Chapter 1: Ramsey Theory

Thomas Britz

1 The Pigeonhole Principle

The Pigeonhole Principle (simple)
If $k+1$ pigeons are put into k pigeonholes, then some pigeonhole contains at least two pigeons.

Example. Roll seven six-sided dice. Let the dice values rolled be "pigeons" and the six possible values be "pigeonholes". The Pigeonhole Principle tells us that at least one value must appear more than once.

Exercise. Six distinct numbers are chosen from the set $\{0,\ldots,9\}$.
Show that at least two of the chosen numbers must be consecutive.

Exercise. $n+1$ distinct numbers are chosen from the numbers $1,\ldots,2n$.
Show that at least two of the chosen numbers are coprime.

Exercise. Can we cover a chessboard with two opposite corners removed using dominoes?

Theorem (new proof by Mixon 2012)
There are infinitely many primes.

Proof. Assume that there is a finite number of primes, $p_1,\ldots,p_N$. Choose some integer K such that $2^K > (K+1)^N$. There are 2^K numbers ("pigeons") in the set $S=\{1,\ldots,2^K\}$ with at most $(K+1)^N$ factorisations $p_1^{k_1},\ldots,p_N^{k_N}$ ("pigeonholes"), since

$$K = \log_2 2^K \geq \log_2 p_1^{k_1}\ldots p_N^{k_N} = \sum_{i=1}^{N} k_i \log_2 p_i \geq \sum_{i=1}^{N} k_i \geq \max_{1\leq i\leq N} k_i\,.$$

By the Pigeonhole Principle, two numbers in S have identical factorisation. They must then be equal, giving a contradiction. □

The next theorem shows that every real number can be approximated by a rational, and that this approximation converges quickly.

Dirichlet's Approximation Theorem (1834)
If $\alpha\in\mathbb{R}$ and $N\in\mathbb{N}$, then $p,q\in\mathbb{N}$ exist so that $q\leq N$ and $|\alpha-\frac{p}{q}|<\frac{1}{Nq}$.

Proof. Replace the latter inequality equivalently by $|q\alpha-p|<\frac{1}{N}$. Partition the unit interval $[0,1)$ into N equal subintervals. Define $N+1$ residues $r_i := i\alpha-\lfloor i\alpha\rfloor$ for $i=0,1,\ldots,N$. These lie in the N subintervals. Let $i<j$. By the Pigeonhole Principle, residues r_i, r_j $(i<j)$ lie in some common subinterval. Set $p:=\lfloor j\alpha\rfloor-\lfloor i\alpha\rfloor$ and $q:=j-i$. Then $0<q\leq N$, and

$$|q\alpha-p| = |\lfloor j\alpha\rfloor-\lfloor i\alpha\rfloor-(j-i)\alpha| = |(i\alpha-\lfloor i\alpha\rfloor)-(j\alpha-\lfloor j\alpha\rfloor)| = |r_i-r_j| < \frac{1}{N}\,.$$ □

Lemma. If $p \equiv 1 \pmod 4$ where p is prime, then $a^2 \equiv -1 \pmod p$ for some a.

Proof. Suppose that $p \equiv 1 \pmod 4$ and set $a := \left(\frac{p-1}{2}\right)!$. Note that $p - i \equiv -i \pmod p$ and apply Wilson's Theorem, which states that $(p-1)! \equiv -1 \pmod p$. Then

$$\begin{aligned} a^2 = \left(\left(\frac{p-1}{2}\right)!\right)^2 &\equiv \left(\left(\frac{p-1}{2}\right)!\right)^2 (-1)^{(p-1)/2} \pmod p \\ &\equiv 1 \times 2 \times \cdots \times \left(\frac{p-1}{2}\right)\left(-\frac{p-1}{2}\right) \times \ldots \times (-2)(-1) \pmod p \\ &\equiv 1 \times 2 \times \cdots \times \left(\frac{p-1}{2}\right)\left(\frac{p+1}{2}\right) \times \ldots \times (p-2)(p-1) \pmod p \\ &\equiv 1 \times 2 \times \cdots \times (p-1) \pmod p \\ &\equiv -1 \pmod p . \end{aligned}$$

□

Theorem (Fermat 1640, Euler 1747)
Each prime $p \equiv 1 \pmod 4$ can be written as a sum of squares $p = x^2 + y^2$.

Proof. By the above lemma, we can find some a with $a^2 \equiv -1 \pmod p$.

Now, consider the pairs of integers (x, y) with $0 \le x, y < \sqrt{p}$. There are $(\lfloor\sqrt{p}\rfloor + 1)^2 > p$ such pairs and at most p values $ax - y \pmod p$. By the Pigeonhole Principle, two distinct pairs (x_1, y_1) and (x_2, y_2) must satisfy $ax_1 - y_1 \equiv ax_2 - y_2 \pmod p$. Set $x := x_1 - x_2$ and $y : y_1 - y_2$. Then

$$x^2 + y^2 \equiv -a^2x^2 + y^2 \equiv -(a(x_1 - x_2))^2 + (y_1 - y_2)^2 \equiv 0 \pmod p .$$

Since (x_1, y_1) and (x_2, y_2) are distinct and $0 \le x_i, y_i < \sqrt{p}$, we see that

$$0 < x^2 + y^2 < (\sqrt{p})^2 + (\sqrt{p})^2 = 2p .$$

Hence, $x^2 + y^2 = p$. □

The following theorem states that no lossless compression algorithm works 100% of the time!

Theorem
Let f be a *lossless digital data compression algorithm.*
If $|f(B)| < |B|$ for some data B, then $|f(A)| > |A|$ for some data A.

Proof. Represent data as binary strings and let $|A|$ now denote the string length of A.
Suppose that $|f(B)| < |B|$ for at least one string B and assume that $|f(A)| \le |A|$ for all strings A.

Let n be the smallest numbers with $|f(B')| < |B'| = n$ for some string B'. Set $m := |f(B')| (< n)$ and let $\mathcal{F}$ be the set of strings B with $|f(B)| = m$. Each of the 2^m binary strings A of length m must obey $|f(A)| \ge |A| = m$. By the assumption, we have that $|f(A)| = |A| = m$ and so $A \in \mathcal{F}$.

Also, B' is also in $\mathcal{F}$ so $|\mathcal{F}| \ge 2^m + 1$. So, by the Pigeonhole Principle, some distinct B_1 and B_2 satisfy $f(B_1) = f(B_2)$. But f is lossless and therefore injective, which raises a contradiction. □

The Pigeonhole Principle (general)
If $km+1$ pigeons are put into k pigeonholes, then some pigeonhole contains at least $m+1$ pigeons.

Example. If there are 3 pigeonholes and 7 pigeons, then some pigeonhole must contain at least 3 pigeons. (Here, $m = 2$ and $k = 3$.)

Exercise. 19 points are drawn in a square with side lengths 1.
Show that at least 3 points lie in a circle with radius less than $\frac{1}{4}$.

The Erdős-Szekeres Theorem (1935)
Each sequence $a_1, \ldots, a_{n^2+1}$ of $n^2 + 1$ distinct integers has a subsequence with $n + 1$ elements that is either increasing or decreasing.

Proof. Suppose that there is no increasing subsequence that has $n + 1$ terms. Let n_i be the longest length of an increasing subsequence starting in a_i, and note that the "pigeons" n_i can only be one of the "pigeonhole" values $1, \ldots, n$.

By the Pigeonhole Principle, there are at least $n + 1$ identical n_i values. Suppose that $n_i = n_j$ for some $i < j$. Then $a_i \not< a_j$, so $a_i > a_j$. We thus have a decreasing subsequence of at least $n+1$ a_i's. □

Example. Consider the following sequence of $n^2 + 1$ integers where $n = 3$:

7 8 9 4 5 6 1 2 3 0

There is no increasing subsequence of size $n + 1 = 4$. However, there are 27 decreasing subsequences of size 4: any of the first three, followed by any of the second three, followed by any of the third three, followed by 0.

Exercise. Let C be a binary code of length $n = 6$ and minimum distance $d = 3$.
Use the Pigeonhole Principle to show that C has at most 8 codewords.

The Pigeonhole Principle (strong)
If $(n_1 - 1) + \cdots + (n_k - 1) + 1$ pigeons are put into k pigeonholes,
then some i^{th} pigeonhole contains at least n_i pigeons.

2 Ramsey's Theorem

Ramsey Theory makes repeated use of the Pigeonhole Principle to show that order exists whenever (certain) random and seemingly unordered structures are sufficiently large.

Exercise. Colour each edge of the complete graph K_6 either red or blue.
Show that there must be a red triangle or a blue triangle.

Ramsey's Theorem (1930, simple)
If $k, m \in \mathbb{N}$ and n is sufficiently large, then each k-colouring of the edges of K_n gives a complete subgraph K_m with monochromatic edges.

Proof. Choose a vertex v_0 in $V_0 := V(K_n)$. Let c_1 be a most-occurring edge colour from v_0, and let V_1 be the vertices with c_1-coloured edge to v_0. Then choose $v_1 \in V_1$ and let c_2 be the most-occurring edge colour from v_1 to vertices in V_1. Let V_2 be the V_1 vertices with c_2-coloured edges to v_1.

For big enough n, we can choose $N := k(m-2)+1$ colours $c_1, c_2, \ldots, c_N$ and vertices $v_0, v_1, \ldots, v_N$. By the Pigeonhole Principle, some $m-1$ colours c_i are identical, say colour c, and the m corresponding vertices v_i induces a K_m subgraph with c-coloured edges. □

Remark. The least such n is denoted $R(m; k)$.

Exercise. Show that $R(3; 2) = 6$.

Exercise. Use the method of the proof of Ramsey's Theorem to show that $R(3; 2) \leq 6$.

Ramsey's Theorem (1930, reduced)
If $n_1, \ldots, n_k \in \mathbb{N}$ and n is sufficiently large, then each colouring of the edges of K_n with colours $c_1, \ldots, c_k$ gives a subgraph K_{n_i} with all edges coloured c_i for some i.

Remark. The least such n is denoted $R(n_1, \ldots, n_k)$. Note that $R(m; k) = R(\underbrace{m, \ldots, m}_{k \text{ times}})$.

Example. $R(3, 3) = R(3; 2) = 6$.

The proof of this version of Ramsey's Theorem is identical to the first proof except that uses the strong version of the Pigeonhole Principle.

Exercise. Complete the proof of the theorem above.

Lemma. (Erdős & Szekeres 1935)

$$R(k, \ell) \leq R(k-1, \ell) + R(k, \ell-1)$$

Proof. We use induction on $k+\ell$, noting that $R(k, 1) = R(1, k) = 1$ for all k. Assume that $R(k-1, \ell)$ and $R(k, \ell-1)$ are well-defined and set $N := R(k-1, \ell)+R(k, \ell-1)$. Colour each edge of K_N either red or blue and choose a vertex v. Let R and B be the vertices with red and blue edges to v, respectively. Then

$$|R| + |B| = N - 1 = R(k-1, \ell) + R(k, \ell-1) - 1\,.$$

Therefore, $|R| \geq R(k-1, \ell)$ or $|B| \geq R(k, \ell-1)$. Assume without loss of generality that $|R| \geq R(k-1, \ell)$. If the vertices R induce a blue-edged K_ℓ, then we are done. Otherwise, by definition of $R(k-1, \ell)$, R induces a red-edged K_{k-1}. By adding the v and its red edges, we have a red-edged K_k. □

Example. For any $k \geq 2$, we note that $R(k,2) = R(2,k) = k$. The bound in the above lemma is therefore tight for $R(3,3)$:

$$R(3,3) = 6 \leq 3 + 3 = R(2,3) + R(3,2)\,.$$

Lemma.

$$R(n_1, \ldots, n_k) \leq R(n_1, \ldots, n_{k-2}, R(n_{k-1}, n_k))$$

Proof. We will use induction on $k \geq 2$, noting that the previous lemma gives the case $k = 2$.
Assume that $R(m_1, \ldots, m_{k-1})$ is well-defined for all m_i and set $N := R(n_1, \ldots, n_{k-2}, R(n_{k-1}, n_k))$.

Colour each edge of K_N in colours $c_1, \ldots, c_k$. Now, recolour each c_{k-1}- and c_k-coloured edges by a new colours c'_{k-1}. By definition, K_N contains either a subgraph K_{n_i} with c_i-coloured edges for some $i \leq k-2$ or a c'_{k-1}-coloured $K_{R(n_{k-1},n_k)}$ subgraph.

Suppose that K_N contains the latter subgraph. This subgraph originally had colours c_{k-1}, c_k, so it (and thus K_N) contained either a c_{k-1}-coloured $K_{n_{k-1}}$ or a c_k-coloured K_{n_k}. Hence, K_N had a c_i-coloured K_{n_i} subgraph for some $i \leq n$. □

Alternate proof of Ramsey's Theorem (reduced). Apply the lemma above. □

Definition.

$$[n] := \{1, \ldots, n\} \qquad \text{and} \qquad \binom{S}{k} := \{X \subseteq S : |X| = k\}$$

Example.

$$[3] = \{1,2,3\} \qquad \binom{[3]}{2} = \big\{\{1,2\},\{1,3\},\{2,3\}\big\}$$

We will now use this new notation to rewrite Ramsey's Theorem.

Ramsey's Theorem (1930)
If $n_1, \ldots, n_k, r \in \mathbb{N}$ and n is sufficiently large, then each colouring of $\binom{[n]}{r}$ with colours $c_1, \ldots, c_k$ gives a c_i-coloured subfamily $\binom{S}{r}$ for some i and n_i-subset $S \subseteq [n]$.

Remark. The least such n is denoted $R_r(n_1, \ldots, n_k)$.

Proof. We generalise the original proof, using induction on r, having already proved the case $r = 2$. Assume that the theorem is true for $r-1$ and set $t := R_{r-1}(n_1, \ldots, n_k)$.

Choose $r-2$ elements $v_1, \ldots, v_{r-2}$ and set $V_{r-1} := [n] - \{v_1, \ldots, v_{r-2}\}$. Continue to choose elements $v_{r-1}, v_r, \ldots$ and sets $V_r, V_{r+1}, \ldots$ as follows. Suppose that we have chosen $v_1, \ldots, v_{i-1}$ and V_i. Choose $v_i \in V_i$. Let V_{i+1} be a maximal subsets of $V_i - \{v_i\}$ so that $T \cup \{v\}$ and $T \cup \{w\}$ have the same colour for all $v, w \in V_{i+1}$ and $(r-1)$-subsets $T \subseteq \{v_1, \ldots, v_i\}$. If n is sufficiently large, then it is possible to choose t elements $v_1, \ldots, v_t$.

Now colour each $(r-1)$-subset T of these, giving T the same colour as $T \cup \{v_t\}$ if $v_t \notin V_t$ and colouring it randomly otherwise. By definition of t, there is a n_i-subset S of the v_i's whose $(r-1)$-subsets each have colour c_i for some i.

Let $U \subseteq S$ be an r-subset and let v be the greatest element in U. Then $T := U - \{v\}$ is a c_i-coloured $(r-1)$-sized subset of S and $T \cup \{v_t\}$ must therefore be c_i-coloured. By construction, $T \cup \{v_t\}$ and $U = T \cup \{v\}$ have the same colour, c_i. We see that each r-subset $U \subseteq S$ is c_i-coloured. □

Remark. The reduced version of Ramsey's Theorem occurs when $r = 2$.

3 Arithmetic Progressions

Definition. An *arithmetic progression* is a sequence of integers of the form $a, a+d, a+2d, \ldots, a+kd$.

Van der Waerden's Theorem (1927)
If $k, r \in \mathbb{N}$ and n is sufficiently large, then each k-colouring of $[n]$ gives a monochromatic arithmetic progression of length r.

Remark. The least such n is denoted $W(r,k)$.

Exercise. Explain why $W(2,k) = k+1$.

Example. Let us prove the theorem for $r=3$, $k=2$ by showing that $W(3,2) \leq 325$.

We begin by partitioning the set $[325]$ into blocks $B_1, \ldots, B_{65}$ each of five consecutive numbers. By the Pigeonhole Principle, the first $33 > 2^5$ B_ℓ's contain B_i and $B_{i+d''}$ with the same 2-colouring pattern. If these contain a monochromatic arithmetic progression of length 3, we are done.

Otherwise, note that the Pigeonhole Principle implies that at least two of the first 3 numbers of B_i, say $i+i'$ and $i+i'+d'$ must have the same colour, say red. Now consider the number $N := i+2d''+i'+2d'$ in the block B_{i+2d}. If N is red, then the numbers $a, a+d, a+2d(=N)$ where $a := i+i'$ and $d := d'+d''$ is a red arithmetic progression. Otherwise, N is green. Since we assumed that B_i and $B_{i+d''}$ contained no monochromatic arithmetic progression of length 3, the numbers $i+2d'$ and $i+2d'+d''$ must be green, so the numbers $a, a+d, a+2d(=N)$ where $a := i+i'+2d'$ and $d := d''$ is a green arithmetic progression.

Ramsey theorem often features arguments such as Cantor diagonalisation type argument in the example above. This is also at the core of the proof of Van der Waerden's Theorem, below, applying it to recursively to blocks of blocks of blocks. To express this without unduly cumbersome notation, let us prove a more general theorem whose proof relies on the following notions.

The i^{th} k-equivalence class of $[0,k]^m = \{0,\ldots,k\}^m$ is the set of vectors

$$(x_1, \ldots, x_{m-i}, \underbrace{k, \ldots, k}_{i \text{ times}}),$$

where $x_1, \ldots, x_{m-i} \neq k$.

Example. The three 3-equivalence classes of $[0,3]^2$ are

$$\{(0,0),(0,1),(0,2),(1,0),(1,1),(1,2),(2,0),(2,1),(2,2)\};$$
$$\{(0,3),(1,3),(2,3)\};$$
$$\{(3,3)\}.$$

Let $S(k,m)$ be the statement that

> for each $r \in \mathbb{N}$, there is some $N := N(k,m,r)$ so that for each r-colouring $\chi : [N] \to [r]$ there are $a, d_1, \ldots, d_m \in \mathbb{N}$ so that, for $x_i \in [k-1]$, $\chi(a + \sum_i x_i d_i)$ is well-defined and constant on each k-equivalence class of $[0,k]^m$.

Lemma. If $S(k,m)$ for some $m \geq 1$, then $S(k,m+1)$.

Proof. Suppose that $S(k,m)$ for some $m \geq 1$ and note that this implies $S(k,1)$. We wish to show that $S(k,m+1)$. Letting $r \in \mathbb{N}$, we can find $M := N(k,m,r)$ and $M' := N(k,1,r^M)$. Set $N = MM'$ and let $\chi : [N] \to [r]$ be given. There are at most distinct r^M vectors $(\chi((\ell-1)M+1), \ldots, \chi(\ell M))$ where $\ell \in [M']$, so number these vectors, inducing a colouring $\chi' : [M'] \to [r^M]$. Since $M' = N(k,1,r^M)$, we can find a', d' so that $\chi'(a' + x_1 d')$ is well-defined and constant for all $x_1 \in [0, k-1]$.

Set $d_1 := d'M$. Then $(\chi((a'-1)M + x_1 d_1 + 1), \ldots, \chi(a'M + x_1 d_1))$ is well-defined and constant for all $x_1 \in [0, k-1]$. Fix x_1 and define $\chi'' : [M] \to [r]$ by $\chi''(s) := \chi((a'-1)M + x_1 d_1 + s)$. Since $M = N(k,m,r)$, there are $a'', d_2, \ldots, d_{m+1}$ so that $\chi''(a'' + \sum_{i=2}^{m+1} x_i d_i)$ is well-defined and constant on each k-equivalence class of $[0,k]^m$. Setting $a := a'' + (a'-1)M$, we see that $\chi(a + \sum_{i=1}^{m+1} x_i d_i)$ is well-defined and constant on each k-equivalence class of $[0,k]^{m+1}$.

Hence, $N = MM' \geq N(k,m,r)$, and so $S(k,m+1)$. □

Lemma. If $S(k,m)$ for all $m \geq 1$, then $S(k+1,1)$.

Proof. Suppose that $S(k,m)$ for all $m \leq 1$. We wish to show that $S(k+1,1)$. Therefore, let $r \in \mathbb{N}$ and $\chi : [N(k,r,r)] \to [r]$. Then there are $a', d_1, \ldots, d_r \in \mathbb{N}$ so that $\chi(a' + \sum_i x_i d_i)$ is well-defined and constant on each k-equivalence class of $[0,k]^r$.

There are at most r function values of χ, so by the Pigeonhole Principle, at least two of the $r+1$ values $\chi(a'), \chi(a' + kd_r), \ldots, \chi(a' + \sum_{i=1}^{r} kd_i)$ must be the same, so $\chi(a' + \sum_{i=u}^{r} kd_i) = \chi(a' + \sum_{i=v}^{r} kd_i)$ for $u < v$. Set $a := a' + \sum_{i=v}^{r} kd_i$ and $d := \sum_{i=u}^{v-1} d_i$. We have seen that $\chi(a + xd) = \chi(a' + \sum_{i=1}^{u-1} 0d_i + \sum_{i=u}^{v-1} xd_i + \sum_{i=v}^{r} kd_i)$ is constant for $x \in [0, k-1]$. It also assumes this constant value in $x = k$:

$$\chi(a + kd) = \chi\Big(a' + \sum_{i=u}^{r} kd_i\Big) = \chi\Big(a' + \sum_{i=v}^{r} kd_i\Big) = \chi(a + 0d)$$

Hence, $\chi(a + xd)$ is constant for $x \in [0,k]$. Therefore, $S(k+1,1)$. □

Theorem
$S(k,m)$ for all $k, m \geq 1$.

Proof. The equivalence classes of $[0,1]^1$ are $\{0\}$ and $\{1\}$, so $S(1,1)$ is trivially true. The theorem therefore follows from the two above lemmas. □

Note that $S(k,1)$ is exactly Van der Waerden's Theorem, which we have now proved.

Polynomial Van der Waerden's Theorem (Bergelson & Leibman 1996)
If $k, r, n \in \mathbb{N}, p_1, \ldots, p_k \in \mathbb{Z}[x]$ with $p_i(0) = 0$ and n sufficiently large, then any r-colouring of $[n]$ gives monochromatic $a, a + p_1(d), \ldots, a + p_k(d)$.

The following theorem is one of the jewels of Ramsey Theorem, from which many other Ramsey results follow. We have essentially proved it by proving $S(k,m)$.

The Hales-Jewett Theorem (1963)
If $k, m, r \in \mathbb{N}$ and n is big enough, then any r-colouring of any cube $C = \{a + \sum_{i=1}^{n} x_i d_i : x_i \in [0,k]\}$ of dimension n and length k has a monochromatic subcube $C' \subseteq C$ of dimension m and length k.

In each of the Ramsey Theory results that we have seen so far, certain monochromatic substructures are guaranteed - but their particular colour is not given. Szemerédis Theorem (1975) goes beyond these theorems by guaranteeing monochromatic substructures of some *particular colour* if it has sufficient density. It is one of the most celebrated - and most difficult to prove - results in Ramsey Theory, and it and its proof form the basis of other great theorems, including the Green-Tao Theorem (2004).

4 Equations

Schur's Theorem (1916)
If $\mathbb{N}$ is finitely coloured, then $a+b=c$ for some same-coloured $a,b,c\in\mathbb{N}$.

Proof. Suppose that $\chi:\mathbb{N}\to[r]$ is an r-colouring of $\mathbb{N}$. By Ramsey's Theorem, we may define $N:=R(3;r)$. Define an edge r-colouring χ' of K_N by $\chi'(\{i,j\}):=\chi(|j-i|)$. By definition of N, there is a χ'-monochromatic triangle in K_N with vertices $i<j<k$. Namely, $\chi'(\{j,k\})=\chi'(\{i,j\})=\chi'(\{i,k\})$. Define $a:=k-j$, $b:=j-i$ and $c:=k-i$. Then $a+b=c$ and $\chi(a)=\chi(b)=\chi(c)$. □

From the proof, we also get the following finite version.

Schur's Theorem (finite version)
If $[n]$ is finitely coloured for sufficiently large n, then $a+b=c$ for some same-coloured $a,b,c\in[n]$.

In fact, all infinite Ramsey theory theorems have finite versions, by the *Compactness Principle*, a diagonal argument using the Axiom of Choice.

Theorem (Schur 1916)
For each $m\in\mathbb{N}$ and sufficiently large prime p, there are solutions $x,y,z\in\mathbb{N}^+$ to

$$x^m+y^m\equiv z^m \pmod p\,.$$

Proof. By the finite version of Schur's theorem, there is a large prime p so that if $[p-1]$ is m-coloured, then same-coloured $a,b,c\in[p]$ exist with $a+b=c$. Consider $\mathbb{Z}_p^*$, the multiplicative group of units in $\mathbb{Z}_p$. Define $H:=\{x^m:x\in\mathbb{Z}_p^*\}$. This is a subgroup of $\mathbb{Z}_p^*$ of index $r:=|\mathbb{Z}_p^*:H|=\gcd(p-1,m)\le m$. The r cosets gH partition $\mathbb{Z}_p^*$, defining an r-colouring χ of $\mathbb{Z}_p^*$: $\chi(a)=\chi(b)$ by if and only if $a,b\in gH$ for some g, i.e. when $ab^{-1}\in H$. Since $r\le m$, χ as an m-colouring of $\mathbb{Z}_p^*$ (or, equivalently, $[p-1]$). There are same-coloured $a,b,c\in\mathbb{Z}_p^*$ with $a+b=c$ or $ac^{-1}+bc^{-1}=1$. Thus, $ac^{-1},bc^{-1}\in H$, so $ac^{-1}\in x^m$, $bc^{-1}=y^m$, $1=z^m$ for $x,y,z\in\mathbb{Z}_p^*$. Then $x^m+y^m=z^m$ in $\mathbb{Z}_p$. □

The following theorem generalises for Van der Waerden's Theorem and Schur's Theorem.

Theorem
For each $k,r,s\ge1$, there is $n:=n(k,r,s)$ so that if $[n]$ is r-coloured, then for some $a,d>0$, the following set is in $[n]$ and is monochromatic:

$$\{a,a+d,\dots,a+kd\}\cup\{sd\}\,.$$

Proof. Use induction on r. Certainly, $n(k,1,s)$ exists and equals $\max\{k+1,s\}$. Assume that the theorem is true for $r-1$ and set $N:=n(k,r-1,s)$. By Van der Waerden's Theorem, we may set $W:=W(kN,r)$ and $n:=sW$. Let χ be an r-colouring of $[n]$. Then $[W]$ contains an arithmetic progression $\{a,a+d',\dots,a+(kN)d'\}$ of constant colour c. If $\chi(sjd')=c$ for some $j\in[N]$, then set $d:=jd'$ and we are done: $\{a,a+d,\dots,a+kd\}\cup\{sd\}$ has constant colour c.

Suppose that $\chi(sjd')\ne c$ for all $j\in[N]$. Define a colouring χ' on $[N]$ by $\chi'(j):=\chi(sjd')$.
Since $N=n(k,r-1,s)$, there is a χ'-monochromatic set $\{a',a'+d'',\dots,a'+kd''\}\cup\{sd''\}$ in $[N]$.
Set $a:=sa'd'$ and $d:=sd''d'$. Then $\{a,a+d,\dots,a+kd\}\cup\{sd\}$ in $[sNd']\subseteq[n]$ is monochromatic. □

The following corollary illustrates how to shift the arithmetic progression with respect to sd.

Corollary. For each $k, r, s \geq$, there is $n := n(k, r, s)$ so that if $[n]$ if r-coloured, then for some $a, d > 0$, the following set is in $[n]$ and is monochromatic:

$$\{a + \lambda d : |\lambda| \leq k\} \cup \{sd\}\,.$$

Proof. Applying $k' = 2k$ to the preceding theorem gives $a', d > 0$ for which $\{a' + \lambda d : \lambda \in [0, 2k]\} \cup \{sd\}$ in $[n]$ is monochromatic. Now relabel $a := a' + kd$. □

Rado's Theorem (1933, simple)
Let $c_1, \ldots, c_n \in \mathbb{Z}$. Then for any finite colouring of $\mathbb{N}$,

$$c_1 x_1 + \cdots + c_n x_n = 0$$

has a monochromatic solution $x_1, \ldots, x_n \in \mathbb{N}$ if and only if $\sum_{i \in I} c_i = 0$ for some subset $I \subseteq [n]$.

The full version of Rado's Theorem extends this to homogenous systems of linear equations.

Example. The numbers $a, a+d, \ldots, a+kd, sd$ form a solution $x_1, \ldots, x_{k+1}$ to the homogenous system of linear equations

$$\begin{aligned} x_1 - x_0 &= x_2 - x_1 \\ &\vdots \\ x_{k-1} - x_{k-2} &= x_k - x_{k-1} \\ x_{k+1} &= s(x_1 - x_0) \end{aligned} \quad ; \text{ that is, } \quad \begin{aligned} -x_0 + 2x_1 - x_2 &= 0 \\ &\vdots \\ -x_{k-2} + 2x_{k-1} - x_k &= 0 \\ sx_0 - sx_1 + x_{k+1} &= 0\,. \end{aligned}$$

Definition. For any $S \subseteq \mathbb{N}$, the *sum set* $\Sigma(S)$ is the set of all finite sums of elements of S.

Example.

$$\Sigma(\{1, 2, 4\}) = \{1, 2, 3, 4, 5, 6, 7\}$$

The following theorem follows as a special case of Rado's Theorem.

Folkman's Theorem (Rado 1970, Sanders 1968)
For $c, k \in \mathbb{N}$, there is some sufficiently large $M \in \mathbb{N}$ so that for any c-colouring of $\mathbb{N}$, there is a k-subset $S \subseteq [M]$ with monochromatic $\Sigma(S)$.

Definition. For any set $S \subseteq \mathbb{N}$, the *product set* $\Pi(S)$ is the set of all finite products of elements of S.

Example.

$$\Pi(\{1, 2, 4\}) = \{1, 2, 4, 8\}$$

Theorem
For $c, k \in \mathbb{N}$, there is some sufficiently large $M \in \mathbb{N}$ so that for any c-colouring of $\mathbb{N}$, there is a k-subsets $S \subseteq [M]$ with monochromatic $\Pi(S)$.

Proof. This result follows from Folkman's Theorem since any sum set in $[N]$ will induce a product set in $\{2^n : n \in [N]\}$. □

The Finite Union Theorem
If the finite subsets of $\mathbb{N}$ are finitely coloured, then there are arbitrarily large families of disjoint subsets $\mathcal{D} = \{D_i\}_{i \in I}$ whose union sets $\{\bigcup_{i \in J} D_i : J \subseteq I, 0 < |J| < \infty\}$ are monochromatic.

It is not hard to show that the finite version of this theorem is equivalent to Folkman's Theorem, via the correspondence between sets S and binary (characteristic) vectors 1_S and thus numbers $\sum_{i \in S} 2^{i-1}$.

Hindman's Theorem (1974)
For any finite colouring of $\mathbb{N}$, there is an infinite subsets $S \subseteq \mathbb{N}$ whose sum set $\Sigma(S)$ is monochromatic.

It is an open question whether arbitrarily large finite sets $S \subseteq \mathbb{N}$ can exist with monochromatic $\Sigma(S) \cup \Pi(S)$ when $\mathbb{N}$ is finitely coloured. Infinitely large such sets do not always exist: Hindman has shown this. This question is open for even the smallest sets $S = \{a, b\}$ and their sets $\Sigma(S) \cup \Pi(S) = \{a, b, a + b, ab\}$.

5 Graphs and Geometry

Theorem (Ramras 2002)
If $k, m \in \mathbb{N}$ such that $k \geq 2$ and N is sufficiently large, then among any N N-subsets $X \subseteq [k(N-1)+1]$, some m sets have at least m elements in common.

Proof. Let $m \in \mathbb{N}$ and assume that the theorem is false. Set $M := k(N-1)+1$. Then there is a family $\mathcal{X}$ of N N-subsets $X \in \binom{[M]}{N}$, each m sets of which have at most $m-1$ elements in common. Then

$$N\binom{N}{m} = \sum_{X\in\mathcal{X}} \sum_{Y\in\binom{X}{m}} 1 = \sum_{Y\in\binom{[M]}{m}} \sum_{X\in\mathcal{X}:Y\subset X} 1 \leq \sum_{Y\in\binom{[M]}{m}} (m-1) = \binom{M}{m}(m-1)\,.$$

Now, $N\binom{N}{m}$ is a polynomial in N of degree $m+1$, whereas $\binom{M}{m}(m-1) = \binom{k(N-1)+1}{m}(m-1)$ is a polynomial in N of degree m, so for sufficiently large N, we have a contradiction. □

The Bipartite Ramsey Theorem (Beineke & Schwenk 1976)
If $k, m \in \mathbb{N}$ and n is sufficiently large, then each k-colouring of the edges of $K_{n,n}$ gives a complete monochromatic subgraph $K_{m,m}$.

Proof. Let N be sufficiently large as the theorem above, and set $n := k(N-1)+1$. Colour the edges of $K_{n,n}$ with k colours and let A, B be the vertex parts. Each vertex $a \in A$ is adjacent to $n = k(N-1)+1$ edges. By the Pigeonhole Principle, a is incident to N edges with the same colour $c(a)$.

There are n vertices $a \in A$, each incident to N edges with colour $c(a)$. Again by the Pigeonhole Principle, N of these n colours $c(a)$ are the same, say c. So, some N vertices in A are each adjacent to N vertices in B by edges of colour c. By definition of N, some m vertices in A are each adjacent to a common set of m vertices in B via edges of colour c. This gives us a c-coloured subgraph $K_{m,m}$. □

Theorem (Zarankiewicz 1952)
Let $m \in \mathbb{N}$ and $\epsilon > 0$. If n is sufficiently large and G is a subgraph of $K_{n,n}$ with at least ϵn^2 edges, then G has $K_{m,m}$ as a subgraph.

Definition. For graphs G, H,

- $\chi(G)$ is the least number of colours in a proper vertex colouring of G. This is the *chromatic number* of G.
- $c(H)$ is the largest size of a connected component of H.
- $r(G,H)$ is the smallest n so that if $E(K_n)$ is coloured red and blue, then K_n either has a red G or a blue H as a subgraph.

Example. $\chi(K_n) = c(K_n) = n$ whereas $\chi(P_n) = 2$ for any path P_n on n vertices, and $c(mP_n) = n$ where mP_n denotes a disjoint set of m such paths. We have previously seen that $r(K_3, K_3) = 6$.

Theorem (Chvátal & Harary 1972)

$$r(G,H) \geq (c(G)-1)(\chi(H)-1)+1$$

Proof. Set $n := (c(G)-1)(\chi(H)-1)$ and consider K_n. We can find $\chi(H)-1$ disjoint $K_{c(G)-1}$ subgraphs of K_n. Colour the edges of these red and all other edges blue. Then K_n has no red G subgraph nor any blue H subgraph. Hence, $r(G,H) \geq n+1$. □

Theorem (Chvátal 1977)
If T_m is a tree on m vertices, then

$$r(T_m, K_m) = (m-1)(n-1)+1\,.$$

Example. For $m=2$, it is clear that $r(T_m, K_n) = n$, as claimed. Similarly for $n=2$, $r(T_m, K_n) = m$.

Proof. By the above theorem,

$$r(T_m, K_n) \geq (c(T_m)-1)(\chi(K_n)-1)+1 = (m-1)(n-1)+1\,.$$

We therefore wish to prove that $r(T_m, K_n) \leq (m-1)(n-1)+1$. Assume that the theorem is true for all m', n' with $m'+n' < m+n$.

Now, arbitrarily colour the edges of K_N red and blue, where $N := (m-1)(n-1)+1$. Let T_{m-1} be a tree obtained by removing an end-vertex from T_m and let v be the vertex that was adjacent to that removed vertex. By the induction assumption, K_N contains either a blue K_n subgraph, and we are done, or it has a red T_{m-1} subgraph. Suppose the latter.

Let $K_{N'}$ be the graph obtained by deleting the vertices of T_{m-1} from K_N. Then $K_{N'}$ has $N-(m-1) = (m-1)(n-2)+1$ vertices. By the induction assumption, $K_{N'}$ has either a red T_m subgraph, in which case we are done, or it has a blue K_{n-1} subgraph. Suppose the latter.

We have supposed that red T_{m-1} and blue K_{n-1} are subgraphs of K_N. Now consider the edges from the vertex v in T_{m-1} to the vertices of K_{n-1}. If one of these edges is red, then add it to v to get a red T_m subgraph. Otherwise, all edges are blue, so add them to K_{n-1} to get a blue K_n. Thus, K_N has either a red T_m subgraph or a blue K_n subgraph.

Hence, $r(T_m, K_n) \leq (m-1)(n-1)+1$, and we are done. □

Surprisingly often, we find that combinatorial results on subsets have analogues on vector subspaces. This is also true of Ramsey's Theorem which has the following vector space analogues.

The Affine Ramsey Theorem (Spencer 1975)
Let $\mathbb{F}$ be a finite field. If $k, r, t \in \mathbb{N}$ and n is sufficiently large, then each r-colouring of the t-dimensional subspaces of $\mathbb{F}^n$ gives a k-dimensional affine subspace of $\mathbb{F}^n$ whose t-dimensional subspaces have the same colour.

Vector Space Ramsey Theorem (Graham, Leeb & Rothschild 1972)
Let $\mathbb{F}$ be a finite field. If $k, r, t \in \mathbb{N}$ and n is sufficiently large, then each r-colouring of the t-dimensional subspaces of $\mathbb{F}^n$ gives a k-dimensional vector subspace of $\mathbb{F}^n$ whose t-dimensional subspaces have the same colour.

The Gallai-Witt Theorem
For each finite subset $V \subseteq \mathbb{R}^n$ and each finite colouring of $\mathbb{R}^n$, there is a monochromatic subset $W \subseteq \mathbb{R}^n$ that can be obtained from V by translating and scaling.

When scaling is not allowed, then it is easy to find finite sets $V \subseteq \mathbb{R}^n$ and colourings of $\mathbb{R}^n$ so that no translation of V is monochromatic.

Definition. A finite set of points V is *Ramsey* if for each $r \in \mathbb{N}$ there is a sufficiently large n such that any r-colouring $\mathbb{R}^n$ has a monochromatic subsets W that is congruent to V, i.e., translating, rotating, and reflecting is allowed.

Theorem
The following hold:

1. All rectangular parallelepipeds ("bricks") are Ramsey.
2. In particular, all equilateral simplexes are Ramsey.
3. Indeed, all regular polygons and polyhedra are Ramsey.
4. Each Ramsey set can be placed on some sphere.

6 Applications

The Happy Ending Problem (Erdős & Szekeres 1935)
If $k \in \mathbb{N}$ and n is sufficiently large, then among any n points in the plane, no three collinear, there are k that form a convex polygon.

Proof. Set $n := R_3(k,k)$. Then, given any n points, label them $1, \ldots, n$ in any order. Colour every triple $\{i,j,\ell\}$ green if $i<j<\ell$ is a clockwise sequence, and red if they form an anti-clockwise sequence. Then there are k points whose triples are all green or all red. Without loss of generality, we may assume that this colour is green. In other words, each triangle among these k points is oriented clockwise. Then these k points form a convex polygon, and we are done. □

Theorem (Erdős 1964)
Let $A \subseteq \mathbb{N}$ so that for each $n \in \mathbb{N}$, there are $a, b \in A$ with $n = ab$. Then for each $k \in \mathbb{N}$, there is some integer $n \in \mathbb{N}$ so that the equation $n = ab$ has at least k solutions with $a, b \in A$.

Proof. Consider square-free integers n only. For such n, let $M(n)$ be the set of prime factors of N. We can find a partition $M(n) = M' \cup M''$ so that $a := \prod M' := \prod_{p \in M'} p$ and $b := \prod M''$ both belong to A.

By Ramsey's theorem, there is a finite set $X \subseteq \mathbb{N}$ with $|X| \geq 2k^2$ so that for each $M \in \binom{X}{k}$, the partition $M = M' \cup M''$ is of the same 'type'. Hence, there exists $Y \subseteq X$, $|Y| \geq 2k$, such that $\binom{Y}{m} \subseteq A$ for an $m \geq \frac{1}{2}$. Thus, for every $M \in \binom{M}{2m}$, $n := \prod M$ has $\binom{2m}{m}$ solutions $n = ab$ in A. □

Challenge. Rewrite this proof so that the cursive part is correct and comprehensible!

Euclid's Lemma or **Bezout's Identity**
For any coprime $a, b \in \mathbb{Z}$, there are $x, y \in \mathbb{Z}$ such that $ax + by = 1$.

Proof. Consider the $b-1$ remainders of $a, 2a, \ldots, (b-1)n$ modulo b. Since a and b are coprime, none of the remainders equal 0. Now assume that no remainder equals 1. Then the $b-1$ remainders must be among the numbers $2, \ldots, b-1$. By the Pigeonhole Principle, two of the remainders must be equal, so $ma \equiv na \pmod{b}$ for distinct $m, n > 0$. Hence, $(m-n)a \equiv 0 \pmod{b}$. Since a and b are coprime, $m - n \equiv 0 \pmod{b}$, or $m \equiv n \pmod{b}$, a contradiction. Hence, $xa \equiv 1 \pmod{b}$ for some $x \in \mathbb{Z}$ and so $ax + by = 1$ for some $y \in \mathbb{Z}$. □

Modified slightly, this proof gives the Chinese Remainder Theorem.

Proizvolov's Identity
Bipartition $[2n]$ into sets $A = \{a_1 > \cdots > a_n\}$ and $B = \{b_1 < \cdots < b_n\}$. Then

$$\sum_{i=1}^{n} |a_i - b_i| = n^2 .$$

Proof. Assume that $a_i, b_i \leq n$ for some i. Then at least $n-(i-1)$ of the a_j's are in $[n]$. Also, at least i of the b_j's are in $[n]$. Then at least $n-(i-1)+i = n+1$ of the a_j's and b_j's are in $[n]$.

By the Pigeonhole Principle, two of these are identical, which gives a contradiction. Hence, a_i and b_i cannot both be less than or equal to $[n]$. Similarly, a_i and b_i cannot both be bigger than equal to $[n]$. Hence, one of a_i and b_i is in $[n]$ and the other is in $[2n]-[n] = \{n+1, \ldots, 2n\}$. Then,

$$\begin{aligned}\sum_{i=1}^{n} |a_i - b_i| &= ((n+1) + \cdots + 2n) - (1 + \cdots + n) \\ &= (\underbrace{n + \cdots + n}_{n \text{ times}}) + (1 + \cdots + n) - (1 + \cdots + n) \\ &= n^2\end{aligned}$$

□

The following theorem is a special case of Turán's Theorem stating when G must contain K_p.

Mantel's Theorem (1907)
If $G = (V, E)$ is a simple graph on n vertices with $|E| > \frac{n^2}{4}$ edges, then G has at least one triangle.

Proof. If $n = 3$, then if $|E| > \frac{n^2}{4} = 2.25$, so G itself is a triangle. Assume for induction that the theorem is true for all values up to $n-1$. Choose an edge $\{u, v\}$ of G, and let H be the subgraph of G obtained by deleting u and v.

If $|E(H)| > \frac{(n-2)^2}{4}$ edges, then H and thus G has a triangle, by assumption.
Suppose instead that $|E(H)| \leq \frac{(n-2)^2}{4} = \frac{n^2}{4} - n + 1$. There are

$$|E - \{u, v\}| - |E(H)| > \frac{n^2}{4} - \frac{(n-2)^2}{4} - 1 = n - 2$$

edges between H and $\{u, v\}$; that is, there are at least $n-1$ such edges, and only $n-2$ vertices in H. By the Pigeonhole Principle, some vertex in H is adjacent to both u and v. Thus, G has a triangle. □

References

[1] M. Aigner and G.M. Ziegler, *Proofs from The Book*, 4th edition, Springer-Verlag, Berlin, 2010.

[2] A. Bogomolny, *Pigeonhole Principle*, `http://www.cut-the-knot.org/do_you_know/pigeon.shtml`, 2016–03.

[3] J. Brandt, *Kombinatorik*, Aarhus University, lecture notes, 2001.

[4] R.L. Graham, M. Grötschel, and L. Lovász (eds.), *Handbook of Combinatorics. I–II*, North-Holland, Amsterdam, 1995.

[5] R.L. Graham, B.L. Rothschild, and J. Spencer, *Ramsey Theory*, 2nd edition, John Wiley & Sons, Inc., New York, 1990.

[6] B.M. Landman and A. Robertson, *Ramsey Theory on the Integers*, 2nd edition, AMS, Providence, RI, 2014.

[7] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992.