VI

Ramsey Theory

In a party of six people there is always a group of three who either all know each other or are all strangers to each other. If the edges of the complete graph on an infinite set N are coloured red or blue then for some infinite set $M \subset N$ all the edges joining vertices of M get the same colour. Both of these assertions are special cases of a theorem published by Ramsey in 1930. The original theorems of Ramsey have been extended in many directions, resulting in what has come to be called *Ramsey theory*: a rich theory expressing the deep mathematical principle, vastly extending the pigeon-hole principle, that no matter how we partition the objects of a 'large' structure into a 'few' classes, one of these classes contains a 'large' subsystem. While Dirichlet's pigeon-hole principle guarantees that we have 'many' objects in the same class, without any condition on their relationship to each other, in Ramsey theory we look for a large substructure in the same class: we do not only want infinitely many red edges, say, but we want all the edges joining vertices of an infinite set to be red. Or, in the first example, we do not only want three pairs of acquaintances, but we want these three acquaintances to 'form a triangle', to be the three pairs of acquaintances belonging to three people.

The quintessential result of Ramsey theory dealing with richer mathematical structures than graphs is van der Waerden's theorem, predating the theorems of Ramsey, which states that given k and p, if W is a large enough integer and we partition the set of the first W natural numbers into k classes, then one of the classes contains an arithmetic progression with p terms.

Ramsey theory is a large and beautiful area of combinatorics, in which a great variety of techniques are used from many branches of mathematics, and whose results are important not only in graph theory and combinatorics, but in set theory, logic, analysis, algebra, and geometry as well. In order to demonstrate this, we shall go well beyond graph theory to present several striking and deep results, including the Erdős–Rado canonical theorem, extending Ramsey's original theorem to infinitely many colours; Shelah's theorem, extending the Hales-Jewett theorem (which itself extends van der Waerden's theorem); and the theorems of Galvin, Prikry, and Hindman about Ramsey properties of infinite sequences. Nevertheless, we shall hardly do more than scratch the surface of modern Ramsey theory.

VI.1 The Fundamental Ramsey Theorems

We shall consider partitions of the *edges* of graphs and hypergraphs. For the sake of convenience a partition will be called a *colouring*, but one should bear in mind that a colouring in this sense has nothing to do with the edge colourings considered in Chapter V. Adjacent edges may have the same colour and, indeed, our aim is to show that there are large subgraphs all of whose edges have the same colour. In a 2-colouring we shall often choose red and blue as colours; a subgraph is *red* (*blue*) if all its edges are red (blue).

As we shall see, given a natural number s, there is an integer $R(s)$ such that if $n \geq R(s)$ then every colouring of the edges of K_n with red and blue contains either a red K_s or a blue K_s. The assertion about a party of six people claims precisely that $R(3) = 6$ will do. In order to show the existence of $R(s)$ in general, for any s and t, we define the *Ramsey number* $R(s, t)$ as the smallest value of n for which every red–blue colouring of K_n yields a red K_s or a blue K_t. In particular, $R(s, t) = \infty$ if there is no such n such that in every red–blue colouring of K_n there is a red K_s or a blue K_t. It is obvious that

$$R(s,\ t) = R(t,\ s)$$

for every $s, t \geq 2$ and

$$R(s,\ 2) = R(2,\ s) = s,$$

since in a red–blue colouring of K_2 either there is a blue edge or else every edge is red. The following result, due to Erdős and Szekeres, states that $R(s, t)$ is finite for every s and t, and at the same time it gives a bound on $R(s, t)$. Although qualitatively it is a special case of Ramsey's original theorem, the bound it gives is considerably better than that given by Ramsey.

Theorem 1 *The function $R(s,t)$ is finite for all $s, t \geq 2$. If $s > 2$ and $t > 2$ then*

$$R(s,\ t) \leq R(s-1,\ t) + R(s,\ t-1) \tag{1}$$

and

$$R(s,\ t) \leq \binom{s+t-2}{s-1}. \tag{2}$$

Proof. As we shall prove (1) and (2), it will follow that $R(s,t)$ is finite.

(i) When proving (1) we may assume that $R(s-1,\ t)$ and $R(s,\ t-1)$ are finite. Let $n = R(s-1,\ t) + R(s,\ t-1)$ and consider a colouring of the edges of K_n

with red and blue. We have to show that in this colouring there is either a red K_s or a blue K_t. To this end, let x be a vertex of K_n. Since $d(x) = n - 1 = R(s-1, t) + R(s, t-1) - 1$, either there are at least $n_1 = R(s-1, t)$ red edges incident with x or there are at least $n_2 = R(s, t-1)$ blue edges incident with x. By symmetry we may assume that the first case holds. Consider a subgraph K_{n_1} of K_n spanned by n_1 vertices joined to x by red edges. If K_{n_1} has a blue K_t, we are done. Otherwise, by the definition of $R(s-1, t)$, the graph K_{n_1} contains a red K_{s-1} which forms a red K_s with x.

(ii) Inequality (2) holds if $s = 2$ or $t = 2$ (in fact, we have equality since $R(s, 2) = R(2, s) = s$). Assume now that $s > 2$, $t > 2$ and (2) holds for every pair (s', t') with $2 \leq s' + t' < s + t$. Then by (1) we have

$$R(s, t) \leq R(s-1, t) + R(s, t-1) \leq \binom{s+t-3}{s-2} + \binom{s+t-3}{s-1} = \binom{s+t-2}{s-1}. \qquad \square$$

It is customary to distinguish ***diagonal Ramsey numbers*** $R(s) = R(s, s)$ and ***off-diagonal Ramsey numbers*** $R(s, t)$, $s \neq t$. It is not surprising that the diagonal Ramsey numbers are of greatest interest, and they are also the hardest to estimate. Re calling that a graph is *trivial* if it is either complete or empty, the diagonal Ramsey number $R(s)$ is the minimal integer n such that every graph of order n has a trivial subgraph of order s.

We see from Theorem 1 that

$$R(s) \leq \binom{2s-2}{s-1} \leq \frac{2^{2s-2}}{\sqrt{s}}. \tag{3}$$

Although the proof above is very simple, the bound (3) was hardly improved for over 50 years. The best improvement is due to Thomason, who in 1988 proved that

$$R(s) \leq \frac{2^{2s}}{s} \tag{4}$$

if s is large. Although the improvement over (3) is small, this is a hard result, and we shall not prove it. In Chapter VII we shall show that $R(s)$ does grow exponentially: $R(s) \geq 2^{s/2}$. It is widely believed that there is a constant c, perhaps even $c = 1$, such that

$$R(s) = 2^{(c+o(1))s},$$

but this is very far from being proved.

The result easily extends to colourings with any finite number of colours: given k and $s_1, s_2, \ldots, s_k$, if n is sufficiently large, then every colouring of K_n with k colours is such that for some i, $1 \leq i \leq k$, there is a K_{s_i} coloured with the i th colour. (The minimal value of n for which this holds is usually denoted by $R_k(s_1, \ldots, s_k)$.) Indeed, if we know this for $k-1$ colours, then in a k-colouring of K_n we replace the first two colours by a new colour. If n is sufficiently large (depending on $s_1, s_2, \ldots, s_k$) then either there is a K_{s_i} coloured with the i th colour for some i, $3 \le i \le k$, or else for $m = R(s_1, s_2)$ there is a K_m coloured with the new colour. In other words, in the original colouring this K_m is coloured with the first two (original) colours. In the first case we are home, and in the second, for $i = 1$ or 2 we can find a K_{s_i} in K_m coloured with the i th colour. This shows that

$$R_k(s_1, \dots, s_k) \le R_{k-1}(R(s_1, s_2), s_3, \dots, s_k).$$

In fact, Theorem 1 also extends to hypergraphs, that is, to colourings of the set $X^{(r)}$ *of all r-tuples* of a finite set X with k colours. This is one of the theorems proved by Ramsey. We now turn our attention to this.

Denote by $R^{(r)}(s, t)$ the minimal value of n for which every red–blue colouring of $X^{(r)}$ yields a red s-set or a blue t-set, provided that $|X| = n$. Of course, a set $Y \subset X$ is called *red* (*blue*) if every element of $Y^{(r)}$ is red (blue). Note that $R(s, t) = R^{(2)}(s, t)$. As in the case of Theorem 1, the next result not only guarantees that $R^{(r)}(s, t)$ is finite for all values of the parameters (which is certainly not at all obvious at first), but also gives an upper bound on $R^{(r)}(s, t)$. The proof is an almost exact replica of the proof of Theorem 1. Note that if $r > \min\{s, t\}$ then $R^{(r)}(s, t) = \min\{s, t\}$, and if $r = s \le t$ then $R^{(r)}(s, t) = t$.

Theorem 2 *Let* $1 < r < \min\{s, t\}$. *Then* $R^{(r)}(s, t)$ *is finite and*

$$R^{(r)}(s, t) \le R^{(r-1)}\left(R^{(r)}(s-1, t), R^{(r)}(s, t-1)\right) + 1.$$

Proof. Both assertions follow immediately if we prove the inequality under the assumption that $R^{(r-1)}(u, v)$ is finite for all u, v, and both $R^{(r)}(s-1, t)$ and $R^{(r)}(s, t-1)$ are also finite.

Let X be a set with $R^{(r-1)}(R^{(r)}(s-1, t), R^{(r)}(s, t-1)) + 1$ elements. Given any red–blue colouring c of $X^{(r)}$, pick an $x \in X$ and define a red–blue colouring c' of the $(r-1)$-sets of $Y = X - \{x\}$ by colouring $\sigma \in Y^{(r-1)}$ the colour of $\sigma \cup \{x\} \in X^{(r)}$. By the definition of the function $R^{(r-1)}(u, v)$ we may assume that Y has a red subset Z (for c') with $R^{(r)}(s-1, t)$ elements.

Now let us look at the restriction of c to $Z^{(r)}$. If it has a blue t-set, we are done, since $Z^{(r)} \subset X^{(r)}$, so a blue t-set of Z is certainly also a blue t-set of X. On the other hand, if there is no blue t-set of Z then there is a red $(s-1)$-set. The union of this red $(s-1)$-set with $\{x\}$ is then a red s-set of X, because $\{x\} \cup \sigma$ is red for every $\sigma \in Z^{(r-1)}$. □

It is easily seen that Theorem 2 and the colour-grouping argument described after Theorem 1 imply the following assertion. Given r and $s_1, s_2, \dots, s_k$, then for large enough $|X|$ every colouring of $X^{(r)}$ with k colours is such that for some i, $1 \le i \le k$, there is a set $S_i \subset X$, $|S_i| = s_i$, all of whose r-sets have colour i. The smallest value of $|X|$ for which this is true is denoted by $R_k^{(r)}(s_1, s_2, \dots, s_k)$; thus $R^{(r)}(s, t) = R_2^{(r)}(s, t)$ and $R_k(s_1, s_2, \dots, s_k) = R_k^{(2)}(s_1, s_2, \dots, s_k)$. The upper bound for $R_k^{(r)}(s_1, s_2, \dots, s_k)$ implied (via colour-grouping) by Theorem 2 is not very good. Imitating the proof of Theorem 1 one arrives at a better upper bound (cf. Exercise 8):

$$R_k^{(r)}(s_1, s_2, \ldots, s_k) \le R_k^{(r-1)}(R_k^{(r)}(s_1 - 1, s_2, \ldots, s_k), \ldots,$$
$$R_k^{(r)}(s_1, \ldots, s_{k-1}, s_k - 1)) + 1.$$

Very few of the nontrivial Ramsey numbers are known, even in the case $r = 2$. It is easily seen that $R(3, 3) = 6$, and with some work one can show that $R(3, 4) = 9$, $R(3, 5) = 14$, $R(3, 6) = 18$, $R(3, 7) = 23$ and $R(4, 4) = 18$. Considerably more effort is needed to prove that $R(3, 8) = 28$ and $R(3, 9) = 36$. Furthermore, McKay and Radziszowski proved in 1995 that $R(4, 5) = 25$. These are the only known two-colour Ramsey numbers. For the other ones, all that is known are bounds, as shown in Table VI. 1. The proofs of many of these bounds needed a surprising amount of ingenuity, work and computing time.

At first sight, the paucity of exact Ramsey numbers may well seem surprising. However, there are many reasons why it is unlikely that a large Ramsey number, like $R(6, 6)$, will ever be determined. The two-colourings of K_n without large monochromatic complete subgraphs lack order: they look as if they had been chosen at random. This apparent disorder makes it highly unlikely that a simple induction argument will give a tight *upper bound* for $R(s, t)$. On the other hand, a head-on attack by computers is also doomed to failure, even for $R(5, 5)$. For example, if all we want to prove is that 48 is an upper bound for $R(5, 5)$, we have to examine over 2^{1000} graphs of order 48: a task well beyond the power of computers.

It is not too easy to prove general *lower bounds* for Ramsey numbers either. As the colourings without large complete monochromatic subgraphs are 'disorderly', it is not surprising that random methods can be used to give fairly good lower bounds. In Chapter VII we shall show some beautiful examples of this.

As it is very difficult to find good estimates for $R(s, t)$ as $s, t \to \infty$, it is not surprising that very few fast-growing Ramsey functions have been determined exactly. In fact, Erdős and Szekeres proved that the right-hand side of (2) is exactly 1 smaller than the value of a natural Ramsey function. In order to present this result, we introduce some terminology. Call a set $S \subset \mathbf{R}^2$ *non-degenerate* if any two points of it have different x coordinates. A *k-cup*, or a *convex k-set*, is a non-degenerate set of k points of the form $\{(x_i, h(x_i)) : i = 1, \ldots, k\}$, where h is a convex function. Writing $s(p, p') = (y - y')/(x - x')$ for the *slope* of the line through the points $p = (x, y)$ and $p' = (x', y')$, if $K = \{p_1, \ldots, p_k\}$ with $p_i = (x_i, y_i)$, $x_1 < \cdots < x_k$, then K is a k-cup iff $s(p_1, p_2) \le s(p_2, p_3) \le \cdots \le s(p_{k-1}, p_k)$. An *ℓ-cap*, or a *concave ℓ-set*, is defined analogously.

Here is then the beautiful result of Erdős and Szekeres about k-cups and ℓ-caps. The first part was published in 1935, the second in 1960.

Theorem 3 *For $k, \ell \ge 2$, every non-degenerate set of $\binom{k+\ell-4}{k-2} + 1$ points contains a k-cup or an ℓ-cap. Also, for all $k, \ell \ge 2$, there is a non-degenerate set $S_{k,\ell}$ of $\binom{k+\ell-4}{k-2}$ points that contains neither a k-cup nor an ℓ-cap.*

Proof. Let us write $\phi(k, \ell)$ for the binomial coefficient $\binom{k+\ell-4}{k-2}$.

(i) We shall prove by induction on $k + \ell$ that every non-degenerate set of $\phi(k, \ell) + 1$ points contains a k-cup or an ℓ-cap. Since a non-degenerate set of 2 points is both a 2-cup and a 2-cap, this is clear if $\min\{k, \ell\} = 2$, since $\phi(k, 2) = \phi(2, \ell) = 1$ for all $k, \ell \geq 2$. Suppose then that $k, \ell \geq 3$ and the assertion holds for smaller values of $k+\ell$. Let S be a non-degenerate set of $\phi(k, \ell)+1$ points and suppose that, contrary to the assertion, S contains neither a k-cup nor an ℓ-cap. Let $L \subset S$ be the set of last points of $(k-1)$-cups. Then $S \backslash L$ has neither a $(k-1)$-cup nor an ℓ-cap so, by the induction hypothesis, $|S\backslash L| \leq \phi(k-1, \ell)$. Therefore $|L| \geq \phi(k, \ell) + 1 - \phi(k-1, \ell) = \phi(k, \ell-1) + 1$ so, again by the induction hypothesis, L contains an $(\ell-1)$-cap, say $\{q_1, \ldots, q_{\ell-1}\}$, with first point our set S contains q_1. Since $q_1 \in L$, a $(k-1)$-cup $\{p_1, \ldots, p_{k-1}\}$, whose last point, p_{k-1}, is precisely q_1. Now, if $s(p_{k-2}, p_{k-1}) \leq s(p_{k-1}, q_2)$ then $\{p_1, \ldots, p_{k-1}, q_2\}$ is a k-cup. Otherwise, $s(p_{k-2}, q_1) > s(q_1, q_2)$, so $\{p_{k-2}, q_1, \ldots, q_{\ell-1}|$ is an ℓ-cup. This contradiction completes the proof of the induction step, and we are done.

(ii) We shall construct $S_{k,\ell}$ also be induction on $k+\ell$. In fact, we shall construct $S_{k,\ell}$ in the form $\{(i, y_i) : 1 \leq i \leq \phi(k, \ell)\}$.

If $\min\{k, \ell\} = 2$ then $\phi(k, \ell) = 1$ and we may take $S_{k,\ell} = \{(1, 0)\}$. Suppose then that $k, \ell \geq 3$ and we have constructed $S_{k,\ell}$ for smaller values of $k + \ell$. Set $Y = S_{k-1,\ell}$, $Z = S_{k,\ell-1}$, $m = \phi(k-1, \ell)$ and $n = \phi(k, \ell-1)$, so that $Y = \{(i, y_i) : 1 \leq i \leq m\}$ contains neither a $(k-1)$-cup nor an ℓ-cap, and $Z = \{(i, z_i) : 1 \leq i \leq n\}$ contains neither a k-cup nor an $(\ell-1)$-cup.

For $\varepsilon > 0$, set $Y^{(\varepsilon)} = \{(i, \varepsilon y_i) : 1 \leq i \leq m\}$ and $Z^{(\varepsilon)} = \{(m+i, m+\varepsilon z_i)$: $1 \leq i \leq n\}$. Now, if $\varepsilon > 0$ is small enough then every line through two points of $Y^{(\varepsilon)}$ goes below the entire set $Z^{(\varepsilon)}$, and every line through two points of $Z^{(\varepsilon)}$ goes above the entire set $Y^{(\varepsilon)}$. Hence, in this case, every cup meeting $Z^{(\varepsilon)}$ in at least two points is entirely in $Z^{(\varepsilon)}$, and every cup meeting $Y^{(\varepsilon)}$ in at least two points is entirely in $Y^{(\varepsilon)}$. But then $Y^{(\varepsilon)} \cup Z^{(\varepsilon)}$ will do for $S_{k,\ell}$ since it continues neither a k-cup nor an ℓ-cup. □

As an easy consequence of Theorem 3, we see that every set of $\binom{2k-4}{k-2} + 1$ points in the plane in general position contains the vertices of some convex k-gon. In 1935, Erdős and Szekeres conjectured that, in fact, every set of $2^{k-2} + 1$ points in general position contains a convex k-gon. It does not seem likely that the conjecture will be proved in the near future, but it is known that, if true, the conjecture is best possible (see Exercise 23).

After this brief diversion, let us return to hypergraphs. Theorem 2 implies that every red–blue colouring of the r-tuples of the natural numbers contains arbitrarily large monochromatic subsets; a subset is *monochromatic* if its r-tuples have the same colour. Ramsey proved that, in fact, we can find an *infinite monochromatic* set.

Theorem 4 *Let* $1 \leq r < \infty$ *and let* $c : A^{(r)} \to [k] = \{1, 2, \ldots, k\}$ *be a* k*-colouring of the* r*-tuples of an infinite set* A*. Then* A *contains a monochromatic infinite set.*

l k	3	4	5	6	7	8	9	10	11	12	13	14	15
3	6	9	14	18	23	28	36	40 43	46 51	51 60	59 69	66 78	73 89
4		18	25	35 41	49 61	53 84	69 115	80 149	96 191	106 238	118 291	129 349	134 417
5			43 49	58 87	80 143	95 216	114 316	442					
6				102 165	298	495	780	1171					
7					205 540	1031	1713	2826					
8						1870	282 3583	6090					
9							6625	565 12715					
10								23854	798				

TABLE VI.1. Some values and bounds for two colour Ramsey numbers.

Proof. We apply induction on r. Note that the result is trivial for $r = 1$, so we may assume that $r > 1$ and the theorem holds for smaller values of r.

Put $A_0 = A$ and pick an element $x_1 \in A_0$. As in the proof of Theorem 2, define a a colouring $c_1 : B_1^{(r-1)} \to [k]$ of the $(r-1)$-tuples of $B_1 = A_0 - \{x_1\}$ by putting $c_1(\tau) = c(\tau \cup \{x_1\})$, $\tau \in B_1^{(r-1)}$. By the induction hypothesis B_1 contains an infinite set A_1 all of whose $(r-1)$-tuples have the same colour, say d_1, where $d_1 \in \{1, \dots, k\}$. Let now $x_2 \in A_1$, $B_2 = A_1 - \{x_2\}$ and define a k-colouring $c_2 : B_2^{(r-1)} \to [k]$ by putting $c_2(\tau) = c(\tau \cup \{x_2\})$, $\tau \in B_2^{(r-1)}$. Then B_2 has an infinite set A_2 all of whose $(r-1)$-tuples have the same colour, say d_2. Continuing in this way we obtain an infinite sequence of elements: $x_1, x_2, \dots$, an infinite sequence of colours: $d_1, d_2, \dots$, and an infinite nested sequence of sets: $A_0 \supset A_1 \supset A_2 \supset \cdots$, such that $x_i \in A_{i-1}$, and for $i = 0, 1, \dots$, all r-tuples whose only element outside A_i is x_i have the same colour d_i. The infinite sequence $(d_n)_1^\infty$ must take at least one of the k values $1, 2, \dots, k$ infinitely often, say $d = d_{n-1} = d_{n_2} = \dots$. Then, by the construction, each r-tuple of the infinite set $\{x_{n_1}, x_{n_2}, \dots\}$ has colour d. □

In some cases it is more convenient to apply the following version of Theorem 4. As usual, the set of natural numbers is denoted by $\mathbb{N}$.

Theorem 5 *For each $r \in \mathbb{N}$, colour the set $\mathbb{N}^{(r)}$ of r-tuples of $\mathbb{N}$ with k_r colours, where $k_r \in \mathbb{N}$. Then there is an infinite set $M \subset \mathbb{N}$ such that for every r any two r-tuples of M have the same colour, provided their minimal elements are not less than the r^{th} element of M.*

Proof. Put $M_0 = \mathbb{N}$. Having chosen infinite sets $M_0 \supset \cdots \supset M_{r-1}$, let M_r be an infinite subset of M_{r-1} such that all the r-tuples of M_r have the same colour. This way we obtain an infinite nested sequence of infinite sets: $M_0 \supset M_1 \supset \cdots$. Pick $a_1 \in M_1$, $a_2 \in M_2 - \{1, \dots, a_1\}$, $a_3 \in M_3 - \{1, \dots, a_2\}$, etc. Clearly, $M = \{a_1, a_2, \dots\}$ has the required properties. □

It is interesting to note that Ramsey's theorem for infinite sets, Theorem 3, easily implies the corresponding result for finite sets, although it fails to give bounds on the numbers $R^{(r)}(s_1, s_2, \ldots, s_k)$. To see this, all one needs is a simple compactness argument, a special case of Tychonov's theorem that the product of compact spaces is compact.

We have already formulated this (see Exercise III.30) but here we spell it out again in a convenient form.

Theorem 6 *Let r and k be natural numbers, and for every $n \geq 1$, let C_n be a non-empty set of k-colourings of $[n]^{(r)}$ such that if $n < m$ and $c_m \in C_m$ then the restriction $c_m^{(n)}$ of c_m to $[n]^{(r)}$ belongs to C_n. Then there is a colouring $c : \mathbb{N}^{(r)} \to [k]$ such that, for every n, the restriction $c^{(n)}$ of c to $[n]^{(r)}$ belongs to C_n.*

Proof. For $m > n$, write $C_{n,m}$ for the set of colourings $[n]^{(r)} \to [k]$ that are restrictions of colourings in C_m. Then $C_{n,m+1} \subset C_{n,m} \subset C_n$ and so $\tilde{C}_n = \bigcap_{m=n+1}^{\infty} C_{n,m} \neq \emptyset$ for every n, since each $C_{n,m}$ is finite. Let $c_r \in \tilde{C}_r$, and pick $c_{r+1} \in \tilde{C}_{r+1}$, $c_{r+2} \in \tilde{C}_{r+2}$, and so on, such that each is in the preimage of the previous one: $c_n = c_{n+1}^{(n)}$. Finally, define $c : \mathbb{N}^{(r)} \to [k]$ by setting, for $\rho \in \mathbb{N}^{(r)}$,

$$c(\rho) = c_n(\rho) = c_{n+1}(\rho) = \cdots,$$

where $n = \max \rho$. This colouring c is as required. □

Let us see then that Theorem 5 implies that $R^{(r)}(s_1, s_2, \ldots, s_k)$ exists. Indeed, otherwise for every n there is a colouring $[n]^{(r)} \to [k]$ such that, for each i, there is no s_i-set all of whose r-sets have colour i. Writing C_n for the set of all such colourings, we see that $C_n \neq \emptyset$ and $C_{n,m} \subset C_n$ for all $n < m$, where $C_{n,m}$ is as in the proof of Theorem 5. But then there is a colouring $c : \mathbb{N}^{(r)} \to [k]$ such that every monochromatic set has fewer than $s = \max s_i$ elements, contradicting Theorem 4.

To conclude this section, we point out a fascinating phenomenon. First, let us see an extension of the fact that $R_k^{(r)}(s_1, \ldots, s_k)$ exists.

Theorem 7 *Let r, k and $s \geq 2$. If n is sufficiently large then for every k-colouring of $[n]^{(r)}$ there is a monochromatic set $S \subset [n]$ such that*

$$|S| \geq \max\{s, \min S\}.$$

Proof. Suppose that there is no such n, that is, for every n there is a colouring $[n]^{(r)} \to [k]$ without an appropriate monochromatic set. Let C_n be the set of all such colourings. Then $C_n \neq \emptyset$ and, in the earlier notation, $C_{n,m} \subset C_n$ for all $n < m$. But then there is a colouring $c : \mathbb{N}^{(r)} \to [k]$ such that its restriction $c^{(n)}$ to $[n]^{(r)}$ belongs to C_n. Now, by Theorem 4, there is an infinite monochromatic set $M \subset \mathbb{N}$. Set $m = \min M$, $t = \max\{m, s\}$, and let S consist of the first t elements of M. Then, with $n = \max S$, the colouring $c^{(n)}$ does have an appropriate monochromatic set, namely S, contradicting $c^{(n)} \in C_n$. □

This is a beautiful result but it is not too unexpected. What *is* surprising and deep is that, as proved by Paris and Harrington in 1977, although Theorem 8 is a (fairly simple) assertion concerning *finite* sets, *it cannot* be deduced from the Peano axioms, that is, it cannot be proved within the theory of finite sets. In other words, we actually *need* the notion of a finite set to prove Theorem 7. This theorem of Paris and Harrington became the starting point of an active area connecting combinatorics and logic.

As this is a book on graph theory, we cannot digress too far into logic, so let us return to graphs. Let $R^*(s)$ be the minimal integer n such that for every two-colouring of $[n]^{(2)}$ there is a monochromatic set $S \subset [n]$ with $|S| \geq \max\{s, |S|\}$. Thus $R^*(s)$ is the minimal value of n such that for every graph G with vertex set $[n]$ there is a set $S \subset [n]$ with $|S| \geq \max\{s, |S|\}$ such that $G[S]$ is trivial, that is, either complete or empty. We know from Theorem 7 that $R^*(s)$ exists. Clearly, $R^*(s) \geq R(s)$ but, not surprisingly, $R^*(s)$ is of a greater order of magnitude than $R(s)$: it turns out that there are positive constants c and d such that $2^{2^{cs}} < R^*(s) < 2^{2^{ds}}$.

VI.2 Canonical Ramsey Theorems

Can anything significant be said about colourings of $\mathbb{N}^{(r)}$ with infinitely many colours? Can we guarantee that there is an infinite set $M \subset \mathbb{N}$ such that on $M^{(r)}$ our colouring is particularly 'nice'? In 1950, Erdős and Rado proved that, unexpectedly, this is precisely the case.

In what follows, $M, N, M_1, N_1, \ldots$ denote countable infinite sets, and $r, s, \ldots$ are natural numbers.

We call two colourings $c_1 : N_1^{(r)} \to C_1$ and $c_2 : N_2^{(r)} \to C_2$ *equivalent* if there is a 1-to-1 map ϕ of N_1 onto N_2 such that for $\rho, \rho' \in N_1^{(r)}$ we have $c_1(\rho) = c_1(\rho')$ if and only if $c_2(\phi(\rho)) = c_2(\phi(\rho'))$.

In an ideal world, for every colouring of $N^{(r)}$ (with any number of colours) there would be an infinite set $M \subset N$ on which the colouring is equivalent to one of finitely many colourings. Surprisingly, even more is true.

Call a colouring $c : N^{(r)} \to C$ *irreducible* if for every infinite subset N_1 of N, the restriction of c to $N_1^{(r)}$ is equivalent to c. Also, call a set C of colourings $\mathbb{N}^{(r)} \to \mathbb{N}$ *unavoidable* if for every colouring c of $\mathbb{N}^{(r)}$ there is an infinite set $M \subset \mathbb{N}$ such that the restriction of c to $M^{(r)}$ is equivalent to a member of C. Erdős and Rado proved that for every r there is a finite unavoidable family of irreducible colourings.

What are examples of irreducible colourings of $\mathbb{N}^{(r)}$? Two constructions spring to mind: a monochromatic colouring, in which all r-sets get the same colour, and an all-distinct colouring, in which no two sets get the same colour. After a moment's thought, we can construct more irreducible colourings. Given $N \subset \mathbb{N}$, $\alpha = \{a_1, \cdots, a_r\} \in N^{(r)}$, $a_1 < \ldots < a_r$, and $S \subset [r] = \{1, \ldots, r\}$, $|S| = s$, set $\alpha_S = \{a_i : i \in S\}$. Define the *S-canonical* colouring $c_S : N^{(r)} \to N^{(s)}$, by setting