Ramsey Theory

I.B. Leader

Michaelmas 2000

1 Monochromatic Systems

1.1 Ramsey's Theorem

We write $\mathbb{N}$ for the set $\{1, 2, 3, \dots\}$ of positive integers. For any positive integer n, we write $[n] = \{1, 2, \dots, n\}$. For any set X, we denote the set $\{A \subset X : |A| = r\}$ of subsets of X of size r by $X^{(r)}$.

By a *k-colouring* of $\mathbb{N}^{(r)}$, we mean a map $c \colon \mathbb{N}^{(r)} \to [k]$. We say that a set $M \subset \mathbb{N}$ is *monochromatic* for c if $c|_{M^{(r)}}$ is constant.

Theorem 1 (Ramsey's Theorem). *Whenever $\mathbb{N}^{(2)}$ is 2-coloured, there exists an infinite monochromatic set.*

Proof. Pick $a_1 \in \mathbb{N}$. There are infinitely many edges from a_1, so we can find an infinite set $B_1 \subset \mathbb{N} - \{a_1\}$ such that all edges from a_1 to B_1 are the same colour c_1.

Now choose $a_2 \in B_1$. There are infinitely many edges from a_2 to points in $B_1 - \{a_2\}$, so we can find an infinite set $B_2 \subset B_1 - \{a_2\}$ such that all edges from a_2 to B_2 are the same colour, c_2.

Continue inductively. We obtain a sequence a_1, a_2, a_3, ... of distinct elements of $\mathbb{N}$, and a sequence c_1, c_2, c_3 of colours such that the edge $a_i a_j$ $(i < j)$ has colour c_i. Plainly we must have $c_{i_1} = c_{i_2} = c_{i_3} = \cdots$ for some infinite subsequence. Then $\{a_{i_1}, a_{i_2}, a_{i_3}, \dots\}$ is an infinite monochromatic set. The result follows. □

Remarks. 1. The same proof shows that if $\mathbb{N}^{(2)}$ is k-coloured then we get an infinite monochromatic set. Alternatively, we could view '1' and '2 or 3 or ... or k' as a 2-colouring of $\mathbb{N}^{(2)}$, and then apply Theorem 1 and use induction on k.

2. An infinite monochromatic set is much more than having arbitrarily large finite monochromatic sets. For example, consider the colouring in which

all edges within each of the sets $\{1, 2\}$, $\{3, 4, 5\}$, $\{6, 7, 8, 9\}$, $\{10, 11, 12, 13, 14\}$, $\{15, 16, 17, 18, 19, 20\}$, ... are coloured blue and all other edges are coloured red. Here there is no infinite blue monochromatic set, but there are arbitrarily large finite monochromatic blue sets.

Example. Any sequence $(x_n)_{n \in \mathbb{N}}$ in a totally ordered set has a monotone subsequence: colour $\mathbb{N}^{(2)}$ by giving ij $(i < j)$ colour UP if $x_i < x_j$ and colour DOWN otherwise; the result follows by Theorem 1.

Theorem 2. *Whenever $\mathbb{N}^{(r)}$ is 2-coloured, there exists an infinite monochromatic set.*

Proof. The proof is by induction on r, the case $r = 1$ being trivial.

Suppose the result holds for $r - 1$. Given $c\colon \mathbb{N}^{(r)} \to [2]$, pick $a_1 \in \mathbb{N}$. Define a 2-colouring c' of $(\mathbb{N} - \{a_1\})^{(r-1)}$ by $c'(F) = c(F \cup \{a_1\})$ for all $F \in (\mathbb{N} - \{a_1\})^{(r-1)}$. By induction, there exists an infinite monochromatic set $B_1 \subset \mathbb{N} - \{a_1\}$ for c'; i.e. there exists a colour c_1 such that $c(F \cup \{a_1\}) = c_1$ for all $F \in B_1^{(r-1)}$.

Now choose $a_2 \in B_1$. In exactly the same way, we get an infinite set $B_2 \subset B_1 - \{a_2\}$ and a colour c_2 such that $c(F \cup \{a_2\}) = c_2$ for all $F \in B_2^{(r-1)}$.

Continue inductively: we obtain a sequence a_1, a_2, a_3, ... of distinct elements of $\mathbb{N}$ and colours c_1, c_2, c_3, ... such that for any $i_1 < i_2 < \cdots < i_r$ we have $c(\{a_{i_1}, a_{i_2}, \ldots, a_{i_r}\}) = c_{i_1}$. But we must have $c_{i_1} = c_{i_2} = c_{i_3} = \cdots$ for some infinite subsequence. Then $\{a_{i_1}, a_{i_2}, a_{i_3}, \ldots\}$ is an infinite monochromatic set. The result follows. □

Example. We saw that, given any $(1, x_1)$, $(2, x_2)$, $(3, x_3)$, ... in $\mathbb{R}^2$ we could pick a subsequence inducing a monotone function. In fact we can insist that the induced function is convex or concave: colour $\mathbb{N}^{(3)}$ by giving ijk $(i < j < k)$ the colour convex or concave according as the corresponding points form a convex or concave triple. The result follows by Theorem 2.

We can deduce the finite form of Ramsey's Theorem from Theorem 2.

Corollary 3. *Let m, $r \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that whenever $[n]^{(r)}$ is 2-coloured there is a monochromatic set $M \in [n]^{(m)}$.*

Proof. Suppose not. We construct a 2-colouring of $\mathbb{N}^{(r)}$ without a monochromatic m-set, contradicting Theorem 2.

For each $n \geq r$, we have a colouring $c_n\colon [n]^{(r)} \to [2]$ with no monochromatic m-set. There are only finitely many ways to colour $[r]^{(r)}$ (two in fact) so infinitely many of $c_r|[r]^{(r)}$, $c_{r+1}|[r]^{(r)}$, $c_{r+2}|[r]^{(r)}$, ... agree; say $c_i|[r]^{(r)} = d_r$ for all i lying in some infinite set A_1, where d_r is some colouring of $[r]^{(r)}$. Among the c_i for $i \in A_1$, infinitely many must agree on $[r+1]^{(r)}$; say $c_i|[r+1]^{(r)} = d_{r+1}$ for all $i \in A_2$, where $d_{r+1}\colon [r+1]^{(r)} \to [2]$ and $A_2 \subset A_1$ is infinite.

Continue inductively: we obtain colourings $d_n\colon [n]^{(r)} \to [2]$ for $n = r$, $r+1$, $r+2$, ... such that

(i) no d_n has a monochromatic m-set (as there is some k such that $d_n = c_k|[n]^{(r)}$); and

(ii) for all n, $d_{n+1}|[n]^{(r)} = d_n$.

Define a colouring $c\colon \mathbb{N}^{(r)} \to [2]$ by $c(F) = d_n(F)$ for any $n \geq \max F$. This is well-defined by (ii), and has no monochromatic m-set by (i). So we have our contradiction. The result follows. □

Remarks. 1. This proof gives no information about the minimal possible $n(m, r)$. There are direct proofs which give upper bounds.

2. The above is a *compactness proof:* what we did was (essentially) show that $\{0, 1\}^{\mathbb{N}}$ with the product topology (i.e. the topology derived from the metric $d(f, g) = 1/\min\{n : f(n) \neq g(n)\}$) is compact.

Theorem 4 (The Canonical Ramsey Theorem). *Whenever we have a colouring of $\mathbb{N}^{(2)}$ with an arbitrary set of colours, there exists an inginite set M such that*

(i) c is constant on $M^{(2)}$; or

(ii) c is injective on $M^{(2)}$; or

(iii) $c(ij) = c(kl)$ iff $i = k$ (for all $i, j, k, l \in M$ with $i < j$ and $k < l$); or

(iv) $c(ij) = c(kl)$ iff $j = l$ (for all $i, j, k, l \in M$ with $i < j$ and $k < l$).

Note that this theorem implies Theorem 1: if we have only a finite set of colours then (ii), (iii) and (iv) are impossible.

Proof. First 2-colour $\mathbb{N}^{(4)}$ by giving $ijkl$ (by which we mean henceforth $i < j < k < l$) colour YES if $c(ij) = c(kl)$ and colour NO if $c(ij) \neq c(kl)$. By Ramsey for 4-sets, we have an infinite monochromatic set M. If M is coloured YES then M is monochromatic for c (for given any ij and kl in $M^{(2)}$, choose any $m < n$ in M with $m > i, j, k, l$; then $c(ij) = c(mn) = c(kl)$.) So in this case (i) holds.

Suppose then that M is coloured NO. Now 2-colour $M^{(4)}$ by giving $ijkl$ colour YES if $c(il) = c(jk)$ and colour NO if $c(il) \neq c(jk)$. Again by

Ramsey, there exists an infinite $M' \subset M$ monochromatic for this colouring. If M' is YES, choose $x_1 < x_2 < x_3 < x_4 < x_5 < x_6$ in M'; then $c(x_2x_3) = c(x_1x_6) = c(x_4x_5)$, a contradiction.

So M' is colour NO. Now 2-colour $M'^{(4)}$ by giving $ijkl$ colour YES if $c(ik) = c(jl)$ and colour NO is $c(ik) \neq c(jl)$. By Ramsey, we have an infinite monochromatic set $M'' \subset M'$. If M'' is colour YES then choose $x_1 < x_2 < x_3 < x_4 < x_5 < x_6$ in M''; then $c(x_1x_3) = c(x_2x_5) = c(x_4x_6)$, a contradiction.

So M'' is colour NO. Now 2-colour $M''^{(3)}$ by giving ijk colour LEFT-SAME if $c(ij) = c(ik)$ and colour LEFT-DIFF if $c(ij) \neq c(ik)$. We get an infinite $M''' \subset M''$ monochromatic for this colouring. Then 2-colour $M'''^{(3)}$ by giving ijk colour RIGHT-SAME if $c(ik) = c(jk)$ and colour RIGHT-DIFF if $c(ik) \neq c(jk)$. We get an infinite monochromatic $M'''' \subset M'''$. Finally, 2-colour $M''''^{(3)}$ by giving ijk colour MID-SAME if $c(ij) = c(jk)$ and colour MID-DIFF if $c(ij) \neq c(jk)$. We get an infinite monochromatic $M''''' \subset M''''$.

If M''''' is colour MID-SAME, choose $x_1 < x_2 < x_3 < x_4$ in M'''''; then $c(x_1x_2) = c(x_2x_3) = c(x_3x_4)$, a contradiction. So M''''' is MID-DIFF.

If M''''' is LEFT-SAME and RIGHT-SAME then it would also be MID-SAME, a contradiction.

If M''''' is LEFT-SAME and RIGHT-DIFF then (iii) holds.

If M''''' is LEFT-DIFF and RIGHT-SAME then (iv) holds.

If M''''' is LEFT-DIFF and RIGHT-DIFF then (ii) holds. □

Remark. We could do it all in *one* colouring of $\mathbb{N}^{(4)}$ by colouring $x_1x_2x_3x_4$ with the partition of $[4]^{(2)}$ induced by c on $\{x_1, x_2, x_3, x_4\}$. The number of colours would be the number of partitions of a set of size $\binom{4}{2}$. In the same way, we can show that if we arbitrarily colour $\mathbb{N}^{(r)}$ we get an infinite $M \subset \mathbb{N}$ and a set $I \subset [r]$ such that for any $x_1, x_2, \ldots, x_r, y_1, y_2, \ldots, y_r \in M^{(r)}$ we have

$$c(x_1, x_2, \ldots, x_r) = c(y_1, y_2, \ldots, y_r) \iff x_i = y_i \text{ for all } i \in I.$$

So in Theorem 4, $I = \emptyset$ is (i), $I = \{1, 2\}$ is (ii), $I = \{1\}$ is (iii) and $I = \{2\}$ is (iv). These 2^r colourings are called the *canonical colourings* of $\mathbb{N}^{(r)}$.

1.2 Van der Waerden's Theorem

In this theorem we shall show:

> whenever $\mathbb{N}$ is 2-coloured, for all $m \in \mathbb{N}$ there exists a monochromatic arithmetic progression of length m (i.e. a, $a + d$, $a + 2d$, $\ldots$, $a + (m - 1)d$ all the same colour).

By the familiar compactness argument, this is the same as:

> for all $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that whenever $[n]$ is 2-coloured, there exists a monochromatic arithmetic progression of length m.

In our proof (of the second form above), we use the following key idea: we show more generally that for all k, $m \in \mathbb{N}$, there exists n such that whenever $[n]$ is k-coloured, there exists a monochromatic arithmetic progression of length m. We write $W(m, k)$ for the smallest n if it exists. Note that proving a more general result by induction can actually be *easier*, because the induction hypothesis is correspondingly stronger.

Another idea we use is the following: let A_1, A_2, ..., A_r be arithmetic progressions of length l—say $A_i = \{a_i, a_i + d_i, \ldots, a_i + (l-1)d_i\}$. We say that A_1, A_2, ..., A_r are *focussed* at f if $a_i + ld_i = f$ for all i; for example, $\{1, 4\}$ and $\{5, 6\}$ are focussed at 7. If in addition each A_i is monochromatic and no two are the same colour then we say that they are *colour-focussed* at f (for the given colouring).

Proposition 5. *Let $k \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that whenever $[n]$ is k-coloured, there exists a monochromatic arithmetic progression of length 3.*

Proof. We make the following claim:

> For all $r \leq k$, there exists n such that whenever $[n]$ is k-coloured, *EITHER* there exists a monochromatic arithmetic progression of length 3 *OR* there exist r colour-focussed arithmetic progressions of length 2.

The result will follow immediately from this claim—just take $r = k$; then whatever colour the focus is, we get a monochromatic arithmetic progression of length 3.

We prove the claim by induction on r. Note that the case $r = 1$ is trivial—we may simply take $n = k + 1$. So assume that we are given n suitable for $r - 1$; we will show that $(k^{2n} + 1)2n$ is suitable for r.

Given a k-colouring of $[(k^{2n} + 1)2n]$ not containing a monochromatic arithmetic progression of length 3, break up $[(k^{2n}+1)2n]$ into blocks of length $2n$, namely $B_i = [2n(i-1) + 1, 2ni]$ for $i = 1, 2, \ldots, k^{2n} + 1$. Inside each block, there are $r - 1$ colour-focussed arithmetic progressions of length 2 (by our choice of n), together with their focus (as the length of each block is $2n$). Now there are k^{2n} possible ways to colour a block, so some two blocks, say B_s and B_{s+t}, are coloured identically. Say B_s contains $\{a_i, a_i+d_i\}$, $1 \leq i \leq r-1$, colour-focussed at f. Then B_{s+t} contains $\{a_i+2nt, a_i+d_i+2nt\}$, $1 \le i \le r-1$, colour-focussed at $f + 2nt$, with corresponding colours the same. But now $\{a_i, a_i + d_i + 2nt\}$, $1 \le i \le r-1$, are arithmetic progressions colour-focussed at $f + 4nt$. Also, $\{f, f + 2nt\}$ is monochromatic of a different colour; so we have r arithmetic progressions of length 2 colour-focussed at $f + 4nt$. This completes the induction; the claim, and hence the result, follow. □

Remarks. 1. The idea of looking at the number of ways to colour a block is called a *product argument.*

2. The above proof gives $W(3,k) \le k^{k^{k^{\cdot^{\cdot^{\cdot^{k^{4k}}}}}}} \Big\}{(k-1)}$, a ‘tower-type’ bound.

Theorem 6 (Van der Waerden's Theorem). *Let m, $k \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that whenever $[n]$ is k-coloured, there exists a monochromatic arithmetic progression of length m.*

Proof. The proof is by induction on m. The case $m = 1$ is trivial (for all k).

Now given m, we can assume as our induction hypothesis that $W(m-1, k)$ exists for all k. We make the following claim:

> For all $r \le k$, there exists n such that whenever $[n]$ is k-coloured, there exists either a monochromatic arithmetic progression of length m or r colour-focussed arithmetic progressions of length $m - 1$.

The result will follow immediately from this claim—just put $r = k$ and look at the focus.

The proof of the claim is by induction on r. For $r = 1$ we may simply take $n = W(m-1, k)$. So suppose $r > 1$. If n is suitable for $r - 1$, we will show that $W(m-1, k^{2n})2n$ is suitable for r.

Given a k-colouring of $[W(m-1, k^{2n})2n]$ with no monochromatic arithmetic progression of length m, we can break up $[W(m-1, k^{2n})2n]$ into $W(m-1, k^{2n})$ blocks of length $2n$, namely B_1, B_2, …, $B_{W(m-1,k^{2n})}$ where $B_i = [2n(i-1)+1, 2ni]$. By definition of $W(m-1, k^{2n})$, we can find blocks B_s, B_{s+t}, …, $B_{s+(m-2)t}$ identically coloured.

Now B_s contains $r - 1$ colour-focussed arithmetic progressions of length $m - 1$, together with their focus, say A_1, A_2, …, A_{r-1} colour-focussed at f, where $A_i = \{a_i, a_i + d_i, \ldots, a_i + (m-2)d_i\}$. Now look at the arithmetic progression $A'_i = \{a_i, a_i + (d_i + 2nt), \ldots, a_i + (m-2)(d_i + 2nt)\}$ for $i = 1$, 2, …, $r - 1$. Then A'_1, A'_2, …, A'_{r-1} are colour-focussed at $f + (m-1)2nt$. But $\{f, f+2nt, \ldots, f+(m-2)2nt\}$ is monochromatic and a different colour.

This completes the induction. The claim, and hence the result, follow. □

We define the *Ackermann* (or *Grzegorczyk*) *hierarchy* to be the sequence of functions $f_1, f_2, \ldots : \mathbb{N} \to \mathbb{N}$ defined by

$$\begin{aligned} f_1(x) &= 2x \\ f_{n+1}(x) &= f_n^{(x)}(1) \quad (n \geq 1), \end{aligned}$$

so

$$\begin{aligned} f_2(x) &= 2^x, \\ f_3(x) &= \left.2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}\right\}x, \\ f_4(1) &= 2,\ f_4(2) = 2^2 = 4,\ f_4(3) = 2^{2^{2^2}} = 65536,\ \ldots . \end{aligned}$$

We say that a function $f\colon \mathbb{N} \to \mathbb{N}$ is *of type n* if there exist c and d with $f(cx) \leq f_n(x) \leq f(dx)$ for all x. Our proof above gives a bound of type m on $W(m,k)$, so our bound on $W(m) = W(m,2)$ grows faster than f_n for all n—this is often a feature of such double inductions. Shelah showed that $W(m,k) \leq f_4(m+k)$. Gowers showed that $W(m) \leq 2^{2^{2^{2^{2^{2^{2^{m}}}}}}}$. The best lower bound known is $W(m) \geq 2^m/8m$.

Corollary 7. *Whenever $\mathbb{N}$ is k-coloured, some colour class contains arbitrarily long arithmetic progressions.*

Remark. We cannot guarantee an infinitely long arithmetic progression. Either

(i) colour $\mathbb{N}$ by colouring 1 red, then 2 and 3 blue, then 4, 5 and 6 red then 7, 8, 9 and 10 blue, and so on; or

(ii) enumerate the infinitely long arithmetic progressions as A_1, A_2, A_3, ... (noting that there are only countably many). Choose x_i, $y_i \in A_i$ with $x_i \neq y_i$ for all i and x_i, $y_i < x_{i+1}$, y_{i+1}. Colour each x_i red and each y_i blue.

Theorem 8 (Strengthened Van der Waerden). *Let m, $k \in \mathbb{N}$. Whenever $\mathbb{N}$ is k-coloured, there is an arithmetic progression of length m that, together with its common difference, is monochromatic(i.e. there exist a, $a+d$, $a+2d$, ..., $a+(n-1)d$ and d all the same colour).*

Proof. The proof is by induction on k; the case $k = 1$ is trivial.

Given n suitable for $k-1$ (i.e. such that whenever $[n]$ is $(k-1)$-coloured there exists a monochromatic arithmetic-progression-with-common-difference of length n), we will show that $W(n(m-1)+k)$ is suitable for k. Indeed, given a k-colouring of $[W(n(m-1)+1,k)]$, there exists a monochromatic arithmetic progression of length $n(m-1)+1$, say a, $a+d$, $a+2d$, $\dots$, $a+n(m-1)d$. If d or $2d$ or $\dots$ or nd is the same colour as this arithmetic progression we are done. If not, we have $\{d, 2d, \dots, nd\}$ $(k-1)$-coloured, so we are done by induction. $\square$

Remark. The case $m = 2$ is known as *Schur's Theorem:* whenever $\mathbb{N}$ is k-coloured, we can solve $x + y = z$ in one colour class. We can also prove Schur's Theorem from Ramsey's Theorem: given a k-colouring c of $\mathbb{N}$, define a k-colouring c' of $[n]^{(2)}$ (n large) by $c'(ij) = c(|j-i|)$. By Ramsey, there exists a monochromatic triangle; i.e. there exist $u < v < w$ with $c'(uv) = c'(vw) = c'(wu)$. So $c(v-u) = c(w-v) = c(w-u)$, and since $(v-u) + (w-v) = (w-u)$ we are done.

1.3 The Hales-Jewett Theorem

Let X be a finite set. A subset L of X^n ('the n-dimensional cube on alphabet X') is called a *line* (or *combinatorial line*) if there exists a non-empty set $I = \{i_1, i_2, \dots, i_r\} \subset [n]$ and $a_i \in X$ for each $i \notin I$ such that

$$L = \{x \in X^n : x_i = a_i \text{ for } i \notin I \text{ and } x_{i_1} = x_{i_2} = \cdots = x_{i_r}\}.$$

We call I the set of *active coordinates* for L. For example, in $[3]^2$ the lines are:

- $\{(1,1),(2,1),(3,1)\}$, $\{(1,2),(2,2),(3,2)\}$, and $\{(1,3),(2,3),(3,3)\}$ with $I = \{1\}$;
- $\{(1,1),(1,2),(1,3)\}$, $\{(2,1),(2,2),(2,3)\}$ and $\{(3,1),(3,2),(3,3)\}$ with $I = \{2\}$; and
- $\{(1,1),(2,2),(3,3)\}$ with $I = \{1,2\}$.

Note that the definition of a line does not depend on the ground set X.

Theorem 9 (The Hales-Jewett Theorem). *Let m, $k \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that whenever $[m]^n$ is k-coloured there exists a monochromatic line.*

Remarks. 1. The smallest such n is denoted by $HJ(m,k)$.

2. The Hales-Jewett Theorem implies Van der Waerden's Theorem—we need only embed a Hales-Jewett cube of sufficiently high dimension linearly into $\mathbb{N}$, and so that the embedding is injective on lines. For example, given a k-colouring c of $\mathbb{N}$, induce a k-colouring c' of $[m]^n$ (n large)

by $c'((x_1, x_2, \ldots, x_n)) = c(x_1 + x_2 + \cdots + x_n)$. By Hales-Jewett, there is a monochromatic line, and this corresponds to a monochromatic arithmetic progression of length m in $\mathbb{N}$. So we should regard the Hales-Jewett theorem as an abstract version of Van der Waerden's Theorem.

If L is a line in $[m]^n$ write L^- and L^+ for its first and last points (in the ordering on $[m]^n$ given by $x \le y$ if $x_i \le y_i$ for all i). Lines L_1, L_2, ..., L_k are *focussed* at f if $L_i^+ = f$ for all i. They are *colour-focussed* (for a given colouring) if in addition each $L_i - \{L_i^+\}$ is monochromatic, no two the same colour.

Proof (of Theorem 9). By induction on m; the case $m = 1$ is trivial.

Given $m > 1$, we may assume that $HJ(m-1, k)$ exists for all k. We make the following claim:

> For all $r \le k$, there exists n such that whenever $[m]^n$ is k-coloured, there exists *EITHER* a monochromatic line *OR* r colour-focussed lines.

The result will follow immediately from this claim—put $r = k$ and look at the focus.

The proof of the claim is by induction on r. For $r = 1$ we may take $n = HJ(m-1, k)$.

Given n suitable for r, we shall show that $n + HJ(m-1, k^{m^n})$ is suitable for $r+1$. Write $n' = HJ(m-1, k^{m^n})$.

Given a k-colouring of $[m]^{n+n'}$ with no monochromatic line, identify $m^{n+n'}$ with $[m]^n \times [m]^{n'}$. There are k^{m^n} ways to colour a copy of $[m]^n$. So by our choice of n', we have a line L in $[m]^{n'}$, say with active coordinate set I, such that for all $a \in [m]^n$ and all $b, b' \in L - \{L^+\}$, we have $c(a, b) = c(a, b') = c'(a)$, say. Now by definition of n, there exist r colour-focussed lines for c', say L_1, L_2, ..., L_r, with acrive coordinate sets I_1, I_2, ..., I_r respectively, and focus f. But now let L_i' be the line through the point (L_i^-, L^-) with active coordinate set $I_i \cup I$ ($i = 1, 2, \ldots, r$). Then L_1', L_2', ..., L_r' are colour-focussed at (f, L^+). And the line through (f, L^-) with active coordinate set I gives us $r+1$ colour-focussed lines. Thus our induction is complete and the claim, and hence the result, follow. □

A *d-dimensional subspace* or *d-parameter set* S in X^n is a set of the following form: there exist disjoint non-empty sets I_1, I_2, ..., $I_d \subset [n]$ and $a_i \in X$ for each $i \in [n] - (I_1 \cup I_2 \cup \cdots \cup I_d)$ such that

$$S = \left\{ x \in X^n : \begin{array}{l} x_i = a_i \text{ for all } i \in [n] - (I_1 \cup I_2 \cup \cdots \cup I_d) \\ x_i = x_j \text{ whenever } i, j \in I_l \text{ for some } l \end{array} \right\}.$$

For example in X^3, $\{(a,b,2) : a,b \in X\}$ and $\{(a,a,b) : a,b \in X\}$ are 2-parameter sets.

Theorem 10 (The Extended Hales-Jewett Theorem). *Let m, k, $d \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that whenever $[m]^n$ is k-coloured, there exists a monochromatic d-parameter set.*

Proof. Regard X^{dn} as $(X^d)^n$—a cube on alphabet X^d. Clearly any line in this (on alphabet X^d) is a d-parameter set on alphabet X, so we can take $n = dHJ(n^d, k)$. □

1.4 Gallai's Theorem

Let $S \subset \mathbb{N}^d$ be a finite set. A *homothetic copy* of S is any set of the form $a + \lambda S$ where $a \in \mathbb{N}^d$ and $\lambda \in \mathbb{N}$. For example, in $\mathbb{N}^1$, a homothetic copy of $\{1, 2, \dots, m\}$ is precisely an arithmetic progression of length m.

Theorem 11 (Gallai's Theorem). *For any finite $S \subset \mathbb{N}^d$ and any k-colouring of $\mathbb{N}^d$, there exists a monochromatic homothetic copy of S.*

Proof. Let $S = \{S(1), S(2), \dots, S(m)\}$. Given a k-colouring c of $\mathbb{N}^d$, define a k-colouring c' of $[m]^n$ (n large) by $c'(\mathbf{x}) = c(\sum_i S(x_i))$. By Hales-Jewett, there is a monochromatic line, giving a monochromatic homothetic copy of S (with λ the number of active coordinates). □

Remarks. 1. Or by a product argument and focussing.

2. For $S = \{(x,y) : x,y \in \{0,1\}\}$, Gallai's Theorem tells us that there exists a monochromatic square. Could we have used 2-parameter Hales-Jewett instead?—No, this would only give us a rectangle.

2 Partition Regular Equations

2.1 Partition Regularity

Let A be an $m \times n$ matrix with rational entries. We say that A is *partition regular (PR) (over $\mathbb{N}$)* if whenever $\mathbb{N}$ is finitely coloured, there is always a monochromatic $\mathbf{x} \in \mathbb{N}^n$ with $A\mathbf{x} = \mathbf{0}$.

Examples. 1. Schur states that the matrix $(1\ 1\ {-1})$ is PR.

2. Strengthened Van der Waerden states that the matrix

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

is PR.

We also talk about 'the equation $A\mathbf{x} = \mathbf{0}$' being PR.

Not every matrix is PR: for example, $(2\ -1)$ is not PR; for we can 2-colour $x \in \mathbb{N}$ by the parity of $\max\{i : 2^i | x\}$. Note that A is PR if and only if λA is PR for any $\lambda \in \mathbb{Q} - \{0\}$, so we could restrict our attention to integer matrices if we wished.

Let A have columns $\mathbf{c^{(1)}}, \mathbf{c^{(2)}}, \dots, \mathbf{c^{(n)}} \in \mathbb{Q}^m$, so

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

We say that A has the *columns property* if there is a partition $B_1 \cup B_2 \cup \dots \cup B_r$ of $[n]$ such that

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

(ii) $\sum_{i \in B_s} \mathbf{c^{(i)}} \in \left\langle \mathbf{c^{(j)}} : j \in B_1 \cup B_2 \cup \dots \cup B_{s-1} \right\rangle$ for $s = 2, 3, \dots, r$

where $\langle\rangle$ denotes linear span over $\mathbb{R}$. (Note that we could have equally said 'over $\mathbb{Q}$' here: if a rational vector is a real linear combination of some rational vectors then it is also a rational combination of them.)

Examples. 1. The matrix $(1\ 1\ {-1})$ has the columns property: take $B_1 = \{1, 3\}$ and $B_2 = \{2\}$.

2. The matrix

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & a \\ 4 & -4 & b \end{pmatrix}$$

has the columns property if and only if $(a, b) = (6, 12)$.

3. The matrix $A = (a_1\, a_2\, \dots\, a_n)$ has the columns property if and only if either $A = 0$ or some non-empty subset of the non-zero a_i sums to zero.

We shall prove:

Rado's Theorem. A rational matrix A is PR if and only if A has the columns property.

One strength of this result is that it shows that partition regularity, which does not at first appear to be checkable in finite time, in fact *is* checkable in finite time.

First we show that Rado's Theorem is true for one equation. We may assume without loss of generality that $a_1, a_2, \ldots, a_n \neq 0$; then we must show

$$(a_1\, a_2\, \ldots\, a_n) \text{ is PR} \iff \sum_{i \in I} a_i = 0 \text{ for some non-empty } I \subset [n].$$

Let p be prime. For $x \in \mathbb{N}$, let $d^p(x)$ be the last non-zero digit in the base p expansion of x, i.e. if $x = d_r p^r + d_{r-1} p^{r-1} + \cdots + d_1 p + d_0$, $0 \leq d_i \leq p-1$ for all i, then $d^p(x) = d_{L(x)}$ where $L(x) = \min\{i : d_i \neq 0\}$. For example, if $x = 1002047000$ in base p then $L(x) = 3$ and $d^p(x) = 7$.

Proposition 12. *Let $a_1, a_2, \ldots, a_n$ be non-zero rationals such that the matrix $(a_1\, a_2\, \ldots\, a_n)$ is PR. Then $\sum_{i \in I} a_i = 0$ for some non-empty $I \subset [n]$.*

Proof. We may assume without loss of generality that $a_1, a_2, \ldots, a_n \in \mathbb{Z}$. Fix a prime p with $p > \sum_{i=1}^n |a_i|$, and define a $(p-1)$-colouring of $\mathbb{N}$ by giving x the colour $d^p(x)$. We know that $\sum_{i \in I} a_i x_i = 0$ for some $x_1, x_2, \ldots, x_n$ all of the same colour, d, say. Let $L = \min\{L(x_i) : 1 \leq i \leq n\}$ and let $I = \{i : L(x_i) = L\}$. Considering $\sum_{i \in I} a_i x_i = 0$ performed in base p, we have $\sum_{i \in I} d a_i \equiv 0 \pmod p$ and so $\sum_{i \in I} a_i \equiv 0 \pmod p$. But $p > \sum_{i=1}^n |a_i|$ and so $\sum_{i \in I} a_i = 0$. □

Remark. Or: for each prime p we get a set I with $\sum_{i \in I} a_i \equiv 0 \pmod p$, so some fixed set I has $\sum_{i \in I} a_i \equiv 0 \pmod p$ for infinitely many p, whence $\sum_{i \in I} a_i = 0$.

Lemma 13. *Let $\lambda \in \mathbb{Q}$. Then whenever $\mathbb{N}$ is finitely coloured, there exist monochromatic x, y and z with $x + \lambda y = z$.*

Proof. (cf the proof of Theorem 8.) If $\lambda = 0$ we are done; if $\lambda < 0$ we may rewrite our equation as $z - \lambda y = x$. So we may assume without loss of generality that $\lambda > 0$; say $\lambda = r/s$ with $r, s \in \mathbb{N}$.

So we need to prove that for all k, there exists an n such that, whenever $[n]$ is k-coloured, there exist monochromatic x, y and z with $x + (r/s)y = z$. We shall prove this by induction on k.

For $k = 1$, take $n = \max\{s, r+1\}$ and $(x, y, z) = (1, s, r+1)$.

Suppose $k > 1$. Given n suitable for $k-1$, we shall show that $W(nr+1, k)$ is suitable for k. Indeed, given a k-colouring of $[W(nr+1, k)]$ we have a monochromatic arithmetic progression of length $nr+1$, say a, $a+d$, ..., $a+nrd$, all of colour c. Look at ds, $2ds$, ..., nds. If, say, ids has colour c then we are done, as $a+(r/s)ids = a+idr$ and $(a, ids, a+idr)$ is a monochromatic triple with colour c. So we may assume the set $\{ds, 2ds, \dots, nds\}$ is $(k-1)$-coloured, and we are done by induction. The claim, and hence the result, follow. □

Theorem 14 (Rado's Theorem for single equations). *Let a_1, a_2, ..., a_n be non-zero rationals. Then $(a_1\, a_2\, \dots\, a_n)$ is PR if and only in $\sum_{i\in I} a_i = 0$ for some non-empty $I \subset [n]$.*

Proof. $\Rightarrow$ is Proposition 12.

$\Leftarrow$ Given a finite colouring of $\mathbb{N}$, fix $i_0 \in I$. For suitable monochromatic x, y and z, we shall set

$$x_i = \begin{cases} x & \text{if } i = i_0 \\ y & \text{if } i \notin I \\ z & \text{if } i \in I - \{i_0\} \end{cases} .$$

We require that $\sum a_i x_i = 0$, i.e. that

$$a_{i_0} x + \Bigg(\sum_{i \in I - \{i_0\}} a_i \Bigg) z + \Bigg(\sum_{i \notin I} a_i \Bigg) y = 0,$$

i.e.

$$a_{i_0} x - a_{i_0} z + \Bigg(\sum_{i \notin I} a_i \Bigg) y = 0,$$

i.e.

$$x + \frac{\left(\sum_{i \notin I} a_i \right)}{a_{i_0}} y - z = 0,$$

and such x, y and z do indeed exist by Lemma 13. □

Proposition 15. *Let A be any matrix with entries in $\mathbb{Q}$. If A is PR then it must have the columns property.*

Proof. We may assume without loss of generality that all the entries of A are integers. Let the columns of A be $\mathbf{c^{(1)}}$, $\mathbf{c^{(2)}}$, ..., $\mathbf{c^{(n)}}$. For any prime p, colour $\mathbb{N}$ with the d^p colouring. By assumption, there exists a monochromatic $\mathbf{x} \in \mathbb{Z}^n$ with $A\mathbf{x} = \mathbf{0}$, i.e. $x_1\mathbf{c^{(1)}} + x_2\mathbf{c^{(2)}} + \cdots + x_n\mathbf{c^{(n)}} = \mathbf{0}$. Say all the x_i have colour d.

We have a partition $B_1 \cup B_2 \cup \cdots \cup B_r$ of $[n]$ given by

$$\begin{aligned} L(x_i) = L(x_j) &\implies i, j \in B_s \text{ for some } s; \\ L(x_i) < L(x_j) &\implies i \in B_s, j \in B_t \text{ for some } s < t. \end{aligned}$$

For infinitely many primes p, say all $p \in P$, we get the same $B_1, B_2, \ldots, B_r$.

Considering $\sum x_i \mathbf{c}^{(\mathbf{i})} = \mathbf{0}$ performed in base p, we have

(i) $\sum_{i \in B_1} d\mathbf{c}^{(\mathbf{i})} \equiv \mathbf{0} \pmod{p}$, where by $\mathbf{u} \equiv \mathbf{v} \pmod{p}$ with $\mathbf{u}$, $\mathbf{v} \in \mathbb{Z}^n$ we mean $u_j \equiv v_j \pmod{p}$ for all j; and

(ii) for all $s \geq 2$, $\sum_{i \in B_s} p^t d\mathbf{c}^{(\mathbf{i})} + \sum_{i \in B_1 \cup \cdots \cup B_{s-1}} x_i \mathbf{c}^{(\mathbf{i})} \equiv \mathbf{0} \pmod{p^{t+1}}$ for some t.

From (i), and as d is invertible, we have $\sum_{i \in B_1} \mathbf{c}^{(\mathbf{i})} \equiv \mathbf{0} \pmod{p}$ for infinitely many p, and so $\sum_{i \in B_1} \mathbf{c}^{(\mathbf{i})} = \mathbf{0}$.

From (ii), for all $s \geq 2$ we have

$$p^t \sum_{i \in B_s} \mathbf{c}^{(\mathbf{i})} + \sum_{i \in B_1 \cup \cdots \cup B_{s-1}} y_i \mathbf{c}^{(\mathbf{i})} \equiv \mathbf{0} \pmod{p^{t+1}}$$

(where $y_i = d^{-1} x_i \pmod{p^{t+1}}$).

We now show that $\sum_{i \in B_s} \mathbf{c}^{(\mathbf{i})} \in \langle \mathbf{c}^{(\mathbf{i})} : i \in B_1 \cup B_2 \cup \cdots \cup B_{s-1} \rangle$. Suppose not. Then there exists $\mathbf{u} \in \mathbb{Z}^m$ with $u \cdot \mathbf{c}^{(\mathbf{i})} = 0$ for all $i \in B_1 \cup B_2 \cup \cdots \cup B_{s-1}$ but with $\mathbf{u} \cdot (\sum_{i \in B_s} \mathbf{c}^{(\mathbf{i})}) \neq 0$. So $p^t \mathbf{u} \cdot (\sum_{i \in B_s} \mathbf{c}^{(\mathbf{i})}) \equiv 0 \pmod{p^{t+1}}$, i.e. $\mathbf{u} \cdot (\sum_{i \in B_s} \mathbf{c}^{(\mathbf{i})}) \equiv 0 \pmod{p}$ for infinitely many p, a contradiction. □

Let m, p, $c \in \mathbb{N}$. A set $S \subset \mathbb{N}$ is an (m, p, c)*-set* with generators $x_1, x_2, \ldots, x_m \in \mathbb{N}$ if

$$S = \left\{ \sum_{i=1}^{m} \lambda_i x_i : \exists j \text{ with } \lambda_i = 0 \,\forall i < j,\ \lambda_j = c,\ \lambda_i \in \{-p, -p+1, \ldots, p\} \,\forall i > j \right\}.$$

So S consists of all numbers in the lists:

$$\begin{aligned} cx_1 + \lambda_2 x_2 + \lambda_3 x_3 + \cdots + \lambda_m x_m &\quad (|\lambda_i| \leq p \ \forall i \geq 2) \\ cx_2 + \lambda_3 x_3 + \cdots + \lambda_m x_m &\quad (|\lambda_i| \leq p \ \forall i \geq 3) \\ &\vdots \\ cx_{m-1} + \lambda_m x_m &\quad (|\lambda_m| \leq p) \\ cx_m. & \end{aligned}$$

Examples. 1. A $(2, p, 1)$-set is an arithmetic progression of length $2p + 1$ together with its common difference.

2. A $(2, 2, 3)$-set is an arithmetic progression of length 5, with middle term divisible by 3, together with thrice its common difference.

Theorem 16. *Let m, p, $c \in \mathbb{N}$ and suppose $\mathbb{N}$ is finitely coloured. Then there exists a monochromatic (m, p, c)-set.*

Proof. Let $\mu = k(m-1)+1$.

Given a k-colouring of $B_0 = [n]$ with n large, look at

$$A_1 = \left\{ c, 2c, \dots, \left\lfloor \frac{n}{c} \right\rfloor c \right\}^1.$$

By Van der Waerden, there is a monochromatic arithmetic progression inside A, say

$$P_1 = \{cx_1 - n_1 d_1, cx_1 - (n_1 - 1)d_1, \dots, cx_1, \dots, cx_1 + n_1 d_1\}$$

where n_1 is large and P_1 has colour k_1, say. Now we restrict attention to

$$B_1 = \left\{ d_1, 2d_1, \dots, \frac{n_1}{(\mu - 1)p} d_1 \right\}.$$

Note that for any integers λ_2, λ_3, ..., $\lambda_\mu \in [-p, p]$ and b_2, b_3, ..., $b_\mu \in B_1$, we have

$$cx_1 + \lambda_2 b_2 + \lambda_3 b_3 + \cdots + \lambda_\mu b_\mu \in P_1,$$

so in particular all sums of this form have colour k_1.

Now look at

$$A_2 = \left\{ cd_1, 2cd_1, \dots, \frac{n_1}{(\mu - 1)pc} d_1 \right\}.$$

By Van der Waerden, there is a monochromatic arithmetic progression inside A_2, say

$$P_2 = \{cx_2 - n_2 d_2, cx_2 - (n_2 - 1)d_2, \dots, cx_2, \dots, cx_2 + n_2 d_2\},$$

where n_2 is large and P_2 has colour k_2, say. Now we restrict attention to

$$B_2 = \left\{ d_2, 2d_2, \dots, \frac{n_2}{(\mu - 2)p} d_2 \right\}.$$

Note that for any integers λ_3, λ_4, ..., $\lambda_\mu \in [p, -p]$, and b_3, b_4, ..., $b_\mu \in B_2$, we have

$$cx_2 + \lambda_3 b_3 + \lambda_4 b_4 + \cdots + \lambda_\mu b_\mu \in P_2,$$

so in particular all sums of this form have colour k_2.

---

[1]Henceforth we shall omit the symbols '$\lfloor$' and '$\rfloor$' in expressions such as this; they should be understood where necessary.

Now look at $A_3 = \dots$.

Keep going μ times: we obtain $x_1, x_2, \dots, x_\mu$ such that each row of the (μ, p, c)-set generated by $x_1, x_2, \dots, x_\mu$ is monochromatic. But since $\mu = k(m-1)+1$, some m of these rows are the same colour, and we are done. $\square$

Remark. For $x_1, x_2, \dots, x_m \in \mathbb{N}$, let

$$\mathrm{FS}(x_1, x_2, \dots, x_m) = \left\{ \sum_{i \in I} x_i : \emptyset \neq I \subset [m] \right\}.$$

Then Theorem 16 for $(m, 1, 1)$-sets implies:

> Whenever $\mathbb{N}$ is finitely coloured, there exist $x_1, x_2, \dots, x_m$ with $\mathrm{FS}(x_1, x_2, \dots, x_m)$ monochromatic.

This result is variously known as the *finite sums theorem, Folkman's Theorem* or *Sanders's Theorem.*

Similarly, we can guarantee a monochromatic

$$\mathrm{FP}(x_1, x_2, \dots, x_m) = \left\{ \prod_{i \in I} x_i : \emptyset \neq I \subset [m] \right\}$$

by looking at $\{2^n : n \in \mathbb{N}\}$.

Lemma 17. *If A has the columns property then there exist $m, p, c \in \mathbb{N}$ such that every (m, p, c)-set contains a solution to $A\mathbf{x} = \mathbf{0}$, i.e. we can solve $A\mathbf{x} = \mathbf{0}$ with all x_i in the (m, p, c)-set.*

Proof. Let A have columns $\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(n)}$. As A has the columns property, we have a partition $B_1 \cup B_2 \cup \dots \cup B_r$ of $[n]$ such that

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

(ii) $\sum_{i \in B_s} \mathbf{c}^{(i)} \in \left\langle \mathbf{c}^{(j)} : j \in B_1 \cup B_2 \cup \dots \cup B_{s-1} \right\rangle$ for all $s \geq 2$.

Suppose

$$\sum_{i \in B_s} \mathbf{c}^{(i)} = \sum_{i \in B_1 \cup B_2 \cup \dots \cup B_{s-1}} q_{is} \mathbf{c}^{(i)}.$$

Then for each s we have

$$\sum_{i \in [n]} d_{is} \mathbf{c}^{(i)} = \mathbf{0}$$

where

$$d_{is} = \begin{cases} 0 & \text{if } i \notin B_1 \cup B_2 \cup \cdots \cup B_s \\ 1 & \text{if } i \in B_s \\ -q_{is} & \text{if } i \in B_1 \cup B_2 \cup \cdots \cup B_{s-1} \end{cases}.$$

Given $x_1, x_2, \ldots, x_r \in \mathbb{N}$, put $y_i = \sum_{s=1}^{r} d_{is} x_s$ for $i = 1, 2, \ldots, n$. Then

$$\sum_{i=1}^{n} y_i \mathbf{c}^{(i)} = \sum_{i=1}^{n} \sum_{s=1}^{r} d_{is} x_s \mathbf{c}^{(i)} = \sum_{s=1}^{r} x_s \sum_{i=1}^{n} d_{is} \mathbf{c}^{(i)} = \mathbf{0}.$$

So $A\mathbf{y} = \mathbf{0}$. Now we are done, for we may take $m = r$, take c to be the least common multiple of the denominators of the q_i, and take p to be c times the maximum of the numerators of the q_i. Then $c\mathbf{y}$ is in the (m, p, c)-set generated by $x_1, x_2, \ldots, x_n$ and $A(c\mathbf{y}) = 0$. □

Theorem 18 (Rado's Theorem). *Let A be a rational matrix. Then A is partition regular if and only if A has the columns property.*

Proof. $\Leftarrow$ is Proposition 15.

$\Rightarrow$ follows from Theorem 16 and Lemma 17. □

Corollary 19 (The Consistency Theorem). *If A and B are partition regular then the matrix $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ is also partition regular. In other words, if we can guarantee to solve $A\mathbf{x} = \mathbf{0}$ in some colour class and $B\mathbf{y} = \mathbf{0}$ in some colour class then we can guarantee to solve both in the* same *colour class.*

Remark. This is not obvious by considerations of partition regularity alone.

Corollary 20. *Whenever $\mathbb{N}$ is finitely coloured, some colour class contains solutions to all PR equations.*

Proof. Suppose not. Then we have $\mathbb{N} = D_1 \cup D_2 \cup \cdots \cup D_k$, and, for each i, a PR matrix A_i such that D_i does not contain a solution of $A_i\mathbf{x} = \mathbf{0}$. Then consider the matrix

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}.$$

This is PR, by the Consistency Theorem, but no D_i contains a solution to it, a contradiction. □

We say that $D_i \subset \mathbb{N}$ is partition regular if it contains a solution to every PR equation. So Corollary 20 says that if $\mathbb{N} = D_1 \cup D_2 \cup \cdots \cup D_k$ then some D_i is PR. Rado conjectured, and Deuber proved, that if D is PR and $D = D_1 \cup D_2 \cup \cdots \cup D_k$ then some D_i is PR.

2.2 Filters and Ultrafilters

A *filter* on $\mathbb{N}$ is a non-empty collection $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$ such that

(i) $\emptyset \notin \mathcal{F}$;

(ii) if $A \in \mathcal{F}$ and $B \supset A$ then $B \in \mathcal{F}$ ('$\mathcal{F}$ is an up-set'); and

(iii) if A, $B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$ ('$\mathcal{F}$ is closed under finite intersections').

Intuitively, we think of the sets of our filter as being the 'large' subsets of $\mathbb{N}$.

Examples. The following are filters

(i) $\{A \subset \mathbb{N} : 1, 2 \in A\}$;

(ii) $\{A \subset \mathbb{N} : A^c \text{ finite}\}$, 'the cofinite filter';

(iii) $\{A \subset \mathbb{N} : E - A \text{ finite}\}$ where E is the set of even numbers.

An *ultrafilter* is a maximal filter. For any $x \in \mathbb{N}$, the set $\{A \subset \mathbb{N} : x \in A\}$ is an ultrafilter, the *principal ultrafilter at x*.

Proposition 21. *A filter $\mathcal{F}$ is an ultrafilter if and only if for all $A \subset \mathbb{N}$, either $A \in \mathcal{F}$ or $A^c \in \mathcal{F}$.*

Proof. $\Leftarrow$ This is obvious since we cannot add A to $\mathcal{F}$ if we already have A^c.
$\Rightarrow$ Suppose $\mathcal{F}$ is an ultrafilter and A, $A^c \notin \mathcal{F}$. Then we must have $B \in \mathcal{F}$ with $B \cap A = \emptyset$ (for otherwise $\mathcal{F}' = \{C \subset \mathbb{N} : C \supset A \cap B \text{ for some } B \in \mathcal{F}\}$ is a filter containing $\mathcal{F}$). Similarly, we must have $C \in \mathcal{F}$ with $C \cap A^c = \emptyset$. But then $B \cap C = \emptyset$, a contradiction. □

Note. If $A \in \mathcal{U}$, an ultrafilter, and $A = B \cup C$, then $B \in \mathcal{U}$ or $C \in \mathcal{U}$ (for otherwise B^c, $C^c \in \mathcal{U}$ by Proposition 21 whence $A^c = B^c \cap C^c \in \mathcal{U}$, a contradiction).

Proposition 22. *Every filter is contained in an ultrafilter.*

Proof. By Zorn's Lemma, it is sufficient to check that every non-empty chain $\{\mathcal{F}_i : i \in I\}$ has an upper bound. Indeed, put $\mathcal{F} = \cup_{i \in I} \mathcal{F}_i$. Then

(i) $\emptyset \in \mathcal{F}$;

(ii) if $A \in \mathcal{F}$ and $B \supset A$ then $A \in \mathcal{F}_i$ for some i, so $B \in \mathcal{F}_i$ and so $B \in \mathcal{F}$;

(iii) if A, $B \in \mathcal{F}$ then A, $B \in \mathcal{F}_i$ for some i (as the $\mathcal{F}_i$ form a chain), so $A \cap B \in F_i$ and so $A \cap B \in \mathcal{F}$.

$\square$

Remarks. 1. Any ultrafilter extending the cofinite filter is non-principal. Also, if $\mathcal{U}$ is non-principal then $\mathcal{U}$ contains all cofinite sets; for if $A \in \mathcal{U}$ for some finite A then $\{x\} \in \mathcal{U}$ for some $x \in A$ by our note above.

2. The Axiom of Choice is needed in some form to get non-principal ultrafilters.

The set of all ultrafilters on $\mathbb{N}$ is denoted $\beta\mathbb{N}$. We define a topology on $\beta\mathbb{N}$ by taking as a base all sets of the form

$$C_A = \{\mathcal{U} \in \beta\mathbb{N} : A \in \mathcal{U}\},\ A \subset \mathbb{N}.$$

This *is* a base: it is sufficient to check that $\bigcup C_A = \beta\mathbb{N}$ and that the intersection of any two of the C_A is another set of the putative base. Plainly $\bigcup C_A = \beta\mathbb{N}$, and $C_A \cap C_B = C_{A\cap B}$ as A, $B \in \mathcal{U}$ if and only if $A \cap B \in \mathcal{U}$. Thus open sets are of the form

$$\bigcup_{i\in I} C_{A_i} = \{\mathcal{U} : A_i \in \mathcal{U} \text{ for some } i \in I\}.$$

Note that $\beta\mathbb{N} - C_A = C_{A^c}$. So closed sets are of the form

$$\bigcap_{i\in I} C_{A_i} = \{\mathcal{U} : A_i \in \mathcal{U} \text{ for all } i \in I\}.$$

We have $\mathbb{N}$ inside $\beta\mathbb{N}$ (identifying $n \in \mathbb{N}$ with the principal ultrafilter $\tilde{n}$ at n). Each point of $\mathbb{N}$ is isolated: $C_{\{n\}} = \{\tilde{n}\}$. Also, $\mathbb{N}$ is dense in $\beta\mathbb{N}$—every non-empty open set in $\beta\mathbb{N}$ meets $\mathbb{N}$ as $\tilde{n} \in C_A$ whenever $n \in A$.

Theorem 23. *$\beta\mathbb{N}$ is a compact Hausdorff space.*

Proof. Hausdorff: Given $\mathcal{U} \neq \mathcal{V}$, we have some $A \in \mathcal{U}$ with $A \notin \mathcal{V}$. But then $A^c \in \mathcal{V}$ and so $\mathcal{U} \in C_A$ and $\mathcal{V} \in C_{A^c}$.

Compact: Given closed sets $(F_i)_{i\in I}$ with the finite intersections property (i.e. all finite intersections are non-empty), we need to show that $\bigcap_{i\in I} F_i \neq \emptyset$. Assume without loss of generality that each F_i is basic, i.e. that $F_i = C_{A_i}$ for some $A_i \subset \mathbb{N}$.

We first observe that the sets $(A_i)_{i\in I}$ also have the finite intersections property. For we have $C_{A_{i_1}} \cap C_{A_{i_2}} \cap \cdots \cap C_{A_{i_n}} = C_{A_{i_1} \cap A_{i_2} \cdots \cap A_{i_n}}$ and so $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n} \neq \emptyset$.

So we can define a filter $\mathcal{F}$ generated by the A_i:

$$\mathcal{F} = \{A \subset \mathbb{N} : A \supset A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n} \text{ for some } i_1, i_2, \ldots, i_n \in I\}.$$

Let $\mathcal{U}$ be an ultrafilter extending $\mathcal{F}$. Then $A_i \in \mathcal{U}$ for all i, so $u \in C_{A_i}$ for all i, and so $\cap_{i \in I} C_{A_i} \neq \emptyset$ as desired. □

Remarks. 1. If we view an ultrafilter as a function from $\mathcal{P}(\mathbb{N}) \to \{0,1\}$ then we have $\beta\mathbb{N} \subset \{0,1\}^{\mathcal{P}(\mathbb{N})}$. We can check that the topology on $\beta\mathbb{N}$ is the restriction of the product topology, and also that $\beta\mathbb{N}$ is a closed subset of $\{0,1\}^{\mathcal{P}(\mathbb{N})}$—so compact by Tychonov.

2. $\beta\mathbb{N}$ is the largest compact Hausdorff space in which $\mathbb{N}$ is dense—it is called the *Stone-Čech compactification of* $\mathbb{N}$.

Let $\mathcal{U}$ be an ultrafilter and p a statement. We write $\forall_{\mathcal{U}} x \;\; p(x)$ to mean $\{x : p(x)\} \in \mathcal{U}$, and say that $p(x)$ holds 'for most x' or 'for $\mathcal{U}$-most x'. For example,

(i) for $\mathcal{U}$ non-principal, $\forall_{\mathcal{U}} x \;\; x > 4$;

(ii) for $\mathcal{U} = \tilde{n}$ we have $\forall_{\mathcal{U}} x \; p(x) \iff p(n)$.

Proposition 24. *Let $\mathcal{U}$ be an ultrafilter and p and q statements. Then*

(i) $\forall_{\mathcal{U}} x \; (p(x) \text{ AND } q(x)) \iff (\forall_{\mathcal{U}} x \; p(x)) \text{ AND } (\forall_{\mathcal{U}} x \; q(x))$

(ii) $\forall_{\mathcal{U}} x \; (p(X) \text{ OR } q(x)) \iff (\forall_{\mathcal{U}} x \; p(x)) \text{ OR } (\forall_{\mathcal{U}} x \; q(x))$

(iii) if $\forall_{\mathcal{U}} x \; p(x)$ *does not hold then* $\forall_{\mathcal{U}} x \; (\text{NOT } p(x))$.

Proof. Let $A = \{x : p(x)\}$ and let $B = \{x : q(x)\}$. Then

(i) $A \cap B \in \mathcal{U}$ if and only if $A \in \mathcal{U}$ and $B \in \mathcal{U}$;

(ii) $A \cup B \in \mathcal{U}$ if and only if $A \in \mathcal{U}$ or $B \in \mathcal{U}$;

(iii) if $\forall_{\mathcal{U}} x \; p(x)$ is false then $A \notin \mathcal{U}$ and so $A^c \in \mathcal{U}$.

□

Note. $\forall_{\mathcal{U}} x \; \forall_{\mathcal{V}} y \; p(x,y)$ need not be the same as $\forall_{\mathcal{V}} y \; \forall_{\mathcal{U}} x \; p(x,y)$—even if $\mathcal{U} = \mathcal{V}$. For example, if $\mathcal{U}$ is non principal then $\forall_{\mathcal{U}} x \; \forall_{\mathcal{U}} y \; x < y$ is true, but $\forall_{\mathcal{U}} y \; \forall_{\mathcal{U}} x \; x < y$ is false.

For $\mathcal{U}$, $\mathcal{V} \in \beta\mathbb{N}$, define

$$\begin{aligned} \mathcal{U} + \mathcal{V} &= \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x \; \forall_{\mathcal{U}} y \; x + y \in A\} \\ &= \{A \subset \mathbb{N} : \{x \in \mathbb{N} : \{y : x + y \in A\} \in \mathcal{V}\} \in \mathcal{U}\}. \end{aligned}$$

Proposition 25. *So defined, + is a well-defined map from $\beta\mathbb{N} \times \beta\mathbb{N} \to \beta\mathbb{N}$. It is associative and left-continuous.*

Proof. First, we show that $\mathcal{U} + \mathcal{V}$ *is* an ultrafilter. Clearly $\emptyset \notin \mathcal{U} + \mathcal{V}$, and if $A \in \mathcal{U} + \mathcal{V}$ and $B \supset A$ then $B \in \mathcal{U} + \mathcal{V}$. Suppose now that $A, B \in \mathcal{U} + \mathcal{V}$, i.e.

$$(\forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ x + y \in A) \text{ AND } (\forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ x + y \in B).$$

Then by proposition 24 (i),

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

whence in turn

$$\forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ (x + y \in A \text{ AND } x + y \in B),$$

i.e.

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

i.e.

$$A \cap B \in \mathcal{U} + \mathcal{V}$$

as required. Finally, suppose that $A \notin \mathcal{U} + \mathcal{V}$. Then

$$\{x : \forall_{\mathcal{V}} y\ x + y \in A\} \notin \mathcal{U}$$

so by Proposition 24 (iii),

$$\forall_{\mathcal{U}} x\ (\text{NOT } (\forall_{\mathcal{V}} y\ x + y \in A))$$

whence in turn

$$\forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ (\text{NOT } x + y \in A),$$

i.e.

$$\forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ (x + y \in A^c)$$

and so

$$A^c \in \mathcal{U} + \mathcal{V}.$$

So we have shown that $\mathcal{U} + \mathcal{V}$ is an ultrafilter.

We next observe that $+ : \beta\mathbb{N} \times \beta\mathbb{N} \to \beta\mathbb{N}$ is associative. Indeed, for any $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \beta\mathbb{N}$,

$$\begin{aligned}
\mathcal{U} + (\mathcal{V} + \mathcal{W}) &= \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x\ \forall_{\mathcal{V}+\mathcal{W}} t\ x + t \in A\} \\
&= \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x\ \{t : x + t \in A\} \in \mathcal{V} + \mathcal{W}\} \\
&= \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ \forall_{\mathcal{W}} z\ y + z \in \{t : x + t \in A\}\} \\
&= \{A \subset \mathbb{N} : \forall_{\mathcal{U}} x\ \forall_{\mathcal{V}} y\ \forall_{\mathcal{W}} z\ x + y + z \in A\} \\
&= (\mathcal{U} + \mathcal{V}) + \mathcal{W}.
\end{aligned}$$

Finally, we show that $+$ is left-continuous, i.e. that for fixed $\mathcal{V}$, the map $\mathcal{U} \mapsto \mathcal{U} + \mathcal{V}$ is continuous. So fix V and an open set C_A in our base for $\beta\mathbb{N}$. Then

$$\begin{aligned} \mathcal{U} + \mathcal{V} \in C_A &\iff A \in \mathcal{U} + \mathcal{V} \\ &\iff \{x : \forall_{\mathcal{V}} y \; x + y \in A\} \in \mathcal{U} \\ &\iff \mathcal{U} \in C_{\{x : \forall_{\mathcal{V}} y \; x+y \in A\}}, \end{aligned}$$

so $+$ is indeed left-continuous. □

Fact. The operation $+$ is neither commutative nor right-continuous.

We seek an idempotent ultrafilter, i.e. some $\mathcal{U} \in \beta\mathbb{N}$ such that $\mathcal{U} + \mathcal{U} = \mathcal{U}$. (Note that any such $\mathcal{U}$ must be non-principal, as $\tilde{n} + \tilde{n} = \widetilde{2n}$.)

Lemma 26 (The Idempotent Lemma). *Suppose X is a non-empty compact Hausdorff topological space and $+ \colon X \times X \to X$ is an associative and left-continuous binary operation. Then there is an element $x \in X$ such that $x + x = x$.*

Proof. Consider $S = \{Y \subset X : Y$ compact and nonempty, $Y + Y \subset Y\}$ (where by $Y + Y$ we mean $\{y + y' : y, y' \in Y\}$).

We first show that *S has a minimal element.* Clearly $X \in S$, so $S \neq \emptyset$, so by Zorn's Lemma it is sufficient to show that if $\{Y_i : i \in I\}$ is a chain in S then $Y = \cap_{i \in I} Y_i \in S$. Since in a compact Hausdorff space a set is compact precisely if it is closed, we see that Y is compact; also $Y \neq \emptyset$ since the Y_i are closed sets having the finite intersection property in the compact space X. Also, for y, $y' \in Y$ we have y, $y' \in Y_i$ for all i, so $y + y' \in Y_i + Y_i \subset Y_i$ for all i and so $y + y' \in Y$. Thus $Y \in S$, proving our claim.

Let Y be a minimal element of S and fix $x \in Y$. We will show that $x + x = x$.

We begin by showing that $Y + x \in S$. We see that $Y + x$ is non-empty and compact, since it is the continuous image of a compact set by left-continuity of $+$. Also, by associativity of $+$, $(Y + x) + (Y + x) = (Y + x + Y) + x \subset Y + x$. This shows that $Y + x \subset Y$.

Now, we know $Y + x \subset Y$, so we must have $Y + x = Y$ by minimality of Y. Hence there exists $y \in Y$ with $y + x = x$. Put $Z = \{y \in Y : y + x = x\}$.

We now show that $Z \in S$. By our remarks above, Z is non-empty. Note that $\{x\}$ is compact, and so closed in the compact Hausdorff subspace Y of X. So Z, which is the inverse image in Y of the set $\{x\}$ under the continuous map $y \mapsto y + x$ is closed, and so compact. Also, for y, $y' \in Z$, we have by associativity of $+$ that $(y + y') + x = y + (y' + x) = y + x = x$, and so $y + y' \in Z$. This shows that $Z \in S$.

But $Z \subset Y$ and so $Z = Y$. In particular, $x \in Z$ and so $x + x = x$ as desired. $\square$

Remark. Hence $Y + x = \{x\}$ and so $Y = \{x\}$.

Corollary 27. *There exists $\mathcal{U} \in \beta\mathbb{N}$ such that $\mathcal{U} + \mathcal{U} = \mathcal{U}$.*

Theorem 28 (Hindman's Theorem). *Whenever $\mathbb{N}$ is finitely coloured, there exist x_1, x_2, x_3, … with $\mathrm{FS}(x_1, x_2, x_3, \dots)$ monochromatic.*

Proof. Given $\mathbb{N} = A_1 \cup A_2 \cup \cdots \cup A_k$, choose an idempotent $\mathcal{U} \in \beta\mathbb{N}$. We have $A_i \in \mathcal{U}$ for some i; write $A = A_i$. (Intuitively, we think of A as the largest colour class.) So $\forall_{\mathcal{U}} y\ y \in A$. Also, as $\mathcal{U}$ is idempotent, $\forall_{\mathcal{U}} x \forall_{\mathcal{U}} y\ x + y \in A$. So $\forall_{\mathcal{U}} x \forall_{\mathcal{U}} y\ \mathrm{FS}(x, y) \subset A$ by Proposition 24. Pick x_1 with $\forall_{\mathcal{U}} y\ \mathrm{FS}(x_1, y) \subset A$.

Now suppose inductively that we have found x_1, x_2, …, x_n such that $\forall_{\mathcal{U}} y\ \ \mathrm{FS}(x_1, x_2, \dots, x_n, y) \subset A$. For each $z \in \mathrm{FS}(x_1, x_2, \dots, x_n, y)$ we have $\forall_{\mathcal{U}} y\ z + y \in A$ and so $\forall_{\mathcal{U}} x \forall_{\mathcal{U}} y\ z + x + y \in A$. Thus by Proposition 24, $\forall_{\mathcal{U}} x \forall_{\mathcal{U}} y\ \mathrm{FS}(x_1, x_2, \dots, x_n, x, y) \subset A$. Let x_{n+1} be such an x.

The result follows by induction. $\square$

3 Infinite Ramsey Theory

For infinite $M \subset \mathbb{N}$, write $M^{(\omega)}$ for the collection $\{L \subset M : L \text{ infinite}\}$ of all infinite subsets of M. Motivated by Theorem 2, we ask: if we finitely colour $\mathbb{N}^{(\omega)}$, is there an infinite monochromatic set (i.e. does there exist $M \in \mathbb{N}^{(\omega)}$ such that all $L \in M^{(\omega)}$ have the same colour)?

Proposition 29. *There is a 2-colouring of $\mathbb{N}^{(\omega)}$ without an infinite monochromatic set.*

Proof. We construct a 2-colouring c such that for all $M \in \mathbb{N}^{(\omega)}$ and all $x \in M$ we have $c(M - \{x\}) \neq c(M)$—this is clearly sufficient to prove the proposition.

Define a relation $\sim$ on $\mathbb{N}^{(\omega)}$ by $L \sim M \iff |L \Delta M| < \infty$. This is clearly an equivalence relation. Let the equivalence classes be $\{E_i : i \in I\}$, and for each i choose $M_i \in E_i$. Now define $c(M)$ to be RED if $|M \Delta M_i|$ is even for some $i \in I$ and to be BLUE if $|M \Delta M_i|$ is odd for some $i \in I$. It is easy to check that this colouring has the required property. $\square$

Note that in the above proof we used AC.

A 2-colouring of $\mathbb{N}^{(\omega)}$ corresponds to a partition $\mathbb{N}^{(\omega)} = Y \cup Y^c$ for some $Y \subset \mathbb{N}^{(\omega)}$. A collection $Y \subset \mathbb{N}^{(\omega)}$ is called *Ramsey* if there exists $M \in \mathbb{N}^{(\omega)}$

with $M^{(\omega)} \subset Y$ or $M^{(\omega)} \subset Y^c$. So Proposition 29 says that 'not all sets are Ramsey'.

We can induce the subspace topology on $\mathbb{N}^{(\omega)} \subset \mathcal{P}(\mathbb{N})$, where we identify $\mathcal{P}(\mathbb{N})$ with $\{0,1\}^{\mathbb{N}}$ with the product topology. So a basic open neighbourhood of $M \in \mathbb{N}^{(\omega)}$ is $\{L \in \mathbb{N}^{(\omega)} : L \cap [n] = M \cap [n]\}$ for some n. Writing $M^{(<\omega)}$ for the collection $\{A \subset M : A \text{ finite}\}$ of finite subsets of M, we have a base of open sets for $\mathbb{N}^{(\omega)}$:

$$\{M \in \mathbb{N}^{(\omega)} : A \text{ an initial segment of } M\}, \qquad A \in \mathbb{N}^{(<\omega)}.$$

Equivalently, we have a metric

$$d(L, M) = \begin{cases} 0 & \text{if } L = M \\ 1/\min(L \Delta M) & \text{if } L \neq M \end{cases}.$$

We call this the *τ-topology* or *usual topology* or *product topology* on $\mathbb{N}^{(\omega)}$.

For $A \in \mathbb{N}^{(<\omega)}$ and $M \in \mathbb{N}^{(\omega)}$, write

$$(A, M)^{(\omega)} = \{L \in \mathbb{N}^{(\omega)} : A \text{ is an initial segment of } L \text{ and } L - A \subset M\}.$$

(We think of this as the collection of sets which 'start as A and carry on inside M'.)

For fixed $Y \subset \mathbb{N}^{(\omega)}$, we say that M *accepts* A *(into* Y*)* if $(A, M)^{(\omega)} \subset Y$, and that M *rejects* A if no $L \in M^{(\omega)}$ accepts A.

Notes. 1. If M accepts A then every $L \in M^{(\omega)}$ accepts A as well.

2. If M rejects A then every $L \in M^{(\omega)}$ rejects A as well.

3. If M accepts A and $A \subset B \subset A \cup M$, then M accepts B as long as $\max A \leq \min M$.

4. M need not accept or reject A.

Lemma 30 (The Galvin-Prikry Lemma). *Given $Y \subset \mathbb{N}^{(\omega)}$, there exists a set $M \in \mathbb{N}^{(\omega)}$ such that either*

(i) M accepts $\emptyset$ into Y; or

(ii) M rejects all of its finite subsets.

Proof. Suppose no $M \in \mathbb{N}^{(\omega)}$ accepts $\emptyset$, i.e. that $\mathbb{N}$ rejects $\emptyset$. We shall inductively construct infinite subsets $M_1 \supset M_2 \supset M_3 \supset \cdots$ of $\mathbb{N}$ and $a_1, a_2, a_3, \ldots \in \mathbb{N}$ with $a_i \in M_i$ for all i and such that M_i rejects all subsets of $\{a_1, a_2, \ldots, a_{i-1}\}$. Then we shall be done, for $\{a_1, a_2, a_3, \ldots\}$ rejects all its finite subsets.

Take $M_1 = \mathbb{N}$. Having chosen $M_1 \supset M_2 \supset \cdots \supset M_k$ and $a_1, a_2, \ldots, a_{k-1}$ as above, we seek $a_k \in M_k$ and $M_{k+1} \subset M_k$ such that M_{k+1} rejects all finite subsets of $\{a_1, a_2, \ldots, a_k\}$.

Suppose this is impossible. Fix $b_1 \in M_k$ with $b_1 > a_i$ for $1 \leq i \leq k-1$. We cannot take $a_k = b_1$ and $M_{k+1} = M_K$ so some $N_1 \in M_k^{(\omega)}$ accepts some subset S of $\{a_1, a_2, \ldots, a_{k-1}, b_1\}$. And S must be of the form $E_1 \cup \{b_1\}$ as M_k rejects all subsets of $\{a_1, a_2, \ldots, a_{k-1}\}$. Now pick $b_2 \in N_1$ with $b_2 > b_1$ and try $a_k = b_2$ and $M_{k+1} + N_1$. We get $N_2 \in N_1^{(\omega)}$ accepting a subset of $\{a_1, a_2, \ldots, a_{k-1}, b_2\}$—say $N_2 = E_2 \cup \{b_2\}$. Keep going: we get $M_k \supset N_1 \supset N_2 \supset \cdots$ and $b_1 < b_2 < b_3 < \cdots$ $(b_1 \in N_{i-1})$, together with $E_1, E_2, E_3, \ldots \subset \{a_1, a_2, \ldots, a_{k-1}\}$ such that $E_n \cup \{b_n\}$ is accepted by N_n for all n. Passing to a subsequence if necessary, we may assume without loss of generality that $E_n = E$ for all n. Then E is accepted by $\{b_1, b_2, b_3, \ldots\}$, contradicting the definition of M_k. □

Theorem 31. *If Y is open then Y is Ramsey.*

Proof. By Galvin-Prikry, there exists $M \in \mathbb{N}^{(\omega)}$ with either

(i) M accepting $\emptyset$; or

(ii) M rejecting all its finite subsets.

If (i) then we have $M^{(\omega)} \subset Y$.

If (ii) then we will show $M^{(\omega)} \subset Y^c$. Indeed, suppose we have $L \in M^{(\omega)}$ with $L \in Y$. Since Y is open, we must have $(A, \mathbb{N})^{(\omega)} \subset Y$ for some initial segment A of L. So in particular, we have $(A, M)^{(\omega)} \subset Y$, i.e. M accepts A, a contradiction. □

Remark. A collection Y is Ramsey if and only if Y^c is Ramsey, so Theorem 31 also says that 'closed sets are Ramsey'.

The $\star$-*topology* or *Ellentuck topology* or *Mathias topology* on $\mathbb{N}^{(\omega)}$ has basic open sets $(A, M)^{(\omega)}$ for $A \in \mathbb{N}^{(<\omega)}$ and $M \in \mathbb{N}^{(\omega)}$. This *is* a base for a topology on $\mathbb{N}^{(\omega)}$:

- $\mathbb{N}^{(\omega)} = (\emptyset, N)^{(\omega)}$ so the union of our putative basic sets is indeed $\mathbb{N}^{(\omega)}$;
- if $(A, M)^{(\omega)}$ and $(A', M')^{(\omega)}$ are basic sets then $(A, M)^{(\omega)} \cap (A', M')^{(\omega)}$ is either $(A \cup A', M \cap M')^{(\omega)}$ or $\emptyset$.

Note that the $\star$-topology is stronger (i.e. has more open sets) than the usual topology.

Theorem 32. *If Y is $\star$-open then Y is Ramsey.*

Proof. Choose $M \in \mathbb{N}^{(\omega)}$ as given by Galvin-Prikry.

(i) If M accepts $\emptyset$ then $M^{(\omega)} \subset Y$.

(ii) If M rejects all its finite subsets then we shall show that $M^{(\omega)} \subset Y^c$. Indeed, suppose $L \in M^{(\omega)}$ with $L \in Y$. Since Y is $\star$-open, we must have $(A, L)^{(\omega)} \subset Y$ for some initial segment A of L. So L accepts A, contradicting 'M rejects A'. □

We say $Y \subset \mathbb{N}^{(\omega)}$ is *completely Ramsey* if for all $A \in \mathbb{N}^{(<\omega)}$ and all $M \in \mathbb{N}^{(\omega)}$ there is some $L \in M^{(\omega)}$ such that $(A, L)^{(\omega)}$ is contained in either Y or Y^c.

This is a stronger property than being Ramsey. For example, let Y be the non-Ramsey set from Proposition 29 and set

$$Y' = Y \cup \{M \in \mathbb{N}^{(\omega)} : 1 \notin M\}.$$

Then certainly Y' is Ramsey, as $\{2, 3, 4, \dots\}^{(\omega)} \subset Y^c$. But Y' is not completely Ramsey; $A = \{1\}$ and $M = \mathbb{N}$ yield no L as desired.

Theorem 33. *If Y is $\star$-open then Y is completely Ramsey.*

Proof. Given $A \in \mathbb{N}^{(<\omega)}$ and $M \in \mathbb{N}^{(\omega)}$, we seek $L \in M^{(\omega)}$ with $(A, L)^{(\omega)}$ contained in Y or Y^c. Now view $(A, M)^{(\omega)}$ as a copy of $\mathbb{N}^{(\omega)}$ as follows. We may assume without loss of generality that $\max A < \min M$. Write $M = \{m_1, m_2, m_3, \dots\}$, where $m_1 < m_2 < m_3 < \cdots$, and define a function $f\colon \mathbb{N}^{(\omega)} \to (A, M)^{(\omega)}$ by $N \mapsto A \cup \{M_i : i \in N\}$. Clearly f is a homeomorphism in the $\star$-topology.

Let $Y' = \{N \in \mathbb{N}^{(\omega)} : f(N) \in Y\}$. Then Y' is $\star$-open since Y is $\star$-open. So by Theorem 32, there exists $L \in \mathbb{N}^{(\omega)}$ with $L^{(\omega)}$ contained in either Y or Y^c. Thus $\{f(N) : N \in L^{(\omega)}\}$ is contained in either Y or Y^c, i.e. $(A, f(L))^{(\omega)}$ is contained in either Y or Y^c. □

So we know that, in the $\star$-topology, all 'locally big' (i.e. open) sets are completely Ramsey. Now we consider 'locally small' (i.e. nowhere dense) sets.

Given a space X, we say that $A \subset X$ is *nowhere dense* if A is not dense in any non-empty open subset, equivalently if for any non-empty open O, there is a non-empty open $O' \subset O$ such that $O' \cap A = \emptyset$, equivalently if $\bar{A}$ has empty interior. For example, $\mathbb{N}$ is nowhere dense in $\mathbb{R}$.

Proposition 34. *A set $Y \subset \mathbb{N}^{(\omega)}$ is $\star$-nowhere-dense if and only if for all $a \in \mathbb{N}^{(<\omega)}$ and all $M \in \mathbb{N}^{(\omega)}$, there is some $L \in M^{(\omega)}$ with $(A, L)^{(\omega)} \subset Y^c$.*

Proof. The first statement says that inside $(A, M)^{(\omega)}$ there is some $(B, L)^{(\omega)}$ missing Y while the second says that inside $(A, M)^{(\omega)}$ there is some $(A, L)^{(\omega)}$ missing Y. So it is immediate that the second statement implies the first.

So suppose that Y is $\star$-nowhere-dense. Then $\bar{Y}$ has non-empty interior and so $\bar{Y}$ is $\star$-nowhere-dense (since $\bar{\bar{Y}} = \bar{Y}$). But $\bar{Y}$ is completely Ramsey by Theorem 33 and so inside $(A, M)^{(\omega)}$ there exists some $(A, L)^{(\omega)}$ contained in either $\bar{Y}$ or $(\bar{Y})^c$. But $\operatorname{int} \bar{Y} = \emptyset$ so $(A, L)^{(\omega)} \subset (\bar{Y})^c$ and so $(A, L)^{(\omega)} \subset Y^c$ as required. $\square$

A subset A of a topological space X is called *meagre* or *of first category* if $A = \bigcup_{n=1}^{\infty} A_n$ with each A_n nowhere dense. For example, $\mathbb{Q}$ is meagre in $\mathbb{R}$.

We can usually think of meagre sets as being small: for example, the Baire Category Theorem states that if X is a non-empty complete metric space and A is a meagre subset of X then $A \neq X$.

Theorem 35. *Let Y be $\star$-meagre. Then for all $A \in \mathbb{N}^{(<\omega)}$ and all $M \in \mathbb{N}^{(\omega)}$, there is some $L \in M^{(\omega)}$ such that $(A, L)^{(\omega)} \subset Y^c$. In particular, Y is $\star$-nowhere-dense.*

Proof. Suppose we are given $A \in \mathbb{N}^{(<\omega)}$ and $M \in \mathbb{N}^{(\omega)}$. Write $Y = \bigcup_{n-1}^{\infty} Y_n$ with each Y_n $\star$-nowhere-dense.

By Proposition 34, we have $M_1 \in M^{(\omega)}$ with $(A, M_1)^{(\omega)} \subset Y_1^c$. Choose $x_1 \in M_1$ with $x_1 > \max A$.

Again by Proposition 34, we have $M_2' \in M_1^{(\omega)}$ with $(A, M_2')^{(\omega)} \subset Y_2^c$ and then $M_2 \in M_2'^{(\omega)}$ with $(A \cup \{x_1\}, M_2)^{(\omega)} \subset Y_2^c$. Choose $x_2 \in M_2$ with $x_2 > x_1$.

Applying Proposition 34 four times, once for each subset of $\{x_1, x_2\}$, we get $M_3 \in M_2^{(\omega)}$ such that each of the sets $(A, M_3)^{(\omega)}$, $(A \cup \{x_1\}, M_3)^{(\omega)}$, $(A \cup \{x_2\}, M_3)^{(\omega)}$ and $(A \cup \{x_1, x_2\}, M_3)^{(\omega)}$ is contained in Y_3^c.

Keep going: we obtain $M \supset M_1 \supset M_2 \supset \cdots$ and $\max A < x_1 < x_2 < \cdots$ with $x_n \in M_n$ for all n and $(A \cup F, M_n)^{(\omega)} \subset Y_n^c$ for all $F \subset \{x_1, x_2, \dots x_n\}$. Then $(A, \{x_1, x_2, \dots\})^{(\omega)} \subset Y_n^c$ for all n and so $(A, \{x_1, x_2, \dots\})^{(\omega)} \subset Y^c$. $\square$

A set A in a topological space is a *Baire set*, or *has the property of Baire*, if $A = O \Delta M$ for some open O and meagre M. We can think of A as being 'nearly open'.

Notes. 1. If A is open then A is Baire.

2. If A is closed then A is Baire, since $A = \operatorname{int} A \Delta (A - \operatorname{int} A)$, where $A - \operatorname{int} A$ is nowhere dense as it is closed and contains no non-empty open set.

The Baire sets form a σ-algebra:

(i) X is Baire.

(ii) If Y is Baire, say $Y = O\Delta M$, then

$$\begin{aligned} Y^c &= O^c\Delta M \\ &= (O'\Delta M')\Delta M \qquad \text{(since } O^c \text{ is closed, and using note (ii) above)} \\ &= O'\Delta(M'\Delta M) \end{aligned}$$

so Y^c is Baire.

(iii) If the sets Y_1, Y_2, Y_3, ... are Baire, say $Y_i = O_i\Delta M_i$, then their union $\bigcup_{n=1}^{\infty} Y_n = (\bigcup_{n=1}^{\infty} O_n)\Delta M$ for some $M \subset \bigcup_{n=1}^{\infty} M_n$ and so $\bigcup_{n=1}^{\infty} Y_n$ is Baire.

Since we noted above that open sets are Baire, it follows that any Borel set is Baire.

Theorem 36. *A collection Y is completely Ramsey if and only if it is $\star$-Baire.*

Proof. $\Leftarrow$ Suppose Y is $\star$-Baire, so $Y = W\Delta Z$ with W open and Z meagre.

Given $(A, M)^{(\omega)}$, we have $L \in M^{(\omega)}$ with $(A, L)^{(\omega)}$ contained in either W or W^c (by Theorem 33) and $N \in L^{(\omega)}$ with $(A, N)^{(\omega)} \subset Z^c$ (by Theorem 35). So either

$$(A, N)^{(\omega)} \subset W \cap Z^c \subset Y$$

or

$$(A, N)^{(\omega)} \subset W^c \cap Z^c \subset Y^c$$

and Y is completely Ramsey as required.

$\Rightarrow$ Suppose conversely that Y is completely Ramsey. We can write $Y = \operatorname{int} Y\Delta(Y - \operatorname{int} Y)$. so it will be sufficient for us to show that $Y - \operatorname{int} Y$ is $\star$-nowhere-dense; we show in particular that given any base set $(A, M)^{(\omega)}$ in the $\star$-topology, there is a non-empty open set inside it which is disjoint from $Y - \operatorname{int} Y$—indeed, we have $L \in M^{(\omega)}$ with $(A, L)^{(\omega)}$ contained in either Y or Y^c.

If $(A, L)^{(\omega)} \subset Y$ then $(A, L)^{(\omega)}$ is disjoint from $Y - \operatorname{int} Y$.

If $(A, L)^{(\omega)} \subset Y^c$ then again $(A, L)^{(\omega)}$ is disjoint from $Y - \operatorname{int} Y$.

So Y is $\star$-Baire, as required. □

Thus any $\star$-Borel set is completely Ramsey, and so certainly any τ-Borel set is Ramsey.

Note. Without Theorem 35, we would have shown that Y is completely Ramsey if and only if Y is the symmetric difference of an open set and a nowhere dense set, and we would not have known that the completely Ramsey sets form a σ-algebra.

Typeset in LaTeX 2ε by Paul A. Russell.