements in A.
A set $A \subset \mathbb{N}$ is called **IP** if there is an infinite set D such that $\sum_D \subset A$.

Theorem (Hindman, 1974)

Given a finite colouring of $\mathbb{N}$, there exists an infinite set $A \subset \mathbb{N}$ such that $\sum_A$ is monochromatic; one of the colours forms an IP set.

Theorem (Baumgartner, 1974)

Let F be the set of finite subsets of $\mathbb{N}$. Then given a finite colouring of F, there is a infinite set $A \subset F$ such that the set of all finite unions of elements of A is monochromatic.

Table of Contents

Combinatorial Lines

Definition

The set $\{1, \ldots, m\}^n$ is the set of strings of length n using $\{1, \ldots, m\}$. If we include the symbol $*$ in our alphabet, we get $\{1, \ldots, n\}^n_*$. Any string containing a $*$ is called a **root word**.
For a root word τ, we define τ_i to be the word where each instance of $*$ in τ is replaced with i.
A **combinatorial line rooted at** τ is the set $L_\tau = \{\tau_i : i \in \{1, \ldots, m\}\}$.

Example

$\{1, 2, 3\}^2$ is the set of points on a Tic-Tac-Toe board. The main diagonal can be written as $L_\tau = \{11, 22, 33\}$ where $\tau = **$.

The Hales-Jewett Theorem

Theorem (Hales & Jewett, 1963)

Let $k, m \in \mathbb{N}$. Then there is a number n such that given any k-colouring of $\{1, \ldots, m\}^n$, there is a monochromatic combinatorial line.

Corollary (Hales-Jewett for $k = 2$, $m = 3$)

In a high enough dimension, Tic-Tac-Toe cannot end in a draw.

One can use Hales-Jewett to prove van der Waerden's Theorem on arbitary semigroups and vector spaces (see Gallai's Theorem for Semigroups and the Gallai-Witt Theorem).

Group actions & Combinatorial Cubes/Subcubes

Not every line in $\{1, \ldots, m\}^n$ is a combinatorial line, but we can create the rest of them by considering root words where the wildcards are related via a group action.

We may also use more than 1 wildcard, say $*, \#$. These produce **combinatorial cubes**. These can contain combinatorial cubes of smaller dimension, called **combinatorial subcubes**.

Example

A Tic-Tac-Toe board is the combinatorial square rooted at $\tau = *\#$ in $\{1, 2, 3\}^2_{*,\#}$, and its subcubes are combinatiorial lines and points.

The Graham-Rothschild Theorem

Theorem (Graham & Rothschild, 1972)

Given an alphabet $\{1, \ldots, s\}$, a group G, and positive integers $\ell, m, k \in \mathbb{N}$, there exists an n such that given any ℓ-colouring of the k-dimensional combinatorial cubes in $\{1, \ldots, s\}^n$, there is an m-dimension combinatorial cube in this set such that all of its k-dimensional subcubes are the same colour.

If we set G to be trivial, and $m = 1, k = 0$, we get the Hales-Jewett Theorem.

If we set $s = 2$, G to be non-trivial, and $\ell = 2, m = 2, k = 1$, the n guaranteed by the theorem is bounded by Graham's Number. This is where that number came from!