Part III — Ramsey Theory

Based on lectures by Imre Leader
Notes adapted by Joshua Snyder

Michaelmas 2019

What happens when we cut up a mathematical structure into many 'pieces' ? How big must the original structure be in order to guarantee that at least one of the pieces has a specific property of interest? These are the kinds of questions at the heart of Ramsey theory. A prototypical result in the area is van der Waerden's theorem, which states that whenever we partition the natural numbers into finitely many classes, there is a class that contains arbitrarily long arithmetic progressions.

The course will cover both classical material and more recent developments in the subject. Some of the classical results that I shall cover include Ramsey's theorem, van der Waerden's theorem and the Hales–Jewett theorem; I shall discuss some applications of these results as well. More recent developments that I hope to cover include the properties of non-Ramsey graphs, topics in geometric Ramsey theory, and finally, connections between Ramsey theory and topological dynamics. I will also indicate a number of open problems.

Pre-requisites

There are almost no pre-requisites and the material presented in this course will, by and large, be self-contained. However, students familiar with the basic notions of graph theory and point-set topology are bound to find the course easier.

Contents

0   Introduction

Vaguely, the main idea of Ramsey theory is as follows — we have some object X that comes with some structure, and then we break X up into finitely many pieces. Can we find a piece that retains some of the structure of X? Usually, we phrase this in terms of *colourings*. We pick a list of finitely many colours, and colour each element of X. Then we want to find a monochromatic subset of X that satisfies some properties.
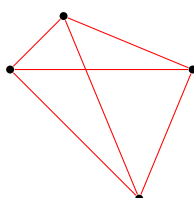
The most classical example of this is a graph colouring problem. We take a graph, say
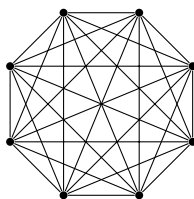
We then colour each edge with either red or blue:
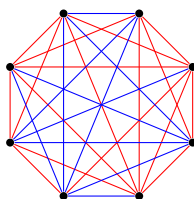
We now try to find a (complete) *subgraph* that is monochromatic, i.e. a subgraph all of whose edges are the same colour. With a bit of effort, we can find a red monochromatic subgraph of size 4:

Are we always guaranteed that we can find a monochromatic subgraph of size 4? If not, how big does the initial graph have to be? These are questions we will answer in this course. Often, we will ask the question in an "infinite" form — given an *infinite* graph, is it always possible to find an infinite monochromatic subgraph? The answer is yes, and it turns out we can deduce results about the finitary version just by knowing the infinite version.

These questions about graphs will take up the first chapter of the course. In the remaining of the course, we will discuss Ramsey theory on the integers, which, for the purpose of this course, will always mean $\mathbb{N} = \{1, 2, 3, \cdots\}$. We will now try to answer questions about the *arithmetic structure* of $\mathbb{N}$. For example, when we finitely colour $\mathbb{N}$, can we find an infinite arithmetic progression? With some thought, we can see that the answer is no. However, it turns out we can find arbitrarily long arithmetic progressions.

More generally, suppose we have some system of linear equations

$$3x + 6y + 2z = 0$$
$$2x + 7y + 3z = 0.$$

If we finitely colour $\mathbb{N}$, can we always find a monochromatic solution to these equations?

Remarkably, there is a *complete characterization* of all linear systems that always admit monochromatic solutions. This is given by *Rado's theorem*, which is one of the hardest theorems we are going to prove in the course.

If we are brave, then we can try to get the multiplicative structure of $\mathbb{N}$ in the picture. We begin by asking the most naive question — if we finitely colour $\mathbb{N}$, can we always find $x, y \in \mathbb{N}$, not both 2, such that $x + y$ and xy are the same colour? This is a very hard question, and the answer turns out to be yes.

1 Graph Ramsey theory

1.1 Infinite graphs

We begin by laying out some notations and definitions.

Definition ($\mathbb{N}$ and $[n]$)**.** We write

$$\mathbb{N} = \{1, 2, 3, 4, \cdots\}.$$

We also write

$$[n] = \{1, 2, 3, \cdots, n\}.$$

Notation. For a set X, we write $X^{(r)}$ for the subsets of X of size r. The elements are known as *r-sets*.

We all know what a graph is, hopefully, but for concreteness we provide a definition:

Definition (Graph)**.** A graph G is a pair (V, E), where $E \subseteq V^{(2)}$.

In particular, our graphs are not directed.

We are mostly interested in graphs with $V = \mathbb{N}$ or $[n]$. In this case, we will write an edge $\{i, j\}$ as ij, and always assume that $i < j$. More generally, in $\mathbb{N}^{(r)}$, we write $i_1 i_2 \cdots i_r$ for an r-set, and implicitly assume $i_1 < i_2 < \cdots < i_r$.

Definition (k-colouring)**.** A k-colouring of a set X is a map $c : X \to [k]$.

Of course, we often replace k with some actual set of colours, e.g. $\{\text{red}, \text{blue}\}$. We make more obvious definitions:

Definition (Monochromatic set)**.** Let X be a set with a k-colouring. We say a subset $Y \subseteq X$ is *monochromatic* if the colouring restricted to Y is constant.

The first question we are interested is in the following — if $\mathbb{N}^{(2)}$ is k-coloured, can we find a complete infinite subgraph that is monochromatic? In other words, is there an infinite subset $X \subseteq \mathbb{N}$ such that $X^{(2)} \subseteq \mathbb{N}^{(2)}$ is monochromatic?

We begin by trying some examples. The correct way to read these examples is to stare at the colouring, and then try to find a monochromatic subset yourself, instead of immediately reading on to see what the monochromatic subset is.

Example. Suppose we had the colouring $c : \mathbb{N}^{(2)} \to \{\text{red}, \text{blue}\}$ by

$$c(ij) = \begin{cases} \text{red} & i + j \text{ is even} \\ \text{blue} & i + j \text{ is odd} \end{cases}.$$

Then we can pick $X = \{2, 4, 6, 8, \cdots\}$, and this is an infinite monochromatic set.

Example. Consider $c : \mathbb{N}^{(2)} \to \{0, 1, 2\}$, where

$$c(ij) = \max\{n : 2^n \mid i + j\} \bmod 3$$

In this case, taking $X = \{8, 8^2, 8^3, \cdots\}$ gives an infinite monochromatic set of colour 0.

**Example.** Let $c : \mathbb{N}^{(2)} \to \{\text{red}, \text{blue}\}$ by

$$c(ij) = \begin{cases} \text{red} & i + j \text{ has an even number of distinct prime factors} \\ \text{blue} & \text{otherwise} \end{cases}$$

It is in fact an open problem to find an explicit infinite monochromatic set for this colouring, or even for which colour do we have an infinite monochromatic set. However, we can prove that such a set must exist!

Theorem (Ramsey's theorem)**.** Whenever we k-colour $\mathbb{N}^{(2)}$, there exists an infinite monochromatic set X, i.e. given any map $c : \mathbb{N}^{(2)} \to [k]$, there exists a subset $X \subseteq \mathbb{N}$ such that X is infinite and $c|_{X^{(2)}}$ is a constant function.

Proof. Pick an arbitrary point $a_1 \in \mathbb{N}$. Then by the pigeonhole principle, there must exist an infinite set $\mathcal{B}_1 \subseteq \mathbb{N} \setminus \{a_1\}$ such that all the a_1-$\mathcal{B}_1$ edges (i.e. edges of the form (a_1, b_1) with $b_1 \in \mathcal{B}_1$) are of the same colour c_1.

Now again arbitrarily pick an element $a_2 \in \mathcal{B}_1$. Again, we find some infinite $\mathcal{B}_2 \subseteq \mathcal{B}_1$ such that all a_2-$\mathcal{B}_2$ edges are the same colour c_2. We proceed inductively.
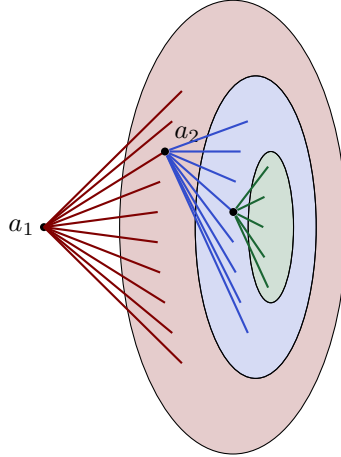
We obtain a sequence $\{a_1, a_2, \cdots\}$ and a sequence of colours $\{c_1, c_2, \cdots\}$ such that $c(a_i, a_j) = c_i$, for $i < j$.

Now again by the pigeonhole principle, since there are finitely many colours, there exists an infinite subsequence $c_{i_1}, c_{i_2}, \cdots$ that is constant. Then $a_{i_1}, a_{i_2}, \cdots$ is an infinite monochromatic set, since all edges are of the colour $c_{i_1} = c_{i_2} = \cdots$. So we are done. $\qquad\square$

This proof exhibits many common themes we will see in the different Ramsey theory proofs we will later do. The first is that the proof is highly non-constructive. Not only does it not give us the infinite monochromatic set; it doesn't even tell us what the colour is.

Another feature of the proof is that we did not obtain the infinite monochromatic set in one go. Instead, we had to first pass through that intermediate structure, and then obtain an infinite monochromatic set from that. In future proofs, we might have to go through *many* passes, instead of just two.

This theorem looks rather innocuous, but it has many interesting applications.

Corollary (Bolzano-Weierstrass theorem)**.** Let $(x_i)_{i \geq 0}$ be a bounded sequence of real numbers. Then it has a convergent subsequence.

Proof. We define a colouring $c : \mathbb{N}^{(2)} \to \{\uparrow, \downarrow\}$, where

$$c(ij) = \begin{cases} \uparrow & x_i < x_j \\ \downarrow & x_j \leq x_i \end{cases}$$

Then Ramsey's theorem gives us an infinite monochromatic set, which is the same as a monotone subsequence. Since this is bounded, it must convergence. $\square$

With a very similar technique, we can prove that we can do this for $\mathbb{N}^{(r)}$ for any r, and not just $\mathbb{N}^{(2)}$.

Theorem (Ramsey's theorem for r sets)**.** Whenever $\mathbb{N}^{(r)}$ is k-coloured, there exists an infinite monochromatic set, i.e. for any $c : \mathbb{N}^{(r)} \to [k]$, there exists an infinite $X \subseteq \mathbb{N}$ such that $c|_{X^{(r)}}$ is constant.

We can again try some concrete examples:

Example. We define $c : \mathbb{N}^{(3)} \to \{\text{red}, \text{blue}\}$ by

$$c(ijk) = \begin{cases} \text{red} & i \mid j + k \\ \text{blue} & \text{otherwise} \end{cases}$$

Then $X = \{2^0, 2^1, 2^2, \cdots\}$ is an infinite monochromatic set.

Proof. We induct on r. This is trivial when $r = 1$. Assume $r > 1$. We fix $a_1 \in \mathbb{N}$. We induce a k-colouring c_1 of $(\mathbb{N} \setminus \{a_1\})^{(r-1)}$ by

$$c_1(F) = c(F \cup \{a_1\}).$$

By induction, there exists an infinite $B_1 \subseteq \mathbb{N} \setminus \{a_1\}$ such that B_1 is monochromatic for c_1, i.e. all a_1-B_1 r-sets have the same colour c_1.

We proceed inductively as before. We get $a_1, a_2, a_3, \cdots$ and colours $c_1, c_2, \cdots$ etc. such that for any r-set F contained in $\{a_1, a_2, \cdots\}$, we have $c(F) = c_{\min F}$.

Then again, there exists $c_{i_1}, c_{i_2}, c_{i_3}, \cdots$ all identical, and our monochromatic set is $\{a_{i_1}, a_{i_2}, a_{i_3}, \cdots\}$. $\square$

Now a natural question to ask is — what happens when we have infinitely many colours? Clearly an infinite monochromatic subset cannot be guaranteed, because we can just colour all edges with different colours.

The natural compromise is to ask if we can find an X such that *either* $c|_X$ is monochromatic, *or* $c|_X$ is injective. After a little thought, we realize this is also impossible. We can construct a colouring on $\mathbb{N}^{(2)}$ as follows: we first colour all edges that involve 1 with the colour 1, then all edges that involve 2 with the colour 2 etc:

It is easy to see that we cannot find an infinite monochromatic subset or an infinite subset with all edges of different colours.

However, this counterexample we came up with still has a high amount of structure — the colour of the edges are uniquely determined by the first element. It turns out this is the only missing piece (plus the obvious dual case).

With this, we can answer our previous question:

Theorem (Canonical Ramsey theorem). For any $c : \mathbb{N}^{(2)} \to \mathbb{N}$, there exists an infinite $X \subseteq \mathbb{N}$ such that one of the following hold:

(i) $c|_{X^{(2)}}$ is constant.

(ii) $c|_{X^{(2)}}$ is injective.

(iii) $c(ij) = c(k\ell)$ iff $i = k$ for all $i, j, k, \ell \in X$.

(iv) $c(ij) = c(k\ell)$ iff $j = \ell$ for all $i, j, k, \ell \in X$.

Recall that when we write ij, we always implicitly assume $i < j$, so that (iii) and (iv) make sense.

In previous proofs, we only had to go through two passes to get the desired set. This time we will need more.

Proof. Consider the following colouring of $X^{(4)}$: let c_1 be a 2-colouring

$$c_1(ijk\ell) = \begin{cases} \texttt{SAME} & c(ij) = c(k\ell) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

Then we know there is some infinite monochromatic set $X_1 \subseteq \mathbb{N}$ for c_1. If X_1 is coloured $\texttt{SAME}$, then we are done. Indeed, for any pair ij and $i'j'$ in X_1, we can pick some huge k, ℓ such that $j, j' < k < \ell$, and then

$$c(ij) = c(k\ell) = c(i'j')$$

as we know $c_1(ijk\ell) = c_1(i'j'k\ell) = \texttt{SAME}$.

What if X_1 is coloured $\texttt{DIFF}$? We next look at what happens when we have edges that are nested each other. We define $c_2 : X_1^{(4)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ defined by

$$c_2(ijk\ell) = \begin{cases} \texttt{SAME} & c(i\ell) = c(jk) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

Again, we can find an infinite monochromatic subset $X_2 \subseteq X_1$ with respect to c_2.

We now note that X_2 cannot be coloured $\texttt{SAME}$. Indeed, we can just look at
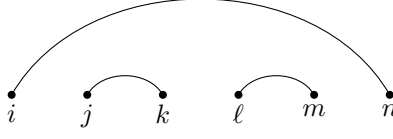
So if X_2 were `SAME`, we would have

$$c(\ell m) = c(in) = c(jk),$$

which is impossible since X_1 is coloured `DIFF` under c_1.

So X_2 is `DIFF`. Now consider $c_3 : X_2^{(4)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ given by

$$c_3(ijk\ell) = \begin{cases} \texttt{SAME} & c(ik) = c(j\ell) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

Again find an infinite monochromatic subset $X_3 \subseteq X_2$ for c_3. Then X_3 cannot be `SAME`, this time using the following picture:
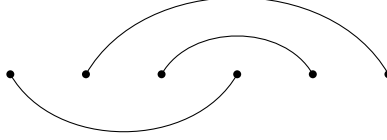
contradicting the fact that c_2 is `DIFF`. So we know X_3 is `DIFF`.

We have now have ended up in this set X_3 such that if we have any two pairs of edges with different end points, then they must be different.

We now want to look at cases where things share a vertex. Consider $c_4 : X_3^{(3)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ given by
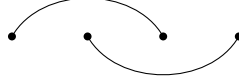
$$c_4(ijk) = \begin{cases} \texttt{SAME} & c(ij) = c(jk) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

Let $X_4 \subseteq X_3$ be an infinite monochromatic set for c_4. Now X_4 cannot be coloured `SAME`, using the following picture:

which contradicts the fact that c_1 is `DIFF`. So we know X_4 is also coloured `DIFF` under c_4.

We are almost there. We need to deal with edges that nest in the sense of (iii) and (iv). We look at $c_5 : X_4^{(3)} \to \{\texttt{LSAME}, \texttt{LDIFF}\}$ given by

$$c_5(ijk) = \begin{cases} \texttt{LSAME} & c(ij) = c(ik) \\ \texttt{LDIFF} & \text{otherwise} \end{cases}$$

Again we find $X_5 \subseteq X_4$, an infinite monochromatic set for c_5. We don't separate into cases yet, because we know both cases are possible, but move on to classify the right case as well. Define $c_6 : X_5^{(3)} \to \{\texttt{RSAME}, \texttt{RDIFF}\}$ given by

$$c_5(ijk) = \begin{cases} \texttt{RSAME} & c(ik) = c(jk) \\ \texttt{RDIFF} & \text{otherwise} \end{cases}$$

Let $X_6 \subseteq X_5$ be an infinite monochromatic subset under c_5.

As before, we can check that it is impossible to get both $\texttt{LSAME}$ and $\texttt{RSAME}$, using the following picture:

contradicting c_4 being $\texttt{DIFF}$.

Then the remaining cases $(\texttt{LDIFF}, \texttt{RDIFF})$, $(\texttt{LSAME}, \texttt{RDIFF})$ and $(\texttt{RDIFF}, \texttt{LSAME})$ corresponds to the cases (ii), (iii) and (iv). □

Note that we could this theorem in one pass only, instead of six, by considering a much more complicated colouring $(c_1, c_2, c_3, c_4, c_5, c_6)$ with values in

$$\{\texttt{SAME}, \texttt{DIFF}\}^4 \times \{\texttt{LSAME}, \texttt{LDIFF}\} \times \{\texttt{RSAME}, \texttt{RDIFF}\},$$

but we still have to do the same analysis and it just complicates matters more.

There is a generalization of this to r-sets. One way we can rewrite the theorem is to say that the colour is uniquely determined by some subset of the vertices. The cases (i), (ii), (iii), (iv) correspond to no vertices, all vertices, first vertex, and second vertex respectively. Then for r-sets, we have 2^r possibilities, one for each subset of the r-coordinates.

Theorem (Higher dimensional canonical Ramsey theorem)**.** Let $c : \mathbb{N}^{(r)} \to \mathbb{N}$ be a colouring. Then there exists $D \subseteq [r]$ and an infinite subset $X \subseteq \mathbb{N}$ such that for all $x, y \in X^{(r)}$, we have $c(x) = c(y)$ if $\{i : x_i = y_i\} \supseteq D$, where $x = \{x_1 < x_2 < \cdots < x_r\}$ (and similarly for y).

1.2 Finite graphs

We now move on to discuss a finitary version of the above problem. Of course, if we finitely colour an infinite graph, then we can obtain a finite monochromatic subgraph of any size we want, because this is just a much weaker version of Ramsey's theorem. However, given $n < N$, if we finitely colour a graph of size N, are we guaranteed to have a monochromatic subgraph of size n?

Before we begin, we note some definitions. Recall again that a graph G is a pair (V, E) where $E \subseteq V^{(2)}$.
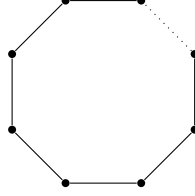
Example. The *path graph* on n vertices P_n is

We can write
$$V = [n], \quad E = \{12, 23, 34, \cdots, (n-1)n\}.$$

Example. The n-cycle C_n is
$$V = [n], \quad E = \{12, 23, \cdots, (n-1)n, 1n\}.$$

Finally, we have

Example. The *complete graph* V_n on n vertices is
$$V = [n], \quad E = V^{(2)}.$$
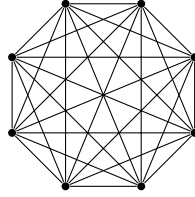
Concretely, the question we want to answer is the following — how big does our (complete) graph have to be to guarantee a complete monochromatic subgraph of size n?

In this section, we will usually restrict to 2 colours. Everything we say will either be trivially be generalizable to more colours, or we have no idea how to do so. It is an exercise for the reader to figure out which it is.

Definition (Ramsey number). We let $R(n) = R(K_n)$ to be the smallest $N \in \mathbb{N}$ whenever we red-blue colour the edges of K_N, then there is a monochromatic copy of K_n.

It is not immediately obvious that $R(n)$ has to exist, i.e. it is a finite number. It turns out we can easily prove this from the infinite Ramsey theorem.

Theorem (Finite Ramsey theorem). For all n, we have $R(n) < \infty$.

Proof. Suppose not. Let n be such that $R(n) = \infty$. Then for any $m \geq n$, there is a 2-colouring c_m of $K_m = [m]^{(2)}$ such that there is no monochromatic set of size n.

We want to use these colourings to build up a colouring of $\mathbb{N}^{(2)}$ with no monochromatic set of size n. We want to say we take the "limit" of these colourings, but what does this mean? To do so, we need these colourings to be nested.

By the pigeonhole principle, there exists an infinite set $M_1 \subseteq \mathbb{N}$ and some fixed 2-colouring d_n of $[n]$ such that $c_m|_{[n]^{(2)}} = d_n$ for all $m \in M_1$.

Similarly, there exists an infinite $M_2 \subseteq M_1$ such that $c_m|_{[n+1]^{(2)}} = d_{n+1}$ for $m \in M_2$, again for some 2-colouring d_{n+1} of $[n+1]$. We repeat this over and over again. Then we get a sequence $d_n, d_{n+1}, \cdots$ of colourings such that d_i is a 2-colouring of $[i]^{(2)}$ without a monochromatic K_n, and further for $i < j$, we have

$$d_j|_{[i]^{(2)}} = d_i.$$

We then define a 2-colouring c of $\mathbb{N}^{(2)}$ by

$$c(ij) = d_m(ij)$$

for any $m > i, j$. Clearly, there exists no monochromatic K_n in c, as any K_n is finite. This massively contradicts the infinite Ramsey theorem. □

There are other ways of obtaining the finite Ramsey theorem from the infinite one. For example, we can use the compactness theorem for first order logic.

Proof. Suppose $R(n) = \infty$. Consider the theory with proposition letters p_{ij} for each $ij \in \mathbb{N}^{(2)}$. We will think of the edge ij as red if $p_{ij} = \top$, and blue if $p_{ij} = \bot$. For each subset of $\mathbb{N}$ of size n, we add in the axiom that says that set is not monochromatic.

Then given any finite subset of the axioms, it mentions only finitely many subsets of $\mathbb{N}$. Suppose it mentions vertices only up to $m \in \mathbb{N}$. Then by assumption, there is a 2-colouring of $[m]^{(2)}$ with no monochromatic subset of size n. So by assigning p_{ij} accordingly (and randomly assigning the remaining ones), we have found a model of this finite subtheory. Thus every finite fragment of the theory is consistent. Hence the original theory is consistent. So there is a model, i.e. a colouring of $\mathbb{N}^{(2)}$ with no monochromatic subset.

This contradicts the infinite Ramsey theorem. □

Similarly, we can do it by compactness of the space $\{1, 2\}^{\mathbb{N}}$ endowed with metric

$$d(f, g) = \frac{1}{2^n} \text{ if } n = \min\{i : f_i \neq g_i\}.$$

By Tychonoff theorem, we know this is compact, and we can deduce the theorem from this.

While these theorems save work by reusing the infinite Ramsey theorem, it is highly non-constructive. It is useless if we want to obtain an actual bound on $R(n)$. We now go back and see what we did when we proved the infinite Ramsey theorem.

To prove the infinite Ramsey theory, We randomly picked a point a_1. Then there are some things that are connected by red to it, and others connected by blue:
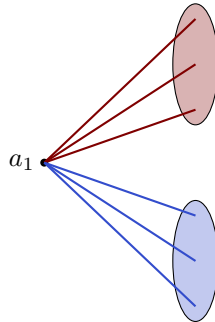
Next we randomly pick a point in one of the red or blue sets, and try to move on from there. Suppose we landed in the red one. Now note that if we find a blue K_n in the red set, then we are done. But on the other hand, we only need a red K_{n-1}, and not a full blown K_n. When we moved to this red set, the problem is no longer symmetric.

Thus, to figure out a good bound, it is natural to consider an asymmetric version of the problem.

* The following section was not lectured but can be examinable since the content is related to Finite Ramsey *

Definition (Off-diagonal Ramsey number)**.** We define $R(n, m) = R(K_n, K_m)$ to be the minimum $N \in \mathbb{N}$ such that whenever we red-blue colour the edges of K_N, we either get a red K_n or a blue K_m.

Clearly we have
$$R(n, m) \leq R(\max\{n, m\}).$$

In particular, they are finite. Once we make this definition, it is easy to find a bound on R.

Theorem. We have

$$R(n, m) \leq R(n - 1, m) + R(n, m - 1).$$

for all $n, m \in \mathbb{N}$. Consequently, we have

$$R(n, m) \leq \binom{n + m - 1}{n - 2}.$$

Proof. We induct on $n + m$. It is clear that

$$R(1, n) = R(n, 1) = 1, \quad R(n, 2) = R(2, n) = n$$

Now suppose $N \geq R(n - 1, m) + R(n, m - 1)$. Consider any red-blue colouring of K_N and any vertex $v \in V(K_n)$. We write

$$v(K_n) \setminus \{v\} = A \cup B,$$

where each vertex A is joined by a red edge to v, and each vertex in B is joined by a blue edge to v. Then

$$|A| + |B| \geq N - 1 \geq R(n - 1, m) + R(n, m - 1) - 1.$$

It follows that either $|A| \geq R(n - 1, m)$ or $|B| \geq R(n, m - 1)$. We wlog it is the former. Then by definition of $R(n - 1, m)$, we know A contains either a blue copy of K_m or a red copy of K_{n-1}. In the first case, we are done, and in the second case, we just add v to the red K_{n-1}.

The last formula is just a straightforward property of binomial coefficients. $\square$

In particular, we find

$$R(n) \leq \binom{2n - 1}{n - 2} \leq \frac{4^n}{\sqrt{n}}.$$

We genuinely have no idea whether $\sim 4^n$ is the correct growth rate, i.e. if there is some ε such that $R(n) \leq (4 - \varepsilon)^n$. However, we do know that for any $c > 0$, we eventually have

$$R(n) \leq \frac{4^n}{n^c}.$$

That was an upper bound. Do we have a lower bound? In particular, does $R(n)$ have to grow exponentially? The answer is yes, and the answer is a very classical construction of Erdös.

Theorem. We have $R(n) \geq \sqrt{2}^n$ for sufficiently large $n \in \mathbb{N}$.

The proof is remarkable in that before this was shown, we had no sensible bound at all. However, the proof is incredibly simple, and revolutionized how we think about colourings.

Proof. Let $N \leq \sqrt{2}^n$. We perform a red-blue colouring of K_N randomly, where each edge is coloured red independently of the others with probability $\frac{1}{2}$.

We let X_R be the number of red copies of K_n in such a colouring. Then since expectation is linear, we know the expected value is

$$
\begin{aligned}
[X_R] &= \binom{N}{n} \left(\frac{1}{2}\right)^{\binom{n}{2}} \\
&\leq \left(\frac{eN}{n}\right)^n \left(\frac{1}{2}\right)^{\binom{n}{2}} \\
&\leq \left(\frac{e}{n}\right)^n \sqrt{2}^{n^2} \sqrt{2}^{-n^2} \\
&= \left(\frac{e}{n}\right)^n \\
&< \frac{1}{2}
\end{aligned}
$$

for sufficiently large n.

Similarly, we have $[X_B] < \frac{1}{2}$. So the expected number of monochromatic K_n is < 1. So in particular there must be some colouring with no monochromatic K_n. $\qquad\square$

Recall the bound
$$R(m, n) \leq \binom{m + n - 2}{m - 1}.$$

If we think of m as being fixed, then this tells us

$$R(m, n) \sim (n + m)^{m-1}.$$

For example, if m is 3, then we have

$$R(3, n) \leq \binom{n + 1}{2} = \frac{n(n + 1)}{2} \sim n^2.$$

We can sort-of imagine where this bound came from. Suppose we randomly pick a vertex v_1. Then if it is connected to at least n other vertices by a red edge, then we are done — if there is even one red edge among those n things, then we

have a red triangle. Otherwise, all edges are blue, and we've found a complete blue K_n.

So this is connected to at most $n - 1$ things by a red edge. So if our graph is big enough, we can pick some v_2 connected to v_1 by a blue edge, and do the same thing to v_2. We keep going on, and by the time we reach v_n, we would have found $v_1, \cdots, v_n$ all connected to each other by blue edges, and we are done. So we have $K(3, n) \sim n^2$.
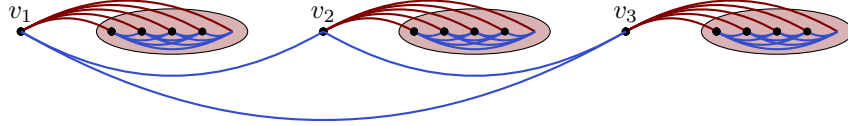
But this argument is rather weak, because we didn't use that large pool of blue edges we've found at v_1. So in fact this time we can do better than n^2.

Theorem. We have
$$R(3, n) \leq \frac{100n^2}{\log n}$$
for sufficiently large $n \in \mathbb{N}$.

Here the 100 is, obviously, just some random big number.

Proof. Let $N \geq 100n^2 / \log n$, and consider a red-blue colouring of the edges of K_N with no red K_3. We want to find a blue K_n in such a colouring.

We may assume that

(i) No vertex v has $\geq n$ red edges incident to it, as argued just now.

(ii) If we have v_1, v_2, v_3 such that $v_1 v_2$ and $v_1 v_3$ are red, then $v_2 v_3$ is blue:
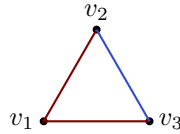
We now let

$$\mathcal{F} = \{W : W \subseteq V(K_N) \text{ such that } W^{(2)} \text{ is monochromatic and blue}\}.$$

We want to find some $W \in \mathcal{F}$ such that $|W| \geq n$, i.e. a blue K_n. How can we go about finding this?

Let W be a uniformly random member of $\mathcal{F}$. We will be done if we can show that that $\mathbb{E}[|W|] \geq n$.

We are going to define a bunch of random variables. For each vertex $v \in V(K_N)$, we define the variable

$$X_v = n\mathbf{1}_{\{v \in W\}} + |\{u : uv \text{ is red and } u \in W\}|.$$

Claim.
$$\mathbb{E}[X_v] > \frac{\log n}{10}$$
for each vertex v.

To see this, let
$$A = \{u : uv \text{ is red}\}$$
and let
$$B = \{u : uv \text{ is blue}\}.$$

then from the properties we've noted down, we know that $|A| < n$ and $A^{(2)}$ is blue. So we know very well what is happening in A, and nothing about what is in B.

We fix a set $S \subseteq B$ such that $S \in \mathcal{F}$, i.e. $S^{(2)}$ is blue. What can we say about W if we condition on $B \cap W = S$?
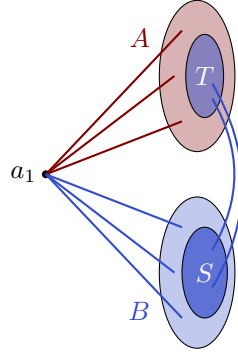
Let $T \subseteq A$ be the set of vertices that are joined to S only by blue edges. Write $|T| = x$. Then if $B \cap W = S$, then either $W \subseteq S \cup T$, or $W \subseteq S \cup \{v\}$. So there are exactly $2^x + 1$ choices of W. So we know

$$\mathbb{E}[X_v \mid W \cap B = S] \geq \frac{n}{2^x + 1} + \frac{2^x}{2^x + 1}(\mathbb{E}[|\text{random subset of T}|])$$
$$= \frac{n}{2^x + 1} + \frac{2^{x-1}x}{2^x + 1}.$$

Now if
$$x < \frac{\log n}{2},$$
then
$$\frac{n}{2^x + 1} \geq \frac{n}{\sqrt{n} + 1} \geq \frac{\log n}{10}$$
for all sufficiently large n. On the other hand, if
$$x \geq \frac{\log n}{2},$$
then
$$\frac{2^{x-1}x}{2^x + 1} \geq \frac{1}{4} \cdot \frac{\log n}{2} \geq \frac{\log n}{10}.$$

So we are done.

Claim.
$$\sum_{v \in V} X_v \leq 2n|W|.$$

To see this, for each vertex v, we know that if $v \in W$, then it contributes n to the sum via the first term. Also, by our initial observation, we know that $v \in W$ has at most n neighbours. So it contributes at most n to the second term (acting as the "u").

Finally, we know that

$$\mathbb{E}[|W|] \geq \frac{1}{2n} \sum \mathbb{E}[X_V] \geq \frac{N}{2n} \frac{\log n}{10} \geq \frac{100n^2}{\log n} \cdot \frac{\log n}{20n} \geq 5n.$$

Therefore we can always find some $W \in \mathcal{F}$ such that $|W| \geq n$. $\qquad\square$

Now of course, there is the following obvious generalization of Ramsey numbers:

Definition $(R(G, H))$**.** Let G, H be graphs. Then we define $R(G, H)$ to be the smallest N such that any red-blue colouring of K_N has either a red copy of G or a blue copy of H.

Obviously, we have $R(G, H) \leq R(|G|, |H|)$. So the natural question is if we can do better than that.

Exercise. Show that
$$R(P_n, P_n) \leq 2n.$$

So sometimes it can be much better.

2   Ramsey theory on the integers

So far, we've been talking about what happens when we finitely colour graphs. What if we k-colour the integers $\mathbb{N}$? What can we say about it?

It is a trivial statement that this colouring has a monochromatic subset, by the pigeonhole principle. Interesting questions arise when we try to take the additive structure of $\mathbb{N}$ into account. So we could ask, can we find a monochromatic "copy" of $\mathbb{N}$.

One way to make this question concrete is to ask if there is an infinite monochromatic arithmetic progression.

The answer is easily a "no"! There are only countably many progressions, so for each arithmetic progression, we pick two things in the progression and colour them differently.

We can also construct this more concretely. We can colour the first number red, the next two blue, the next three red etc. then it is easy to see that it doesn't have an infinite arithmetic progression.

But this is somewhat silly, because there is clearly a significant amount of structure in the sequence there. It turns out the following is true:

Theorem (van der Waerden theorem)**.** Let $m, k \in N$. Then there is some $N = W(m, k)$ such that whenever $[N]$ is k-coloured, then there is a monochromatic arithmetic progression of length n.

The idea is to do induction on m. We will be using colourings with much greater than k colours to deduce the existence of $W(m, k)$.

We can try a toy example first. Let's try to show that $W(3, 2)$ exists. Suppose we have three natural numbers:

By the pigeonhole principle, there must be two things that are the same colour, say

If this is the case, then if we don't want to have an arithmetic progression of length 3, then the fifth number must be blue

We now cut the universe into blocks into 5 things. Again by the pigeonhole principle, there must be two blocks that look the same. Say it's this block again.

Now we have two sequences, and the point at the end belongs to both of the two sequences. And no matter what colour it is, we are done.

For $k = 3$, we can still find such a block, but now that third point could be a third colour, say, green. This will not stop us. We can now look at these big blocks, and we know that eventually, these big blocks must repeat themselves.
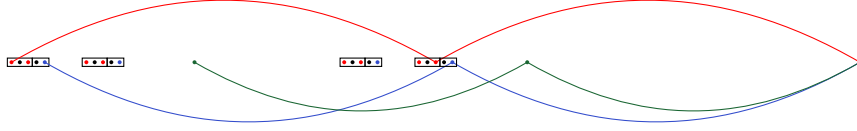
Here we did the case $m = 2$, and we used the pigeonhole principle. When we do $m > 2$, we will use van der Waerden theorem for smaller m inductively.

We now come up with names to describe the scenario we had above.

Definition (Focused progression)**.** We say a collection of arithmetic progressions $A_1, A_2, \cdots, A_r$ of length m with

$$A_i = \{a_i, a_i + d_i, \cdots, a_i + (m-1)d_i\}$$

are *focused* at f if $a_i + md_i = f$ for all $1 \leq i \leq r$.

Example. $\{1, 4\}$ and $\{5, 6\}$ are focused at 7.

Definition (Colour focused progression)**.** If $A_1, \cdots, A_r$ are focused at f, and each A_i is monochromatic and no two are the same colour, then we say they are *colour focused* at f.

We can now write the proof.

Proof. We induct on m. The result is clearly trivial when $m = 1$, and follows easily from the pigeonhole principle when $m = 2$.

Suppose $m > 1$, and assume inductively that $W(m-1, k')$ exists for any $k' \in \mathbb{N}$.

Here is the claim we are trying to established:

Claim. For each $r \leq k$, there is a natural number n such that whenever we k-colour $[n]$, then either

(i) There exists a monochromatic AP of length m; or

(ii) There are r colour-focused AP's of length $m - 1$.

It is clear that this claim implies the theorem, as we can pick $r = k$. Then if there isn't a monochromatic AP of length m, then we look at the colour of the common focus, and it must be one of the colours of those AP's.

To prove the claim, we induct on r. When $n = 1$, we may take $W(m-1, k)$. Now suppose $r > 1$, and some n' works for $r - 1$. With the benefit of hindsight, we shall show that
$$n = W(m-1, k^{2n'})2n'$$
works for r.

We consider any k-colouring of $[n]$, and suppose it has no monochromatic AP of length m. We need to find r colour-focused progressions of length $n - 1$.

We view this k-colouring of $[n]$ as a $k^{2n'}$ colouring of blocks of length $2n'$, of which there are $W(m-1, k^{2n'})$.

Then by definition of $W(m-1, k^{2n'})$, we can find blocks

$$B_s, B_{s+t}, \cdots, B_{s+(m-2)t}$$

which are coloured identically. By the inductive hypothesis, we know each B_s contains $r-1$ colour-focused AP's of length $m-1$, say $A_1, .., A_{r-1}$ with first terms $a_1, \cdots, a_r$ and common difference $d_1, \cdots, d_{r-1}$, and also their focus f, because the length of B_s is $2n'$, not just n'.

Since we assumed there is no monochromatic progression of length n, we can assume f has a different colour than the A_i.

Now consider $A'_1, A'_2, \cdots, A'_{r-1}$, where A'_i has first term a_i, common difference $d_i + 2n't$, and length $m-1$. This difference sends us to the next block, and then the next term in the AP. We also pick A'_r to consist of the all the focus of the blocks B_i, namely

$$A'_r = \{f, f + 2n't, \cdots, f + 2n't(m-2)\}$$

These progressions are monochromatic with distinct colours, and focused at $f + (2n't)(m-1)$. So we are done. $\qquad\square$

This argument, where one looks at the induced colourings of blocks, is called a *product argument*.

The bounds we obtain from this argument are, as one would expect, terrible. We have

$$W(3,k) \le k^{\cdot^{\cdot^{\cdot^{k^{4k}}}}},$$

where the tower of k's has length $k-1$.

A simple corollary of Van der Waerden is the following

Corollary. Whenever $\mathbb{N}$ is finitely coloured there is some colour class containing arbitrarily long arithmetic progressions.

Trivially we cannot guarantee an infinite monochromatic AP. A simple counterexample to this is simply to colour the first element blue, the next two red, the next three blue etc.

Alternatively we can list all countably many APs as $A_1, A_2, A_3, ...$ and choose distinct $x_1, y_1 \in A_1$ and let x_1 be red and y_1 be blue. Now pick distinct x_2, y_2 in A_2 (not x_1 or y_1) and let x_2 be red and y_2 be blue.

How can we strengthen this a bit?

Theorem (Strengthened Van der Waerden)**.** Whenever $\mathbb{N}$ is finitely coloured there exists a monochromatic AP of length m together with it's common difference

Proof. Given m, will do induction on the number of colours k. For $k = 1$ the result is trivially true. Given n' suitable for $k-1$ we will show that $n = W((m-1)n' + 1, k)$ is sufficient for k colours:

Given a k colouring of $[n]$, have a monochromatic AP $a, a+d, a+2d, \ldots, a+(m-1)n'd$ WLOG coloured red.

d red $\implies$ done with $\{a, a+d, \ldots, a+(m-1)d\} \cup \{d\}$

$2d$ red $\implies$ done with $\{a, a + 2d, \ldots, a + (m-1)2d\} \cup \{2d\}$

$\ldots$

$n'd$ red $\implies$ done with $\{a, a + n'd, \ldots, a + (m-1)n'd\} \cup \{n'd\}$

Therefore we may assume that $\{d, 2d, \ldots, n'd\}$ is $(k-1)-$coloured but then we are done by definition of n' $\hspace{2cm}$ □

Henceforth we will no longer care about bounds since they will simply be too large. Note that the case with $m = 2$ here corresponds to Schur's thoerem i.e there is a monochromatic solution to $x + y = z$ whenever $\mathbb{N}$ is finitely coloured.

Now we can generalize in a different way. What can we say about monochromatic structures a k-colouring of $\mathbb{N}^d$? What is the right generalization of van der Waerden theorem?

To figure out the answer, we need to find the right generalization of arithmetic progressions.

Definition (Homothetic copy)**.** Given a finite $S \subseteq \mathbb{N}^d$, a *homothetic copy* of S is a set of the form

$$\ell S + M,$$

where $\ell, M \in \mathbb{N}^d$ and $\ell \neq 0$.

Example. An arithmetic progression of length m is just a homothetic copy of $[m] = \{1, 2, \cdots, m\}$.

Thus, the theorem we want to say is the following:

Theorem (Gallai)**.** Whenever $\mathbb{N}^d$ is k-coloured, there exists a monochromatic (homothetic) copy of S for each finite $S \subseteq \mathbb{N}^d$.

In order to prove this, we prove the beautiful Hales–Jewett theorem, which is in some sense the abstract core of the argument we used for van der Waerden theorem.

We need a bit of set up. As usual, we write

$$[m]^n = \{(x_1, \cdots, x_n) : x_i \in [m]\}.$$
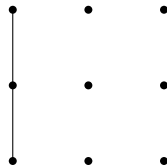
Here is the important definition:

Definition (Combinatorial line)**.** A *combinatorial line* $L \subseteq [m]^n$ is a set of the form

$$\{(x_1, \cdots, x_n n) : x_i = x_j \text{ for all } i, j \in I; x_i = a_i \text{ for all } i \notin I\}$$

for some fixed non-empty set of coordinates $I \subseteq [n]$ and $a_i \in [m]$.
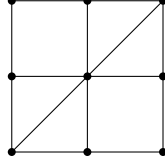
I is called the set of *active coordinates*.

Example. Here is a line in $[3]^2$

This line is given by $I = \{1\}$, $a_2 = 1$.

The following shows all the lines we have:

It is easy to see that any line has exactly $[m]$ elements. We write L^- and L^+ for the first and last point of the line, i.e. the points where the active coordinates are all 1 and m respectively. It is clear that any line L is uniquely determined by L^- and its active coordinates.

Example. In $[3]^3$, we have the line

$$L = \{(1,2,1), (2,2,2), (3,2,3)\}.$$

This is a line with $I = \{1,3\}$ and $a_2 = 2$. The first and last points are $(1,2,1)$ and $(3,2,3)$.

Then we have the following theorem:

Theorem (Hales–Jewett theorem). For all $m, k \in \mathbb{N}$, there exists $N = HJ(m,k)$ such that whenever $[m]^N$ is k-coloured, there exists a monochromatic line.

Note that this theorem implies van der Waerden's theorem easily. The idea is to embed $[m]^N$ into $\mathbb{N}$ linearly, so that a monochromatic line in $[m]^N$ gives an arithmetic progression of length m. Explicitly, given a k-colouring $c : \mathbb{N} \to [k]$, we define $c' : [m]^N \to [k]$ by

$$c'(x_1, x_2, \cdots, x_N) = c(x_1 + x_2 + \cdots + x_N).$$

Now a monochromatic line gives us an arithmetic progression of length m. For example, if the line is

$$L = \{(1,2,1), (2,2,2), (3,2,3)\},$$

then we get the monochromatic progression $4, 6, 8$ of length 3. In general, the monochromatic AP defined has $d = |I|$.

The proof is essentially what we did for van der Waerden theorem. We are going to build a lot of almost-lines that point at the same vertex, and then no matter what colour the last vertex is, we are happy.

Definition (Focused lines). We say lines $L_1, \cdots, L_r$ are *focused* at $f \in [m]^N$ if $L_i^+ = f$ for all $i = 1, \cdots, r$.

Definition (Colour focused lines). If $L_1, \cdots, L_r$ are focused lines, and $L_i \setminus \{L_i^+\}$ is monochromatic for each $i = 1, \cdots, r$ and all these colours are distinct, then we say $L_1, \cdots, L_r$ are *colour focused* at f.

Proof. We proceed by induction on m. This is clearly trivial for $m = 1$, as a line only has a single point.

Now suppose $m > 1$, and that $HJ(m-1, k)$ exists for all $k \in \mathbb{N}$. As before, we will prove the following claim:

Claim. For each $1 \leq r \leq k$, there is an $n \in \mathbb{N}$ such that in any k-colouring of $[m]^n$, either

(i) there exists a monochromatic line; or

(ii) there exists r colour-focused lines.

Again, the result is immediate from the claim, as we just use it for $r = k$ and look at the colour of the focus.

The prove this claim, we induct on r. If $r = 1$, then picking $n = HJ(m-1, k)$ works, as a single colour-focused line of length n is just a monochromatic line of length $n - 1$, and $[m-1]^n \subseteq [m]^n$ naturally.

Now suppose $r > 1$ and n works for $r - 1$. Then we will show that $n + \alpha$ works, where

$$\alpha = HJ(m - 1, \hat{k}\hat{m}\hat{n})$$

Consider a colouring $c : [m]^{n+\alpha} \to [k]$, and we assume this has no monochromatic lines.

We think of $[m]^{n+\alpha}$ as $[m]^n \times [m]^\alpha$. So for each point in $[m]^\alpha$, we have a whole cube $[m]^n$. Consider a k^{m^n} colouring of $[m]^\alpha$ as follows — given any $x \in [m]^\alpha$, we look at the subset of $[m]^{n+\alpha}$ containing the points with last α coordinates x. Then we define the new colouring of $[m]^\alpha$ to be the restriction of c to this $[m]^n$, and there are m^n possibilities.

Now there exists a line L such that $L \setminus L^+$ is monochromatic for the new colouring. This means for all $a \in [m]^n$ and for all $b, b' \in L \setminus L^+$, we have

$$c(a, b) = c(a, b').$$

Let $c''(a)$ denote this common colour for all $a \in [m]^n$. this is a k-colouring of $[m]^n$ with no monochromatic line (because c doesn't have any). So by definition of α, there exists lines $L_1, L_2, \cdots, L_{r-1}$ in $[m]^n$ which are colour-focused at some $f \in [m]^n$ for c''.

In the proof of van der Waerden, we had a progression within each block, and also how we just between blocks. Here we have the same thing. We have the lines in $[m]^n$, and also the "external" line L.

$$L$$
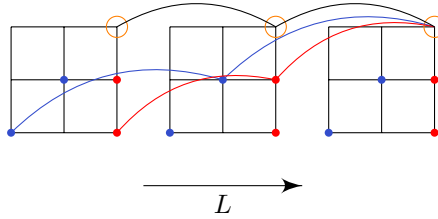
Consider the line $L_1', L_2', \cdots, L_{r-1}'$ in $[m]^{n'+n}$, where

$$(L_i')^- = (L_i^-, L^-),$$

and the active coordinate set is $I_i \cup I$, where I is the active coordinate set of L.

Also consider the line F with $F^- = (f, L^-)$ and active coordinate set I.

Then we see that $L_1', \cdots, L_{r-1}', F$ are r colour-focused lines with focus (f, L^+). $\qquad\square$

We can now prove our generalized van der Waerden.

Theorem (Gallai)**.** Whenever $\mathbb{N}^d$ is k-coloured, there exists a monochromatic (homothetic) copy of S for each finite $S \subseteq \mathbb{N}^d$.

Proof. Let $S = \{S(1), S(2), \cdots, S(m)\} \subseteq \mathbb{N}^d$. Given a k-colouring $\mathbb{N}^d \to [k]$, we induce a k-colouring $c : [m]^N \to [k]$ by

$$c'(x_1, \cdots, x_N) = c(S(x_1) + S(x_2) + \cdots + S(x_N)).$$

By Hales-Jewett, for sufficiently large N, there exists a monochromatic line L for c', which gives us a monochromatic homothetic copy of S in $\mathbb{N}^d$. For example, if the line is (1 2 1), (2 2 2) and (3 2 3), then we know

$$c(S(1) + S(2) + S(1)) = c(S(2) + S(2) + S(2)) = c(S(3) + S(2) + S(3)).$$

So we have the monochromatic homothetic copy $\lambda S + \mu$, where $\lambda = 2$ (the number of active coordinates), and $\mu = S(2)$. $\qquad\square$

Largeness in Ramsey theory

In van der Waerden, we proved that for each k, m, there is some n such that whenever we k-colour $[n]$, then there is a monochromatic AP of length m. How much of this is relies on the underlying set being $[n]$? Or is it just that if we finitely colour $[n]$, then one of the colours must contain a lot of numbers, and if we have a lot of numbers, then we are guaranteed to have a monochromatic AP?

Of course, by "contains a lot of numbers", we do not mean the actual number of numbers we have. It is certainly not true that for some n, whenever we k-colour any set of n integers, we must have a monochromatic m-AP, because an arbitrary set of n integers need not even contain an m-AP at all, let alone a monochromatic one. Thus, what we really mean is *density*.

Definition (Density)**.** For $A \subseteq \mathbb{N}$, we let the *density* of A as

$$\bar{d}(A) = \limsup_{(b-a) \to \infty} \frac{A \cap [a, b]}{|b - a|}.$$

Clearly, in any finite k-colouring of $\mathbb{N}$, there exists a colour class with positive density. Thus, we want to know if merely a positive density implies the existence of progressions. Remarkably, the answer is yes!

Theorem (Szemerédi theorem)**.** Let $\delta > 0$ and $m \in \mathbb{N}$. Then there exists some $N = S(m, \delta) \in \mathbb{N}$ such that any subset $A \subseteq [N]$ with $|A| \geq \delta N$ contains an m-term arithmetic progression.

The proof of this theorem is usually the subject of an entire lecture course, so we are not going to attempt to prove this. Even the particular case $m = 3$ is very hard.

This theorem has a lot of very nice Ramsey consequences. In the case of graph colourings, we asked ourselves what happens if we colour a graph with *infinitely* many colours. Suppose we now have a colouring $c : \mathbb{N} \to \mathbb{N}$. Can we still find a monochromatic progression of length m? Of course not, because c can be injective.

Theorem. For any $c : \mathbb{N} \to \mathbb{N}$, there exists a m-AP on which either

  (i) c is constant; or

 (ii) c is injective.

It is also possible to prove this directly, but it is easy with Szemerédi.

Proof. We set
$$\delta = \frac{1}{m^2(m+1)^2}.$$

We let $N = S(m, \delta)$. We write
$$[N] = A_1 \cup A_2 \cup \cdots \cup A_k,$$

where the A_i's are the colour-classes of $c|_{[N]}$. By choice of N, we are done if $|A_i| \geq \delta N$ for some $1 \leq i \leq k$. So suppose not.

Let's try to count the number of arithmetic progressions in $[N]$. There are more than $N^2/(m+1)^2$ of these, as we can take any $a, d \in [N/m+1]$. We want to show that there is an AP that hits each A_i at most once.

So, fixing an i, how many AP's are there that hit A_i at least twice? We need to pick two terms in A_i, and also decide which two terms in the progression they are in, e.g. they can be the first and second term, or the 5th and 17th term. So there are at most $m^2|A_i|^2$ terms.

So the number of AP's on which c is injective is greater than
$$\frac{N^2}{(m+1)^2} - k|A_i|^2 m^2 \geq \frac{N^2}{(m+1)^2} - \frac{1}{\delta}(\delta N)^2 m^2 = \frac{N^2}{(m+1)^2} - \delta N^2 m^2 \geq 0.$$

So we are done. Here the first inequality follows from the fact that $\sum |A_i| = [N]$ and each $|A_i| < \delta N$. $\qquad\square$

Our next theorem will mix arithmetic and graph-theoretic properties. Consider a colouring $c : \mathbb{N}^{(2)} \to [2]$. As before, we say a set X is monochromatic if $c|_{X^{(2)}}$ is constant. Now we want to try to find a monochromatic set with some arithmetic properties.

The first obvious question to ask is — can we find a monochromatic 3-term arithmetic progression? The answer is no. For example, we can define $c(ij)$ to be the parity of largest power of 2 dividing $j - i$, and then there is no monochromatic 3-term arithmetic progression.

What if we make some concessions — can we find a blue 10-AP, or if not, find an infinite red set? The answer is again no. This construction is slightly more involved. To construct a counterexample, we can make progressively larger red cliques, and take all other edges blue. If we double the size of the red cliques every time, it is not hard to check that there is no blue 4-AP, and no infinite red set.

What if we further relax our condition, and only require an arbitrarily large red set?

Theorem. For any $c : \mathbb{N}^{(2)} \to \{\text{red}, \text{blue}\}$, either

(i) There exists a blue m-AP for each $m \in \mathbb{N}$; or

(ii) There exists arbitrarily large red sets.

Proof. Suppose we can't find a blue m-AP for some fixed m. We induct on r, and try to find a red set of size r.

Say $A \subseteq \mathbb{N}$ is a progression of length N. Since A has no blue m-term progression, so it must contain many red edges. Indeed, each m-AP in A must contain a red edge. Also each edge specifies two points, and this can be extended to an m-term progression in at most m^2 ways. Since there are $N^2/(m+1)^2$. So there are at least
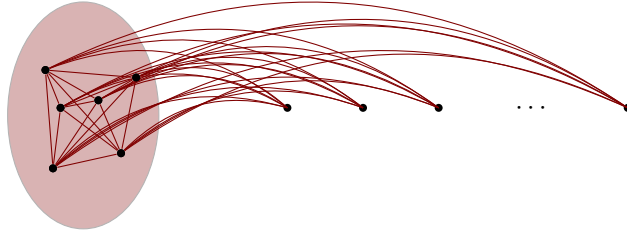
$$\frac{N^2}{m^2(m+1)^2}$$

red edges. With the benefit of hindsight, we set

$$\delta = \frac{1}{2m^2(m+1)^2}.$$

The idea is that since we have lots of red edges, we can try to find a point with a lot of red edges connected to it, and we hope to find a progression in it.

We say $X, Y \subseteq \mathbb{N}$ form an (r, k)-structure if

(i) They are disjoint

(ii) X is red;

(iii) Y is an arithmetic progression;

(iv) All X-Y edges are red;

(v) $|X| = r$ and $|Y| = k$.

We show by induction that there is an (r, k)-structure for each r and k.

A $(1, k)$ structure is just a vertex connected by red edges to a k-point structure. If we take $N = S(\delta, k)$, we know among the first N natural numbers, there are at least $N^2/(m^2(m+1)^2)$ red edges inside $[N]$. So in particular, some $v \in [N]$ has at least δN red neighbours in $[N]$, and so we know v is connected to some k-AP by red edges. That's the base case done.

Now suppose we can find an $(r - 1, k')$-structure for all $k' \in \mathbb{N}$. We set

$$k' = S\left(\frac{1}{2m^2(m+1)^2}, k\right).$$

We let (X, Y) be an $(r - 1, k')$-structure. As before, we can find $v \in Y$ such that v has $\delta|Y|$ red neighbours in Y. Then we can find a progression Y' of

length k in the red neighbourhood of v, and we are done, as $(X \cup \{v\}, Y')$ is an (r, k)-structure, and an "arithmetic progression" within an arithmetic progression is still an arithmetic progression. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Before we end this chapter, we make a quick remark. Everything we looked for in this chapter involved the additive structure of the naturals. What about the multiplicative structure? For example, given a finite colouring of $\mathbb{N}$, can we find a monochromatic geometric progression? The answer is trivially yes. We can look at $\{2^x : x \in \mathbb{N}\}$, and then multiplication inside this set just looks like addition in the naturals.

But what if we want to mix additive and multiplicative structures? For example, can we always find a monochromatic set of the form $\{x + y, xy\}$? Of course, there is the trivial answer $x = y = 2$, but is there any other? This question was answered positively in 2016! We will return to this at the end of the course.

3 Partition Regularity

In the previous chapter, the problems we studied were mostly "linear" in nature. We had some linear system, namely that encoding the fact that a sequence is an AP, and we wanted to know if it had a monochromatic solution. More generally, we can define the following:

Definition (Partition regularity)**.** We say an $m \times n$ matrix A over $\mathbb{Q}$ is *partition regular* if whenever $\mathbb{N}$ is finitely coloured, there exists an $x \in \mathbb{N}^n$ such that $Ax = 0$ and x is monochromatic, i.e. all coordinates of x have the same colour.

Recall that $\mathbb{N}$ does not include 0.

There is no loss in generality by assuming A in fact has entries in $\mathbb{Z}$, by scaling A, but sometimes it is (notationally) convenient to consider cases where the entries are rational.

The question of the chapter is the following — when is a matrix partition regular? We begin by looking at some examples.

Example (Schur's theorem)**.** *Schur's theorem* says whenever $\mathbb{N}$ is finitely coloured, there exists a monochromatic set of the form $\{x, y, x + y\}$. In other words the matrix $\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ is partition regular, since

$$\begin{pmatrix} 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff z = x + y.$$

Example. How about $\begin{pmatrix} 2 & 3 & -5 \end{pmatrix}$. This is partition regular, because we can pick any x, and we have

$$\begin{pmatrix} 2 & 3 & -5 \end{pmatrix} \begin{pmatrix} x \\ x \\ x \end{pmatrix} = 0.$$

This is a trivial example.

How about van der Waerden's theorem?

Example. Van der Waerden's theorem says there is a monochromatic 3-AP $\{x_1, x_2, x_3\}$ whenever $\mathbb{N}$ is finitely-coloured. We know x_1, x_2, x_3 forms a 3-AP iff

$$x_3 - x_2 = x_2 - x_1,$$

or equivalently

$$x_3 + x_1 = 2x_2.$$

This implies that $\begin{pmatrix} 1 & -2 & 1 \end{pmatrix}$ is partition regular. But this is actually not a very interesting thing to say, because $x_1 = x_2 = x_3$ is always a solution to this equation. So this falls into the previous "trivial" case.

On the second example sheet we saw a stronger version of van der Waerden. It says we can always find a monochromatic set of the form

$$\{d, a, a + d, a + 2d, \cdots, a + md\}.$$

By including this variable, we can write down the property of being a progression in a non-trivial format by

$$
\begin{pmatrix} 1 & 1 & 0 & -1 \\ 2 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ a \\ x_2 \\ x_3 \end{pmatrix} = 0
$$

This obviously generalizes to an arbitrary m-AP, with matrix

$$
\begin{pmatrix}
1 & 1 & -1 & 0 & 0 & \cdots & 0 \\
1 & 2 & 0 & -1 & 0 & \cdots & 0 \\
1 & 3 & 0 & 0 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & m & 0 & 0 & 0 & \cdots & -1
\end{pmatrix}.
$$

We've seen three examples of partition regular matrices. Of course, not every matrix is partition regular. The matrix $\begin{pmatrix} 1 & 1 \end{pmatrix}$ is not partition regular, for the silly reason that two positive things cannot add up to zero.

Let's now look at some non-trivial first matrix that is not partition regular.

Example. The matrix $\begin{pmatrix} 2 & -1 \end{pmatrix}$ is not partition regular, since we can put $c(x) = (-1)^n$, where n is the maximum integer such that $2^n \mid x$. Then $\{x, 2x\}$ is never monochromatic.

A similar argument shows that if $\lambda \in \mathbb{Q}$ is such that $\begin{pmatrix} \lambda & -1 \end{pmatrix}$ is partition regular, then $\lambda = 1$.

But if we write down a random matrix, say $\begin{pmatrix} 2 & 3 & -6 \end{pmatrix}$? The goal of this chapter is to find a complete characterization of matrices that are partition regular.

Definition (Columns property)**.** Let

$$
A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ c^{(1)} & c^{(2)} & \cdots & c^{(n)} \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.
$$

We say A has the *columns property* if there is a partition $[n] = B_1 \cup B_2 \cup \cdots \cup B_d$ such that

$$
\sum_{i \in B_s} c^{(i)} \in \operatorname{span}\{c^{(i)} : c^{(i)} \in B_1 \cup \cdots \cup B_{s-1}\}
$$

for $s = 1, \cdots, d$. In particular,

$$
\sum_{i \in B_1} c^{(i)} = 0
$$

What does this mean? Let's look at the matrices we've seen so far.

Example. $\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ has the columns property by picking $B_1 = \{1, 3\}$ and $B_2 = \{2\}$.

Example. $\begin{pmatrix} 2 & 3 & -5 \end{pmatrix}$ has the columns property by picking $B_1 = \{1, 2, 3\}$.

Example. The matrix

$$\begin{pmatrix} 1 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & -1 & 0 & \cdots & 0 \\ 1 & 3 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & m & 0 & 0 & 0 & \cdots & -1 \end{pmatrix}$$

has the column property. Indeed, we take $B_1 = \{1, 3, \cdots, m+2\}$, and since $\{c^{(3)}, \cdots, c^{(m+2)}\}$ spans all of $\mathbb{R}^m$, we know picking $B_2 = \{2\}$ works.

Example. $\begin{pmatrix} \ell & -1 \end{pmatrix}$ has the columns property iff $\ell = 1$. In particular, $\begin{pmatrix} 1 & 1 \end{pmatrix}$ does not have a columns property.

Given these examples, it is natural to conjecture the following:

Theorem (Rado's theorem)**.** A matrix A is partition regular iff it has the column property.

This is a remarkable theorem! The property of being partition regular involves a lot of quantifiers, over infinitely many things, but the column property is entirely finite, and we can get a computer to check it for us easily.

Another remarkable property of this theorem is that neither direction is obvious! It turns out partition regularity implies the column property is (marginally) easier, because if we know something is partition regular, then we can try to cook up some interesting colourings and see what happens. The other direction is harder.

To get a feel of the result, we will first prove it in the case of a single equation. The columns property in particular becomes much simpler. It just means that there exists a non-empty subset of the non-zero a_i's that sums up to zero.

Theorem. If $a_1, \cdots, a_n \in \mathbb{Q} \setminus \{0\}$, then $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is partition regular iff there exists a non-empty $I \subseteq [n]$ such that

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

For a fixed prime p, we let $d(x)$ denote the last non-zero digit of x in base p, i.e. if

$$x = d_n p^n + d_{n-1} p^{n-1} + \cdots + d_0,$$

then

$$L(x) = \min\{i : d_i \neq 0\}$$

and $d(x) = d_{L(x)}$. We now prove the easy direction of the theorem.

Proposition. If $a_2, \cdots, a_n \in \mathbb{Q} \backslash \{0\}$ and $\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ is partition regular, then

$$\sum_{i \in I} a_i = 0$$

for some non-empty I.

Proof. We wlog $a_i \in \mathbb{Z}$, by scaling. Fix a suitably large prime

$$p > \sum_{i=1}^{n} |a_i|,$$

and consider the $(p-1)$-colouring of $\mathbb{N}$ where x is coloured $d(x)$. We find $x_1, \cdots, x_n$ such that

$$\sum a_i x_i = 0.$$

and $d(x_i) = d$ for some $d \in \{1, \cdots, p-1\}$. We write out everything in base p, and let

$$L = \min\{L(x_i) : 1 \leq i \leq n\},$$

and set

$$I = \{i : L(x_i) = L\}.$$

Then for all $i \in I$, we have

$$x_i \equiv d \pmod{p^{L+1}}.$$

On the other hand, we are given that

$$\sum a_i x_i = 0.$$

Taking mod p^{L+1} gives us

$$\sum_{i \in I} a_i d = 0 \pmod{p^{L+1}}.$$

Since d is invertible, we know

$$\sum_{i \in I} a_i = 0 \pmod{p^{L+1}}.$$

But by our choice of p, this implies that $\sum_{i \in I} a_i = 0$. $\square$

Here we knew what prime p to pick ahead. If we didn't, then we could consider *all* primes p, and for each p we find $I_p \subseteq [n]$ such that

$$\sum_{i \in I_p} a_i = 0 \pmod{p}.$$

Then some I_p has to occur infinitely often, and then we are done. Note that this is merely about the fact that if a fixed number is 0 mod n for arbitrarily large n, then it must be zero. This is not some deep local-global correspondence number theoretic fact.

It turns out this is essentially the only way we know for proving this theorem. One possible variation is to use the "first non-zero digit" to do the colouring, but this is harder.

Let's now try and do the other direction. Before we do that, we start by doing a warm up. Last time, we proved that if we had $\begin{pmatrix} 1 & \lambda \end{pmatrix}$, then this is partition regular iff $\lambda = -1$. We now prove a three element version.

Proposition. The equation $\begin{pmatrix} 1 & \lambda & -1 \end{pmatrix}$ is partition regular for all $\lambda \in \mathbb{Q}$.

Proof. We may wlog $\lambda > 0$. If $\lambda = 0$, then this is trivial, and if $\lambda < 0$, then we can multiply the whole equation by -1.

Say
$$\lambda = \frac{r}{s}.$$

The idea is to try to find solutions of this in long arithmetic progressions.

Suppose $\mathbb{N}$ is k-coloured. We let
$$\{a, a + d, \cdots, a + nd\}$$

be a monochromatic AP, for n sufficiently large.

If sd were the same colour as this AP, then we'd be done, as we can take
$$x = a, \quad y = sd, \quad z = a + \frac{r}{s}sd.$$

In fact, if any of $sd, 2sd, \cdots, \left\lfloor \frac{n}{r} \right\rfloor sd$ have the same colour as the AP, then we'd be done by taking
$$x = a, \quad y = isd, \quad z = a + \frac{r}{s}isd = a + ird \leq a + nd.$$

If this were not the case, then $\left\{ sd, 2sd, \cdots, \left\lfloor \frac{n}{r} \right\rfloor sd \right\}$ is $(k-1)$-coloured, and this is just a scaled up copy of $\mathbb{N}$. So we are done by induction on k. $\qquad\square$

Note that when $\lambda = 1$, we have $\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ is partition regular, and this may be proved by Ramsey's theorem. Can we prove this more general result by Ramsey's theorem as well? The answer is, we don't know.

It turns out this is not just a warm up, but the main ingredient of what we are going to do.

Theorem. If $a_1, \cdots, a_n \in \mathbb{Q}$, then $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is partition regular iff there exists a non-empty $I \subseteq [n]$ such that
$$\sum_{i \in I} a_i = 0.$$

Proof. One direction was done. To do the other direction, we recall that we had a really easy case of, say,
$$\begin{pmatrix} 2 & 3 & -5 \end{pmatrix},$$
because we can just make all the variables the same?

In the general case, we can't quite do this, but we may try to solve this equation with the least number of variables possible. In fact, we shall find some monochromatic x, y, z, and then assign each of $x_1, \cdots, x_n$ to be one of x, y, z.

We know
$$\sum_{i \in I} a_i = 0.$$

We now partition I into two pieces. We fix $i_0 \in I$, and set
$$x_i = \begin{cases} x & i = i_0 \\ y & i \in I \setminus \{i_0\} \\ z & i \notin I \end{cases}.$$

We would be done if whenever we finitely colour $\mathbb{N}$, we can find monochromatic x, y, z such that

$$a_{i_0} x + \left(\sum_{i \in I \setminus \{i_0\}} a_i \right) z + \left(\sum_{i \notin I} a_i \right) y = 0.$$

But, since

$$\sum_{i \in I} a_i = 0,$$

this is equivalent to

$$a_{i_0} x - a_{i_0} z + (\text{something}) y = 0.$$

Since all these coefficients were non-zero, we can divide out by a_{i_0}, and we are done by the previous case. $\qquad\square$

Note that our proof of the theorem shows that if an equation is partition regular for all last-digit base p colourings, then it is partition regular for all colourings. This might sound like an easier thing to prove that the full-blown Rado's theorem, but it turns out the only proof we have for this implication is Rado's theorem.

We now prove the general case. We first do the easy direction, because it is largely the same as the single-equation case.

Proposition. If A is an $m \times n$ matrix with rational entries which is partition regular, then A has the columns property.

Proof. We again wlog all entries of A are integers. Let the columns of A be $c^{(1)}, \cdots, c^{(n)}$. Given a prime p, we consider the $(p-1)$-colouring of $\mathbb{N}$, where x is coloured $d(x)$, the last non-zero digit in the base p expansion. Since A is partition regular, we obtain a monochromatic solution.

We then get a monochromatic $x_1, \cdots, x_n$ such that $Ax = 0$, i.e.

$$\sum x_i c^{(i)} = 0.$$

Any such solution with colour d induces a partition of $[n] = B_1 \cup B_2 \cup \cdots \cup B_r$, where

- For all $i, j \in B_s$, we have $L(x_i) = L(x_j)$; and

- For all $s < t$ and $i \in B_s$, $j \in B_t$, the $L(x_i) < L(x_j)$.

Last time, with the benefit of hindsight, we were able to choose some large prime p that made the argument work. So we use the trick we mentioned after the proof last time.

Since there are finitely many possible partitions of $[n]$, we may assume that this particular partition is generated by infinitely many primes p. Call these primes $\mathcal{P}$. We introduce some notation. We say two vectors $u, v \in \mathbb{Z}^m$ satisfy $u \equiv v \pmod{p}$ if $u_i \equiv v_i \pmod{p}$ for all $i = 1, \cdots, m$.

Now we know that

$$\sum x_i c^{(i)} = 0.$$

Looking at the first non-zero digit in the base p expansion, we have

$$\sum_{i \in B_1} dc^{(i)} \equiv 0 \pmod{p}.$$

From this, we conclude that, by multiplying by d^{-1}

$$\sum_{i \in B_1} c^{(i)} \equiv 0 \pmod{p},$$

for all $p \in \mathcal{P}$. So we deduce that

$$\sum_{i \in B_1} c^{(i)} = 0.$$

Similarly, for higher s, we find that for each base p colouring, we have

$$\sum_{i \in B_s} p^t dc^{(i)} + \sum_{i \in B_1 \cup \ldots \cup B_s} x_i c^{(i)} \equiv 0 \pmod{p^{t+1}}$$

for all $s \geq 2$, and some t dependent on s and p. Multiplying by d^{-1}, we find

$$\sum_{i \in B_s} p^t c^{(i)} + \sum_{i \in B_1 \cup \ldots \cup B_{s-1}} (d^{-1} x_i) c^{(i)} \equiv 0 \pmod{p^{t+1}}. \qquad (*)$$

We claim that this implies

$$\sum_{i \in \mathcal{B}_s} c^{(i)} \in \langle c^{(i)} : i \in B_1 \cup \cdots \cup B_{s-1} \rangle.$$

This is not exactly immediate, because the values of x_i in $(*)$ may change as we change our p. But it is still some easy linear algebra.

Suppose this were not true. Since we are living in a Euclidean space, we have an inner product, and we can find some $v \in \mathbb{Z}^m$ such that

$$\langle v, c^{(i)} \rangle = 0 \text{ for all } i \in B_1 \cup \cdots \cup B_{s-1},$$

and

$$\left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \neq 0.$$

But then, taking inner products with gives

$$p^t \left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \equiv 0 \pmod{p^{t+1}}.$$

Equivalently, we have

$$\left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \equiv 0 \pmod{p},$$

but this is a contradiction. So we showed that A has the columns property with respect to the partition $[n] = B_1 \cup \cdots \cup B_r$. $\qquad \square$

We now prove the hard direction. We want an analogous gadget to our $\begin{pmatrix} 1 & \lambda & -1 \end{pmatrix}$ we had for the single-equation case. The definition will seem rather mysterious, but it turns out to be what we want, and its purpose becomes more clear as we look at some examples.

Definition $((m, p, c)$-set)**.** For $m, p, c \in \mathbb{N}$, a set $S \subseteq \mathbb{N}$ is an (m, p, c)-set with generators $x_1, \cdots, x_m$ if S has the form

$$S = \left\{ \sum_{i=0}^{m} \lambda_i x_i : \begin{array}{c} \lambda_i = 0 \text{ for all } i < j \\ \lambda_j = c \\ \lambda_i \in [-p, p] \end{array} \right\}.$$

In other words, we have

$$S = \bigcup_{j=1}^{m} \{ cx_j + \lambda_{j+1} x_{j+1} + \cdots + \lambda_m x_m : \lambda_i \in [-p, p] \}.$$

For each j, the set $\{ cx_j + \lambda_{j+1} x_{j+1} + \cdots + \lambda_m x_m : \lambda_i \in [-p, p] \}$ is called a *row* of S.

Example. What does a $(2, p, 1)$ set look like? It has the form

$$\{ x_1 - px_2, x_1 - (p-1)x_2, \cdots, x_1 + px_2 \} \cup \{ x_2 \}.$$

In other words, this is just an arithmetic progression with its common difference.

Example. A $(2, p, 3)$-set has the form

$$\{ 3x_1 - px_2, \cdots, 3x_1, \cdots, 3x_1 = px_2 \} \cup \{ 3x_2 \}.$$

The idea of an (m, p, c) set is that we "iterate" this process many times, and so an (m, p, c)-set is an "iterated AP's and various patterns of their common differences".

Our proof strategy is to show that that whenever we finitely-colour $\mathbb{N}$, we can always find an (m, p, c)-set, and given any matrix A with the columns property and any (m, p, c)-set (for suitable m, p, c), there will always be a solution in there.

Proposition. Let $m, p, c \in \mathbb{N}$. Then whenever $\mathbb{N}$ is finitely coloured, there exists a monochromatic (m, p, c)-set.

Proof. It suffices to find an (m, p, c)-set all of whose rows are monochromatic, since when $\mathbb{N}$ is k-coloured, and (m', p, c)-set with $m' = mk + 1$ has m monochromatic rows of the same colour by pigeonhole, and these rows contain a monochromatic (m, p, c)-set, by restricting to the elements where a lot of the λ_i are zero. In this proof, whenever we say (m, p, c)-set, we mean one all of whose rows are monochromatic.

We will prove this by induction. We have a k-colouring of $[n]$, where n is very very very large. This contains a k-colouring of

$$B = \left\{ c, 2c, \cdots, \left\lfloor \frac{n}{c} \right\rfloor c \right\}.$$

Since c is fixed, we can pick this so that $\frac{n}{c}$ is large. By van der Waerden, we find some set monochromatic

$$A = \{cx_1 - Nd, cx_1 - (N-1)d, \cdots, cx_1 + Nd\} \subseteq B,$$

with N very very large. Since each element is a multiple of c by assumption, we know that $c \mid d$. By induction, we may find an $(m-1, p, c)$-set in the set $\{d, 2d, \cdots, Md\}$, where M is large. We are now done by the (m, p, c) set on generators $x_1, \cdots, x_n$, provided

$$cx_1 + \sum_{i=2}^{m} \lambda_i x_i \in A$$

for all $\lambda_i \in [-p, p]$, which is easily seen to be the case, provided $N \geq (m-1)pM$. $\qquad\square$

Note that the argument itself is quite similar to that in the $\begin{pmatrix} 1 & \lambda & -1 \end{pmatrix}$ case.

Recall that Schur's theorem said that whenever we finitely-colour $\mathbb{N}$, we can find a monochromatic $\{x, y, x+y\}$. More generally, for $x_1, x_2, \cdots, x_m \in \mathbb{N}$, we let

$$FS(x_1, \cdots, x_m) = \left\{\sum_{i \in I} x_i : I \subseteq [n], I \neq \emptyset\right\}.$$

The existence of a monochromatic $(m, 1, 1)$-sets gives us

Corollary (Finite sum theorem)**.** For every fixed m, whenever we finitely-colour $\mathbb{N}$, there exists $x_1, \cdots, x_m$ such that $FS(x_1, \cdots, x_m)$ is monochromatic.

This is since an $(m, 1, 1)$-set contains more things than $FS(x_1, \cdots, x_m)$. This was discovered independently by a bunch of different people, including Folkman, Rado and Sanders.

Similarly, if we let

$$FP(x_1, \cdots, x_m) = \left\{\prod_{i \in I} x_i : I \subseteq [n], I \neq \emptyset\right\},$$

then we can find these as well. For example, we can restrict attention to $\{2^n : n \in \mathbb{N}\}$, and use the finite sum theorem. This is the same idea as we had when we used van der Waerden to find geometric progressions.

But what if we want both? Can we have $FS(x_1, \cdots, x_m) \cup FP(x_1, \cdots, x_m)$ in the same colour class? The answer is actually not known! Even the case when $m = 2$, i.e. finding a monochromatic set of the form $\{x, y, x+y, xy\}$ is open. Until mid 2016, we did not know if we can find $\{x+y, xy\}$ monochromatic $(x, y > 2)$.

To finish he proof of Rado's theorem, we need the following proposition:

Proposition. If A is a rational matrix with the columns property, then there is some $m, p, c \in \mathbb{N}$ such that $Ax = 0$ has a solution inside any (m, p, c) set, i.e. all entries of the solution lie in the (m, p, c) set.

In the case of a single equation, we reduced the general problem to the case of 3 variables only. Here we are going to do something similar — we will use the columns property to reduce the solution to something much smaller.

Proof. We again write

$$
A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ c^{(1)} & c^{(2)} & \cdots & c^{(n)} \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.
$$

Re-ordering the columns of A if necessary, we assume that we have

$$
[n] = B_1 \cup \cdots \cup B_r
$$

such that $\max(B_s) < \min(B_{s+1})$ for all s, and we have

$$
\sum_{i \in B_s} c^{(i)} = \sum_{i \in B_1 \cup \ldots \cup B_{s-1}} q_{is} c^{(i)}
$$

for some $q_{is} \in \mathbb{Q}$. These q_{is} only depend on the matrix. In other words, we have

$$
\sum d_{is} c^{(i)} = 0,
$$

where

$$
d_{is} = \begin{cases} -q_{is} & i \in B_1 \cup \cdots B_{s-1} \\ 1 & i \in B_s \\ 0 & \text{otherwise} \end{cases}
$$

For a fixed s, if we scan these coefficients starting from $i = n$ and then keep decreasing i, then the first non-zero coefficient we see is 1, which is good, because it looks like what we see in an (m, p, c) set.

Now we try to write down a general solution with r many free variables. Given $x_1, \cdots, x_r \in \mathbb{N}^r$, we look at

$$
y_i = \sum_{s=1}^{r} d_{is} x_s.
$$

It is easy to check that $Ay = 0$ since

$$
\sum y_i c^{(i)} = \sum_i \sum_s d_{is} x_s c^{(i)} = \sum_s x_s \sum_i d_{is} c^{(i)} = 0.
$$

Now take $m = r$, and pick c large enough such that $cd_{is} \in \mathbb{Z}$ for all i, s, and finally, $p = \max\{cd_{is} : i, s \in \mathbb{Q}\}$.

Thus, if we consider the (m, p, c)-set on generators $(x_m, \cdots, x_1)$ and y_i as defined above, then we have $Ay = 0$ and hence $A(cy) = 0$. Since cy is integral, and lies in the (m, p, c) set, we are done! □

We have thus proved Rado's theorem.

Theorem (Rado's theorem)**.** A matrix A is partition regular iff it has the column property.

So we have a complete characterization of all partition regular matrices.

Note that Rado's theorem reduces Schur's theorem, van der Waerden's theorem, finite sums theorem etc. to just checking if certain matrices have the columns property, which are rather straightforward computations.

More interestingly, we can prove some less obvious "theoretical" results.

Corollary (Consistency theorem)**.** If A, B are partition regular in independent variables, then

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

is partition regular. In other words, we can solve $Ax = 0$ and $By = 0$ simultaneously in the same colour class.

Proof. The matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

has the columns property if A and B do. $\qquad\qquad\qquad\qquad\qquad\square$

In fact, much much more is true.

Corollary. Whenever $\mathbb{N}$ is finitely-coloured, one colour class contains solutions to *all* partition regular systems!

Proof. Suppose not. Then we have $\mathbb{N} = D_1 \cup \cdots \cup D_k$ such that for each D_i, there is some partition regular matrix A_i such that we cannot solve $A_i x = 0$ inside D_i. But this contradicts the fact that $\operatorname{diag}(A_1, A_2, \cdots, A_k)$ is partition regular (by applying consistency theorem many times). $\qquad\square$

Where did this whole idea of the (m, p, c)-sets come from? The original proof by Rado didn't use (m, p, c)-sets, and this idea only appeared only a while later, when we tried to prove a more general result.

In general, we call a set $D \subseteq \mathbb{N}$ *partition regular* if we can solve any partition regular system in D. Then we know that $\mathbb{N}, 2\mathbb{N}$ are partition regular sets, but $2\mathbb{N} + 1$ is not (because we can't solve $x + y = z$, say). Then what Rado's theorem says is that whenever we finitely partition $\mathbb{N}$, then one piece of $\mathbb{N}$ is partition regular.

In the 1930's, Rado conjectured that there is nothing special about $\mathbb{N}$ to begin with — whenever we break up a partition regular set, then one of the pieces is partition regular. This conjecture was proved by Deuber in the 1970s who introduced the idea of the (m, p, c)-set.

It is not hard to check that D is partition regular iff D contains an (m, p, c) set of each size. Then Deuber's proof involves showing that for all $m, p, c, k \in \mathbb{N}$, there exists $n, q, d \in \mathbb{N}$ such that any k-colouring of an (n, q, d)-set contains a monochromatic (m, p, c) set. The proof is quite similar to how we found (m, p, c)-sets in the naturals, but instead of using van der Waerden theorem, we need the Hales–Jewett theorem.

We end by mentioning an open problem in this area. Suppose A is an $m \times n$ matrix A that is not partition regular. That is, there is some k-colouring of $\mathbb{N}$ with no solution to $Ax = 0$ in a colour class. Can we find some bound $f(m, n)$, such that every such A has a "bad" colouring with $k < f(m, n)$? This is an open problem, first conjectured by Rado, and we think the answer is yes.

What do we actually know about this problem? The answer is trivially yes for $f(1, 2)$, as there aren't many matrices of size 1×2, up to rescaling. It is a non-trivial theorem that $f(1, 3)$ exists, and in fact $f(1, 3) \leq 24$. We don't know anything more than that.

4 Ultrafilters and Hindman's Theorem

Recall the finite sum theorem — for any fixed k, whenever $\mathbb{N}$ is finitely coloured, we can find $x_1, \cdots, x_k$ such that $FS(x_1, x_2, \cdots, x_k)$ is monochromatic. One natural generalization is to ask the following question — can we find an infinite set A such that

$$FS(A) = \left\{ \sum_{x \in B} x : B \subseteq A \text{ finite non-empty} \right\}$$

is monochromatic? The first thought upon hearing this question is probably either "this is so strong that there is no way it can be true", or "we can deduce it easily, perhaps via some compactness argument, from the finite sum theorem".

Both of these thoughts are wrong. It is true that it is always possible to find these $x_1, x_2, \cdots$, but the proof is hard. This is *Hindman's theorem.*

We will venture into the world of topology to prove this theorem for reasons that only become apparent once we are on our way, so let's begin our journey.

Definition (Filter). A filter is $\mathcal{F} \subset P(\mathbb{N})$ such that

i) $\emptyset \notin \mathcal{F}$, $\mathbb{N} \in \mathcal{F}$

ii) $A \in \mathcal{F}$, $B \subset A \implies B \in \mathcal{F}$ *"closed under superset"*

iii) $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$ *"closed under finite intersections"*

Example.   – $\{A \subset \mathbb{N} : 3 \in A\}$

  – $\{A \subset \mathbb{N} : 3, 4 \in A\}$

  – NOT $\{A \in \mathbb{N} : A \text{ is infinite }\}$ since $E \cap O = \emptyset \notin \mathcal{F}$ where $E = evens$ and $O = odds$

  – $\{A \subset \mathbb{N} : A^c \text{ finite }\}$ *"cofinite filter"*

  – $\{A \in \mathbb{N} : E - A \text{finite}\}$

A stronger notion of filters is needed to define the topological space required, for this we define ultrafilters.

Definition (Ultrafilter). An ultrafilter is a maximal filter, i.e it cannot be extended.

Example.   1: This is an ultrafilter since any set not in $\mathcal{F}$ does not contain 3, thus it's intersection with $\{3\} \in \mathcal{F}$ is $\emptyset$.

  2: Not an ultrafilter since 1 extends it.

  4: Not an ultrafilter since 5 extends it.

  5: Not an u;trafilter since $\{A \subset \mathbb{N} : M - A \text{ finite }\}$ extends it where $M = \{n : 4|n\}$

Definition (Principal Ultrafilter). The principal ultrafilters are $\tilde{n} = \{A \subset \mathbb{N} : n \in A\}$

The following proposition gives us a helpful characterisation of ultrafilters.

Proposition. A filter $\mathcal{F}$ is an ultrafilter iff $\forall A \subset \mathbb{N}, A \in \mathcal{F}$ or $A^c \in \mathcal{F}$

Proof. ($\Leftarrow$) If $A \in \mathcal{F}$ then cannot add A^c else we have $A \cap A^c = \emptyset$

($\Rightarrow$) Given $A \notin \mathcal{F}$: Want $A^c \in \mathcal{F}$. Must have $B \cap A = \emptyset$ for some $B \in \mathcal{F}$ else the filter $\mathcal{F}' = \{C \subset \mathbb{N} : C \supset A \cap B \text{ some } B \in F \text{ would extend } \mathcal{F}$. Thus $B \subset A^c$, whence $A^c \in \mathcal{F}$

$\square$

Remark. If $\mathcal{U}$ is an ultrafilter and $A \in \mathcal{U}$ with $A = B \cup C$. Then $B \in \mathcal{U}$ or $C \in \mathcal{U}$, else $B^c \in \mathcal{U}$ and $C^c \in \mathcal{U}$ so $B^c \cup C^c = A^c \in \mathcal{U}$ which is a contradiction.

We now prove that all filters are housed an ultrafilter.

Theorem. Every filter is contained in an ultrafilter

Use Zorn's Lemma

Proof. Given a filter $\mathcal{F}_0$, we seek a maximal filter extending $\mathcal{F}_0$. By Zorn's Lemma, since $\mathbb{P}(\mathbb{N})$ is a poset, it is enough to show that every non-empty chain of filters $(\mathcal{F}_i : i \in I)$ extending $\mathcal{F}_0$ has an upper bound.
As in every Zorn proof ever seen, we let $\mathcal{F} =_{i \in I} \mathcal{F}_i$ and claim that this is an upper bound. Clearly $\mathcal{F} \supset \mathcal{F}_i$ and $\mathcal{F} \supset \mathcal{F}_0$ so just need to show that $\mathcal{F}$ is a filter:

i) $\emptyset \notin \mathcal{F}$, $\mathbb{N} \in \mathcal{F}$ since $\emptyset \notin \mathcal{F}_i$ for any i and $\mathbb{N} \in \mathcal{F}_i$ for all i.

ii) Given $A \in \mathcal{F}$ and $B \subset A$ have $A \in \mathcal{F}_i$ for some i and so $B \in \mathcal{F}_i$ since $\mathcal{F}_i$ is closed under superset. Therefore $B \in \mathcal{F}$.

iii) Given $A, B \in \mathcal{F}$: Have $A, B \in \mathcal{F}_i$ for some i as the $\mathcal{F}_i$ are nested since they form a chain. Therefore $A \cap B \in \mathcal{F}_i$ since $\mathcal{F}_i$ is closed under finite intersection. Thus $A \cap B \in \mathcal{F}$ $\square$

Remark. – From this we get that some filter extends the cofinite filter. Any ultrafilter extending the cofinite filter is non-principal. Conversely, if $\mathcal{U}$ is a non-prinipal ultrafilter then $\mathcal{U}$ extends the cofinite filter, else $\mathcal{U}$ has some finite $A \in \mathcal{U}$, so by the remark regarding '$B \cup C \in \mathcal{U}$' we get $\{n\}'in\mathcal{U}$ for some n.

– We do need some form of Axiom of Choice (Here we have used the AOC through using Zorn, since the two are equivalent) to get a non-principal ultrafilter.

Definition ($\beta\mathbb{N}$)**.** We define $\beta\mathbb{N} \subset \mathbb{P}(\mathbb{N})$ as follows

$$\beta\mathbb{N} = \{\mathcal{U} : \mathcal{U} \text{ is an ultrafilter}\}$$

We want to define a topology on this set. We can define the topology in afew ways.
First we can give the base of open sets $C_A = \{\mathcal{U} : A \in \mathcal{U}\}$ for each $\in \mathbb{N}$. This is a base for a topology, since:

- $_AC_A = \beta\mathbb{N}$
- $C_A \cap C_B = C_{A\cap B}$ since $A, B \in \mathcal{U} \equiv A \cap B \in \mathcal{U}$. This is since $\mathcal{U}$ is closed under finite intersection and under superset.

Thus the open sets are all $_{i\in I}C_{A_i} = \{\mathcal{U} : \exists u \text{ with } A_i \in \mathcal{U}\}$.
The base of closed sets are also the C_A, $A \subset \mathbb{N}$ because

$$(C_A)^c = \{\mathcal{U} : A \notin \mathcal{U}\} = \{\mathcal{U} : A^c \in \mathcal{U}\} = C_{A^c}$$

and a set is closed if it's complement is open.
Therefore the closed sets are $\bigcap_{i\in I} C_{A_i} = \{\mathcal{U} : A_i \in \mathcal{U} \text{ for all } i\}$.

We can view $\mathbb{N} \subset \beta\mathbb{N}$ by identifying $n \in \mathbb{N}$ with $\tilde{n} \in \beta\mathbb{N}$. Note that each $\tilde{n}$ is an isolated point since $\{\tilde{n} = C_{\{n\}}\}$ and so $\{\tilde{n}\}$ is open.
Also, $\mathbb{N}$ is dense in $\beta\mathbb{N}$ (every non-empty open set meets $\mathbb{N}$). Indeed, for any $A \neq \emptyset$ have $\tilde{n} \in \mathbb{C}_A$ for all $n \in$.

Theorem. $\beta\mathbb{N}$ is a compact Hausdorff space

Proof. We prove both results seperately:

<u>Hausdorff</u> Given $\mathcal{U}, \mathcal{V} \in \beta\mathbb{N}$ with $\mathcal{U} \neq \mathcal{V}$: $\exists A \subset \mathbb{N}$ with $A \in \mathcal{U}$, $A^c \in \mathcal{V}$ so $\mathcal{U} \in C_A$ and $\mathcal{V} \in C_{A^c}$ and C_A, C_{A^c} are disjoint (since cannot have A, A^c in an ultrafilter).
Remember: Hausdorff $\equiv$f two non-equal points have non-empty sets around them that do not intersect.

<u>Compact</u> Given closed sets $(F_i : i \in I)$ with the finite intersection property (all finite intersetions are non-empty) we want to show that $\bigcap_{i\in I} F_i \neq \emptyset$ since this is the contrapositive of the normal definition of compactness.
WLOG each F_i is basic, say $F_i = C_{A_i}$ (If not, can express F_i as the intersection of basic sets). Note that the A_i, $i \in I$ themselves have the finite intersection property: Indeed,

$$C_{A_{i_1}} \cap C_{A_{i_2}} \cap \cdots \cap C_{A_{i_n}} = C_{A_{i_1} \cap \ldots A_{i_n}}$$

and so $A_{i_1} \cap \cdots \cap A_{i_n} \neq \emptyset$ since $\emptyset$ is not in any filter. So we can define a filter generated by the A_i:

$$\mathcal{F} := \{A \subset \mathbb{N} : A \supset A_{i_1} \cap \cdots \cap A_{i_n} \text{ for some } i_1, \ldots, i_n \in I\}$$

Let the ultrafilter $\mathcal{U}$ extend $\mathcal{F}$. Then each $A_i \in \mathcal{U}$, so $\mathcal{U} \in C_{A_i}$ for all i and hence $\mathcal{U} \in \bigcap_{i\in I} C_{A_i}$.

$\square$

(i) If we view an ultrafilter $\mathcal{U} \subset \mathbb{P}(\mathbb{N})$ as a function from $\mathbb{P}(\mathbb{N}) \to \{0, 1\}$ i.e as a point in $\{0,1\}^{\mathbb{P}(\mathbb{N})}$, then we have $\beta\mathbb{N} \subset \{0,1\}^{\mathbb{P}(\mathbb{N})}$. Can check that the product topology on $\{0,1\}^{\mathbb{P}(\mathbb{N})}$ restricts to our topology on $\beta\mathbb{N}$ and that $\beta\mathbb{N}$ is closed in $\{0,1\}^{\mathbb{P}(\mathbb{N})}$ - so $\beta\mathbb{N}$ is compact by Tychonov.

(ii) Why is $\beta\mathbb{N}$ interesting? It is the largest compact Hausdorff space in which $\mathbb{N}$ is dense. More precisely, for any compact Hausdorff X and $f : \mathbb{N} \to X$ there exists a unique continuous extension $\hat{f} : \beta\mathbb{N} \to X$. $\beta\mathbb{N}$ is called the Stone-Cech compactification of $\mathbb{N}$.

4.1   Ultrafilter Quantifiers

Definition ($\forall_{\mathcal{U}}x$)**.** Let $\mathcal{U}$ be an ultrafilter, p a statement. Write $\forall_{\mathcal{U}}x\ p(x)$ if $\{x : p(x) \text{ true}\} \in \mathcal{U}$. This is read "for $\mathcal{U}$-most x, $p(x)$" or just "for most x, $p(x)$".

Example.   – $\forall_{\mathcal{U}}x, x > 4$ ($\mathcal{U}$ non-principal)

– If $\mathcal{U} = \tilde{n}$ then $\forall_{\mathcal{U}}xp(x) \equiv p(n)$

We have that $\forall_{\mathcal{U}}$ "meshes" perfectly with logical connectives.

Proposition. Let $\mathcal{U}$ be an ultrafilter and p, q be statements. Then

  (i) $\forall_{\mathcal{U}}x\ (p(x) \wedge q(x)) \equiv (\forall_{\mathcal{U}}x\ p(x))$ AND $(\forall_{\mathcal{U}}x : q(x))$

 (ii) $\forall_{\mathcal{U}}x\,(p(x) \vee q(x)) \equiv (\forall_{\mathcal{U}}x : p(x))$ OR $(\forall_{\mathcal{U}}x : q(x))$

(iii) $(\forall_{\mathcal{U}}x : p(x))$ FALSE $\equiv \forall_{\mathcal{U}}x\,(\neg p(x))$

Proof. Let $A = \{x \in \mathbb{N} : p(x)\}$, $B = \{x \in \mathbb{N} : q(x)\}$

  (i) Says $A \cap B \in \mathcal{U} \equiv A \in \mathcal{U}$ and $B \in \mathcal{U}$

 (ii) Says $A \cup B \in \mathcal{U} \equiv A \in \mathcal{U}$ and $B \in \mathcal{U}$

(iii) Says $A \notin \mathcal{U} \equiv A^c \in \mathcal{U}$

$\square$

– Conditions (ii) and (iii) are almost rediculous and far to much to ask for.

– It is not true that:

$$\forall_{\mathcal{U}}x\forall_{\mathcal{V}}yp(x,y) \equiv \forall_{\mathcal{V}}y\forall_{\mathcal{U}}xp(x,y)$$

even if $\mathcal{U} = \mathcal{V}$. For example, if $\mathcal{U}$ is non-principal, then $\forall_{\mathcal{U}}x\forall_{\mathcal{U}}yx < y$ since for all x, most y satisfy $x < y$. However, $\forall_{\mathcal{U}}y\forall_{\mathcal{U}}xx < y$ does not hold since no y has $\forall_{\mathcal{U}}xx < y$.

Definition (+)**.**

$$\mathcal{U} + \mathcal{V} = \{A \subset \mathbb{N} : \forall_{\mathcal{U}}x\forall_{\mathcal{U}}y\ x + y \in A\}$$

Equivalently, without quantifiers we have:

$$\mathcal{U} + \mathcal{V} = \{A \subset \mathbb{N} : \{x \in \mathbb{N} : \{y \in \mathbb{N} : x + y \in \mathbb{N}\}\}\}$$

Example. $\tilde{n} + \tilde{m} = n \tilde{+} m$ so + extends the usual + on $\mathbb{N}$.

Note that $\mathcal{U} + \mathcal{V}$ is an ultrafilter, indeed:

  (i) $\emptyset \notin \mathcal{U} + \mathcal{V}$, $\mathbb{N} \in \mathcal{U} + \mathcal{V}$

 (ii) If $A \in \mathcal{U} + \mathcal{V}$ and $B \supset A$ then $B \in \mathcal{U} + \mathcal{V}$

(iii) If $A, B \in \mathcal{U} + \mathcal{V}$, then

$\forall_{\mathcal{U}}x\forall_{\mathcal{V}}y\ x+y \in A$ and $\forall_{\mathcal{U}}x\forall_{\mathcal{V}}y\ x+y \in B \implies \forall_{\mathcal{U}}x\forall_{\mathcal{V}}y\ (x+y \in A$ and $x+y \in \mathcal{B})$

Which implies $\forall_{\mathcal{U}}x\forall_{\mathcal{V}}y\ (x + y \in A \cap B)$. The first implication follows from (i) of the previous proposition applied twice.

(iv) If $A \notin \mathcal{U} + \mathcal{V}$ then

$$\text{NOT } \forall_{\mathcal{U}} x \forall_{\mathcal{V}} y \ x+y \in A \implies \forall_{\mathcal{U}} x \forall_{\mathcal{V}} y \text{ NOT } x+y \in A \equiv \forall_{\mathcal{U}} x \forall_{\mathcal{V}} y \ x+y \in A^c$$

where the first implication follows from (iii) of the previous proposition applied twice.

We note that $+$ is associative:

$$\mathcal{U} + (\mathcal{V}+) = \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x \forall_{\mathcal{V}} y \forall_{\mathcal{W}} z \ x + y + z \in A\} = (\mathcal{U} + \mathcal{V})+$$

This section needs to be completed

5 Infinite Ramsey Theory

We know that if $\mathbb{N}^{(r)}$ is 2-coloured for any fixed $r = 2, 3, 4, \ldots$ then there exists a monochromatic set M. What if we 2-colour the infinite subsets of M?

Definition $(M^{(\omega)})$**.** For an infinite set $M \subset \mathbb{N}$ we write $M^{(\omega)}$ for the set

$$M^{(\omega)} = \{L \subset M : |L| = \infty\}$$

Question: Given a 2-colouring of $\mathbb{N}^{(\omega)}$, must there be an $M \in \mathbb{N}^{(\omega)}$ such that $M^{(\omega)}$ is monochromatic?

Example. Colouring $M \in \mathbb{N}^{(\omega)}$ RED if $\sum_{x \in M} \frac{1}{x}$ converges and BLUE otherwise, we get such an $M = \{1, 2, 4, 8, \ldots\}$ with $M^{(\omega)}$ RED. Here we could take M to be any convergent sequence.

Sadly, the answer to our question is no:

Proposition. $\exists$ 2-colouring c of $\mathbb{N}^{(\omega)}$ such that no $M \in \mathbb{N}^{(\omega)}$ is monochromatic.

Proof. We'll find a 2-colouring c such that

$$\forall M \in \mathbb{N}^{(\omega)} \ \forall x \in M \ c(M - x) \neq c(M)$$

We are then done as $M^{(\omega)}$ is not monochromatic. This is massive overkill.
Define a relation $\sim$ on $\mathbb{N}^{(\omega)}$ by $L \sim M$ if $|L \triangle M| \infty$.
This is an equivalence relation with equivalence classes $E_i : i \in I$.
Pick $M_i \in E_i$ for each i. Given $M \in \mathbb{N}^{(\omega)}$, say $M \in E_i$, colour M RED if $|M \triangle M_i|$ is even and BLUE if $|M \triangle M_i|$ is odd. $\qquad\square$

We used the Axiom of Choice in the proof above when we chose an M_i for each E_i.
A 2-colouring of $\mathbb{N}^{(\omega)}$ corresponds to a partition $Y \cup Y^c$ of $\mathbb{N}^{(\omega)}$. We say that Y is *Ramsey* if $\exists M \in \mathbb{N}^{(\omega)}$ such that $M^{(\omega)} \subset Y$ or $M^{(\omega)} \subset Y^c$. Therefore Proposition 1 says that not all sets Y are Ramsey, which sets are Ramsey? Hopefully most of them.
We have $\mathbb{N}^{(\omega)} \subset \mathbb{P}(\mathbb{N}) \to \{0,1\}^{\mathbb{N}}$ which has the product topology. So a basic neighbourhood of $M \in \mathbb{N}^{(\omega)}$ is $\{L \in \mathbb{N}^{(\omega)} : L \cap [n] = M \cap [n]\}$. Equivalently, we have a metric $d(L, M)$:

$$d(L, M) = \begin{cases} 0 \text{ if } L = M \\ \frac{1}{\min(L \triangle M)} \text{ if } L \neq M \end{cases}$$

In words, two sets are close if the first element they disagree on is large. So the basic open sets are $\{M : A \text{ is an initial segment of M}\}$ for each finite $A \subset \mathbb{N}$. This is sometimes called the τ-*topology* or the *product topology* or the *usual topology*. Our aim is to show that open sets are Ramsey according to the τ-topology.

Definition $(M^{(<\omega)})$**.** $M^{(<\omega)} = \{A \subset M : A \text{ finite}\}$

Definition $((A, M)^{(\omega)})$**.** For $M \in \mathbb{N}^{(\omega)}$, $A \in \mathbb{N}^{(<\omega)}$, $(A, M)^{(\omega)}$ is defined as

$$(A, M)^{(\omega)} = \{L \in \mathbb{N}^{(\omega)} : A \text{ is initial segment of } L, L \setminus A \in M^{(\omega)}\}$$

In words, this is the set of infinite things starting in A and continuing in M.

Definition (Accepts/Rejects)**.** For $M \in \mathbb{N}^{(\omega)}$, $A \in \mathbb{N}^{(<\omega)}$, we say that M *accepts* A (into Y) if $(A, M)^{(\omega)} \subset Y$. Say that M rejects A if no $L \in M^{(\omega)}$ accepts A

- If M accepts/rejects A then every $L \in M^{(\omega)}$ also accepts/rejects A.
- If M accepts A and $\max A < \min M$, then M accepts $A \cup A'$ for all $A' \in \mathbb{N}^{(<\omega)}$.
- M need not accept or reject A.

[Galvin-Prikry] Let $Y \subset \mathbb{N}^{(\omega)}$ be fixed. Then $\exists M \in \mathbb{N}^{(\omega)}$ such that either M accepts $\emptyset$ or M rejects all of it's finite subsets.

Proof. We may assume that no $M \in \mathbb{N}^{(\omega)}$ accepts $\emptyset$. i.e $\mathbb{N}$ rejects $\emptyset$. We want to construct infinite sets $M_1 \supset M_2 \supset ...$ and points $a_1 < a_2 < ...$ with $a_n \in M_n$ $\forall n$ such that M_n rejects all subsets of $\{a_1, \ldots, a_{n-1}\}$. Then we are done as $\{a_1, a_2, \ldots, \}$ rejects all of it's finite subsets. Let $M_1 = \mathbb{N}$ then M_1 rejects $\emptyset$. Inductively, suppose we have chosen $M_1 \supset M_2 \cdots \supset M_n$ and $a_1, \ldots, a_{n-1}$ such that $a_i \in M_i$ for all $1 \le i \le n-1$ and M_i rejects all finite subsets of $(a_1, \ldots, a_{i-1})$ for each $1 \le i \le n$. Suppose for contradiction that we cannot choose a_n, M_{n+1} suitably.
Pick $b_1 > a_{n-1}, b_1 \in M_n$. Cannot take $a_n = b_1, M_{n+1} = M_n$ and so some infinite $N_1 \subset M_n$ accepts some subset of $\{a_1, \ldots, a_{n-1}, b_1\}$ - say the subset $E_1\{b_1\}$.
Pick $b_2 \in \mathbb{N}_1$, $b_2 > b_1$. Cannot take $a_n = b_2, M_{n+1} = N_1$, so some infinite $N_2 \subset N_1$ accepts some subset of $(a_!, \ldots, a_{n-1}, b_2)$ - say the subset $E_2\{b_2\}$.
Continuing this we obtain $b1 < b_2 < \ldots$ and $M_n \supset N_1 \supset N_2 \supset \ldots$ with $b_{k+1} \in N_k$ for all k and the subsets $E_1, E_2, \ldots$ of $\{a_1, \ldots, a_{n-1}\}$ such that N_k accepts $E_k\{b_k\}$. Passing to a subsequence we may assume that all the E_k are the same, say $E_k = E$ for all k and so $\{b_1, b_2, \ldots\}$ accepts E - contradicting M_n rejecting E. $\square$

The above proof is in the style of waiting until we fail and when we think we have, showing we actually haven't.

Theorem. Let Y be open. Then Y is Ramsey.

Proof. Choose M as given by the Galvin-Prikry Lemma. If M accepts $\emptyset$ then we have $M^{(\omega)} \subset Y$. Thus we may assume that M rejects all of it's finite subsets.
<u>Claim:</u> $M^{(\omega)} \subset Y^c$
Suppose not, then some $L \in M^{(\omega)}$ has $L \in Y$. Since Y is open, there exists an initial segment A of L such that $(A, \mathbb{N})^{(\omega)} \subset Y$. In particular, $(A, M)^{(\omega)} \subset Y$ i.e M accepts A. However M rejects A since it rejects all of it's finite subsets. $\square$

The above proof is a very overkill proof, the second such of this chapter. Note that Y ramsey $\equiv Y^c$ Ramsey so the closed sets are also Ramsey.

5.1 Borel Sets are Ramsey

A stronger topology (more open sets) is the $* - topology$ (also called the Ellen-Tuch/Matthias topology). It's basic open sets are the $(A, M)^{(\omega)}$ for $A, M \in \mathbb{N}^{(\omega)}$. This is the base for a topology: They cover $\mathbb{N}^{(\omega)}$ since $\mathbb{N}^{(\omega)} = (\emptyset, \mathbb{N})^{(\omega)}$ and

$$(A, M)^{(\omega)} \wedge (A', M')^{(\omega)} = \emptyset \text{ or } (A \cup A', M \cup M')^{(\omega)}$$

Theorem. Let $Y \subset \mathbb{N}^{(\omega)}$. Then Y *-open $\implies Y$ ramsey.

Proof. Let M be as given by Galvin-Prikry. If M accepts $\emptyset$ then we have $M^{(\omega)} \subset Y$ so we may assume M rejects all of it's finite subsets.
<u>Claim:</u> $M^{(\omega)} \subset Y^c$
Suppose some $L \in M^{(\omega)}$ has $L \in Y$. AS Y is *-open, have $(A, L)^{(\omega)} \subset Y$ for some initial segment A of L. Thus L accepts A. This contradicts the fact that M rejects A since it rejects all of it's finite subsets. $\square$

Note that the above contradiction, unlike the previous ones, is not overkill.

Definition (Completely Ramsey). Say $Y \subset \mathbb{N}^{(\omega)}$ is completely ramsey if

$$\forall A \in \mathbb{N}^{(<\omega)} \ M \in \mathbb{N}^{(\omega)} \ \exists L \in M^{(\omega)} \text{ with } (A, L)^{(\omega)} \subset Y \text{ or } (A, L)^{(\omega)} \subset Y^c$$

Note that completely ramsey is strictly stronger than Ramsey. Indeed, let Y be the non-ramsey set from the first proposition and set $Z = Y \cup \{2, 3, \dots\}^{(\omega)}$. Then certainly Z is Ramsey e.g $\{2, 3, \dots\}^{(\omega)} \subset Z$ is Ramsey

Index