Delft University of Technology

Ramsey's theorem

Björn Titulaer

August 31, 2020

Contents

1 Abstract

In this report we will take a look at various proofs of Ramsey's theorem, some of the bounds that result from those proofs and applications of Ramsey's theorem. We will consider the proof of Ramsey himself, the proof of Skolem, the proof given by Erdős and Szekeres and the proof of again Erdős and Rado. The best upper bound for higher order Ramsey numbers is obtained by following the proof of Erdős and Rado and this bound has only marginally improved since then. We will also state and prove the lower bound given by Erdős and Hajnal. In the end we will apply Ramsey's theorem to both the Happy Ending problem and the monotone subsequence problem. The bounds we get for both problems using Ramsey's theorem, however, are quite weak compared to bounds that do not use Ramsey's theorem, so the theorem is more useful in proving the existence of a certain number than it is to find strong bounds for it.

2 Introduction

The Pigeonhole Principle is a simple statement that says that if we put a large number of pigeons in a finite number of holes, then one of those holes is going to end up with a large number of pigeons inside. In 1928 Frank Plumpton Ramsey published an article in which he proved an extension of the pigeonhole Principle which is know known as Ramsey's theorem.

In this report we will first explain and prove Ramsey's theorem. We will start with the simplest form of the theorem and extend it from there.

After we have established Ramsey's theorem we will look at various different proofs and compare the upper bounds found by those proofs. We stop at the bound found by Erdős and Rado because from then the upper bound has only marginally improved. We will also prove and discuss two lower bounds.

Lastly we will look at two applications of Ramsey's theorem, both proven by again Erdős and Szekeres with the help of Klein. Those two applications are the monotone subsequence problem, which asks for the smallest possible size of a sequence of real numbers such that it has to contain a monotone subsequence of a certain length, and the Happy Ending problem. This problem is about how many points we have to put in a plane to guarantee the existence of a convex n-gon. Both problems will be solved multiple times, at least once using Ramsey's theorem and at least once without using Ramsey's theorem.

3 Ramsey's theorem

Ramsey's theorem exists in many variants, both finite and infinite. In this section we will focus mostly on one of the finite variants, which we will slowly expand. The variant of Ramsey's theorem that we will first consider is about the coloring of a graph.

Suppose we have a complete graph $G = (V, E)$. We are interested in specific smaller parts of the graph, so-called subgraphs. A subgraph $S \subseteq G$ is a graph $S = (V', E')$ such that $V' \subseteq V$ and $E' = \{(x, y) \in E : x, y \in V'\}$. In essence we take a subset of the set of our original vertices and only consider the edges between those vertices. Now we will color all the edges of our graph G in two colors: red and blue. While the complete graph is arbitrarily colored, there might be specific parts of the graph that are ordered more neatly than the complete graph. In particular, we might find a subgraph that only contains edges of one color. We will call this monochromatic.

Right now one could ask himself how large of a monochromatic subgraph we could find if we started with n vertices. Ramsey looked at this problem and he found the following result:

Theorem 1. *For all integers k and l there exists a least integer $R(k, l)$ such that whenever $n \geq R(k, l)$, any red-blue coloring of a complete graph on n vertices has either a blue monochromatic subgraph A of size k or a red monochromatic subgraph B of size l.*

The least integer $R(k, l)$ is often called a Ramsey number. Only a few non-trivial Ramsey numbers are known. Even though Ramsey proved his theorem in his original paper himself, we will look here at the proof given by Erdős and Szekeres [2], both because it gives better bounds for Ramsey numbers and because it is easier to follow.

Proof. We will prove the theorem by induction on k and l.

For the base case, it is not hard to see that $R(2, k) = R(k, 2) = k$.

So now assume that for a certain k and l both $R(k-1, l)$ and $R(k, l-1)$ exist. We will show that $R(k, l) \leq R(k-1, l)+R(k, l-1)$. Suppose we have a complete graph $G = (V, E)$ with $|V| = R(k-1, l)+R(k, l-1)$. Take any vertex x. Define $S = \{y \in V : \{x, y\} \text{ is blue}\}$ and $T = \{y \in V : \{x, y\} \text{ is red}\}$.

Every vertex, except for x, is now either in S or in T. This means that S and T together have $R(k-1, l) + R(k, l-1) - 1$ vertices. Therefore, either $|S| \geq R(k-1, l)$ or $|T| \geq R(k, l-1)$.

Suppose $|S| \geq R(k-1, l)$. We then either obtain a blue monochromatic subset A of S of size $k-1$ or a red monochromatic subset B of S of size l. In the first case we have that x is not in A, so from the definition of our set S we see that adding x to the vertices of A results in a blue monochromatic subgraph of size k. If we had a red monochromatic subgraph B of size l instead we would be done immediately.

If we had $|T| \geq R(k, l-1)$ instead, the same argument holds.

In both cases the inequality holds, which proves the theorem by induction. □

It is not hard to show that $R(k, l) \leq \binom{k+l-2}{k-1}$ using induction. For $R(k, 2)$ and $R(2, k)$ it holds, so suppose it holds for $R(k-1, l)$ and $R(k, l-1)$. Then from the inequality of the proof we find that

$$R(k, l) \leq R(k-1, l) + R(k, l-1) \leq \binom{k+l-3}{k-2} + \binom{k+l-3}{k-1} = \binom{k+l-2}{k-1}. \quad (3.1)$$

In the theorem we only allowed two colors: red and blue. A very similar theorem exists for an arbitrary number of colors. We will have to introduce some new notation for this.

When the number of colors matters, we will talk about an r-coloring c to indicate that the coloring uses r colors. Furthermore, we will refer to colors as numbers from now on. So an r-coloring assigns a natural number smaller or equal to r to every edge. In other words, an r-coloring on a graph $G = (V, E)$ is a function $c : E \to \{1, \ldots, r\}$.

Instead of talking about a red monochromatic graph if all its edges are red, we will now talk about an i-monochromatic graph if all its edges are of color i. If we allow multiple colors, we can restate Ramsey's theorem as follows:

Theorem 2. *For all integers r and $l_1, \ldots, l_r$ there exists a least integer $R(l_1, \ldots, l_r)$ such that for any r-coloring c there is an i-monochromatic subgraph with l_i vertices for some i.*

Proof. Just like the 2-color case, the proof uses induction.

For the base case, note that $R(2, \ldots, 2, k, 2, \ldots, 2) = k$, no matter where we put the k.

Now assume that for all i, $R(l_1, \ldots, l_i - 1, \ldots, l_r)$ exists. We will show that

$$R(l_1, \ldots, l_r) \leq 2 + \sum_{i=1}^{r} (R(l_1, \ldots, l_i - 1, \ldots, l_r) - 1).$$

Note that in the case $r = 2$ this gives the same inequality as obtained before. Let $n = 2 + \sum_{i=1}^{r} (R(l_1, \ldots, l_i - 1, \ldots, l_r) - 1)$.

Let c be any r-coloring on a graph $G = (V, E)$ with n vertices and take any vertex $x \in V$ arbitrarily. If we denote the edge between two vertices x and y as $\{x, y\}$ we can define $A_i = \{y \in V : c(\{x, y\}) = i\}$. We must have that $|A_i| \geq R(l_1, \ldots, l_i - 1, \ldots, l_r)$ for some i, though this is no longer trivial.

To see this, assume the contrary, so $|A_i| < R(l_1, \ldots, l_i - 1, \ldots, l_r)$ for all i. Since $|A_i|$ is integer, we find that $|A_i| \leq R(l_1, \ldots, l_i - 1, \ldots, l_r) - 1$. Every element of V is in exactly one A_i, except for x itself. So

$$\sum_{i=1}^{r} |A_i| = n - 1 = 1 + \sum_{i=1}^{r} (R(l_1, \ldots, l_i - 1, \ldots, l_r) - 1).$$

However using our assumption we find

$$\sum_{i=1}^{n} |A_i| \leq \sum_{i=1}^{n} (R(l_1, \ldots, l_i - 1, \ldots, l_r) - 1),$$

which is a contradiction.

So take one j such that $|A_j| \geq R(l_1, \ldots, l_j - 1, \ldots, l_r)$. Let S be the subgraph of G with A_j as its vertex set. Because of the definition of the Ramsey number we either have for some i an i-monochromatic subgraph of S of size l_i, $1 \leq i \neq j \leq r$, or a j-monochromatic subgraph T of size $l_j - 1$.

In the first case, we are done, so assume we can find a j-monochromatic subgraph T of size $l_j - 1$. Just like in the 2-color case, because of how we chose A_j, adding x to T results in a j-monochromatic subgraph of size l_j. By induction the theorem follows. □

We could have proven the multicolored Ramsey theorem differently, using induction on the number of colors. That proof however gives a much weaker bound on the Ramsey numbers. The alternative proof is given below.

Alternative proof of Theorem 2. As stated before, we will use induction on the number of colors.

The base case would be Ramsey's theorem for two colors, which we proved earlier.

Now assume that for all $r' < r$ and all integers $l_1, \ldots, l_{r'}$ the number $R(l_1, \ldots, l_{r'})$ exists. Let $p_1, \ldots, p_r$ be integers. We will show that $R(p_1, \ldots, p_r)$ exists as well. Let G be a complete graph on $R(p_1, \ldots, p_{r-2}, R(p_{r-1}, p_r))$ vertices. This number exists because the Ramsey number only has $r - 1$ colors. Let c be any r-coloring on G.

We will define a new coloring c' with $r - 1$ colors by 'merging' the last two colors of c. More formally, if e is an edge, $c'(e) = c(e)$ if $c(e) \neq r$ and $c'(e) = r - 1$ if $c(e) = r$. Now we either have an i-monochromatic subgraph S on c' of size p_i, $i \leq r - 2$ or we have a $(r - 1)$-monochromatic subgraph T on c' of size $R(p_{r-1}, p_r)$. In the first case we are done since on that subset $c' = c$, so assume the second part holds.

All edges of T are either of color $r - 1$ or of color r under c, so c restricted to S only has 2 colors. Then it follows from the definition of $R(p_{r-1}, p_r)$ that we either have a $(r - 1)$-monochromatic subset of size p_{r-1} or an r-monochromatic subset of size p_r.

So we find that $R(p_1, \ldots, p_r)$ exists, which completes the induction. □

We are now in a position to generalize Ramsey's theorem even further. In essence a coloring is a function that maps every edge, which in the case of a complete graph is just a subset of any two vertices, to a 'color', or in our case, a number. It would also be possible to take subsets of more than two vertices and assign a number to those.

In this case, it makes little sense to talk about traditional graphs, since an edge is always between two vertices. Instead, we can look at the set $\{1, 2, \ldots, n\}$. A traditional r-coloring then becomes a function that maps all subsets of 2 elements of $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, r\}$. It is easy to extend this to allow an arbitrary number of elements, so we define an r-coloring on subsets of k elements from $\{1, 2, \ldots, n\}$ as a function that maps all subsets of $\{1, 2, \ldots, n\}$ of k elements to $\{1, 2, \ldots, r\}$.

This is closely related to a hypergraph, which is a vertex set V combined with an 'edge' set E, where every element of E is a subset of V. If we would number our vertices between 1 and n and only take the edges with k elements we would be able to apply our extended definition of a coloring to graphs as well, but we will not do so in this report and we will stick to the previous notation.

We will use both the set $\{1, 2, \ldots, n\}$ and its subsets of k elements a lot, so we will shorten $\{1, 2, \ldots, n\}$ to $[n]$ and denote the family of all its subsets of k elements by $[n]^k$. For colorings, to prevent confusion we will denote with a subscript on how many elements the coloring takes place, so c_k always denotes a coloring on subsets of k elements. We will either say that c_k is a coloring of $[n]^k$ or, to emphasize the number of elements of the coloring, we will say that c_k is a coloring of $[n]$ on k elements.

With this extended definition of a coloring we will now restate and prove Ramsey's theorem one last time.

Theorem 3 (Ramsey's theorem)**.** *For all integers k, r and $l_1, \ldots, l_r$ there exists a least integer $R_k(l_1, \ldots, l_r)$ such that, whenever $n \geq R_k(l_1, \ldots, l_r)$, for any arbitrary r-coloring* *of* $[n]$ *on* k *elements there exist an* i *and a set* $S \subseteq [n]$ *such that* S *is* i*-monochromatic with* $|S| = l_i$*.*

Proof. The proof works slightly different if we allow colorings on more than 2 elements. We will prove the theorem by using induction on the number of elements k of the coloring.

For the base case, we take $k = 2$, which we proved earlier.

Assume that for a certain k and for all integers r and $l_1, \ldots, l_r$ the number $R_k(l_1, \ldots, l_r)$ exists. To show that for all r and $l_1, \ldots, l_r$ the number $R_{k+1}(l_1, \ldots, l_r)$ exists as well, we will use another induction argument.

It is once again easy to see that $R_{k+1}(k+1, k+1, \ldots, l, \ldots, k+1) = l$, no matter where we put the l.

Now assume that, apart from the assumption made earlier, $R_{k+1}(l_1, \ldots, l_i - 1, \ldots, l_r)$ exists as well for all i. We now claim that

$$R_{k+1}(l_1, \ldots, l_r) \leq R_k(R_{k+1}(l_1-1, l_2, \ldots, l_r), R_{k+1}(l_1, l_2-1, \ldots, l_r), \ldots, R_{k+1}(l_1, l_2, \ldots, l_r-1))+1.$$

The right hand side exists because we assumed that R_k always exists. Suppose we have n equal to the right hand side above. Let c_{k+1} be an arbitrary coloring of $[n]^{k+1}$.

Take any element s of $[n]$. Define a new coloring d_k by

$$d_k(x_1, \ldots, x_k) = c_{k+1}(s, x_1, \ldots, x_k).$$

Note that $[n] \setminus \{s\}$ has $R_k(R_{k+1}(l_1 - 1, l_2, \ldots, l_r), \ldots, R_{k+1}(l_1, \ldots, l_r - 1))$ elements. So we know that for some i there exists a subset T of $[n] \setminus \{s\}$ such that T is i-monochromatic on d_k and $|T| = R_{k+1}(l_1, \ldots, l_i - 1, \ldots, l_r)$.

Going back to c_{k+1}, we know that T either has a j-monochromatic subset on c_{k+1} of size l_j, for some $1 \leq j \neq i \leq r$, or an i-monochromatic subset on c_{k+1} of size $l_i - 1$.

In the former case we are done immediately, so assume the latter is true. Call the subset A. Note that $A \subseteq T$, so we immediately find that $A \cup \{s\}$ is an i-monochromatic subset of size $l_i - 1 + 1 = l_i$.

Now using the second induction, for all $l_1, \ldots, l_r$ we have proven that $R_{k+1}(l_1, \ldots, l_r)$ exists.

The theorem now follows by the first induction. □

It is worth noting that $k = 1$ can be solved directly. In this case we would be coloring single elements. It is not hard to see then that $R_1(l_1, \ldots, l_r) = \sum_{i=1}^{r}(l_i - 1) + 1$. This is a more general form of the pigeonhole principle. Here we want to know how many pigeons we need in order to have one of the r numbered holes contain at least a certain number of pigeons, in this case our l_i. We could have taken this as our induction base as well.

4 Bounds for Ramsey numbers

As stated before, the exact Ramsey numbers are still unknown with the exception of the trivial cases and a handful of small numbers. In this section we will look at upper and lower bounds for the Ramsey numbers. We have already seen an upper bound for the Ramsey numbers for pairs in Equation 3.1.

We will look at three other proofs that each give their own bounds, namely the proof of Ramsey himself (1928), the proof of Skolem (1933) and the proof given by Erdős and Rado (1952). The first two proofs give upper bounds that are way larger than the bound we already found, but they are still interesting from a historical point of view. The last proof was found much later than all the other proofs, but it results in significantly better upper bounds when we consider Ramsey numbers with colorings on more than 2 elements.

Lastly we will prove a lower bound to get an indication how close our upper bound is to the actual Ramsey number.

4.1 Ramsey's proof

Ramsey himself used a different proof than the proof of Erdős and Szekeres [4]. Although this proof does not give the best upper bound it is still of importance because it was the first proof of the theorem. Ramsey did not prove his theorem directly. Instead he stated a different, equivalent problem and proved that one instead. He only considered the symmetric case.

Theorem 4 (An equivalent Ramsey theorem)**.** *For all integers k, l and h there exists a least integer $m_k(l, h)$ such that, whenever $n \geq m_k(l, h)$, any arbitrary 2-coloring c_k of $[n]^k$ is such that there exist two subsets S and T with empty intersection such that $|S| = l$ and $|T| = h$ and all k-element subsets of $S \cup T$ with at least one element from S have the same color.*

It is not hard to see that this is indeed equivalent to Ramsey's theorem. We have that $m_k(l, h)$ has to be larger than $R_k(l, l)$, since $m_k(l, 0) = R_k(l, l)$ and it is increasing if we would consider it as a function of h. Similarly we have that $m_k(l, h) < R_k(l+h, l+h)$, because if we can find a subset of size $l + h$ of which all k-element subsets are of the same color then we can arbitrarily divide those elements over two sets S and T with the desired cardinality. Since any k-element subset of $S \cup T$ is of the same color, certainly all k-element subsets containing at least one element of S would be of the same color. Summarized we have that

$$R_k(l, l) \leq m_k(l, h) \leq R_k(l + h, l + h). \tag{4.1}$$

From this equation it follows immediately that both theorems are indeed equivalent.

Proof. The proof uses an induction argument on k.

The base case is $k = 1$. One can easily verify that $m_1(l, h) = \max(2l - 1, l + h)$.

So, assume the theorem holds up to a certain k. First, we will show that $m_k(1, h)$ exists and after we will use induction again on l. We claim that $m_k(1, h) = 1 + m_{k-1}(h, 0)$ works.

Indeed, suppose $n = 1+m_{k-1}(h,0)$ and let c_k be a 2-coloring of $[n]^k$. Take any element x of $[n]$. Now define a new coloring c'_{k-1} on $k-1$ subsets $\{i_1,\dots,i_{k-1}\}$ of $[n]\setminus\{x$ by$\}$ $c'_{k-1}(i_1,\dots,i_{k-1}) = c_k(i_1,\dots,i_{k-1},x)$. It follows that c'_{k-1} only has two colors as well.

Since $[n]\setminus\{x\}$ has $m_{k-1}(h,0)$ elements, we can find a subset A with $|A| = h$ that is monochromatic under c'_{k-1}. By taking $S = \{x\}$ and $T = A$, the claim follows.

Now assume that, still for the same k as before, the theorem holds up to a certain l, so $m_k(i,h)$ exists for all $i < l$. If we now define $F(x) = m_k(1,x)$, we will prove that $m_k(l,h) = m_k(l-1, F^l(\max(k-1,h))$ works. Here, $F^n(x) = F(F(F(\cdots(x)\cdots)$, so we apply F n times to x. In particular this means by the definition of F that any 2-coloring on $F^n(x)$ elements has an element s and a subset T with $|T| = F^{n-1}(x)$ such that all k-subsets of $\{s\}\cup T$ that contain s are monochromatic under this coloring. This turns out to be a useful property for the proof.

Let c_k be any 2-coloring of $[n]^k$ with $n = m_k(l-1, F^l(\max(k-1,h))$. Because of the definition of n we can find a subset S with $|S| = l-1$, a subset T_0 with $|T_0| = F^l(\max(k-1,h))$ and a number z such that any k-subset of $S\cup T_0$ containing at least one element of S is of color z under c_k. We will assume that $z = 1$, though if $z = 2$ the same argument holds.

Because of how F was defined we can find an element t_1 of T_0 and a set $T_1 \subseteq T_0$ with $|T_1| = F^{l-1}(\max(k-1,h))$ such that all k-subsets of $\{t_1\}\cup T_1$ containing t_1 are monochromatic. If this color is 1, then we could add t_1 to S to find the desired set S with any subset of size h of T_1, which is possible since $|T_1|$ is much larger than h. So assume the color is 2.

Once again there exists an element t_2 and a set T_2 with $|T_2| = F^{l-2}(\max(k-1,h))$ such that any k-element subset of $\{t_2\}\cup T_2$ is of the same color. Following the same argumentation as before, if this color is 1, we add t_2 to S and we are done in a similar way, so the color has to be 2.

We can do this l times in total to obtain elements $t_1,\dots,t_l$ and sets $T_1,\dots,T_l$ with $|T_l| = \max(h,k-1)$ such that for all $1 \le i \le l$, we must have that all k-subsets of $\{t_i\}\cup T_i$ that contain t_i are of color 2. $|T_l| \ge h$, so let T be any h-element subset of T_l. We claim that $S = \{t_1,\dots,t_l\}$ and T are such that all k-subsets of $S\cup T$ containing at least one elements of S are of color 2.

Indeed, let $x_1,\dots,x_k$ be any k-subset of $S\cup T$ containing at least one t_i for some i. Let j be the smallest integer such that $t_j = x_p$ for some $1 \le p \le k$. Now all the other elements of $x_1,\dots,x_k$ must be in T_j, since if x_i was from S, it would be equal to t_v for some $v > j$ and thus be in $T_{v-1} \subseteq T_j$.

If instead x_i was from T, x_i is in T_l and since $T_l \subseteq T_j$ we can conclude that $x_i \in T_j$. Now from the definition of T_j we immediately find that $x_1,\dots,x_k$, which contains t_j, is of color 2.

The theorem now follows by induction. $\square$

The upper bounds from this proof are very large. Note that

$$R_k(l,l) = m_k(l,0) = m_k(l-1,1) = \cdots = m_k(l-k+1,l-1),$$

since all k-subsets automatically include an element from the first set, since the second set has less than k elements. The proof gives the best bound for $m_l(l-k+1,l-1)$. From the proof we have that

$$R_k(l,l) = m_k(l-k+1, k-1) \leq m_k(l-1, F^l(k-1)). \tag{4.2}$$

It is not hard to see from the definition of F that $F(x) > x$, so we can conclude that $F^l(k-1) > k-1$. By applying equation 4.2 $l-1$ times we now find

$$R_k(l,l) \leq m_k(1, F^{l+l-1+l-2+\cdots+2}(k-1)) = F^C(k-1),$$

With $C = \frac{(l-k+2)(l-k+1)}{2}$. We can quickly find for $k=2$ that

$$F(x) = 1 + m_1(x,0) = 1 + 2x - 1 = 2x.$$

This gives that $F^l(x) = 2^l \cdot x$. So we see that

$$R_2(l,l) = m_2(l-1,1) \leq F^{\frac{(l-1)l}{2}}(1) = 2^{\frac{(l-1)l}{2}}. \tag{4.3}$$

Ramsey stated a better bound immediately after in the same paper: He found that

$$m_2(l,h) \leq h \cdot (l+1)! \tag{4.4}$$

also holds, which gives

$$R_2(l,l) = m_2(l-1,1) \leq l!.$$

To prove equation 4.4, we use induction on l.

For $l=1$, suppose we have $2h$ elements. If c is a coloring of those elements and we take any element x, it follows that there are at least h of the $2h-1$ elements that are of the same color when combined with x, which proves the base case.

Now suppose it holds for $l-1$. We will show it also hold for l. Let $n = h \cdot (l+1)! = h(l+1) \cdot l!$ and c be any coloring of $[n]^2$. Then since it holds for l we can find a set S and a set T with $|S| = l-1$ and $|T| = h(l+1)$ such that all pairs containing an element from S are of the same color, say 1. We will now consider two cases.

First, suppose there is an x in T such that there are at least h elements of T that are of color 1 together with x. In this case, we are done when we add x to S and take those h elements as our new T.

Suppose now instead that this x does not exist. This means that all elements in T have at most $h-1$ other elements that are of color 1 with this element. Take x_1 from T. Then we can find a set T_1 from T that are all of color 2 together with x with $|T_1| \geq h(l+1) - 1 - (h-1) = hl$. Now take an element x_2 from T_1. We once again find a set T_2 such that all elements of T_2 with x_2 are of color 2 with $|T_2| \geq (l-1)h$. Continue this to find elements $x_1, \ldots, x_l$ and a set T_l with $|T_l| \geq h$. Then taking $S = \{x_1, \ldots, x_l\}$ and $T = T_l$ we indeed find the right monochromatic set, which proves the bound.

If we allow multiple colors we can use induction on the number of colors, just like in the alternate proof of the multicolored Ramsey theorem with $k=2$, that

$$R_2(l, \ldots, l) \leq l! \ldots !, \tag{4.5}$$

where we have r colors and $r-1$ factorials.We just proved this for $r=2$.

Now suppose it holds up to some r. Then let $n = l!\ldots!$ with $r-1$ factorials and let c be any r-coloring of $[n]^2$. Then if we let c' be c with the last two colors 'merged', just like before, then we find by the induction hypothesis a monochromatic subset of c' of size $n!$. If this monochromatic subset is not of the last color, we immediately find a monochromatic subset of c.

If it were the last color, we find from the base case that we have a monochromatic subset of size n of one of those two colors we merged, so the bound indeed holds.

It is evident that this bound grows incredibly quickly, especially if we allow multiple colors.

4.2 Skolem's proof

In 1933 Skolem found a different proof for Ramsey's theorem that provided better upper bounds than Ramsey's proof [6]. His proof is interesting because he did not try to find $R_k(n, n)$, the smallest integer that always allows a monochromatic subset of size n. Instead he started with $[m]$ and tried to find the largest monochromatic subset of $[m]$ that was guaranteed to exist. The proof of Erdős and Rado, which we will consider after this one, also uses this principle. He proved the following theorem:

Theorem 5 (Skolem's variant of Ramsey)**.** *Suppose we have a set $[m]$ and let k and r be integers. Let c_k be any r-coloring of $[m]^k$. Then there exists a monochromatic subset under c_k of size at least $f(k, r, m)$.*

How large this $f(k, r, m)$ can be will follow from the proof. We will first state the proof and then derive a condition for f.

Skolem's proof of Ramsey's theorem. Skolem also used an induction argument on the number of elements in the coloring k.

The base case is $k = 1$. It is easy to verify that $f(1, r, m) = \left\lceil \frac{m}{r} \right\rceil$.

So suppose the theorem holds up to a certain k, so for all r-colorings c_{k-1} of $[m]^{k-1}$ we can find a monochromatic subset of size at least $f(k-1, r, m)$. Suppose c_k is an r-coloring on k elements of $[m]^k$. We will give a procedure to obtain a set that is monochromatic under c_k.

Start by taking any element a_1 from $[m]$. Define a new coloring c^1_{k-1} on $[[m] \setminus \{a_1\}]^{k-1}$ by $c^1_{k-1}(x_1, \ldots, x_{k-1}) = c_k(x_1, \ldots, x_{k-1}, a_1)$. By the induction hypothesis we can find a monochromatic subset S_1 of size $m_1 = f(k-1, r, m-1)$ of color r_1.

From S_1, pick a new element a_2 and define a coloring c^2_{k-1} by $c^2_{k-1}(x_1, \ldots, x_{k-1}) = c_k(x_1, \ldots, x_{k-1}, a_2)$. Once again we obtain a monochromatic subset S_2 under c^2_{k-1} of size $f(k-1, r, m_1 - 1)$ and color r_2.

Continue this until we find $a_1, \ldots, a_t$ and $S_1, \ldots, S_t$ with $t = r \cdot (f(k, r, m) - 1) + 1$ for which all subsets of k elements from $\{a_i\} \cup S_i$ containing a_i are of color r_i. Those exist as long as all the S_i are nonempty. The r_i only take r values so there must exist $r_{j_1}, r_{j_2}, \ldots, r_{j_{f(k,r,m)}}$ such that all r_j are the same. We claim that $\{a_{j_1}, a_{j_2}, \ldots, a_{j_{f(k,r,m)}}\}$ is monochromatic.

Indeed, take any k-element subset $\{a_{x_1}, \ldots, a_{x_k}\}$ where we can assume the elements are in increasing order. Now we must have that $a_{x_2}, \ldots, a_{x_k}$ are all in S_{x_1}. Because of how we chose S_{x_1} it now immediately follows that $c(a_{x_1}, \ldots, a_{x_k}) = r_{x_k}$, so the theorem holds. $\square$

The proof works as long as all the S_i are nonempty. So if we define

$$g(0, k, r, m) = m,$$
$$g(z, k, r, m) = f(k, r, g(z-1, k, r, m) - 1),$$

Then $g(z, k, r, m)$ is the size of set S_z following the procedure above. Now we define $f(k, r, m)$ for $k > 1$ to be the largest integer such that

$$g(r(f(k, r, m) - 1) + 1, k - 1, r, m) \geq 1.$$

This exists, because g is decreasing in its first argument. This is the same as requiring $S_{r(f(k,r,m)-1)+1}$ to be nonempty.

Skolem provided an explicit bound in the case $k = 2$. Finding an upper bound for $R_2(l, \ldots, l)$ with r colors is equivalent to finding a lower bound for m such that $f(2, r, m) \geq l$, since this m would be an upper bound for $R_2(l, \ldots, l)$. We will show that

$$m = \frac{r^{rl-r+2} - 1}{r - 1}$$

will give that $f(2, r, m) \geq l$.

Note that $f(2, r, m) \geq l$ is the same as requiring $g(r(l - 1) + 1, 1, r, m) \geq 1$. We will show that $g(z, 1, r, m) \geq \frac{r^{rl-r+2-z}-1}{r-1}$ using induction.

For $z = 0$, this is trivial from the definition of m.

So suppose it holds up to a certain z. We then have that

$$g(z, 1, r, m) = f(1, r, g(z-1, 1, r, m) - 1). \tag{4.6}$$

From how we defined f we get that

$$f(1, r, g(z-1, 1, r, m) - 1) = \left\lceil \frac{g(z-1, 1, r, m) - 1}{r} \right\rceil \geq \frac{g(z-1, 1, r, m) - 1}{r}.$$

Using the induction hypothesis we thus find

$$g(z, 1, r, m) \geq \frac{\frac{r^{rl-r+3-z}-1}{r-1} - 1}{r} = \frac{r^{rl-r+3-z} - 1 - r + 1}{r(r-1)} = \frac{r^{rl-r+2-z} - 1}{r - 1}, \tag{4.7}$$

as required. We thus find that

$$g(r(l-1) + 1, 1, r, m) \geq \frac{r^{rl-r+2-(rl-r+1)} - 1}{r - 1} = \frac{r-1}{r-1} = 1. \tag{4.8}$$

So this m is indeed an upper bound for $R_2(l, \ldots, l)$ with r colors. We see that this bound is a lot smaller than the one Ramsey found. Even for $r = 2$ we find $R_2(l, l) \leq 2^{2l} - 1$ which is considerably better than $l!$.

If we take $k > 2$, finding a bound for $g(z, k-1, r, m)$ becomes more difficult and the bounds become less precise, since we would have to estimate $f(k-1, r, m)$ as well.

4.3 Proof by Erdős and Rado

A nice bound on the Ramsey numbers was found by Erdős and Rado in 1952 [1]. It uses a different proof of Ramsey's theorem than the one mentioned earlier. The proof goes as follows:

Proof. We will once again use induction on the number of elements in the coloring k.

Our base case remains $k = 2$.

So, assume that for a certain k we have that $R_{k-1}(l_1 - 1, l_2 - 1, \ldots, l_r - 1)$ exists. We want to show that $R_k(l_1, l_2, \ldots, l_r)$ exists as well. We will first explain the procedure to find a monochromatic set of sufficient size, which can be carried out if we start with a large enough starting set, and then derive an upper bound for the size of this starting set.

So, assume n is large enough, and c_k is an r-coloring of $[n]^k$. First, take $k-2$ distinct elements $a_1, a_2, \ldots, a_{k-2}$. Let S_{k-1} be the set $[n] \setminus \{a_1, \ldots, a_{k-2}\}$.

Take any arbitrary element a_{k-1} of S_{k-1}. We color the elements of $S_{k-1} \setminus \{a_{k-1}\}$ by a new coloring (on one element) which we call c_1^{k-1} as follows:

$$c_1^{k-1}(x) = c_k(a_1, \ldots, a_{k-1}, x).$$

We now define S_k as the largest monochromatic subset of S_{k-1} under c_1^{k-1}. Note that c_1^{k-1} has at most r colors, so we know that $|S_k| \geq \frac{|S_{k-1}|-1}{r}$.

We take a_k from S_k arbitrary, and define a new coloring c_1^k on 1-elements of $S_k \setminus \{a_k\}$ by the elements $a_1, \ldots, a_k$, by coloring two elements x, y in the same color if for all $k-2$-subsets $a_{i_1}, \ldots, a_{i_{k-2}}$ of $a_1, \ldots, a_{k-1}$ we have that

$$c_k(a_{i_1}, \ldots, a_{i_{k-2}}, a_k, x) = c_k(a_{i_1}, \ldots, a_{i_{k-2}}, a_k, y).$$

One possibility for such a coloring would be

$$c_1^k(x) = c_k(a_1, \ldots, a_{k-2}, a_k, x) + r c_k(a_1, \ldots, a_{k-3}, a_{k-1}, a_k, x) + \cdots + r^{k-1} c_k(a_2, \ldots, a_k, x).$$

It is not important for the proof to have an explicit formula for c_1^k, the formula just shows that such a coloring indeed exists.

There are $k-1$ possible subsets of $k-2$ elements of $a_1, \ldots, a_{k-1}$, so c_1^k has at most r^{k-1} colors. We set S_{k+1} as the largest monochromatic subset of $S_k \setminus \{a_k\}$ under c_1^k. Since we took the largest subset, we know that $|S_{k+1}| \geq \frac{|S_k|-1}{r^{k-1}}$.

In general, suppose we have defined $(a_1, \ldots, a_j)$ and a subset S_{j+1}. Take a_{j+1} from S_{j+1} arbitrarily and define c_1^{j+1} the same as c_1^k: Two elements x, y are of the same color under c_1^{j+1} if for all $(k-2)$-subsets$\{a_{i_1}, \ldots, a_{i_{k-2}}\}$ of $\{a_1, \ldots, a_{j-1}\}$ we have that

$$c_k(a_{i_1}, \ldots, a_{i_{k-2}}, a_j, x) = c_k(a_{i_1}, \ldots, a_{i_{k-2}}, a_j, y).$$

Define S_{j+2} as the largest monochromatic subset of $S_{j+1} \setminus \{a_{j+1}\}$ under c_1^{j+1}. There are at most $r^{\binom{j}{k-2}}$ different colors, so we have that

$$|S_{j+2}| \geq \frac{|S_{j+1}| - 1}{r^{\binom{j}{k-2}}}.$$

Continue this procedure until we reach a_t, $t = R_k(l_1 - 1, l_2 - 1, \ldots, l_r - 1) + 1$. We now define a last coloring d_{k-1} on $k-1$ elements of $\{a_1, \ldots, a_{t-1}\}$ by

$$d_{k-1}(a_{i_1}, \ldots, a_{i_{k-1}}) = c_k(a_{i_1}, \ldots, a_{i_{k-1}}, a_t).$$

Because of how we chose t we know that for some i there exists an i-monochromatic set of size $l_i - 1$, say $B = \{b_1, \ldots, b_{l_i - 1}\}$. We claim that $B \cup \{a_t\}$ is also i-monochromatic under c_k.

For that, suppose $b_{j_1}, \ldots, b_{j_k} \in$ B. We can assume that $(j_1, j_2, \ldots, j_k)$ is increasing. If $b_{j_k} = a_t$, we immediately find $c(b_{j_1}, \ldots, b_{j_k}) = i$. So, assume $b_{j_k} \neq a_t$. Then, because both b_{j_k} and a_t are in $S_{j_{k-1}}$, we know that $c_k(b_{j_1}, \ldots, b_{j_{k-1}}, b_{j_k}) = c_k(b_{j_1}, \ldots, b_{j_{k-1}}, a_t)$. Since $d_{k-1}(b_{j_1}, \ldots, b_{j_{k-1}}) = i$, we find by our definition of d_{k-1} that $c_k(b_{j_1}, \ldots, b_{j_{k-1}}, a_t) = i$ as well, so $c_k(b_{j_1}, \ldots, b_{j_k}) = i$ as requested. So B is indeed i-monochromatic under c_k, which proves the theorem by induction. $\square$

We assumed at the start that we had a 'large enough' set. By specifying how large 'large enough' is, we find an upper bound for the Ramsey numbers. Our proof holds as long as we can keep finding a_i from the set S_i, up to a_t. In other words, all the S_i need to be nonempty. Since those sets are decreasing, this reduces to S_t needs to be nonempty. We have already derived recurrence relations for the cardinalities of our sets S_i in our proof. If we set $s_i = |S_i|$, we find:

$$s_k = \frac{n-k+1}{r} \tag{4.9}$$

$$s_{i+1} \geq \frac{s_i - 1}{r^{\binom{i-1}{k-2}}}, \qquad i = k, \ldots, r. \tag{4.10}$$

We now want to find n in such a way that $s_t > 0$. For ease of notation we define $m_i = r^{-\binom{i}{k-2}}$. By using our recurrence relation 4.10 repeatedly, we get

$$s_t \geq s_{t-1} m_{t-2} - m_{t-2} \tag{4.11}$$

$$s_t \geq s_{t-2} m_{t-2} m_{t-3} - m_{t-2} m_{t-3} - m_{t-2} \tag{4.12}$$

$$\vdots \tag{4.13}$$

$$s_t \geq s_k m_{k-1} m_k \ldots m_{t-2} - m_{k-1} m_k \ldots m_{t-2} - m_k m_{k+1} \ldots m_{t-2} - \ldots - m_{t-2} \tag{4.14}$$

$$s_t \geq (n-k+1) m_{k-2} m_{k-1} \ldots m_{t-2} - m_{k-1} m_k \ldots m_{t-2} - m_k m_{k+1} \ldots m_{t-2} - \ldots - m_{t-2}. \tag{4.15}$$

For the last inequality we used that $m_{k-2} = \frac{1}{r}$. We now require the right-hand side of equation 4.15 to be positive to find

$$(n-k+1) m_{k-2} m_{k-1} \ldots m_{t-2} - m_{k-1} m_k \ldots m_{t-2} - m_k m_{k+1} \ldots m_{t-2} - \ldots - m_{t-2} > 0 \tag{4.16}$$

$$n > k - 1 + \frac{1}{m_{k-2}} + \frac{1}{m_{k-2} m_{k-1}} + \ldots + \frac{1}{m_{k-2} m_{k-1} \cdots m_{t-3}} \tag{4.17}$$

$$n \geq k + r^{\binom{k-2}{k-2}} + r^{\binom{k-1}{k-2} + \binom{k-2}{k-2}} + \ldots + r^{\binom{t-3}{k-2} + \binom{t-2}{k-2} + \ldots + \binom{k-2}{k-2}}. \tag{4.18}$$

We now claim that $\sum_{k=r}^{n} \binom{k}{r} = \binom{n+1}{r+1}$. This follows immediately by induction on n. The base case $n = 0$ is trivial, so assume it holds for n. Then

$$\sum_{k=r}^{n+1} \binom{k}{r} = \binom{n+1}{r} + \binom{n}{r+1} = \binom{n+1}{r+1}.$$

So equation 4.18 reduces to

$$n \geq k + \sum_{j=k-1}^{t-2} r^{\binom{j}{k-1}}. \tag{4.19}$$

Since n is an upper bound for $R_k(l_1, \ldots, l_r)$, we want n to be as small as possible, so we choose n equal to the right-hand side. Equation 4.19 is not yet an explicit upper bound since the bound depends on t, which we defined as $R_{k-1}(l_1 - 1, l_2 - 1, \ldots, l_r - 1) + 1$. We will first estimate the binomial coefficient to obtain

$$k + \sum_{j=k-1}^{t-2} r^{\binom{j}{k-1}} \leq k + \sum_{j=k-1}^{t-2} r^{(j^{k-1})}. \tag{4.20}$$

Powers of powers can get a bit cumbersome to work with. Instead we shall use $k \uparrow n$ to denote k^n following Knuth's up-arrow notation. Then we can define

$$k \ \uparrow\ \uparrow n = \ k \ \uparrow\ (k \ \uparrow\ (\ldots \uparrow\ k) \ldots), \tag{4.21}$$

where we have $n - 1$ arrows. This needs to be resolved from right to left, since otherwise we would just be multiplying the powers and $k \uparrow\uparrow n$ would just be $k^{(k^{n-1})}$. We could continue this procedure, but this is not necessary for this report.

We would like to get rid of the sum in Equation 4.20. To do this we use that for r and k greater than 2

$$r \uparrow (j+1) \uparrow (k-1) \geq (r \uparrow j \uparrow (k-1)) \cdot (r \uparrow 1 \uparrow (k-1)) \geq (r \uparrow j \uparrow (k-1)) \cdot 2.$$

This gives that

$$(r \uparrow j \uparrow (k-1)) \leq (r \uparrow (j+1) \uparrow (k-1)) - (r \uparrow j \uparrow (k-1)).$$

If we use this in equation 4.20 we obtain

$$k + \sum_{j=k-1}^{t-2} r \uparrow j \uparrow (k-1) \leq k + \sum_{j=k-1}^{t-2} ((r \uparrow (j+1) \uparrow (k-1)) - (r \uparrow j \uparrow (k-1)))$$

$$= k + (r \uparrow (t-1) \uparrow (k-1)) - (r \uparrow (k-1) \uparrow (k-1)).$$

Since $r \uparrow (k-1) \uparrow (k-1)$ is larger than k, this results to

$$k + (r \uparrow (t-1) \uparrow (k-1)) - r \uparrow (k-1) \uparrow (k-1)) \leq r \uparrow (t-1) \uparrow (k-1)$$

$$= r \uparrow R_{k-1}(l_1 - 1, l_2 - 1, \ldots, l_r - 1) \uparrow (k-1).$$

As a result we find

$$R_k(l_1, l_2, \dots, l_r) \leq r \uparrow R_{k-1}(l_1 - 1, l_2 - 1, \dots, l_r - 1) \uparrow (k-1). \tag{4.22}$$

We can write equation 4.22 in a way that is a little easier to apply continuously:

$$R_k(l_1, l_2, \dots, l_r) \uparrow k \leq (r \uparrow k) \uparrow R_{k-1}(l_1 - 1, l_2 - 1, \dots, l_r - 1) \uparrow (k-1). \tag{4.23}$$

Note the brackets around $r \uparrow k$. By continually applying equation 4.23 we obtain

$$R_k(l_1, \dots, l_r) \leq (r \uparrow k) \uparrow (r \uparrow (k-1)) \uparrow \dots \uparrow (r \uparrow 2) \uparrow R_1(l_1 - k + 1, \dots, l_r - k + 1).$$

If we set $C = \sum_{i=1}^{r} l_i - k + 1$, then we get the following final explicit bound:

$$R_k(l_1, \dots, l_r) \leq (r \uparrow k) \uparrow (r \uparrow (k-1)) \uparrow \dots \uparrow (r \uparrow 2) \uparrow (C - r + 1). \tag{4.24}$$

So we approximately find for r colors that $R_k(l, \dots, l) \leq (r \uparrow\uparrow k)^{C'l}$, where C' is a constant depending on r and k.

5 Lower bounds

So far we have seen various proofs of Ramsey's theorem, all of which give their own upper bound. In this section we will look at a lower bound given by Erdős and Hajnal. We will closely follow the proof given in [3]. We will restrict ourselves to the symmetric case with two colors. The proof does not construct a coloring that is such that there is no monochromatic coloring of a certain size, as one might initially expect. Instead it merely shows the existence of such a coloring using a probabilistic argument. As such it is sometimes also called an existence argument.

Theorem 6. *If $\binom{n}{l}2^{1-\binom{l}{k}} < 1$, then we have that $R_k(l,l) > n$.*

Proof. Suppose $\binom{n}{l}2^{1-\binom{l}{k}} < 1$. Now create a random coloring c of $[n]^k$ by setting for all distinct elements $x_1, \ldots, x_k$,

$$P(c(x_1, \ldots, x_k) = 1) = P(c(x_1, \ldots, x_k) = 2) = \frac{1}{2}.$$

We color all different subsets independently. Now we define for all subsets S of $[n]$ with $|S| = l$ the event A_S to be the event that S is monochromatic. Then

$$P(A_S) = P(\text{S is 1-monochromatic}) + P(\text{S is 2-monochromatic}) = 2 \cdot 2^{-\binom{l}{k}} = 2^{1-\binom{l}{k}}.$$

Now if we look at the event B of *some* subset of l elements to be monochromatic we see that

$$P(B) \leq \sum_{|S|=l} P(A_S) = \binom{n}{l} 2^{1-\binom{l}{k}} < 1. \tag{5.1}$$

Since $P(B) < 1$ there must exist a coloring without a monochromatic subset of l elements. □

We see that all that is left is finding an upper bound for n such that equation 5.1 holds. We can bound the first binomial by

$$\binom{n}{l} < n^l,$$

as long as $l \geq 2$ this inequality is strict (we will see that n is much larger than l). Similarly $\binom{l}{k} \geq \frac{(l-k+1)^k+1}{k!}$ as long as $l \neq k$, which was the trivial case. So if we take

$$n \leq 2^{\frac{1}{k!}\cdot(l-k+1)^{(k-1)}}, \tag{5.2}$$

we see that

$$2\binom{n}{l} < 2n^l \leq 2^{\frac{1}{k!}l(l-k+1)^{k-1}+1} \leq 2^{\frac{l}{k}\binom{l-1}{k-1}} = 2^{\binom{l}{k}}. \tag{5.3}$$

For this n we thus found that $2\binom{n}{l} < 2^{\binom{l}{k}}$, which is exactly the same as equation 5.1. The right hand side of equation 5.2 is thus a lower bound for $R_k(l,l)$. This lower bound does not grow nearly as fast as the upper bound. From $k = 3$, however, there is a way to obtain a much better lower bound from $R_k(l,l)$ for $R_{k+1}(l,l)$.

Lemma 1 (Stepping-up lemma)**.** *If $k \geq 3$ and $R_k(l,l) > n$ then we must have that $R_{k+1}(2l+k-4, 2l+k-4) > 2^n$.*

Proof. Suppose $k \geq 3$ and let n be as above. Then there exists a coloring c_k on $[n]^k$ that does not contain a monochromatic subset of size l. We will construct a new coloring c'_{k+1} on $[2^n]^{k+1}$ that does not contain a monochromatic subset of size $2l+k-4$, effectively 'stepping up' the bound from k to $k+1$.

Instead of considering $[2^n]$ we will consider a set $S = \{(a_1, a_2, \ldots, a_n) : a_i \in \{0,1\}\}$. If the context is not clear we will denote elements of S with a superscript and elements of an element of S, like the a_i in the previous sentence, with a subscript. It is evident that S has 2^n elements.

Now we will order two elements $x = (a_1, \ldots, a_n)$ and $x' = (a'_1, \ldots, a'_n)$ of S by looking at the last different element of x and x'. Suppose that for some i we have that $a_i \neq a'_i$ and for $j > i$ we have that $a_j = a'_j$. We then say that $x > x'$ if $a_i > a'_i$ and similarly $x < x'$ if $a_i < a'_i$. Since all the a_i are either zero or one, this reduces to $x > x'$ if $a_i = 1$ and $a'_i = 0$ and $x < x'$ if $a_i = 0$ and $a'_i = 1$ with i the same as before. Furthermore we then define $\delta(x, x') = i$. Note that δ then takes values in $[n]$.

This is a proper ordering, since it coincides with the regular ordering of natural numbers if we consider members of S as binary numbers with $x = \sum_{i=1}^{n} a_i 2^{i-1}$. Take any $(k+1)$-subset $\{x^1 < \cdots < x^{k+1}\}$ of S. Define $y_i = \delta(x^i, x^{i+1})$. Note that $y_i \neq y_{i+1}$, since if this were the case, $x^i < x^{i+1}$ would give $x^{i+1}_{y_i} = 1$ while $x^{i+1} < x^i$ would give $x^{i+1}_{y_i} = 0$, which is obviously a contradiction. Furthermore we have that

$$\delta(x^1, x^{k+1}) = \max_{1 \leq i \leq k} \delta(x^i, x^{i+1}). \tag{5.4}$$

To see this, let $z = \max_{1 \leq i \leq k} \delta(x^i, x^{i+1})$. Then for some j we have that $\delta(x^j, x^{j+1}) = z$. Now $x^j_z = 0$ and we claim that $x^1_z = 0$ as well. This is not hard to see, since if $x^1_z = 1$ it has to change from 1 to 0 in some other x^i with $i < j$. Then since z has to be the last number that is different we obtain $x^{i-1} > x^i$ which is a contradiction. Similarly we have that $x^{k+1}_z = 1$. Now for every $p > z$ we have that $x^1_p = x^{k+1}_p$, since otherwise we get a contradiction with z being the maximum, so equation 5.4 indeed holds.

Now we color the subsets of T as follows.

First, if $y_1 < \cdots < y_n$ or $y_1 > \cdots > y_n$ we set $c'_{k+1}(x^1, \ldots, x^{k+1}) = c_k(y_1, \ldots, y_n)$. This is well defined since $y_i = \delta(x^i, x^{i+1})$ takes values in $[n]$ and all the y_i are different from the assumption.

Next, if we have that $y_1 < y_2 > y_3$, we set $c'_{k+1}(x^1, \ldots, x^{k+1}) = 1$. If instead we had $y_1 > y_2 < y_3$, then we define $c'_{k+1}(x^1, \ldots, x^{k+1}) = 2$. All the other subsets are colored arbitrary.

Now take any subset $T = \{x^1, \ldots, x^{2l+k-4}\}$ of S with $|T| = 2l+k-4$. We will assume that T is 1-monochromatic and derive a contradiction.

We once again define $y_i = \delta(x^i, x^{i+1})$. We claim that we can always find a monotonic subsequence of the form $y_i, y_{i+1}, \ldots, y_{i+l-1}$, so either $y_i < \cdots < y_{i+l-1}$ or $y_i > \cdots > y_{i+l-1}$. Now for $i \leq 2l-3$ we can not have that $y_{i-1} > y_i < y_{i+1}$, since in that case we have that $c(x^{i-1}, x^i, \ldots, x^{i+k-1}) = 2$. Note that this is the part we had to assume $k \geq 3$ for. This means that if we look at $y_1, \ldots, y_{2l-3}$ we have at most one j for which $y_{j-1} < y_j > y_{j+1}$. If this j does not exist or $j \geq l$ then $y_1, \ldots, y_l$ suffices.

If instead $j \leq l-1$ then $j+l-1 \leq 2l-2$ and thus $y_j, \ldots, y_{j+l-1}$ is of the desired form, so we can always find a monotonic subsequence $y_i, \ldots, y_{i+l-1}$.

Now since we have l elements of $[n]$ there must be by assumption k of them, say $y_{j_1}, \ldots, y_{j_k}$, for which $c_k(y_{j_1}, \ldots, y_{j_k}) = 2$. We will assume $j_1 < \cdots < j_k$. If $y_i > \cdots > y_{i+l-1}$ we claim that $c'_{k+1}(x^{j_1}, \ldots, x^{j_k}, x^{j_k+1}) = 2$. From the monotonicity and equation 5.4 it follows that

$$\delta(x^{j_i}, x^{j_{i+1}}) = \delta(x^{j_i}, x^{j_i+1}) = y_{j_i}.$$

It is also clear that $\delta(x^{j_k}, x^{j_k+1}) = y_{j_k}$. Since the y_{j_i} are monotonic we immediately find that $c'_{k+1}(x^{j_1}, \ldots, x^{j_k}, x^{j_k+1}) = 2$ since they are colored in the first way described.

If we had $y_i < \cdots < y_{i+l-1}$ instead then we would obtain $c'_{k+1}(x^{j_1}, x^{j_1+1}, \ldots, x^{j_k+1}) = 2$ instead.

So we see that T is not 1-monochromatic. If we assumed that T was 2-monochromatic the same argument would hold. □

From the first proof we found $R_3(l,l) > 2^{cl^2}$. The lemma gives us a way better lower bound for $k \geq 4$. In general we can prove using induction that for $k \geq 3$,

$$R_k(l,l) > (2\uparrow\uparrow(k-2))^{c(k)l^2}. \tag{5.5}$$

Here $c(k)$ denotes a constant that depends on k. For $k = 3$ this is clear, so suppose it holds up to a certain k. Then from the lemma we find

$$R_{k+1}(l,l) > 2\uparrow R_k(\left\lfloor \frac{2l-4+k}{2} \right\rfloor, \left\lfloor \frac{2l-4+k}{2} \right\rfloor) > (2\uparrow 2\uparrow\uparrow k-2)^{c\left\lfloor \frac{2l-4+k}{2} \right\rfloor^2} >$$

$$(2\uparrow\uparrow k-1)^{c'(k)l^2}.$$

If we compare this lower bound with the upper bound found by Erdős and Rado, we see that the lower bound is one exponent lower than the upper bound. If we had $(R_3(l,l) > 2\uparrow\uparrow 2)^{cl^2}$ instead or if we could somehow improve the lemma to also work for the case $k = 2$ we would have a lower bound of the same 'size' as the upper bound. This is however still an open problem.

6 Applications of Ramsey's theorem

It should come as no surprise that a theorem as broadly useable as Ramsey's theorem has many applications. In this section we will look at two applications of Ramsey's theorem, namely the Happy Ending problem and the monotone subsequence problem[2].

6.1 Monotone subsequence problem

The monotone subsequence problem is quite simple in nature. Suppose we have a finite subsequence $x_1, \ldots, x_n$ of real numbers. Now we want to know how large of a monotone subsequence we are guaranteed to get. This problem was solved by Erdős and Szekeres in 1935, but the proof we will look at is given much later, in 1959, by Seidenberg [5].

Theorem 7 (Monotone subsequence problem)**.** *Let n and m be integers. Suppose we have a sequence $a_1, \ldots, a_{(n-1)(m-1)+1}$ of real numbers. Then we must either have a increasing subsequence of length n or a decreasing subsequence of length m.*

In this case we will also call a sequence monotonic of it remains constant somewhere. So for a sequence $a_1, \ldots, a_n$ monotonically decreasing means that for $i < j$ we have that $a_i \leq a_j$ and monotonically increasing means that for $i < j$ we must have $a_i \geq a_j$.

Proof. Suppose $a_1, \ldots, a_{(n-1)(m-1)+1}$ is a sequence of real numbers and suppose there is no decreasing subsequence of length m. To every number a_i of our sequence we assign a pair (x_i, y_i) such that x_i is the maximum length of an increasing subsequence starting from a_i and similarly y_i is the maximum length of a decreasing subsequence starting at a_i. Now for $i < j$ we claim that we can not assign the same pair to both a_i and a_j, so either $x_i \neq x_j$ or $y_i \neq y_j$.

Indeed, if $a_i \leq a_j$ then we can add a_i to the maximum increasing subsequence starting from a_j of length x_j to find that $x_i \geq x_j + 1$. Following the same argument, if $a_i > a_j$ we see that $y_i \geq y_j + 1$. We can always find a monotone subsequence of length 1 from a_i by taking the subsequence containing only a_i, so we must have that $x_i \geq 1$ and $y_i \geq 1$. Furthermore, from the assumption it follows that $y_i \leq m - 1$ for all i.

Now since $1 \leq y_i \leq m - 1$ we see that y_i only takes $m - 1$ values. Since we have $(n-1)(m-1)+1$ pairs, there must be at least one y_0 assigned to n numbers, so we have $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ with $y_{i_j} = y_0$. Now we must have that $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$is an increasing subsequence. Remember that we derived earlier for $i < j$ that if $a_i > a_j$ then we had that $y_i \geq y_j + 1$, which clearly does not hold here, so the subsequence is indeed increasing. □

It is easy to see that $(n-1)(m-1)+1$ is indeed the minimum size of the sequence for which we are guaranteed to find a monotone subsequence of the desired size. If we have $(n-1)(m-1)$ real numbers as follows: $x_1 = m - 1, x_2 = m - 2, \ldots, x_m - 1 = 1, x_m = 2(m-1), x_{m+1} = 2(m-1) - 1 \ldots$ and so on, so we have $n - 1$ times a decreasing subsequence of size $m - 1$. The reader can verify that this does indeed not contain an increasing subsequence of size n nor a decreasing subsequence of size m.

It is also possible to prove the monotone subsequence problem using Ramsey's theorem.

Theorem 8. *If $x_1, \ldots, x_{R(m,n)}$ is a sequence of real numbers then we must either have an increasing subsequence of length m or a decreasing subsequence of length n.*

Proof. The proof is very straightforward. Suppose $x_1, \dots, x_{R(m,n)}$ is a sequence of real numbers. Now we define a coloring on subsets of those numbers as follows:

For $i < j$ we let $c(\{x_i, x_j\}) = 1$ if $x_i \leq x_j$ and $c(\{x_i, x_j\}) = 2$ if $x_i > x_j$.

Now under this coloring we are guaranteed to find either m numbers that are 1-monochromatic or n numbers that are 2-monochromatic. In the first case those m numbers are a monotonically increasing subsequence and in the second case the n numbers are a monotonically decreasing subsequence. □

Certainly $R_2(n, m)$ is much larger than $(n-1)(m-1)+1$. We found as an upper bound $R_2(n, m) \leq \binom{n+m-2}{n-1}$ and even the lower bound we considered for the symmetric case was still exponential. So this proof certainly gives numbers that are far larger than the actual solution. Part of that is due to the fact that the Ramsey numbers need to be estimated themselves, but even the actual Ramsey numbers would likely give bounds far larger than the solution.

6.2 Happy Ending problem

Just like the monotone subsequence problem the Happy Ending problem was formulated by Erdős and Szekeres in 1935. It was named the Happy Ending problem because two of the mathematicians that worked on the problem, George Szekeres and Esther Klein, later became married. They found that as long as you put enough points in a finite plane you can always find a convex n-gon. This is an extension to the problem solved by Klein, who found that from five points in the plane there are always four that form a convex 4-gon. The exact number of points you need is to this date still unknown, but just as with the Ramsey numbers upper bounds are known.

The Happy Ending problem is very similar to Ramsey's theorem: We start with a very large number of points of which we know rather little, in Ramsey's case an arbitrary coloring and here arbitrarily placed points in a plane, and go back to a smaller subset of which we know a lot more. In Ramsey's theorem this was a monochromatic subset, in the Happy Ending problem it is a convex n-gon.

We will first state some definitions about convexity, making what we said before more precise.

Definition 1 (Convex combination)**.** *If we have n points $x_1, \dots, x_n$ in the plane then a convex combination of those n points is a point $y = \sum_{i=1}^{n} \lambda_i x_i$, with $\sum_{i=1}^{n} \lambda_i = 1$ with $\lambda_i \geq 0$ for all i.*

The set of all convex combinations of a set of points is called the *convex hull* of those points. For two points, the convex hull is just the line segment between the points. For three points we obtain two cases. If the three points lie on the same line then the convex hull is the line segment from the first endpoint to the last (with one point somewhere in between). If the three points are not on the same line then the convex hull is the triangle and its interior formed by those three points. We see that we obtain different cases, depending on if the three points lie on the same line or not. In the case all the points lie on the same line one of the three points is a convex combination of the others, in the other case this does not hold. This motivates the following definition.

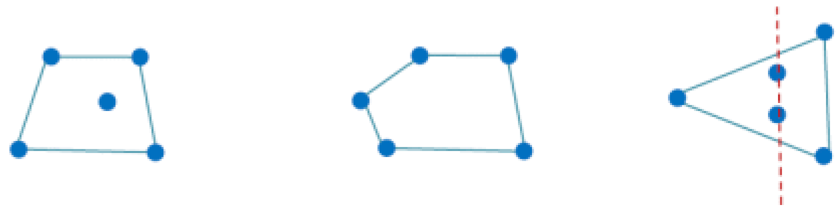


Figure 1: The proof of Esther Klein that any five points in the plane with no three on one line contain four points that form a convex 4-gon. With five points, the convex hull has either four, five or three corners. In the first case (left), the corners of the convex hull form a convex 4-gon. In the second case (middle), the points form a convex 5-gon and any four points will do. In the last case(right) the two points in the interior combined with the two points on the right of the dotted line form a convex 4-gon.

Definition 2. *We call a set of n points a convex n-gon if none of the n points is a convex combination of the other points.*

This definition is the same as one would intuitively expect a convex n-gon to be.

We will now formally state the Happy Ending problem.

Theorem 9 (Happy Ending problem)**.** *For every integer n there exists a larger integer N such that, whenever there are at least N points placed in the 2d plane in such a way, that no three points lie on the same line, then n of those points form a convex n-gon.*

There are two proofs of the theorem using Ramsey Numbers. The original one uses Ramsey numbers with a 2-coloring on subsets of four elements, while a newer proof uses a 2-coloring on subsets of three elements, but with higher Ramsey numbers.

Proof using $R_4(n,5)$. We will show that $R_4(n,5)$ is an upper bound for the number of points needed to guarantee a convex n-gon. So assume we have $N = R_4(n,5)$ points placed in the plane. Now color the subsets of four elements of $[N]$ as follows: $c(x_1,x_2,x_3,x_4) = 1$ if x_1,x_2,x_3,x_4 form a convex 4-gon and $c(x_1,x_2,x_3,x_4) = 2$ otherwise. We will show two things:

1. Of five points in a plane we can always find a subset of four points that are convex (the result that Klein proved). Since this makes it impossible to find 5 points of which all 4-subsets are not convex, this guarantees that if we put $R_4(n,5)$ points in a plane we can find n points for which all its 4-subsets are convex.

2. if all its 4-subsets are convex, those n points form a convex n-gon, proving the theorem. The second statement actually goes both ways, but we only need the forward direction for this proof.

To prove (1), assume we have five points in a plane, with no three on one line. The convex hull of those five points has either five, four or three corners. In the first two cases, we are done, so assume the latter is true. Call the three points which form the corners of the convex hull a, b and c. The remaining two points, say d and e, lie inside the triangle because we assumed that no three points lie on the same line. Draw a line through d and e. This line splits the triangle in two, so on one side of the line, there must be two points. We can safely assume those points are a and b. Now it is easy to see that (a,b,d,e) is a convex 4-gon, see also Figure 1. So we find that (1) is indeed true.

To prove (2), we will show its contrapositive is true, so whenever n points do not form a convex n-gon then four of those points do not form a convex 4-gon. Assume we have n points that are not convex. Then one of those points, say y, lies in the interior of the convex hull of the other points which we will call $x_1, \dots, x_{n-1}$. For simplicity we will assume that x_1 is next to x_2 and x_{n-1}, x_2 is next to x_1 and x_3 and so on. We can now redivide the convex hull in triangles, $T_1 = (x_1, x_2, x_3), T_2 = (x_1, x_3, x_4), \dots, T_{k-3} = (x_1, x_{n-2}, x_{n-1})$. Since the triangles subdivide the convex hull, there is a triangle that has y in its interior, say (x_1, x_i, x_{i+1}). Recall that we assumed no three points lie on one line, so it is not possible for y to lie on the boundary of a triangle. Then (x_1, y, x_i, x_{i+1}) is a subset of four points that is not convex. This proves (2), which also proves the theorem. □

A different proof was found later. This was found by Micheal Tarsy. He took a course on combinatorics where he was asked to prove this theorem on a exam. The theorem was proven during class, but he could not attend that class, and thus came up with a different proof himself.

Proof using $R_3(n, n)$. We will show that $R_3(n, n)$ is an upper bound as well.

To see this, suppose we have $R_3(n, n)$ points in a plane. First order these points by assigning them a number from 1 to $R_3(n, n)$. Throughout the proof, whenever we talk about points $x_1, x_2, \dots, x_n$ or points $y_1, y_2, \dots, y_n$ then the indices will indicate the ordering, so $x_1 < x_2 < x_3$ and so on.

Now define a coloring c as follows: $c(x_1, x_2, x_3) = 1$ if going from x_1 to x_2 to x_3 goes in clockwise direction and $c(x_1, x_2, x_3) = 2$ if it goes counter-clockwise. Because of the definition of $R_3(n, n)$, we know that we can find n points such that either all subsets of three points travel clockwise or all travel counter-clockwise. Assume we have n numbered points $x_1, \dots, x_n$ for which any ordered subset of three elements is travelled in the clockwise direction. We claim that those n points are convex.

To prove this, assume the contrary. Then there is a point x_i that is in the interior of the convex hull of $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$. To simplify the proof we will rename all points of $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$. For $j < i$ we set $y_j = x_j$ and for $j > i$ we set $y_j = x_{j+1}$.

Now subdivide the convex hull in triangles in the same way as in the previous proof, so with triangles $T_1 = (y_1, y_2, y_3)$, $T_2 = (y_1, y_3, y_4)$ and so on. Then x_i is in the interior of one of those triangles, say (y_1, y_j, y_{j+1}).

We know from the assumption that these three points are ordered clockwise. From (y_1, x_i, y_{j+1}) it follows that x_i is in between y_1 and y_{j+1}. From (y_1, y_j, x_i) we find that x is either lower than y_1 (not possible from the first requirement) or higher than y_j, so x_i is in between y_j and y_{j+1}. This gives a contradiction, since then (y_j, x, y_{j+1}) is ordered counter-clockwise. See also Figure 2. □

There is another proof of the Happy Ending problem, one that does not use Ramsey's theorem.

Happy Ending problem. Suppose we have N points in a plane with no three points on one line. First, note that there is a line such that no line between two of those points is parallel or perpendicular to this line. This is easy to see, since there are only a finite number of lines between two of those N points. Now we rotate all the points until this line lies parallel to the x-axis. Now we order the points by their x coordinate. Since no two

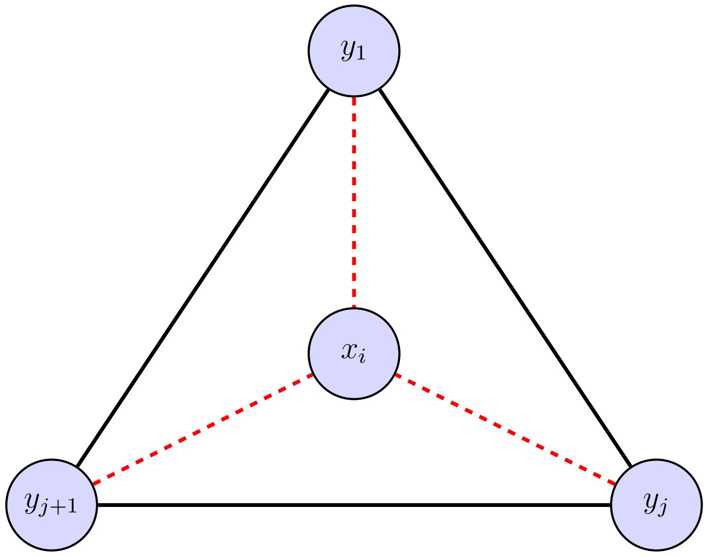


Figure 2: The triangle with x_i in its interior. From the assumption all triplets are ordered clockwise, so y_1, y_j and y_{j+1} have to be as above (up to a rotation). No matter the value of x_i we can find a triplet that is ordered counterclockwise, which is a contradiction.

points lie perpendicular to the line we found, no two points have the same x coordinate and this is well defined.

Now suppose we take k points $x_1 < \cdots < x_k$. If we look at the gradient of the line from x_1 to x_2 and compare it to the gradient of the line going from x_2 to x_3, we either have that the second gradient is larger than the first or the second gradient is smaller than the first. Equal is not possible, since we would obtain three points on one line.

We will call those k points ordered convexly if their subsequent gradients decrease monotonically and ordered concavely if they increase monotonically. We call the least number of points such that we can find either k points that are ordered convexly or l points that are ordered concavely $f(k, l)$. We claim now that

$$f(k, l) \geq f(k-1, l) + f(k, l-1) - 1. \tag{6.1}$$

So suppose we have $f(k-1, l) + f(k, l-1) - 1$ points in the plane. Order them as above.

Now for the first $f(k-1, l)$ points we either have l that are ordered concavely or $k-1$ that are ordered convexly. In the first case we are done, so assume we have $k-1$ convexly ordered points. Call the last point a_1 and take the next point with all previous points excluding a_1.

We once again have $f(k-1, l)$ points, so we find l concavely ordered points or $k-1$ convexly ordered points. Once again assume we have $k-1$ convexly ordered points. Call the last point a_2 and continue this until we have $a_{f(k,l-1)}$. It follows quickly that all a_i are distinct points. Now from those points that are the endpoint of $k-1$ convexly ordered points, we have either k convexly ordered points or $l-1$ concavely ordered points.

In the first case we are done, so assume the second part holds. Call the concavely ordered points $b_1, b_2, \ldots, b_{l-1}$. b_1 is the endpoint of $k-1$ convexly ordered points, so there exist $c_1, \ldots, c_{k-2}, b_1$ that are convexly ordered. Now if the gradient of $a_{k-2}b_1$ is smaller than the gradient of b_1b_2 then $a_1, b_1, b_2, \ldots, b_{l-1}$ are l points that are concavely ordered.

If instead the gradient of $a_{k-2}b_1$ is larger than the gradient of b_1b_2 then we find that $a_1, \dots, a_{k-2}, b_1, b_2$ are k convexly orderde points, so the theorem holds.

We will prove that $f(n,n)$ points is enough to guarantee a convex n-gon. We can first rotate the points such that no two points have the same x coordinate. It is easy to see that if we find a convex n-gon after the rotation, rotating the points back will still result in a convex n-gon.

From the definition of $f(n,n)$ we find either n convexly ordered points or n concavely ordered points. We will show that those n points form a convex n-gon.

To prove this, we use the contrapositive. So if from N points we can not find a convex n-gon, then we can not find n points that are ordered convexly or concavely.

Take any arbitrary n points. Since they do not form a convex n-gon, one of the points is a convex combination of the other points, say y. By using the triangle procedure again we find a triangle $(x_1, x_j.x_{j+1})$ that has y in its interior. We will have four different cases, depending on if the gradient of x_1x_j is larger than the gradient of x_jx_{j+1} and if y is to the right of x_j or to the left.

All cases are shown in Figure 3 on the next page. The red dashed line is the ordering. We always find both two subsequent gradients that are increasing and two subsequent gradients that are decreasing. □

From Equation 6.1 it is possible to derive an explicit upper bound. We claim that

$$f(k,l) \leq \binom{k+l-4}{k-2} + 1, \tag{6.2}$$

which we will prove by induction.

For the base case we have that

$$f(3,k) = f(k,3) = k = \binom{k-1}{k-2} + 1.$$

Now suppose it holds for $f(k-1,l)$ and $f(k,l-1)$. Then we obtain that

$$f(k,l) \leq f(k-1,l) + f(k,l-1) \leq \binom{k+l-3}{k-3} + 1 + \binom{k+l-3}{k-2} + 1 - 1 = \binom{k+l-4}{k-2} + 1.$$

So we find that $f(n,n) \leq \binom{2n-4}{n-2} + 1$, which was an upper bound to guarantee a convex n-gon.

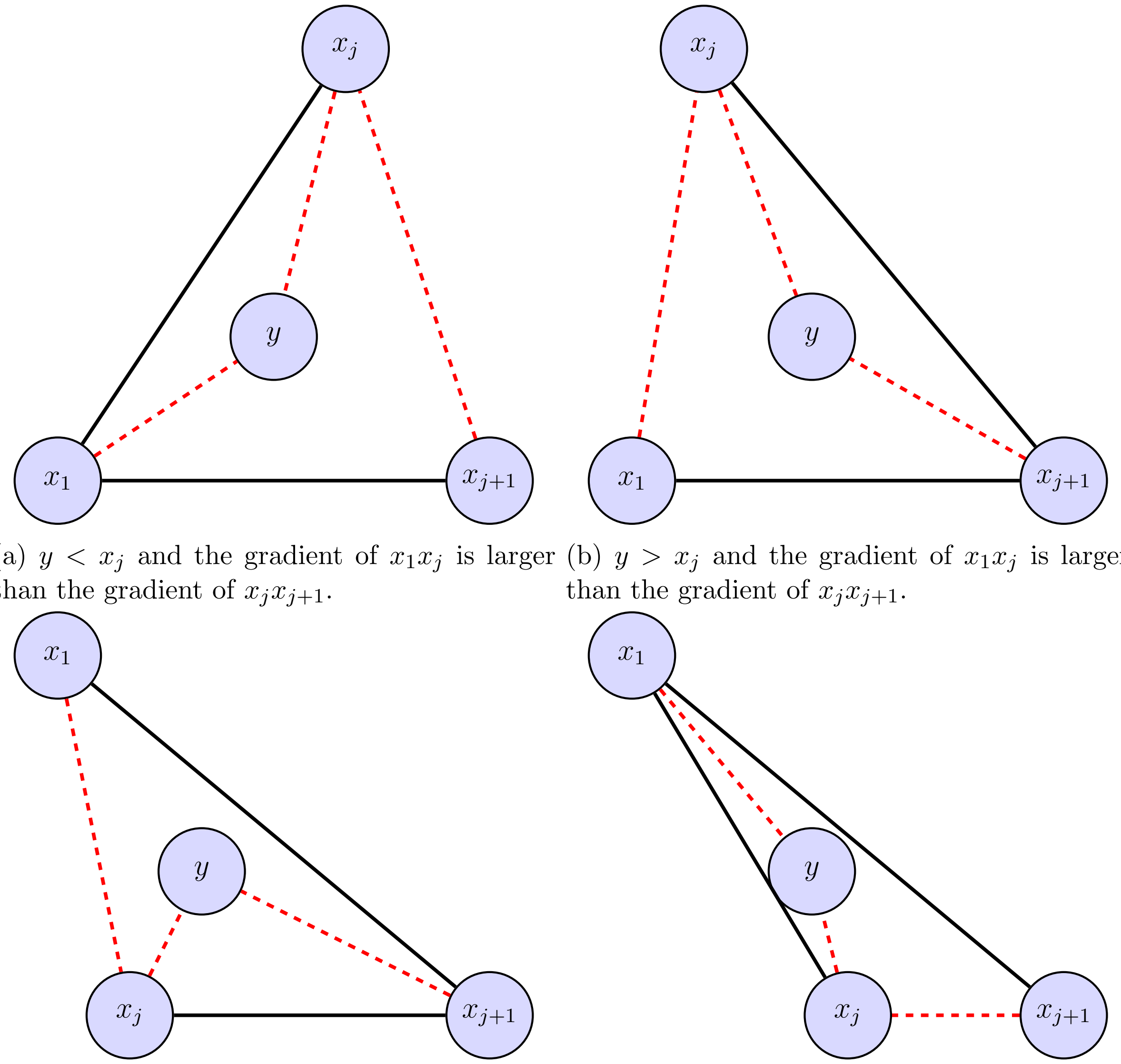
From the proof using Ramsey's theorem we found that $R_3(n,n)$ and $R_4(n,5)$ are upper bounds as well. From Equation (xxx) we find that

$$R_3(n,n) \leq (2^3)^{(2^2)^{2n-1}}$$

and similarly

$$R_4(n,5) \leq (2^4)^{(2^3)^{(2^2)^{n-2}}}.$$

Our upper bound for $R_3(n,n)$ is better than the bound for $R_4(n,5)$, but this is still much larger than $\binom{2n-4}{n-2} + 1$. So we see that in this case our proofs using Ramsey's theorem give way larger upper bounds than the proof that did not use Ramsey's theorem.



(a) $y < x_j$ and the gradient of $x_1 x_j$ is larger than the gradient of $x_j x_{j+1}$.

(b) $y > x_j$ and the gradient of $x_1 x_j$ is larger than the gradient of $x_j x_{j+1}$.

(c) $y < x_j$ and the gradient of $x_1 x_j$ is smaller than the gradient of $x_j x_{j+1}$.

(d) $y > x_j$ and the gradient of $x_1 x_j$ is smaller than the gradient of $x_j x_{j+1}$.

Figure 3: The four possibilities for the position of y. The dashed red lines are the ordering of the numbers. In every case we have both an increasing gradient and a decreasing gradient. Note that even though x_1 does not have to directly precede y or x_j it still makes it impossible for the n points to all be ordered convexly or concavely.

7 Conclusion

In this report we have seen various proofs of Ramsey's theorem which all gave their own upper bound. We have also seen two applications of Ramsey's theorem.

In both cases, however, the proof using Ramsey's theorem gives way larger upper bounds than the proof that did not use Ramsey's theorem.

For the monotone subsequence problem, the minimum size of a sequence such that we either have a subsequence of size m that is monotonically increasing or a subsequence of size n that is monotonically decreasing is $(n-1)(m-1)+1$. Ramsey's proof however gave $R_2(n,m)$ as upper bound. Even the lower bound of $R_2(n,m)$ grows exponentially, while the actual solution only grows quadratically.

For the Happy Ending problem we found that $\binom{2n-4}{n-2}+1$ is an upper bound for the number of points we need in a plane to guarantee a convex n-gon. This solution grows about as fast as the upper bound for $R_2(n,n)$, but both proofs using Ramsey's theorem needed higher order Ramsey numbers. It is evident that the bounds for the higher order Ramsey numbers far exceed our binomial coefficient.

Erdős and Szekeres conjectured that you would need a minimum of $2^{n-2}+1$ is the least number of points you need to guarantee a convex n-gon. They conjectured this based on the minimum number of points needed for convex 3-gons (3), 4-gons (5) and 5-gons (9). Very recently, in 2006, it was found that you need 17 points in the plane to always obtain a convex 6-gon, which is still in line with the conjecture.

In both problems, the fact that Ramsey's theorem is so broadly applicable is both a blessing and a curse. The theorem is incredibly useful in proving that certain things exist, in the case of the monotone subsequence problem the length of the sequence such that we would always find a monotone subsequence of the desired length and in the case of the Happy Ending problem a number of points in the plane such that we can always find a convex n-gon.

However, because the theorem can be applied in all sorts of situations, many specific parts of the two different problems are simply not used when we use Ramsey's theorem. This in turn gives very weak bounds, which only become weaker when we also have to estimate the Ramsey numbers.

Lastly it is not very likely that we will ever find an explicit formulae for the Ramsey numbers. The bound from Erdős and Rado, which was found in 1952, roughly 25 years after Ramsey published his paper has only very slightly improved in the 68 years after that. To paraphrase a statement made by Erdős:

If a vastly more powerful alien force comes to earth and demands the value of $R_2(5,5)$ or else they destroy the planet, it would be wise to use every mathematician and computer available to find it. If they instead want to know $R_2(6,6)$, it is probably easier to try to destroy the aliens.

References

[1] P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London mathematical society*, 2:417–439, 1952.

[2] P. Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[3] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2013.

[4] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 1928.

[5] A. Seidenberg. A simple proof of a theorem of Erdős and Szekeres. *The Journal of the London Mathematical Society*, 34:352, 1959.

[6] Th. Skolem. Ein kombinatorischer Satz mit Anwendung auf ein logisches Entscheidungsproblem. *Fundamenta Mathematicae*, 20:245–261, 1933.