Chapter 1

Ramsey Theory

Pigeonhole Principle *Given n pigeons in q pigeonholes, there has to be*

- *a pigeonhole with* **at least** $\left\lceil \frac{n}{q} \right\rceil$ *pigeons, and*
- *a pigeonhole with* **at most** $\left\lfloor \frac{n}{q} \right\rfloor$ *pigeons.*[1]

Of course the Pigeonhole Principle (PP) just formulates simple general properties of any decent "average"-concept: there should always be an instance that is at least the average and an instance that is at most the average.

The more general form of the PP allows that the pigeonholes are labeled and that we designate an individual "target" number of pigeons for each of them. Formally: **General Pigeonhole Principle** *Let $n_1, \dots, n_q \in \mathbb{N}$ be arbitrary positive integers and let P be a set which is partitioned into q pairwise disjoint subsets $P_1, \dots, P_q$, that is $P = \cup_i = 1^q P_i$. If $|P| \geq \sum_{i=1}^{q} n_i - q + 1$ then there exists an index $i \in [q]$ such that $|P_i| \geq n_i$.*

For a formal proof, say of the first statement of the PP, one can note that the negation is simply saying that *all* pigeonholes have *strictly less* than $\left\lceil \frac{n}{q} \right\rceil$, i.e. at most $\left\lceil \frac{n}{q} \right\rceil - 1$ pigeons. This leads to a contradiction to all pigeons appearing in one of these q pigeonholes, as $q \cdot \left(\left\lceil \frac{n}{q} \right\rceil - 1 \right) < n$.

In applications of the PP a set of objects/individuals/entities (pigeons) are classified into categories (pigeonholes) according to some characteristic. The categories are *exclusive*, so no object/person/entity (pigeon) can be put in two categories, but each of them belongs to one of them.

In our treatment we will mostly think of the classification into categories as a *coloring* of the base set.

In the first part of our course we will take the Pigeonhole Principle to a whole new level while studying both the quantitative and the qualitative aspects of Ramsey theory.

[1] Or saying the same more formally: if the elements of a set Q are classified into q pairwise disjoint subsets (i.e. Q is the disjoint union of the sets Q_i, $i = 1, \dots q$), then there is a subset Q_j with $|Q_j| \geq \left\lceil \frac{|Q|}{q} \right\rceil$ elements and there is a subset Q_ℓ with $|Q_\ell| \leq \left\lfloor \frac{|Q|}{q} \right\rfloor$ elements.

1.1 Ramsey's theorem for graphs

1.1.1 Two-colour Ramsey numbers for cliques

- Warm-up problem from sociology
  - How many people can be at a party without three mutual friends or three mutual strangers?
  - Make a graph: vertices := people, red edge := friends, blue edge := strangers $\Rightarrow$ how large can a two-coloured complete graph without monochromatic triangles be?
  - Answer, part 1: at least 5: red graph is C_5
  - Answer, part 2: at most 5:
    * Suppose we have six vertices, and consider the edges incident to the first one
    * wlog (at least) three of these are red (where $3 = \lceil \frac{5}{2} \rceil$; PP is used with the 5 incident edges (pigeons) classified into 2 classes (pigeonholes) according to their color)
    * if any two such endpoints share a red edge $\to$ red triangle, done
    * therefore the endpoints of the three red edges span a blue triangle, done

Definition 1.1.1 (Ramsey numbers). *Given $s \in \mathbb{N}$, let $R(s)$ be the minimum $n \in \mathbb{N}$ such that every red-blue colouring of the edges of K_n contains a subgraph isomorphic to K_s the edges of which all have the same color (refered to as being* monochromatic *(or* m.c.*, for short)).*

- Observations
  - We have just proved $R(3) = 6$
  - Upper bound proof: **finding** a monochromatic clique in an **arbitrary** colouring
  - Lower bound proof: construction of a **specific** colouring **without** monochromatic cliques
  - Often convenient to only consider red subgraph: cliques $\leftrightarrow$ red cliques, independent sets $\leftrightarrow$ blue cliques
  - Finiteness of $R(42)$ for example is totally unclear at this point

Theorem 1.1.2 (Ramsey [8], 1930). *For every $s \in \mathbb{N}$, $R(s)$ is finite.*

- Philosophy: "every large system, no matter how chaotic, contains ordered subsystems"
- Quintessential Ramsey result — find monochromatic substructures in large coloured structures
- Ramsey: British logician, primarily interested in existence of $R(s)$

Claim 1.1.3. *For every $s \in \mathbb{N}$, $R(s) \leq 4^s$.*

Proof. Let $n = 2^{2s}$, and fix an arbitrary $\texttt{red}/\texttt{blue}$ edge-colouring $c : E(K_n) \to \{\texttt{red}, \texttt{blue}\}$ of K_n. We will find a monochromatic K_s.

To this end we first will find a sequence of vertices $v_1, v_2, \ldots, v_{2s-2} \in V := V(K_n)$, which is *right-monochromatic*, by which we mean that for any fixed index $i = 1, 2, \ldots, 2s-3$, the edges going from v_i to a vertex v_j with a larger index j have the same color. In other words for any $i \in \{1, 2, \ldots, 2s-3\}$, there exists a color $c^*(i) \in \{\texttt{red}, \texttt{blue}\}$, such that $c(v_i v_j) = c^*(i)$ for every j, $i < j \leq 2s-2$. Once we find such a right monochromatic sequence, we will be done. Indeed, the PP provides us with a subsequence $v_{i_1}, \ldots, v_{i_{s-1}}$ of length $\lceil \frac{2s-3}{2} \rceil = s-1$, such that

$c^*(i_1) = \cdots = c^*(i_{s-1})$ and then the vertices $v_{i_1}, \ldots, v_{i_{s-1}}$, together with the last vertex v_{2s-2} form a monochromatic clique of order s (in color $c^*(i_1)$).

So to complete the proof we just need to find this long enough right-monochromatic sequence. We do this in a quite greedy fashion, using again the PP. We will keep picking the next vertex arbitrarily from the set of vertices still under consideration, then deleting all neighbours whose edges are coloured with the less frequently appearing colour, and note that we have at least half of the vertices remaining. Formally, let us set $S_0 := V$ and for every $i \in \{0, 1, \ldots, 2s-3\}$ do the following. Given a set S_i of size 2^{2s-i}, we select an arbitrary vertex in S_i, name it v_{i+1}, and let B_{i+1} and R_{i+1} denote the sets of those neighbors of v_{i+1} in S_i which are connected to it via a `blue` and a `red` edge, respectively. Then obviously $|B_{i+1}| + |R_{i+1}| = |S_i| - 1$. We choose S_{i+1} to be the larger of B_{i+1} and R_{i+1}, so for its size we have

$$|S_{i+1}| \geq \left\lceil \frac{|B_i| + |R_i|}{2} \right\rceil = \left\lceil \frac{2^{2s-i} - 1}{2} \right\rceil = 2^{2s-(i+1)},$$

as desired. To complete the proof we just need to check that this process can go on long enough, i.e. v_{2s-2} can actually be selected. For that we need S_{2s-3} to be non-empty, which is the case since $|S_{2s-3}| = 2^{2s-(2s-3)} = 8$. (So in fact in the theorem we could have claimed the upper bound $4^s/8$ instead.) □

Even though this upper bound is getting close to being a century old, the order 4^s is still essentially the best known. We will return to the question of how good these bounds are when we discuss lower bounds in the next section; for now we see a couple of generalisations.

Hungarian mathematicians Paul Erdős and George Szekeres came across the problem independently (see their motivation two sections later), and obtained slightly better quantitative bounds. For the improvement one can observe that the proof above was quite "wasteful" in the sense that we always followed greedily the immediately best option, towards the larger monochromatic degree, and then we completely ignored the fact that once we did that in some color, in that color it is enough to find a clique of one smaller order. This makes the problem asymmetric after the first step of the proof, because in the other color we still need to find a clique of same order as before. To accommodate this asymmetry, the following definition is necessary.

Definition 1.1.4 ((not necessarily symmetric) Ramsey numbers). *Given $s, t \in \mathbb{N}$, let $R(s,t)$ be the minimum $n \in \mathbb{N}$ such that every red-blue colouring of the edges of K_n contains either a red K_s or a blue K_t.*

- Observations
  - Swapping `red`/`blue`: $\Rightarrow R(s,t) = R(t,s)$
  - $R(s,1) = 1$, $R(s,2) = s$.

The following upper bound of Erdős and Szekeres will be proved on the homework as a guided exercise.

Theorem 1.1.5 (Erdős–Szekeres [4], 1935). *For every $s, t \in \mathbb{N}$, $R(s,t) \leq \binom{s+t-2}{s-1}$. In particular,*

$$R(s) = O\left(\frac{4^s}{\sqrt{s}}\right).$$

1.1.2 Generalization 1: Ramsey's theorem for infinite graphs

- What happens if we colour the edges of an infinite graph, instead of a large finite graph?
- Infinite graphs
  - Vertex set $\mathbb{N}$, Edge set $\binom{\mathbb{N}}{2}$

  - Colour every edge *red* or *blue*
- Finite monochromatic cliques
  - In particular, for any $t \in \mathbb{N}$ by considering the restriction of the colouring to the edges between the first $R(t, t)$ numbers, we are guaranteed to find a monochromatic clique of size t.
  - Thus we definitely have arbitrarily large monochromatic cliques
- Infinite monochromatic cliques
  - This is **NOT** the same as an infinite monochromatic clique
    * These large finite cliques can be bounded and far apart
  - Question: Do we get an infinite monochromatic clique?

Theorem 1.1.6 (Ramsey [8], 1930)**.** *For any two-colouring of $\binom{\mathbb{N}}{2}$, there exists an infinite set $S \subset \mathbb{N}$ for which $\binom{S}{2}$ is monochromatic.*

Proof. One can repeat the vertex selection procedure in the proof of Claim 1.1.3 infinitely often and hence create an infinite right-monochromatic sequence. The proof of this is identical to the one there with the obvious adaptation that $S_i = B_{i+1} \cup R_{i+1}$ being infinite implies S_{i+1} being infinite. And the infinite right-monochromatic sequence gives rise to an infinite monochromatic clique (as at least one of the colors must occur infinitely many times among the c^*-values). □

Homework: infinite Ramsey Theorem $\Rightarrow$ finite Ramsey Theorem

1.1.3 Generalization 2: Multicolour Ramsey numbers

In many applications the relation between people (or other entities) are not necessarily binary. After all, there must be more to human (or other) relations than love and hate. For this reason the following definition arises quite naturally.

Definition 1.1.7 (Multicolour Ramsey numbers)**.** *Given integers $r \geq 2$ and $t_1, t_2, \dots, t_r \in \mathbb{N}$, let $R_r(t_1, t_2, \dots, t_r)$ be the minimum $n \in \mathbb{N}$ such that for any colouring of the edges of K_n with colours from $[r]$, there is some index i for which there is a monochromatic K_{t_i} of colour i.*

Formally, by an r-coloring of the edges we mean a function $c : E(K_n) \to [r]$. Note that we had to forget our nice habit of using actual colors in our coloring and retreat to the (probably more boring and definitely less colorful) realm of naming our colors by integers. This is purely for practical purposes, as statements about more than two colors become quite cumbersome to write down when using not only `red` and `blue`, but also `yellow`, *green*, *orange*, *purple*, etc ... You get the picture(!)

Theorem 1.1.8. *For any $r \geq 2$ and $t_1, t_2, \dots, t_r \in \mathbb{N}$, $R_r(t_1, t_2, \dots, t_r)$ is finite.*

Proof. Proof by induction on r, the number of colours. Base case, $r = 2$, is Theorem 1.1.2.

For the induction step, suppose $r \geq 3$, and we have numbers $t_1, t_2, \dots, t_r$. We will take a large enough n, the formula given later in the proof, and fix an arbitrary r-colouring c of the edges of K_n.

The idea is to go "colorblind", combine the last two colors together and use the finiteness of the Ramsey numbers for $r - 1$ colors. Of course this will guarantee what we want only in the first $r - 2$ colours. In order to have what we want in the last two colors as well, we will ask our $(r - 1)$-color Ramsey number to deliver a large enough clique in the last colour, so we can use that to take both of the colorblinded orginal colors.

Let us now formalize this idea. We define coloring $c^* : E(K_n) \to [r-1]$ from c. Let $c^*(xy) = r-1$ if $c(xy) = r$ and $c^*(xy) = c(xy)$ otherwise. By the induction hypothesis, $R_{r-1}(t_1, t_2, \ldots, t_{r-2}, R(t_{r-1}, t_r))$ is finite, and we choose $n = R_{r-1}(t_1, t_2, \ldots, t_{r-2}, R(t_{r-1}, t_r))$. Note that here we use that we use that we already can assume the finiteness of the Ramsey number for *any* large value of clique orders if the number of colors is only $r-1$. Now the definition of the Ramsey number provides us an appropriate monochromatic clique in one of the $r-1$ colors. If this monochromatic clique is in one of the first $r-2$ colours, then we are done, as we then have a monochromatic clique of size t_i in colour i, $1 \leq i \leq r-2$. Otherwise we have a clique of size $R(t_{r-1}, t_r)$ that uses the combined colour. We now restore the original colouring, so that all of these edges are coloured either $r-1$ or r. By definition of $R(t_{r-1}, t_r)$, we also find the desired monochromatic clique in this case. □

Remarks.

- What kind of upper bound does this give?
  - Following the argument in the proof, we get
  $$R_r(t_1, t_2, \ldots, t_r) \leq R(t_1, R(t_2, R(t_3, \ldots R(t_{r-1}, t_r) \ldots))),$$
  - Applying Theorem 1.1.5 and the simplification that $\binom{s+t-2}{s-1} < 2^{s+t}$, this shows that we have
  $$R_r(t_1, t_2, \ldots, t_r) \leq 2^{t_1 + 2^{t_2 + 2^{\cdot^{\cdot^{\cdot^{2^{t_{r-1}+t_r}}}}}}}$$
  - In particular, $R_r(t, t, \ldots, t) \leq 2^{2^{\cdot^{\cdot^{\cdot^{2^{2t+1}}}}}}$ (tower of height r)
- Can we do better?
  - By splitting colours evenly and merging them simultaneously in the above argument, one can reduce the upper bound to a tower of height $\log r$.
  - In the homework you are asked to give an upper bound of the form $r^{\sum_i t_i}$ (which is much better!).

1.2 Lower bounds for Ramsey's theorem

Recall that to **lower bound** $R(s, t)$ one needs to **provide a colouring of a large complete graph without a `red` monochromatic** K_s **and a `blue` monochromatic** K_t.
For example for $R(3, 3)$ we were "lucky" to have the C_5-construction that complements our upper bound of 6 perfectly and hence proves that $R(3, 3) = 6$. The value of $R(4, 4)$ is known (it is 18) mainly because we are again lucky enough to have an incredibly nice coloring on 17 vertices which does the deed. Starting from $s \geq 5$ however, it is unclear how to generalize this construction the "right way". Or rather, the obvious generalization does not anymore match the upper bounds we have available from our various PP-based arguments. For $R(5, 5)$ all what is known is that

$$43 \leq R(5, 5) \leq 48.$$

The upper bound was improved from 49 to 48 just recently (in 2017), with heavy use of computer checking. It is worthwhile to think over what such a proof must deal with. There are $2^{\binom{48}{2}} > 10^{338}$ `red`/`blue`-colourings of the complete graph on 48 vertices. The program must consider all of them and verify that they all contain a monochromatic K_5. Now, there are about 10^{80} particles in the (observable) universe and the age of the universe is thought of being about 10^{26} nanoseconds. So every single particle in the universe has to check at least 10^{232} of these cases in every single nanosecond of its existence and then they have a chance to be finished by now ... This indicates the enormous numbers involved in this simple combinatorial problem and maybe explains our futility in solving it. And it also indicates that the recent verification must do something clever besides pure brute-force checking.

1.2.1 A first idea: Dense K_s-free graphs

The first idea one might have for a construction is to be greedy. This sometimes works, greedy algorithms are often effective in computer science. Here one could argue with the following heuristic.

Heuristic. *We need two K_s-free graphs complementing each other, that is together they should occupy all the $\binom{n}{2}$ edges of K_n. Let us first focus on the* `red` *graph and make sure that it uses up as many of these edges as possible, and deal with the* `blue` *graph later.*

This approach leads us to a natural extremal graph theory problem, asking for the maximum number of edges a K_s-free graph on n vertices can have. Let us first see what happens when $s = 3$, that is, in the case of triangle-free graphs. After some trial and error with examples of triangle-free graphs on a small number of vertices, one convinces oneself that the complete bipartite graph $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ seems to be *a* triangle-free graph with many edges. The result that indeed one cannot do better, i.e. that every graph with

$$e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right) + 1 = \left\lceil \frac{n}{2} \right\rceil \cdot \left\lfloor \frac{n}{2} \right\rfloor + 1$$

edges does have a triangle, is one of the first theorems of Extremal Graph Theory.

Theorem 1.2.1 (Mantel, 1907). *If G is K_3-free then $e(G) \leq e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right)$.*

Proof. Consider a vertex w of maximum degree in a triangle-free graph G, i.e. let $d(w) = \Delta(G) =: \Delta$. Recall that $N(w)$ is the neighborhood of w, and let us denote by $R(w) = V(G) \setminus N(w)$ the rest. We bound from above the number of edges of G by adding up all the degrees of vertices in $R(w)$. Indeed, by adding up the degrees of vertices in $R(w)$ we account for each edge of G at least once, since G is triangle-free, hence $N(w)$ contains no edge. Consequently,

$$e(G) \leq \sum_{v \in R(w)} d(v) \leq \sum_{v \in R(w)} \Delta = |R(w)| \cdot \Delta = (n - \Delta)\Delta \leq \left(n - \left\lfloor \frac{n}{2} \right\rfloor\right) \cdot \left\lfloor \frac{n}{2} \right\rfloor = e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right),$$

as required. Here we used that $|R(w)| = n - |N(w)| = n - \Delta$, and then maximized the quadratic function $x \mapsto (n - x)x$ over the integers. □

Remark. When adding up the degrees in $R(w)$ we accounted for each edge between $R(w)$ and $N(w)$ exactly once, and for each inside $R(w)$ exactly twice. The reason we did not worry so much because of this overcount is our firm belief in our construction being optimal. In the complete bipartite graph there are no edges inside $R(w)$, so if it is indeed optimal we do not lose anything by this estimation.

The construction of complete bipartite graphs easily generalizes when instead of K_3 we want to forbid K_{s+1}. Then we can take a graph with a vertex set partitioned into s parts inlcude all edges between parts and no edges inside the parts. These graphs are called *complete s-partite graphs* and can be parametrized by the sizes of its parts $t_1, \ldots, t_s$. Complete s-partite graphs do not contain K_{s+1}, since two of the $s + 1$ vertices of any copy of a K_{s+1} would have to be in the same part (by the PP), but vertices in the same part are not adjacent, contradiction. Among complete s-partite graphs the most edges are contained in the one where the parts are as equal as possible, so any two parts have sizes differing by at most one. Indeed, otherwise we can move a vertex from a bigger part to smaller part and increase the number of edges. This graph is called the *Turán-graph* and is denoted by $T_{n,s}$.

Turán has shown in 1941 (and we will shown in a couple of weeks) that the Turán graph $T_{n,s}$ is indeed the K_{s+1}-free graph with the most number of edges on n vertices.

Let us now return to our original problem of constructing an appropriate 2-coloring. As the `red` graph, we decided to take the K_s-free Turán graph $T_{n,s-1}$ which uses up the most edges from K_n. What is then the `blue` graph? It is the disjoint union of $s-1$ cliques of order roughly $\frac{n}{s-1}$. In order to ensure that the `blue` graph also has no K_s, we better make sure that $\frac{n}{s-1} < s$, that is $n \leq (s-1)^2$. In other words, with this method we can constructed Ramsey graphs on $(s-1)^2$ vertices, but no more. Hence

$$R(s,s) \geq (s-1)^2 + 1,$$

pretty pathetic when compared to the best known upper bound, which stands close to 4^s.

1.2.2 The right idea: random construction

The coloring of the previous subsection is pretty simple, yet it is surprisingly hard to improve. For a short period of time Turán himself believed his construction to be optimal. Erdős massively destroyed this belief in 1947 via an equally simple, but fundamentally different idea.

Heuristic. *We want the same from the* `red` *and the* `blue` *graph (they should be K_s-free). Their roles are symmetric. Each edge has as much reason to be* `red` *than to be* `blue`*. Let us choose the color of each edge uniformly at random, independently from each other.*

Theorem 1.2.2 (Erdős, 1947). $R(t,t) \geq (1-o(1))\frac{t}{e\sqrt{2}}2^{\frac{t}{2}}$.

Proof. The idea of this proof is to prove the *existence* of a large Ramsey colouring without actually presenting it. Colour each edge of K_n by `red` or `blue` with probability 1/2, such that these random choices are mutually independent of each other. In other words, our probability space consists of the set of all `red`/`blue`-colourings of $E(K_n)$ with all colorings being equally likely.

We want to avoid a monochromatic K_t. So for each $R \in \binom{[n]}{t}$, i.e. each set R of t vertices, we define E_R be the event that the induced subgraph of K_n on R is monochromatic. The probability that E_R happens is: $\mathbb{P}(E_R) = 2(\frac{1}{2})^{\binom{t}{2}}$ and we have $\binom{n}{t}$ such events. The probability that there exists a monochromatic K_t can then be estimated by the union bound

$$\mathbb{P}\left(\bigcup_{R\in\binom{[n]}{t}} K_t\right) \leq \sum_{R\in\binom{[n]}{t}} \mathbb{P}(R_K) = \binom{n}{t}\cdot 2\cdot\left(\tfrac{1}{2}\right)^{\binom{t}{2}} \leq 2\left(\tfrac{en}{t}\right)^t\left(\tfrac{1}{2}\right)^{\binom{t}{2}}.$$

If this expression is less than 1, then there exists a `red`/`blue`-coloring of $E(K_n)$ without a monochromatic K_t. Taking the tth root and rearranging we obtain that if $n < \frac{t}{2^{\frac{1}{t}}\cdot\sqrt{2}e}2^{\frac{t}{2}}$, then $\mathbb{P}($ there is a m.c. $K_t) < 1$. Therefore, there exists a `red`/`blue`-colouring without a monochromatic K_t on

$$n = \left\lfloor \frac{t}{2^{\frac{1}{t}}\cdot\sqrt{2}e}2^{\frac{t}{2}} \right\rfloor$$

vertices. It exists not with positive probability, or 99% probability, but with absolute, 100% certainty, SURELY THERE IS ONE. And hence, $R(t,t) \geq (1-o(1))\frac{t}{e\cdot\sqrt{2}}2^{\frac{t}{2}}$, as claimed. □

Let us remark that this proof in fact shows that *almost every colouring* of a K_n is a good colouring to avoid cliques of order two more. However, we cannot explicitly find one. (see the Constructive Combinatorics course next semester.)
Recall where we stand:

$$2^{\frac{t}{2}} \leq R(t,t) \leq 4^t.$$

So both bounds are exponential now, but they are still very far apart. A relatively recent improvement (about a decade old) by a factor which is superpolynomial (if ever so slightly) is considered a great breakthrough and appeared in the *Annals of Mathematics*. But there are no improvements to the bases. In particular, it would be a fantastic advance to prove that $R(t,t) < 3.9999^t$ holds.

In the above proof the use of probability is not essential, one could simply *count* bad colorings among all colorings and conclude that there must be a good one left even after taking out all the bad ones. Ultimately this is true about every statement in discrete probability. However, the idea of introducing randomness is a major paradigm shift. It directs our attention to the various tools of probability theory, some of which would really be problematic to say, not to mention find, through just counting. The improvement of the next section is a initial step in this direction.

1.2.3 A twist on the method: improving the constant factor

It is worthwhile to note that one can prove[2] that with probability tending to 1, the random coloring *will* contain monochromatic cliques of order t, so in a way the crude analysis through the union bound is essentially best possible.

Using some alterations to the random construction however, we can improve the Erdős lower bound by a constant factor $\sqrt{2}$. By the above, in this regime it is simply not anymore true that the random coloring is a good one, still there *is* a good one.

Theorem 1.2.3. $R(t,t) \geq (1-o(1))\frac{t}{e}2^{\frac{t}{2}}$.

Proof. Like in the previous theorem, let us colour the edges of K_n uniformly at random by either `red` or `blue` with probability $\frac{1}{2}$. As we mentioned before this proof, if we raise n above what we have worked with in Theorem 1.2.2, it is inevitable that with overwhelming probability there *will be* (many) monochromatic K_t. Our plan is to destroy each of these by deleting a vertex from them and hope that the remaining two-colored clique, now without any monochromatic K_t, has retained most of the original vertices. In other words, we need to show that the number of monochromatic K_t is of smaller order than the number of vertices.

To this end, let X be the random variable that equals the number of monochromatic K_t's in this two-colouring. To have an idea about this seemingly complicated random variable, we express it as the sum of many simple ones and apply a simple yet surpringly powerful general property of expectation of variables: its linearity. For each t-element set K, let X_K denote the indicator random variable of the event that K induces a monochromatic K_t. Then $X = \sum_{K\in\binom{[n]}{t}} X_K$ and by the linearity of expectation

$$\mathbb{E}[X] = \sum_{K\in\binom{[n]}{t}} \mathbb{E}[X_k] = \sum_{K\in\binom{[n]}{t}} \mathbb{P}(X_K) = \binom{n}{t}\cdot 2 \cdot \left(\tfrac{1}{2}\right)^{\binom{t}{2}}.$$

Therefore, there exists a colouring c such that the number of monochromatic K_t's is at most $\binom{n}{t}\cdot 2\cdot\left(\frac{1}{2}\right)^{\binom{t}{2}}$. Fix such a colouring and delete one vertex from each monochromatic K_t. This gives us a `red`/`blue`-coloring on at least $n - \binom{n}{t}\cdot 2\cdot\left(\frac{1}{2}\right)^{\binom{t}{2}}$ vertices without any monochromatic K_t. Hence

$$R(k,k) > n - \binom{n}{t}\cdot 2\cdot\left(\frac{1}{2}\right)^{\binom{t}{2}} \geq n - \left(\frac{ne}{t}\cdot 2^{-\frac{t-1}{2}+\frac{1}{t}}\right)^t \tag{1.1}$$

[2]We do not do it here. One must use the second moment method

where we estimated $\binom{n}{t} \le \left(\frac{ne}{t}\right)^t$. Substituting $n = \sqrt{2}^t \cdot \frac{t}{e}$, we obtain a `red`/`blue`-coloring on

$$\sqrt{2}^t \cdot \frac{t}{e} - \left(2^{\frac{1}{2}+\frac{1}{t}}\right)^t = \sqrt{2}^t \cdot \frac{t}{e}(1 - o(1))$$

vertices.[3] This shows the promised lower bound on $R(t,t)$. □

1.3 Hypergraph Ramsey theory

1.3.1 A motivation: the Happy Ending Problem

> Our present problem has been suggested by Miss Esther Klein in connection with the following proposition.
>
> From 5 points of the plane of which no three lie on the same straight line it is always possible to select 4 points determining a convex quadrilateral.
>
> We present E. Klein's proof here because later on we are going to make use of it. If the least convex polygon which encloses the points is a quadrilateral or a pentagon the theorem is trivial. Let therefore the enclosing polygon be a triangle ABC. Then the two remaining points D and E are inside ABC. Two of the given points (say A and C) must lie on the same side of the connecting straight line $\overline{DE}$. Then it is clear that $AEDC$ is a convex quadrilateral.
>
> Miss Klein suggested the following more general problem. *Can we find for a given n a number N such that from any set containing at least N points it is possible to select n points forming a convex polygon?*
>
> There are two particular questions: (1) does the number N corresponding to n exist? (2) If so, how is the least $N(n)$ detremined as a function of n?

The text above is from the introduction of a paper of Paul Erdős and George Szekeres from 1935. We put it here in original quote, because in retrospect this paper turned out be pioneering in two different, and at the time completely new fields: *combinatorial geometry* and *Ramsey theory.*

Two years before, in 1933, the three main protagonists, along with a group of other young mathematically inclined, like Turán (whose name we will also hear a lot this semester), were meeting regularly after university in the main park of Budapest and taking long walks in the wood to discuss, what else, mathematics. Apparently, already then, this is what the cool kids were doing in their free time.

It was at one of these meetings when Esther confronted the boys with her proof for a convex quadrangle and the general question. Paul and George immediately jumped on the topic with great enthusiasm, they got really excited by what they felt was a completely new type of geometric problem. They gave constructions of point sets of size 2^{t-2} in general position, without containing a convex t-gon. They also answered (1) in two different ways and in the process they rediscovered Ramsey's Theorem independently.

To motivate how the connection might have come about let us first make a few things precise from the introduction. First of all we will be dealing with point sets P that are *in general position*, i.e., no three points of P are on the same line. With forsight we will denote by $HE(t)$ (and not $N(t)$) the minimum integer n such that any n points in the plane in general position contain t points spanning a convex t-gon.

All four-element point sets in general position in the plane are exactly one of two kinds: either they form a convex 4-gon or not, depending on whether their convex hull is a quadrilateral or

[3] To optimize the choice of n (in terms of t) might be difficult to do precisely, because of the binomial coefficient involved in the expression. It is easy to see however (and it is worthwhile to actually do!), that substituting a constant factor larger n, say $n = \sqrt{2}^t \cdot \frac{t}{e-\varepsilon}$, would make the expression on the right hand side of (1.1) negative. So we are at the asymptotical optimum.

a triangle. The first easy, but important, observation is that the convexity of a t-element set in general position could be characterized through the convexity of its four-element subsets.

Proposition 1.3.1. *A t-element subset in general position forms a convex t-gon if and only if all $\binom{t}{4}$ of its four-subsets form a convex 4-gon.*

Proof. If $P \subseteq \mathbb{R}^2$ forms a convex t-gon, then no point is the convex combination of the other $n-1$. In particular no point is the convex combination any other three points, so every four-subset is convex.

In the other direction, suppose that every four-subset of P is a convex 4-gon. If a point $p \in P$ would be a convex combination of the others, then it is also a convex combination of just three of them: the vertices of the triangle which contains it, from an arbitrary triangulation of the convex hull of P. This provides non-convex four-subset of P, a contradiction. □

The second important observation is the proposition of Klein, which says that it is impossible that for some five-element set *none of* the $\binom{5}{4}$ four-element subsets are convex.

Proposition 1.3.2. *It cannot happen that for some 5-element point set in general position none of the four-element subset forms a convex 4-gon.*

The natural classification of four-element point sets and the relation of these classes to larger point sets lead Erdős and Szekeres to the idea to *color the four-element subsets* of n points in general position by `red` or `blue` given whether they are in convex position or not, respectively. Then Proposition 1.3.1 translates to a t-element subset being in convex position if and only if all its $\binom{t}{4}$ four-element subset are `red`. Klein's proposition on the other hand forbids the presence of a five-element set with all its four-subsets being `blue`.

So, let's do some wishful thinking. If we were to know that there exists an integer, however large, but finite, denoted mysteriously by $R^{(4)}(t,5)$, such that for *any* `red`/`blue`-coloring $c : \binom{[R]}{4} \to \{\texttt{red}, \texttt{blue}\}$ of the 4-element subsets of the $R^{(4)}(t,5) =: R$-element set $[R]$, there exists a t-element subset $T \subseteq [R]$ with all its 4-subsets `red` *or* a 5-element subset $T \subseteq [R]$ with all its 4-subsets `blue`, so if we know all this, then we would be done! Because then, we claim, $HE(t) \leq R^{(4)}(t,5)$, so $HE(t)$ was also finite. Indeed, should such miraculous $R := R^{(4)}(t,5)$ existed for some t, then taking an arbitrary set $P \subseteq \mathbb{R}^2$ of R points in general position and creating the coloring described above, this coloring cannot contain a 5-subset with all its four-subsets having color `blue`! But then, by the magic property of the number R, there must be a t-subset $T \subseteq [R]$ for which *every* 4-subset is `red`. And that, via Proposition 1.3.1, implies that T is in convex position!

All we need is the existence of such a magic number $R^{(4)}(t,5)$. This motivated Erdős and Szekeres,[4] and motivates us as well, to introduce a Ramsey number for colorings, where instead of edges (i.e. 2-element subsets), we color k-element subsets.

1.3.2 The hypergraph Ramsey theorem

What is a hypergraph? It is a generalization of the concept of graphs, where instead of just 2-element vertex sets, as edges, we consider arbitrary subsets of a vertex set V. Formally, a *hypergraph* is defined as a pair $(V, \mathcal{F})$ of a vertex set V and edge set $\mathcal{F}$, where $\mathcal{F} \subseteq 2^V$. Often, if it is not ambiguous, we omit referring to the vertex set and identify the hypergraph with its edge set $\mathcal{F}$. A hypergraph is called *k-uniform*, for some positive integer k, if all its edges have size k, that

[4]There might also have been other motivating factors ... But this is just speculation .... Anyway, Esther Klein and George Szekeres were married a couple of years after the initiation of the problem by the former and its extension (together with Erdős) by the later. This prompted Paul Erdős to coin the term *Happy End Problem.* Klein and Szekeres escaped persecution of Jews in Hungary before the second world war and settled in Australia afterwards. They died within an hour of each other at the age of 95 and 94, respectively. A good example of how far an innocent-looking math problem might lead you ...

is if $\mathcal{F} \subseteq \binom{V}{k}$. A k-uniform hypergraph is sometimes called a *k-graph*. The edges of a hypergraph are sometimes called hyperedges, and the edges of a k-graph are sometimes called k-edges.

Examples.

(1) For $k = 2$, we get back our good old graph concept: a 2-graph is just a graph.

(2) The analogue of complete graphs: the *complete k-graph* on t vertices contains all k-subsets of the t-element vertex set $[t]$ and is denoted by $K_t^{(k)}$. In other words, $K_t^{(k)} = \left([t], \binom{[t]}{k}\right)$.

(3) There are various analogues of many graph theoretic concepts, like path and cycles: tight paths/cycles, loose paths/cycles, ℓ-tight paths/cycles, Berge-cycles, etc ...

(4) Projective planes: $V =$ points := 1-dimensional subspaces of K^3, $\mathcal{F} =$ lines := 2-dimensional subspace of K^3, where K is an arbitrary field. When K is the finite field $\mathbb{F}_q$, we get a $(q+1)$-uniform hypergraph with $q^2 + q + 1$ vertices and equally many hyperedges. E.g. the Fano plane is a 3-uniform hypergraph on 7 vertices with 7 edges obtained from the projective plane defined over $\mathbb{F}_2$.

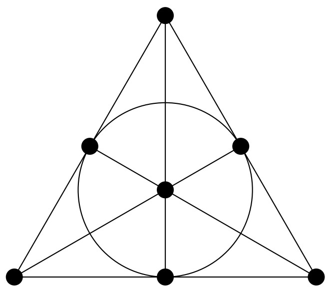

Figure 1.1: The Fano plane

For the rest of this section we will only be concerned with the complete k-uniform hypergraph. We will define hypergraph Ramsey number as the straightforward generalization of graph Ramsey numbers.

Definition 1.3.3 (Hypergraph Ramsey Number). *Given $k \in \mathbb{N}$, and $s, t \geq k$, $R^{(k)}(s,t)$ is the minimum $n \in \mathbb{N}$ such that for every coloring $c : \binom{[n]}{k} \to$ `red`/`blue` there exists a set $T \in \binom{[n]}{t}$ such that $c(S) =$ `red` for every $S \in \binom{T}{k}$ or there exists a set $T \in \binom{[n]}{s}$ such that $c(S) =$ `blue` for every $S \in \binom{T}{k}$.*

Sometimes we refer to the property of the subset T in the definition that T hosts a `red` $K_s^{(k)}$ or that it hosts a `blue` $K_t^{(k)}$. We call a hypergraph with all its edges colored with the same color *monochromatic*.

Before going on on, let us make some simple observations:

(1) $R^{(2)}(s,t) = R(s,t)$.

(2) $R^{(1)}(s,t) = s + t - 1$ (think over the detailed proof!)

(3) $R^{(k)}(s,t) = R^{(k)}(t,s)$.

(4) $R^{(k)}(k,t) = t$.

Analogous to the graph case, the first question we should ask ourselves is whether $R^{(k)}(s,t)$ is finite.

Theorem 1.3.4. *For arbitrary positive integers k, and $t, s \geq k$, the value $R^{(k)}(s,t)$ is finite.*

Remark. In the homework exercise you will be asked to define hypergraph Ramsey numbers for more than two colors and prove their finiteness. Moreover, the theorem also extends to the infinite setting analogously to the infinite Ramsey theorem for graphs.

First proof. We will follow the idea of the proof of Claim 1.1.3 and build a sequence where the color of every edge depends only on its smallest vertex in the sequence, and this way one can naturally identify a color with each vertex of the sequence. Then the 1-uniform Ramsey theorem will be used to select a subsequence where all these colors are the same, hence providing us with the monochromatic clique we want.

A sequence $v_1, \dots, v_\ell \in U$ of vertices will be called right-neighborhood-monochromatic in the set U, if the color of a k-set contained in U depends only on its element from the sequence with the smallest index (provided such an element exists). Formally, a sequence $v_1, \dots, v_\ell \in U$ is called *right-neighborhood-monochromatic in* U if there exists a coloring $c^* : [\ell] \to \{\texttt{red}, \texttt{blue}\}$ such that for every k-set $T \in \binom{U}{k}$ with $T \cap \{v_1, \dots, v_\ell\} \neq \emptyset$, we have $c(T) = c^*(v_{\min T})$, where we adopted the notation $\min T = \min\{j : v_j \in T\}$.

Similarly to Claim 1.1.3, our goal is to build a long enough right-neighborhood-monochromatic sequence $v_1, \dots, v_\ell$ in some set V_ℓ of large enough size. For us the length $\ell = t + s - 2k + 1$ will suffice with $|V_{t+s-2k+1}| = t + s - k$. Indeed, if we succeed to build such a sequence, then the 1-uniform Ramsey theorem will provide us with a subsequence of length $t - k + 1$ which is c^*-monochromatic in $\texttt{red}$ or a subsequence of length $s - k + 1$ which is c^*-monochromatic in $\texttt{blue}$. Such a sequence unioned with the remaining $k - 1$ vertices in $V_{t+s-2k+1} \setminus \{v_1, \dots, v_{t+s-2k+1}\}$ forms a c-monochromatic subset T of the size required for its color. It is easy to check that a set T obtained this way is always monochromatic. For this just note that every k-subset $S \subseteq T$ does contain at least one element of the sequence (since the rest has only $k - 1$ elements). Then the c-color of this S is the c^*-value of the minimum sequence-index appearing in S. But all these colors are the same by the way we selcted the elements of T from the sequence.

Let us see now how can we build recursively a right-neighborhood sequence $v_1, \dots, v_i$ in some set V_i and the appropriate coloring $c^* : [i] \to \{\texttt{red}, \texttt{blue}\}$.

For v_1 we choose an arbitrary vertex in $V_0 := [n]$. To find V_1, we consider the coloring of the complete $(k-1)$-uniform hypergraph on $V_0 \setminus \{v_1\}$ induced by the c-colors of the k-sets containing v. Formally, let $\tilde{c}(Q) := c(Q \cup \{v_1\})$ for every $Q \in \binom{V_0 \setminus \{v_1\}}{k-1}$. The $(k-1)$-uniform Ramsey theorem (true by induction) will provide us with a (large) $\tilde{c}$-monochromatic subset N_1 of size n_1 and we put $V_1 = N_1 \cup \{v_1\}$. Then v_1 is right-neighborhood-monochromatic in V_1 by the definition of N_1, because we can simply define $c^*(1)$ to be the $\tilde{c}$-color of the $(k - 1)$-subsets of N_1.

Given a sequence $v_1, \dots, v_i$ that is right-neighborhood-monochromatic in some V_i with an appropriate function $c^* : [i] \to \{\texttt{red}, \texttt{blue}\}$ we choose v_{i+1} arbitrarily from $N_i := V_i \setminus \{v_1, \dots, v_i\}$. We find V_{i+1} by considering the coloring $\tilde{c}$, defined by $\tilde{c}(Q) := c(Q \cup \{v_{i+1}\})$ for every $Q \in \binom{N_i \setminus \{v_{i+1}\}}{k-1}$, and (hoping to) use the $(k-1)$-uniform Ramsey theorem to provide us with a (large) $\tilde{c}$-monochromatic subset N_{i+1} of size n_{i+1}. Then we put $V_{i+1} = N_{i+1} \cup \{v_1, \dots, v_{i+1}\}$. The vertex v_{i+1} is right-neighborhood-monochromatic in $N_{i+1} \cup \{v_{i+1}\}$ because N_{i+1} is $\tilde{c}$-monochromatic and consequently the whole sequence $v_1, \dots v_{i+1}$ is right-neighborhood-monochromatic in V_{i+1}. Indeed, we can simply extend the already existing $c^* : [i] \to \{\texttt{red}, \texttt{blue}\}$ to $i + 1$ by defining $c^*(i+1)$ to be the $\tilde{c}$-color of the $(k - 1)$-subsets of N_{i+1}.

Now the only thing left to do is to make sure that we are able to build a long enough sequence and have the last set $V_{t+s-2k+1}$ of size at least $k-1$. We secure this by choosing the sizes n_i large enough with respect to n_{i+1}. We set $n_{t+s-2k+1} = k - 1$. To make sure that there is a monochromatic set N_{i+1} of size n_{i+1} in an arbitrary $\texttt{red}/\texttt{blue}$-coloring of the $(k - 1)$-subsets of a set of size n_i we recursively choose $n_i = 1 + R^{(k-1)}(n_{i+1}, n_{i+1})$ for every $i = s + t - 2k, \dots, 2, 1, 0$. We can do this because by induction all these $(k-1)$-uniform Ramsey numbers are finite. Consequently we succed in creating the appropriate right-neighborhood-monochromatic sequence, provided $n \geq n_0 + 1$. □

A (very-very) slight improvement of these bounds can be given by induction, where we do not "forget" what color are the right-edges of the newly found element of the sequence. It is also possible to extend the theorem to arbitrary (finite) number of colors.

Homework.

1. *Prove the recursion* $R^{(k)}(t,s) \leq R^{(k-1)}(R^{(k)}(t-1,s), R^{(k)}(t,s-1))$ *and conclude Theorem 1.3.4.*
2. *Define the r-color Ramsey-number* $R_r^{(k)}(t_1,\dots,t_r)$ *and prove that it is finite.*

An immediate question after verifying any kind of finiteness should be "How large is finite?" Let us see. Using the recursion $n_i = R^{(k-1)}(n_{i+1}, n_{i+1}) + 1 =: R^{(k-1)}(n_{i+1}) + 1$ and the final value $n_{s+t-2k+1} = k-1$, we obtain

$$n_0 = R^{(k-1)}(R^{(k-1)}(...(R^{(k-1)}(k-1)+1)...)+1)+1,$$

where the function $R^{(k-1)}$ appears $s+t-2k+1$ times.
For $k=2$ this resolves to $R^{(2)}(s,t) \leq R^{(1)}(R^{(1)}(...(R^{(1)}(1)+1)...)+1)+1 = 2^{s+t-3}$, indicating that the above proof is indeed the generalization of the argument in the proof of Claim 1.1.3.

For $k=3$ we obtain

$$R^{(3)}(s,t) \leq R^{(2)}(R^{(2)}(...(R^{(2)}(2)+1)...)+1)+1 \leq 2^{2\cdot R^{(2)}(...(R^{(2)}(2)+1)...)+1)-1}$$
$$\leq 2^{2^{2\cdot R^{(2)}(...(R^{(2)}(2)+1)...))}-1} \leq \cdots \leq 2^{2^{2^{\cdot^{\cdot^{2}}}}},$$

a tower of height $t+s-3$.

For uniformity $k=4$ things get completely out of hand: we have $t+s-3$ 3-uniform Ramsey functions, each being a tower function, embedded inside one another ... To see a concrete example, let us note that the bound we get for $R^{(4)}(5,5)$, which is our upper bound on the Happy-Ending number $HE(5)=9$, is a tower of 29 twos. For an upper bound on $17 = HE(6) \leq R^{(4)}(6,5)$ we have to take a tower of 2s of the height, which is a tower of 2s of height 29. This might question your intuitive understanding of the word finite...

Before proving the theorem again, let us analyse the proof given above and try to see why the bound became so incredibly large. We built our "right-monochromatic" sequence $v_1, v_2, ..., v_{s+t-2k+1}$ to this particular length so at the end we could use the 1-uniform Ramsey numbers for the sequence. For each new element of the sequence we had to apply the $(k-1)$-uniform Ramsey numbers. Our upper bound on the $(k-1)$-uniform Ramsey numbers are very large, their repeated application forces them to be embedded inside the arguments of the previous one and this causes the bound to become very-very-very-$\cdots$ -very large.

In contrast to this, the 1-uniform Ramsey bound, which we used only once at the end is rather small: only the sum of the two arguments (minus one. In fact this is not only a bound, but the exact value.) Erdős and Rado turned the proof idea on its head and tried to use the 1-uniform Ramsey numbers many times, but in exchange reduce the use of $(k-1)$-Ramsey numbers. They decided to build a much longer sequence by using the 1-dimensional Ramsey numbers in each round, so they have to use the $(k-1)$-uniform Ramsey numbers only once, at the end, for the sequence. This reduces the bound to a function we can actually write down. This function will still be large, but is relatively "close" to the truth.

Second proof of Theorem 1.3.4. Again we will apply induction on the uniformity k. The base case $k=1$ was proved already.

Let $n \in \mathbb{N}$ be chosen large enough with respect to k, t, and s and let us be given an arbitrary two-coloring $c : \binom{[n]}{k} \to \{\texttt{red}, \texttt{blue}\}$ of the k-subsets of $[n]$. Our goal is to find either a monochromatic

t-element subset $T \subseteq [n]$ in $\texttt{red}$ (i.e. $c(Q) = \texttt{red}$ for every $Q \in \binom{T}{k}$) or a monochromatic s-element subset $T \subseteq [n]$ in $\texttt{blue}$ (i.e. $c(Q) = \texttt{blue}$ for every $Q \in \binom{T}{k}$).

Our goal is to build a sequence $v_1, \ldots, v_\ell$ of vertices that is *right-$(k-1)$-neighborhood-monochromatic.* By this we mean that the color of a k-set contained in $\{v_1, \ldots, v_\ell\}$ should only depend on its $k-1$ smallest indices. Formally, by this we mean that there exists a coloring $c^* : \binom{[\ell-1]}{k-1} \to \{\texttt{red}, \texttt{blue}\}$ such that for every k-subset $T \subseteq \{v_1, \ldots, v_\ell\}$, we have $c(T) = c^*(T_{\min}(k-1))$, where we adopted the notation $T_{\min}(k-1) \in \binom{[\ell-1]}{k-1}$ for the set of the $k-1$ smallest indices j of vertices $v_j \in T$.

How long a sequence should we build? We observe that if we manage to build a right-$(k-1)$-neighborhood monochromatic sequence of length $\ell = R^{(k-1)}(t-1, s-1) + 1$, then we are done. The $(k-1)$-subsets of $[\ell-1]$ are colored according to c^*. By the property of the Ramsey number $R^{(k-1)}(t-1, s-1)$, we find an index subset $I \subseteq [\ell-1]$ of size $t-1$ which is monochromatic $\texttt{red}$ under c^* or an index subset $I \subseteq [\ell-1]$ of size $s-1$ which is monochromatic $\texttt{blue}$ under c^*. Then we claim that the subset $T := \{v_i : i \in I\} \cup \{v_\ell\}$ is c-monochromatic in the same color as I is c^*-monochromatic in. And then of course it also has the required size (t in case of color $\texttt{red}$ and s in case of color $\texttt{blue}$). To see this, let us take an arbitrary k-element subset $Q \subseteq T$. By the right-$(k-1)$-neighborhood monochromatic property of the coloring c^* we have that $c(Q) = c^*(Q_{\min}(k-1))$, where $Q_{\min}(k-1) \subseteq I$ is a $(k-1)$-element subset of the c^*-monochromatic index subset I.

So to complete the proof of our theorem we need to construct the right-$(k-1)$-neighborhood monochromatic sequence of the required length. Our plan is to construct a sequence $v_1, \ldots v_i$ recursively. We will maintain a set N_i of vertices that are still "eligible" to be added to the sequence, that is, the c-color of any k-subset of vertices of $V_i := \{v_1, \ldots v_i\} \cup N_i$ with at least $k-1$ vertices among $v_1, \ldots, v_i$ indeed only depends on the $k-1$ smallest indices. We will pick the next vertex v_{i+1} arbitrarily from N_i and then reduce N_i to create N_{i+1} in order for the right-$(k-1)$-neighborhood monochromatic property also to hold for k-subsets involving the new vertex v_{i+1}.

To start let us select arbitrary vertices $v_1, \ldots v_{k-2} \in [n]$ and set $N_{k-2} = [n] \setminus \{v_1, \ldots, v_{k-2}\}$. Suppose that $i \geq k-2$ and we are given a sequence $v_1, \ldots, v_i$ and a set N_i disjoint from it, such that there exists a coloring $c^* : \binom{[i]}{k-1} \to \{\texttt{red}, \texttt{blue}\}$ with the property that for every k-subset $T \subseteq \{v_1, \ldots, v_i\} \cup N_i$ with $|T \cap \{v_1, \ldots, v_i\}| \geq k-1$, we have $c(T) = c^*(T_{\min}(k-1))$. Note that such a sequence $v_1, \ldots, v_i$ is always right-$(k-1)$-neighborhood monochromatic and that our initial choices vacuously satisfy the condition.

We choose the next vertex $v_{i+1} \in N_i$ arbitrarily. In order to designate $N_{i+1} \subseteq N_i \setminus \{v_{i+1}\}$, we define a function $\tilde{c} : N_i \setminus \{v_{i+1}\} \to \{\texttt{red}, \texttt{blue}\}^{\binom{i}{k-2}}$, such that the components of $\tilde{c}(w)$ for a vertex $w \in N_i \setminus \{v_{i+1}\}$ correspond to the $(k-2)$-subsets of $[i]$, and for the component corresponding to a subset $L \in \binom{[i]}{k-2}$ we have

$$\tilde{c}(w)_L := c(L \cup \{v_{i+1}, w\}).$$

This function is so defined that if we chose N_{i+1} to be the $\tilde{c}$-inverse image of any $\texttt{red}/\texttt{blue}$-vector, then we ensure that the desired property of the sequence $v_1, \ldots, v_i$ and the function c^* extends to the $(k-1)$-element index subsets containing $i+1$. Indeed, choosing N_{i+1} to be the inverse image of the fixed $\texttt{red}/\texttt{blue}$-vector $\alpha \in \{\texttt{red}, \texttt{blue}\}^{\binom{i}{k-2}}$, we can extend c^* to $\binom{[i+1]}{k-1}$ as follows. The function c^* is already defined for sets in $\binom{[i]}{k-1}$, now choose an index set $I \in \binom{[i+1]}{k-1}$ that contains $i+1$ and define $c^*(I) := \alpha_{I \setminus \{i+1\}}$. To check that this definition is in line with what is desired from c^*, let us take any k-subset $Q \subseteq \{v_1, \ldots, v_{i+1}\} \cup N_{i+1}$, with $v_{i+1} \in Q$ and having $|Q \cap \{v_1, \ldots v_{i+1}\}| = k-1$. Then Q is of the form $Q = \{v_j : j \in J\} \cup \{v_{i+1}, w\}$, where $J \subseteq \binom{[i]}{k-2}$ and $w \in N_{i+1}$. In particular we have $Q_{\min}(k-1) = J \cup \{i+1\}$. By definition of $\tilde{c}$, and c^*, and since $w \in \tilde{c}^{-1}(\alpha)$, we obtain

$$c(Q) = \tilde{c}(w)_J = \alpha_J = c^*(J \cup \{i+1\}) = c^*(Q_{\min}(k-1)),$$

verifying the desired property of c^*.

To make N_{i+1} large we choose the largest possible $\tilde{c}$-inverse image. Hence N_{i+1} can be chosen so its size is at least the average size of an inverse image, that is

$$|N_{i+1}| \geq \left\lceil \frac{|N_i| - 1}{2^{\binom{i}{k-2}}} \right\rceil, \tag{1.2}$$

for every $i \geq k - 2$.

Now let us estimate the size of n required in this proof. To make the sequence long enough, that is, to be able to choose the ℓth element of the sequence, we need that the set $N_{\ell-1}$ has at least one element. To ensure that $|N_{\ell-1}| \geq 1$ we use (1.2) repeatedly and choose $n = |N_{k-2}| + k - 2$ large enough. Namely, if $|N_{k-2}| \geq 2^{\binom{\ell}{k-1}}$ then we can choose the subsets $N_{k-2} \supseteq N_{k-1} \supseteq \cdots \supseteq N_{\ell-1}$, such that

$$2^{\binom{\ell-1}{k-1}} \leq |N_{k-2}| \leq 2^{\binom{k-2}{k-2}}|N_{k-1}| \leq 2^{\binom{k-2}{k-2}} \cdot 2^{\binom{k-1}{k-2}}|N_k| \leq \cdots \leq \prod_{j=k-2}^{\ell-2} 2^{\binom{j}{k-2}}|N_{\ell-1}| =$$

$$= 2^{\sum_{j=k-2}^{\ell-2} \binom{j}{k-2}} \cdot |N_{\ell-1}| = 2^{\binom{\ell-1}{k-1}} \cdot |N_{\ell-1}|,$$

implying that $N_{\ell-1}$ is not empty and v_ℓ can be chosen.

In conclusion, the choice

$$n = 2^{\binom{\ell-1}{k-1}} + k - 2 = O\left(2^{\ell^{k-1}}\right)$$

with

$$\ell = R^{(k-1)}(t, s)$$

is sufficiently large for our selection process to go through and thus provides an upper bound on the Ramsey number $R^{(k)}(t, s)$. □

By repeated application of this theorem we obtain a greatly improved upper bound compared to the first proof. We highlight this here by explicitly writing out the bound for the symmetric case $t = s$.

Corollary 1.3.5. *$R^{(k)}(t, t)$ is upper bounded by a tower function of t of height k.*

This upper bound is actually not that far from the truth. There is a construction of a `red`/`blue`-coloring of the k-sets without a monochromatic t-clique on a vertex set of size that is a tower function of t of height $k - 1$. To decide which height is the truth, even just for the 3-uniform Ramsey function, is worth a $500 reward (by Erdős).

1.3.3 The Canonical Ramsey Theorem

In this section, we temporarily abandon our pursuit of the various bounds in "quantitative" Ramsey theory and return to the "qualitative" philosophical orgins of "complete disorder is impossible". The fundamental question of Ramsey theory is: given a classification (i.e., a coloring) of the elements of some structure, what sort of "order" can one necessarily find in it? We have seen many examples where a structure is colored with an arbitrary finite number of colors and we concluded the existence of a large "orderly" substructure (where by "orderly" we meant a substructure that is monochromatic).

An instance of this was the infinite Ramsey theorem (Theorem 1.1.6). This can be generalized for hypergraphs and arbitrary finite number of colors.

Theorem 1.3.6 (HW). *For any positive integer r and any r-colouring of $\binom{\mathbb{N}}{k}$, there exists an infinite set $S \subseteq \mathbb{N}$ for which $\binom{S}{k}$ is monochromatic.*

In this subsection we will ask ourselves whether complete disorder would still be impossible if we colored our structire by *infinitely* many colors. We immediately realize that we must revise our notion of "orderly" substruture, as coloring each pair in $\binom{\mathbb{N}}{2}$ by a different color will not even leave us a monochromatic subset of size three![5]

In light of this example it seems necessary to include the situation when *all* pairs of elements of a set have distinct colors among orderly structures. This motivates the following definition.

Definition 1.3.7. *Given a colouring $c : \binom{\mathbb{N}}{2} \to C$, a set $S \subseteq \mathbb{N}$ is called* c-rainbow *if no two pairs of S have the same color.*

In the above coloring example the whole $\mathbb{N}$ is a rainbow set. Is this enough for the concept of "orderly"? Are we always guaranteed to find either an infinite monochromatic set or an infinite rainbow set? The answer is still no. To see this, simply colour each pair $\{i, j\}$ with its minimal element $\min\{i, j\}$. In this coloring we still do not find a monochromatic set of size three, but neither find a rainbow set of size three. This example motivates the following definition.

Definition 1.3.8. *Given a colouring $c : \binom{\mathbb{N}}{2} \to C$, a set $S \subseteq \mathbb{N}$ is called* c-left-injective *if there is an injective map $c^* : \mathbb{N} \to C$, such that $c(ij) = c^*(\min\{i, j\})$.*

The name *left-injective* subset originates in its property that the colour of an edge is uniquely determined by its *left* endpoint. Note that the a left-injective subset forms a right-monochromatic sequence (from the last subsections).

Of course there is nothing special about the minimum, we could also define a coloring of $\binom{\mathbb{N}}{2}$ by setting the color of every edge to be the maximum of its endpoints. Then there is no monochromatic, no rainbox, and no right-injective set of size three. Hence analogously we define the notion of *right-injective* colouring.

Definition 1.3.9. *Given a colouring $c : \binom{\mathbb{N}}{2} \to C$, a set $S \subseteq \mathbb{N}$ is called* c-right-injective *if there is an injective map $c^* : \mathbb{N} \to C$, such that $c(ij) = c^*(\max\{i, j\})$.*

Surprisingly, it is not only necessary but also sufficient that we extend our notion of "orderly" subset to include these four cases: one of them will occur! This is stated in the next Canonical Ramsey Theorem.

Theorem 1.3.10 (Erdős-Rado, 1950). *Let $c : \binom{\mathbb{N}}{2} \to C$ be a coloring. Then there is some infinite set $S \subseteq \mathbb{N}$ such that either*

(i) S is c-monochromatic, or

(ii) S is c-left-injective, or

(iii) S is c-right-injective, or

(iv) S is c-rainbow.

Remarks

- This is a strengthening of Ramsey's Theorem for finitely many colors from the Homework (Theorem 1.3.6). Indeed, when the number of colors used is finite, then options (*ii*)-(*iv*) are impossible.)
- The colorings appearing on these four types of sets are called the *canonical* colorings. In case (*iv*) the color of an edge is determined injectively by both enpoints, in case (*ii*) it is determined by the left endpoint, in case (*iii*) it is determined by the right endpoint, and in case (*i*) it is just determined (by no endpoint). The theorem states that every colouring contains an infinite canonically coloured clique.

[5] For notational simplicity we restrict ourselves further to the 2-uniform case; analogous results hold for arbitrary uniformity k.

Proof. The idea of the proof, just like in the approach to the Happy-Ending Problem, is to try to use local information to deduce something for the global structure. Since we will need to compare the colors on *pairs* of edges, we should be interested in the coloring of 4-element subsets. One of the main questions is how to reduce the number of colors to finite, so that we are able to use the 4-uniform Ramsey Theorem. To this end we will colour the 4-subsets of $\mathbb{N}$ such that we encode the information about the color pattern on the edges between these four integers. This coloring will use only finitely many colors since we will only be interested in the colour pattern i.e. keeping track of which edges have the same colour and which do not, but we will **not** care exactly which particular colors we use to create this pattern. It might sound surprising at first that this information, the color pattern on 4-element sets, is sufficient to deduce the existence of an infinite set with a canonical coloring.

Formally, let us define a coloring $\hat{c} : \binom{\mathbb{N}}{4} \to \mathcal{B}\left(\binom{[4]}{2}\right)$ where $\mathcal{B}\left(\binom{[4]}{2}\right)$ is the set of all set partitions[6] of the six-element set $\binom{[4]}{2}$. The value $\hat{c}(\{i_1, i_2, i_3, i_4\})$ for some 4-subset $\{i_1, i_2, i_3, i_4\}$ with $i_1 < i_2 < i_3 < i_4$ is just the set partition that is induced by the inverse images of c on the set $\binom{\{i_1,i_2,i_3,i_4\}}{2}$ and hence in turn on the set $\binom{[4]}{2}$.

For example, the value of $\hat{c}$,

- for a rainbow 4-set is $\{\{12\}, \{13\}, \{14\}, \{23\}, \{24\}, \{34\}\}$,
- for a monochromatic 4-set is $\{\{12, 13, 14, 23, 24, 34\}\}$,
- for a left-injective subset is $\{\{12, 13, 14\}, \{23, 24\}, \{34\}\}$,
- for a right-injective subset is $\{\{12\}, \{13, 23\}, \{14, 24, 34\}\}$.

We use the 4-uniform Ramsey Theorem for 203 colors (Theorem 1.3.6) and find an infinite $\hat{c}$-monochromatic subset $S = \{s_1 < s_2 < \cdots < s_i < \cdots\} \subseteq \mathbb{N}$. In other words there is a set partition $\mathbf{p} \in \mathcal{B}\left(\binom{[4]}{2}\right)$ such that for every 4-subset $T = \{i_1 < i_2 < i_3 < i_4\} \subseteq S$ we have $c(i_u i_v) = c(i_w i_z)$ for some $uv, wz \in \binom{[4]}{2}$ if and only if uv and wz are in the same class of the set partition $\mathbf{p}$. Now we have a little case distinction based on how $\mathbf{p}$ looks like.

Case 1. $\mathbf{p} = \{\{12\}, \{13\}, \{14\}, \{23\}, \{24\}, \{34\}\}$. In this case the whole S is rainbow. Indeed, for any two edges $s_1 s_2$ and $s_3 s_4$, there exists a 4-element subset T containing both of these edges. Since $\hat{c}(T) = \mathbf{p}$, all edges, in particular also $s_1 s_2$ and $s_3 s_4$ have distinct c-colors.

Case 2. There is a partition class of $\mathbf{p}$ with at least two pairs.

We divide further into subcases, depending on whether there is a class containing a disjoint pair.

Case 2a. There are two pairs that are disjoint and are contained in the same partition class of $\mathbf{p}$.

If 12 and 34 are in the same class then S is an infinite c-monochromatic subset, as the color of any two edges $s_x s_y$ and $s_u s_v$ are equal since they are both equal to the color of the edge $s_a s_{a+1}$, say with $a = \max\{x, y, u, v\} + 1$.

If 14 and 23 are in the same class then $S \setminus \{s_1\}$ is an infinite c-monochromatic subset, since the color of any two edges $s_x s_y$ and $s_u s_v$, with $x, y, u, v \geq 2$, are both equal to that of $s_1 s_a$, where $a = \max\{x, y, u, v\} + 1$.

If 13 and 24 are in the same class, then $S_{\mathsf{even}} = \{s_{2i} : i \in \mathbb{N}\}$ is an infinite c-monochromatic subset. Indeed, suppose we have two pairs $s_x s_y$ and $s_u s_v$ with $x, y, u, v \in 2\mathbb{N}$ and, without loss of generality, $x < y$, $u < v$ and $x < u$. Then, if $u < y < v$, these two edges must be the same color. If $y < u$, then both edges are the same color as the edge $s_{x+1} s_{u+1}$. Finally, if $v < y$, then both edges are the same color as $s_{u+1} s_{y+1}$.

[6]The cardinality of $\mathcal{B}\left(\binom{[4]}{2}\right)$ is the sixth Bell number $B_6 = 203$, which is the sum of the Stirling numbers $S(6, k)$of the second kind with the summation running till $k = 6$.

Case 2b. Every two pairs that are disjoint are in different partition classes of $\mathbf{p}$.

Consequently there must be two pairs xz and $yz \in \binom{[4]}{2}$ that are in the same partition class of $\mathbf{p}$. Without loss of generality $x < y$. Using a similar argument as before, one can show the following:

If $x < z < y$, then $S_{\mathtt{even}}$ is c-monochromatic.

If $z < y < x$, then $S_{\mathtt{even}}$ is c-left-injective.

If $y < x < z$, then $S_{\mathtt{even}}$ is c-right-injective. □

Remark. The Canonical Ramsey Theorem extends to the k-uniform setting. There we will have to admit 2^k canonical colourings.