Ramsey Theory on Graphs

Hernán Ibarra Mejia

November 28, 2023

Chapter 1

Introduction and Classical Results

This is a set of lecture notes taken by me from the Part III course "Ramsey Theory on Graphs", lectured by Dr J. Sahasrabudhe in Michaelmas, 2023. I take full responsibility for any mistakes in these notes.

We fix some notation first.

- Define, for all positive integers n,

$$[n] \coloneqq \{1, 2, \ldots, n\}.$$

- K_n denotes the complete graph on n vertices, i.e.

$$V(K_n) = [n] \text{ and } E(K_n) = \{\{i, j\} \colon i \neq j \text{ and } i, j \in [n]\}.$$

- If X is a set and $r \in \mathbb{N}$, let

$$X^{(r)} \coloneqq \{S \subseteq X \colon |S| = r\}.$$

- Functions $\chi \colon E(K_n) \to [k]$ are called k-colourings of (the edges of) K_n. If $k = 2$ we sometimes instead give names to the colours, i.e. we consider functions $E(K_n) \to \{\text{red}, \text{blue}\}$.

1 The Ramsey Numbers

1.1 Definitions

Here is the big picture question

What can we say about the structure of an arbitrary 2-colouring of $E(K_n)$?

A priori it seems that nothing can be said: an arbitrary 2-colouring is disordered and random. However, we will see that order arises even in these circumstances.

Say that a colouring $\chi\colon E(K_n) \to [r]$ contains a monochromatic clique[1] of size k if there is a subgraph G of K_n, isomorphic to K_k, such that $\chi|_{E(G)}$ is constant. In other words, inside K_n there is a smaller complete graph K_k that χ colours with only colour. When $k = 2$ sometimes we speak of the colouring χ containing a red (or blue) clique or size k; this just means that χ contains a monochromatic clique G so that $\chi|_{E(G)}$ is constant with value red (respectively blue). Usually we will say a colouring of K_n when we mean a colouring of $E(K_n)$, for the sake of simplicity (all of our colourings will be edge-colourings).

Consider the following example.

Example 1. In any party of six people, there are three people who are (pairwise) already acquainted or are (pairwise) meeting for the first time. Indeed, call one of the people A. Then there are five other people at the party who either A knows already or A is meeting for first time. By the pigeonhole principle, A knew 3 people or A didn't know 3 people; call them X, Y and Z. Without loss of generality, assume A knew them (the other case is similar).

Then if there are two people from the set $\{X, Y, Z\}$ that know each other then, together with A, they would for a group of three people that know each other. Otherwise, X, Y and Z are mutual strangers and we are done.

Examine this argument. We have shown that a complete graph on six vertices always has a monochromatic clique of size three. Remarkable, isn't it?

So does this always happen? Can we always find a monochromatic clique of size k in a colouring of K_n? If n is large enough, Ramsey says yes. Define the *diagonal Ramsey numbers* as

$$R(k) \coloneqq \min\{n \in \mathbb{Z}^+ \colon \text{All 2-colourings of } K_n \text{ have a monochromatic clique of size } k\},$$

for all $k \in \mathbb{Z}^+$. We showed that all colourings of K_6 contain a monochromatic clique of size 3, i.e. that $R(3) \leq 6$. We actually have $R(3) = 6$. To see this, note that if all colourings of K_n contain a monochromatic clique of size k then all colourings of K_m contain a monochromatic clique of size k for all $m \geq n$ (just remove $m - n$ vertices and apply the hypothesis). Contrapositively, if a colouring of K_5 does not contain a clique of size 3 then nor does a colouring of K_n for all $n \in [4]$. And indeed this is the case, as the following Figure shows. It follows that $R(3) = 6$.

Figure: K_5 coloured with no monochromatic clique of size 3

Before we move one, we will generalize our definition slightly. For positive integers k and l, define the *Ramsey numbers* to be

$$R(l, k) = \min\{n \in \mathbb{Z}^+ \colon \text{All 2-colourings of } K_n \text{ contain either a blue clique of size } l \text{ or a red clique of size } k\}.$$

[1] "Clique" is another word for a complete graph.

Clearly $R(l,k) = R(k,l)$ by symmetry and $R(k,k) = R(k)$ for all l, k.

1.2 Ramsey's Theorem

If you haven't noticed already, you should know that there is a problem with out definition of Ramsey numbers and diagonal Ramsey numbers. Both are defined as a minimum of a subset of the naturals, yet we did not prove that the subsets are non-empty. In other words, how can we be sure that there isn't some k such that for all n we have that there is a colouring of K_n with no monochromatic k-clique? (similarly for the more general Ramsey numbers). Thus, our first task will be to verify that these numbers are well-defined. It suffices to focus on the general Ramsey numbers, since the diagonal Ramsey numbers are special case of them.

We will do so recursively: first we show that $R(l,k)$ well-defined for small values of k and l, and then we will use a recursive upper bound to show that all this is true for all values of k and l.

Obviously $R(1,1) = 1$. Furthermore, $R(2,1) = R(1,2) = 1$ which you can easily check. Finally, $R(2,2) = 2$. This is enough to show the following.

Theorem 2 (Ramsey's Theorem)**.** *For all positive integers k and l, the Ramsey number $R(l,k)$ is well-defined. Furthermore, if $l, k \geq 2$ we have*

$$R(l,k) \leq R(l-1,k) + R(l,k-1).$$

Proof. Induction on $l + k$. We know $R(l,k)$ is well-defined for $l + k = 2$ so the base case (when $l = k = 1$) is done. By inductive hypothesis both $a := R(l-1,k)$ and $b = R(l,k-1)$ are well-defined. Let $n := a + b$. We will show that any colouring of a graph with n vertices has either a blue l-clique or a red k-clique. This will prove that $R(l,k)$ is well-defined and further $R(l,k) \leq n$, which is what we want.

So, let $\chi\colon E(K_n) \to \{\text{red}, \text{blue}\}$ be a colouring and let x be an arbitrary vertex of K_n. By $N_B(x)$ we denote the set of vertices that are adjacent to x by a blue edge, and similarly for $N_R(x)$.

Note that $|N_B(x)|+|N_R(x)|$ is just the degree of x, i.e. $n-1$. It follows that either $|N_B(x)|\geq a$ or $|N_R(x)|\geq b$ for if we had both $|N_B(x)|\leq a-1$ and $|N_r(x)|\leq b-1$ then

$$n - 1 = |N_B(x)|+|N_R(x)|\leq (a-1) + (b-1) = n-2,$$

a contradiction. There are two cases then.

Case 1

The number of blue neighbours of x is at least a, i.e. $|N_B(x)|\geq a$. By definition of a, the subgraph induced by $|N_B(a)|$ contains a red k-clique or a blue $(l-1)$-clique. In the former case we are done: we have found a red k-clique. In the latter case, we note that if we adjoin the vertex x to the blue $(l-1)$-clique in $N_B(x)$ we get a blue l-clique in K_n.

Case 2

The number of red neighbours of x is at least b. This case is done completely analogously to Case 1. $\square$

2 Bounds on Ramsey Numbers

2.1 Upper Bounds

Note that $R(l, 2) = l$ and $R(2, k) = k$ for all l and k. We can use the recursive bound to give us a non-recursive bound.

Theorem 3 (Erdős-Szekeres, 1935)**.** *For all positive integers l, k we have*

$$R(l, k) \leq \binom{l+k-2}{l-1}$$

Proof. Again, we do induction on $l + k$. The base case is trivial. By Theorem 2 and the inductive hypothesis

$$\begin{aligned} R(l, k) &= R(l-1, k) + R(l, k-1) \\ &\leq \binom{l+k-3}{l-2} + \binom{l+k-3}{l-1} \\ &= \binom{l+k-2}{l-1}, \end{aligned}$$

where in the last equality we used the identity $\binom{n+1}{m+1} = \binom{n}{m+1} + \binom{n}{m}$. $\square$

Corollary 4. *For all positive integers k,*

$$R(k) \leq \binom{2k-2}{k-1} \leq c\frac{4^k}{\sqrt{k}}$$

for some constant $c > 0$.

Proof. The first inequality is just Theorem 3 for the case $l = k$. For the second one we can approximate using Stirling's formula or by clever formulae-manipulation, see e.g. this link. $\square$

So roughly we can bound $R(k)$ by something that grows like 4^k. Is this close to how $R(k)$ really grows?

2.2 Lower bounds

For some time people thought 4^k was far from the truth and that the growth of $R(k)$ should be polynomial. Indeed, it is easy to get a quadratic lower bound for $R(k)$. The rough idea is to take $(k-1)$ copies of a blue K_{k-1} and connect all the remaining edges in red such that we end up with a $K_{(k-1)^2}$ with no

monochromatic k-clique. Erdős surprised everyone by providing an exponential lower bound, and by furthermore introducing a non-standard (back then) proof technique, now known as the "probabilistic method".

We will use binomial coefficients quite a bit, so the following bounds will be useful.

Proposition 5. *For all $n, k \in \mathbb{N}$ with $k \leq n$ we have*

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$$

Proof. For the lower bound we note that

$$\begin{aligned}\binom{n}{k} &= \frac{n}{k} \cdot \frac{n-1}{k-1} \cdots \frac{n-k+1}{1} \\ &= \prod_{m=0}^{k-1} \frac{n-m}{k-m} \qquad (1.1)\end{aligned}$$

We would like to show that all of the factors in the product are bounded below by n/k. Indeed, as $k \leq n$ we have $km \leq nm$ for all $0 \leq m \leq k-1$, and so $km - nk \leq nm - nk$. Rewriting the latter we get $k(m-n) \leq n(m-k)$ which, negating both sides and reordering we get

$$\frac{n-m}{k-m} \geq \frac{n}{k} \text{ for all } 0 \leq m \leq k-1.$$

This implies the lower bound.

For the upper bound take the representation in equation (1.1). As $n-m \leq n$ for $0 \leq m \leq k-1$ we get

$$\binom{n}{k} \leq \frac{n^k}{k!}$$

Now notice that

$$e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!} \geq \frac{k^k}{k!}.$$

Therefore $\frac{1}{k!} \leq \left(\frac{e}{k}\right)^k$ and so we get

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)$$

as desired. □

Now we can prove the result by Erdős. I know the bound seems intimidating but the idea of the proof is simple enough: just pick a random colouring and analyse what is the probability of getting (or not getting) the colouring you want. The bound follows naturally from this approach.

Theorem 6 (Erdős, 1948)**.** *For all $0 < \varepsilon < 1$ there is some k large enough so that*

$$R(k) \geq (1-\varepsilon)\frac{k}{e\sqrt{2}}2^{k/2}.$$

Proof. Let $0 < \varepsilon < 1$. We will later see how big k needs to be, so for now assume it to be arbitrary. Define

$$n := \left\lfloor (1-\varepsilon)\frac{k2^{k/2}}{e\sqrt{2}} \right\rfloor.$$

(Again, this n looks like it comes out of nowhere but it is picked such that most everything we get at the end is cancelled out). Let χ be a random red/blue colouring of $E(K_n)$, i.e. each edge has equal probability of being red and blue, and furthermore the choices for colours between different edges are independent. We will show that

$$P := \mathbb{P}(\chi \text{ contains a monochromatic } k\text{-clique}) < 1.$$

This implies that there *is* a colouring of $E(K_n)$ such that there is no monochromatic k-clique, i.e. that $R(k) > n$.

By definition,

$$P = \mathbb{P}\left(\bigcup_{S \in [n]^{(k)}} \{S \text{ is monochromatic}\}\right)$$

We can apply the union bound, which says that the "probability of the union is at most the sum of the probabilities". For each $S \in [n]^(k)$ the probability that it is monochromatic is the sum of the probability that it is completely red and the probability that it is completely blue; both of these cases have probability $\frac{1}{2^{\binom{k}{2}}}$. Thus we have

$$P \leq \binom{n}{k} 2^{1-\binom{k}{2}}$$

Using the standard bound for the binomial we get

$$\begin{aligned} P &\leq 2\left(\frac{en}{k}\right)^k \cdot 2^{\frac{-k(k-1)}{2}} \\ &= 2\left(\frac{e\sqrt{2}n}{k}2^{-k/2}\right)^k \\ &\leq 2(1-\varepsilon)^k, \end{aligned}$$

where we used the definition of n in the last inequality. Hence it is clear that P can be made arbitrarily small as for large enough k; so definitely we can pick k with $P < 1$. Thus, we have determined that $R(k) > n$ which implies, since $R(k)$ is an integer, that

$$R(k) \geq (1-\varepsilon)\frac{k2^{k/2}}{e\sqrt{2}}.$$

□

In essence we have proven that, very roughly, $(\sqrt{2})^k \leq R(k) \leq 4^k$. What is the right base in the exponent? This is the big open question in Ramsey Theory.

Chapter 2

On $R(3, k)$

Using the bounds on the previous chapter, we quickly deduce that

$$R(3, k) \leq \binom{k+1}{2} \leq (k+1)^2.$$

A quadratic growth happens to be indeed very close to the truth. We will try to get a quadratic lower bound in the next section.

1 (Ab)using the the probabilistic method

As we are going to keep the value of l fixed at 3 we will change how we think about colourings. Note that to find a lower bound $f(k) \leq R(3, k)$ it suffices to prove that a graph G with $f(k)$ vertices exists that does not contain a triangle nor an independent set of size k. For if we had such a G then we can construct a colouring of $K_{f(k)}$ that colours a copy of G in $K_{f(k)}$ completely blue and all other edges red. Such a colouring will not have a blue triangle, since G contains no triangles, and no red k-clique since that would imply G has an independent k-set. Hence $f(k) \leq R(3, k)$ as desired.

To find G we could try to use the probabilistic method as we have before. That is, having $f(k)$ vertices, connect them randomly, with each edge having probability $1/2$ to appear. Then we would show that such a random graph (up to certain modification, cf. the actual proof) has a non-zero probability of having no triangles and no independent k-sets. The problem is that the probability turns out to be so close to zero it is difficult to prove that it must be non-zero using approximations like e.g., the union bound. The solution is to consider a random graph that is more likely to have fewer edges but not too few so as to have independent k-sets. This motivates the following definition.

Definition 7 (Binomial Random Graph)**.** For $n \in \mathbb{N}$ and $p \in [0, 1]$ we define the probability space $G(n, p)$ to be that on all graphs $[n]$ where each edge is independently included with probability p.

So far we have considered $G(n, 1/2)$. With this more general definition, it is possible to prove an OK lower bound. If G is a graph, denote by $\alpha(G)$ the maximum k so that G contains an independent set of size k.

Theorem 8 (Erdős)**.** *There exists some constant $c > 0$ such that for all k we have*

$$R(3, k) \geq c \left(\frac{k}{\log k} \right)^{3/2}.$$

Proof. Define $n := \left(\frac{k}{\log k}\right)^{3/2}$ and $p := n^{-2/3} = \frac{\log k}{k}$. Forget about these values for now since it is not clear why they were chosen.

Sample a graph G from the distribution $G(n, p)$ (henceforth we abbreviate this as $G \sim G(n, p)$). Define $\tilde{G}$ to be the graph G but deleting all vertices that are part of a triangle or an independent set of size k. Then $\tilde{G}$ has no triangles and $\alpha(\tilde{G}) < k$ by construction. How many vertices are there left in $\tilde{G}$? We will show that

$$\mathbb{E}(|\tilde{G}|) \geq cn,$$

for some constant $c > 0$. Denote by X the number of triangles in G and Y the number of independent sets of size k in G. Then we clearly have

$$|\tilde{G}| \geq n - X - Y.$$

This is not necessarily equality because we can get "lucky" in that e.g., a vertex may be both part of a triangle and of an independent k-set, so we only delete it once (we can't delete it twice). The bound $n - X - Y$ is the worst possible scenario. We can take expectations on both sides of the inequality and use linearity to conclude

$$\mathbb{E}(|\tilde{G}|) \geq n - \mathbb{E}(X) - \mathbb{E}(Y). \tag{2.1}$$

We will now obtain bounds on $\mathbb{E}(X)$ and $\mathbb{E}(Y)$.

Observe that

$$X = \sum_{T \in [n]^{(3)}} \mathbf{1}_{\{T \text{ forms a triangle in } G\}}.$$

Take expectations:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{T \in [n]^{(3)}} \mathbb{E}(\mathbf{1}_{\{T \text{ forms a triangle in } G\}}) \\
&= \sum_{T \in [n]^{(3)}} \mathbb{P}(\{T \text{ forms a triangle in } G\}) \\
&= \sum_{T \in [n]^{(3)}} p^3 = \binom{n}{3} p^3.
\end{aligned}$$

It immediately follows that $\mathbb{E}(X) \leq \frac{1}{6}(np)^3$. A similar method works for Y. First, write it as a sum of indicator functions:

$$Y = \sum_{S \in [n]^{(k)}} \mathbf{1}_{\{S \text{ is an independet } k\text{-set in } G\}}.$$

Take expectations and we get

$$\begin{aligned}
\mathbb{E}(Y) &= \binom{n}{k}(1-p)^{\binom{k}{2}} \\
&\leq \left(\frac{en}{k}\right)^k (1-p)^{\binom{k}{2}} \\
&\leq \left(\frac{en}{k}\right)^k e^{-p\binom{k}{2}} \\
&= \left(\frac{en}{k} e^{-p\left(\frac{k-1}{2}\right)}\right)^k \\
&= \left(\frac{en}{k} e^{\frac{-pk}{2}} e^{\frac{p}{2}}\right)^k
\end{aligned}$$

where the inequalities came from the standard bound on the binomial and the fact that $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$. As $0 < p < 1$ it follows that $e^{p/2} < \sqrt{e} < 2$. Hence,

$$\mathbb{E}(Y) \leq \left(\frac{2en}{k} e^{\frac{-pk}{2}}\right)^k.$$

Therefore we get

$$\mathbb{E}(|\tilde{G}|) \geq n - \frac{1}{6}(np)^3 - \left(\frac{2en}{k} e^{\frac{-pk}{2}}\right)^k$$

It is clear now that we need $(np)^3$ to grow like n; hence why the value of p was chosen. A similar heuristic was used to choose the value of n.

In any case, replacing the value of p (both in terms of n and in terms of k) yield

$$\mathbb{E}(|\tilde{G}|) \geq n - \frac{n}{6} - \left(\frac{2en}{k^{3/2}}\right)^k$$

Replacing the value of n in terms of k yields

$$\mathbb{E}(|\tilde{G}|) \geq \frac{5n}{6} - \left(\frac{2e}{(\log k)^{3/2}}\right)^k.$$

And so it is clear that for large enough k we have $\left(\frac{2e}{(\log k)^{3/2}}\right)^k$ being arbitrarily small, so definitely $\mathbb{E}(|\tilde{G}|) \geq cn$ for some $c > 0$ and for large enough k. Thus, if k is large enough, there is a graph $\tilde{G}$ with the desired properties and at least cn vertices and so

$$R(3, k) \geq cn$$

for large enough k. To get the statement for all k decrease c while keeping it positive until it applies for all k. □

This was the "vertex deletion method": we deleted some vertices from the random graph to get one with the desired properties. However this is not efficient: our bound worsens when we delete vertices.

What about deleting edges instead? This is much harder since one may inadvertently create independent sets while trying to delete triangles. Our idea will be to delete edges from collections of edge-disjoint triangles in $G \sim G(n,p)$ and remove all edges from these triangles.

Theorem 9 (Erdős)**.** *There exists some $c > 0$ such for all k we have*

$$R(3,k) \geq c\left(\frac{k}{\log k}\right)^2$$

To prove this, we will need a couple of lemmata.

Lemma 10. *Let $\mathcal{F} = \{A_1, \ldots, A_m\}$ be a family of events in a probability space. Let $\mathcal{E}_t$ be the event that at least t independent events from $\mathcal{F}$ occur. Then*

$$\mathbb{P}(\mathcal{E}_t) \leq \frac{1}{t!}\left(\sum_{i=1}^{m} \mathbb{P}(A_i)\right)^t.$$

Proof. Note that

$$\mathbf{1}_{\mathcal{E}_t} = \frac{1}{t!} \sum_{\substack{(i_1,\ldots,i_t)\in[m]^t,\\ A_{i_1},\ldots,A_{i_t}\text{ are}\\ \text{distinct independent}\\ \text{events}}} \mathbf{1}_{A_{i_1}}\mathbf{1}_{A_{i_2}}\cdots\mathbf{1}_{A_{i_t}}.$$

Take expectations on both sides. Note that the equality turns into an inequality if we delete the long condition on the sum.

$$\begin{aligned}\mathbb{P}(\mathcal{E}_t) &\leq \frac{1}{t!} \sum_{(i_1,\ldots,i_t)\in[m]^t} \mathbb{P}(A_{i_1})\cdots\mathbb{P}(A_{i_t}) \\ &\leq \frac{1}{t!}\left(\sum_{i=1}^{m}\mathbb{P}(A_i)\right)^t \qquad \square\end{aligned}$$

From now on we say a property P holds of $G \sim G(n,p)$ with high probability (sometimes abbreviated as w.h.p) if

$$\mathbb{P}(G \text{ has property } P) \to 1 \text{ as } n \to \infty.$$

Lemma 11. *Let $n, k \in \mathbb{Z}^+$ and $p \in (0,1)$ be such that $pk \geq 16\log n$. Then with high probability every subset of size k of $G \sim G(n,p)$ contains at least $\frac{pk^2}{8}$ edges.*

Assume the preceding lemma for a moment. Then we can prove the result by Erdős.

Proof of Theorem 9. Here I can motivate the choices of n and p a bit more. Note that the expected number of triangles is $p^3\binom{n}{3}$ which is asymptotically

p^3n^3, while the number of edges is asymptotically n^2p. So that the expected number of edges is the same as the expected number of triangles we would need p to grow like $n^{-1/2}$. Indeed, we would like to delete just enough edges to get rid of all of the triangles; otherwise we might get large independent sets.

Now we begin the proof. Let

$$n = \left(\frac{c_1 k}{\log k}\right)^2 \quad \text{and} \quad p = c_2 n^{-1/2} = \frac{c_2}{c_1}\frac{\log k}{k},$$

where $c_1, c_2 > 0$ are constants to be chosen later. Take $G \sim G(n,p)$ and let $\mathcal{T}$ be a maximal collection of edge-disjoint triangles in G. Define $\tilde{G}$ to be G with all edges contained in a triangle of $\mathcal{T}$ removed. Then $\tilde{G}$ has no triangles. Indeed, if it had a triangle then, by maximality of $\mathcal{T}$, it would have to be a triangle that shares an edge with a triangle in $\mathcal{T}$; but then this triangle would disappear in $\tilde{G}$ (even though it is not part of $\mathcal{T}$).

We will show that

$$\mathbb{P}(\alpha(\tilde{G}) \geq k) < 1$$

since this suffices.

Let Q be the event that every set of k vertices of G contains at least $\frac{pk^2}{8}$ edges. By Lemma 11 (with the right choice of constants—we verify this later) we know that $\mathbb{P}(Q) = 1 - o(1)$ where $o(1) \to 0$ as $n \to \infty$.

Now note that

$$\mathbb{P}(\alpha(\tilde{G}) \geq k) \leq \mathbb{P}(\{\alpha(\tilde{G}) \geq k\} \cap Q) + \mathbb{P}(Q^{\mathrm{c}}),$$

where $\mathbb{P}(Q^{\mathrm{c}})$ tends to zero. Hence it suffices to show that $\mathbb{P}(\{\alpha(\tilde{G}) \geq k\} \cap Q)$ is strictly less than one.

But consider the event $\{\alpha(\tilde{G}) \geq k\} \cap Q$. If this is the case there is some set $S \subset [n]^{(k)}$ of k vertices of G so that $\mathcal{T}$ removes at least $\frac{pk^2}{8}$ edges of S (so that S turns independent in $\tilde{G}$). In other words, $\mathcal{T}$ has at least $\frac{pk^2}{8\cdot 3}$ triangles meeting S in at least two vertices (here we use the assumption that triangles in $\mathcal{T}$ are edge-disjoint). As this is an existential statement (there exists S such that...) it is a union over all k-element subsets of $[n]$. Apply the union bound and we get that $\mathbb{P}(\alpha(\tilde{G}) \geq k)$ is at most $\binom{n}{k}P + o(1)$ where P is defined to be

$$\mathbb{P}\left(\mathcal{T} \text{ has at least } \frac{pk^2}{24} \text{ triangles meeting } [k] \text{ in at least 2 vertices}\right).$$

Let $\{T_i\}_i$ be the collection of triangles in K_n that meet $[k]$ in at least two vertices, and let $A_i \coloneqq \{T_i \text{ is in } G\}$. Note that if $T_{i_1}, \ldots, T_{i_k}$ are edge-disjoint then $A_{i_1}, \ldots, A_{i_k}$ are independent.

Next, let $t = \lceil \frac{pk^2}{24} \rceil$ and take $\mathcal{E}_t$ to be as in Lemma 10, where $\mathcal{F} = \{A_i\}_i$.

Then we clearly have

$$\begin{aligned} P \leq \mathbb{P}(\mathcal{E}_t) &\leq \frac{1}{t!}\left(\sum_i \mathbb{P}(T_i \subseteq G)\right)^t \\ &= \frac{1}{t!}\left(\sum_i p^3\right)^t \\ &= \frac{1}{t!}\left(\binom{k}{2}(n-2)p^3\right)^t \\ &\leq \frac{1}{t!}(k^2np^3)^t. \end{aligned}$$

Notice that $t! > \left(\frac{t}{e}\right)^t$ as can be easily seen by the Taylor expansion of e^t. Hence,

$$\begin{aligned} P &\leq \left(\frac{ek^2np^3}{t}\right)^t \\ &= (24enp^2)^t \\ &= (24ec_2^2)^t. \end{aligned}$$

Thus, choosing c_2 to be $e^{-1}24^{-1/2}$ we get that $P \leq e^{-t}$. We might as well choose the other constant before finishing off.

Lemma 11 needs us to verify that $pk \geq 16 \log n$. We claim that $c_1 := \frac{c_2}{48}$ suffices. Indeed,

$$pk = \frac{c_2}{c_1}\log k \geq 48 \log k > 16 \log n$$

since $n < k^2$ by definition.

Finally, we had

$$\begin{aligned} \mathbb{P}(\alpha(\tilde{G}) \geq k) &\leq \binom{n}{k}P + o(1) \\ &\leq \binom{n}{k}e^{-t} + o(1) \\ &\leq \left(\frac{en}{k}\right)^k n^{-k} + o(1), \end{aligned}$$

where we have used the fact that

$$t \geq \frac{pk^2}{24} \geq 2k \log k \geq k \log n.$$

It follows that

$$\mathbb{P}(\alpha(\tilde{G}) \geq k) \leq \left(\frac{e}{k}\right)^k + o(1) \to 0$$

as k is large. We can get the statement for all k by multiplying by a small constant. $\square$

2 Large Deviation inequalities

Let Z be a Gaussian random variable. Then,

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq t) \leq e^{\frac{-t^2}{2\operatorname{Var} Z}}.$$

Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli random variables. We done this by $X_i \sim \operatorname{Bern}(p)$. Let $S_n = \sum_{i=1}^n X_i$. Morally, often the tails of random variables S_n look like Gaussian tails. Note that $\mathbb{E}(S_n) = np$ and $\operatorname{Var}(S_n) = p(1-p)n$

Theorem 12 (Chernoff inequality)**.** *Let $X_1, \ldots, X_n \sim \operatorname{Bern}(p)$ and take S_n to be their sum. Then*

$$\mathbb{P}(|S_n - pn| \geq t) \leq 2\exp\left(\frac{-t^2}{2pn} + \frac{t^3}{(pn)^2}\right)$$

We take this for granted. Then we can prove Lemma 11.

Proof of Lemma 11. Consider the probability P that there is some subset of size k of $G \sim G(n,p)$ that contains strictly less than $\frac{pk^2}{8}$ edges. As this is an existential statement we can apply the union bound, so

$$P \leq \binom{n}{k}\mathbb{P}\left(e(G[k]) < \frac{pk^2}{8}\right),$$

where $e(G[k])$ denotes the number of edges in the subgraph of G induced by $[k]$. Let $P' := \mathbb{P}\left(e(G[k]) < \frac{pk^2}{8}\right)$ so that $P \leq \binom{n}{k}P'$. Note that $e(G[k])$ is a sum of $\binom{k}{2}$ i.i.d. Bernoulli random variables with parameter p. So

$$\mathbb{E}(e(G[k])) = p\binom{k}{2}$$

What follows? The proof given in lectures doesn't seem to work. □

3 The Local Lemma

Definition 13. Let $\mathcal{F} = \{A_1, \ldots, A_n\}$ be a family of events in a probability space. A *dependency graph* Γ for $\mathcal{F}$ is a graph with a vertex for each even of $\mathcal{F}$ such that for all $i \in [n]$ the event A_i is independent of all events it's not adjacent to.

Remark 14. A dependency graph is not necessarily unique. For example, the complete graph is always a dependency graph.

Theorem 15 (The Local Lemma, Symmetric version)**.** *Let $\mathcal{F} = \{A_1, \ldots, A_n\}$ be a family of events in a probability space. Suppose Γ is a dependency graph for $\mathcal{F}$ with maximum degree Δ.*

If

$$\mathbb{P}(A_i) \leq \frac{1}{e(\Delta+1)},$$

then

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) > 0$$

By taking complements in the conclusion, we see that this is much better than the union bound, if the hypotheses are satisfied of course. Intuitively, we think of $\mathcal{F}$ as a family of "bad" events: the Local Lemma says that there is some (maybe small but positive) chance that no bad event happens. This is useful in situations where the probabilistic method is hard to apply. Indeed, the probabilistic method is good "for finding the hay in the haystack", i.e. if some event is likely to happen then we are good. The Local Lemma is good for finding the needle in the haystack, as we will see shortly.

Theorem 16 (Spencer)**.** *For all k we have*

$$R(k) \geq (1-o(1))\frac{\sqrt{2}}{e}k2^{k/2}$$

Proof. Let $\varepsilon > 0$ and define $n := \lfloor (1-\varepsilon)\frac{\sqrt{2}}{e}k2^{k/2}\rfloor$. Suppose χ is a random colouring of $E(K_n)$ (like before, we assume each edge is coloured red or blue with equal probability). Define for $S \in [n]^{(k)}$ the event

$$A_S := \{S \text{ is monochromatic in } \chi\}.$$

If we can show that there is a non-zero chance that none of the events in A_S happens then we are done. This is exactly the conclusion of the Local Lemma. If $\mathcal{F} := \{A_S\}_{S\in[n]^{(k)}}$ then define a dependency graph on $\mathcal{F}$ by declaring that A_S is adjacent to A_T iff they share at least one edge. In other words

$$A_S \sim A_T \iff 2 \leq |S \cap T| \leq k-1.$$

Simple counting reveals that the maximum degree of Γ is given by

$$\Delta := \sum_{t=2}^{k-1} \binom{k}{t}\binom{n-k}{k-t}.$$

On the other hand we clearly have

$$\mathbb{P}(A_S) = 2^{1-\binom{k}{2}}$$

for all $S \in [n]^{(k)}$. Therefore the problem is reduced to checking that

$$2^{1-\binom{k}{2}} \leq \frac{1}{e(\Delta+1)}$$

TODO: Complete this check. It's nontrivial, but apparently the first term in the Δ sum dominates all other terms. We have to use the fact that $n >> k$ somehow.

□

How to prove the Local Lemma? It helps to consider a slight generalization, which also has applications on its own

Theorem 17 (Lopsided Local Lemma)**.** *Let $\mathcal{F} = \{A_1, \ldots, A_n\}$ be a family of events in a probability space. Suppose Γ is a dependency graph for $\mathcal{F}$. Let $0 \leq x_1, \ldots, x_n < 1$ be real numbers satisfying*

$$\mathbb{P}(A_i) \leq x_i \prod_{j : A_j \sim A_i \text{ in } \Gamma} (1 - x_j).$$

Then

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) \geq \prod_{i=1}^{n} (1 - x_i) > 0$$

Note that the restriction on $\mathbb{P}(A_i)$ gets stronger as A_i is dependent on more things. If A_i is independent of all other events then we can take it as an isolated vertex in Γ and pick $x_i = \mathbb{P}(A_i)$ (assuming this is strictly less than 1). We give one other application before proving the Local Lemmmata.

Theorem 18 (Erdős)**.** *There is some constant $c > 0$ such that for all k*

$$R(3, k) > c\left(\frac{k}{\log k}\right)^2$$

Proof. Let $\varepsilon > 0$ and set $n \coloneqq \lfloor \varepsilon^4 \left(\frac{k}{\log k}\right)^2 \rfloor$. Also, define $p \coloneqq \frac{\varepsilon}{\sqrt{n}}$. Consider $G \sim G(n, p)$. For all $T \in [n]^{(3)}$ define the event

$$A_T \coloneqq \{T \text{ is contained in } G\}.$$

Similarly, for all $I \in [n]^{(k)}$ define

$$B_I \coloneqq \{I \text{ is independent in } G\}.$$

We would like to apply the Lopsided Local Lemma to these events. Let $\mathcal{A} \coloneqq \{A_T\}_{T \in [n]^{(3)}}$ and $\mathcal{B} \coloneqq \{B_I\}_{I \in [n]^{(k)}}$. If $\mathcal{F} \coloneqq \mathcal{A} \cup \mathcal{B}$ then we can define a dependency graph on $\mathcal{F}$ by again stating that two events are adjacents if the corresponding graphs share at least one edge. More formally, we declare that

$$\begin{aligned} A_T \sim A_{T'} &\iff |T \cap T'| = 2 \\ A_T \sim B_I &\iff 2 \leq |I \cap T| \leq 3 \\ B_I \sim B_{I'} &\iff 2 \leq |I \cap I'| < k. \end{aligned}$$

We now note that any A_T has $\binom{3}{2}(n-3) = 3(n-3)$ neighbours in $\mathcal{A}$; so at most 3n neighbours in $\mathcal{A}$. Similarly, A_T has $\binom{3}{2}\binom{n-3}{k-2} + \binom{n-3}{k-3}$ neighbours in β, which is at most $4n^{k-2}$.

We do the same calculations for B_I. In $\mathcal{A}$ there are $\binom{k}{3} + \binom{k}{2}(n-k)$ neighbours of B_I, so at most $nk^2 + k^3$. Similarly, TODO: I can't get the bounds to work in this one $\square$

Proof of Lopsided Local Lemma. Note that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) = \prod_{i=1}^{n} \mathbb{P}(A_i^c \mid A_1^c \cap \cdots \cap A_{i-1}^c),$$

which can be seen by induction on n and applying the identity $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B \mid A)$. It follows that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) = \prod_{i=1}^{n} \left(1 - \mathbb{P}(A_i \mid A_1^c \cap \cdots \cap A_{i-1}^c)\right)$$

since we have the identity $\mathbb{P}(A \mid B) = 1 - \mathbb{P}(A^c \mid B)$. Then it suffices to show that

$$\mathbb{P}(A_i \mid A_1^c \cap \cdots \cap A_{i-1}^c) \leq x_i$$

for all i. So let $i \in [n]$. More generally, we prove that

$$\mathbb{P}\left(A_i \mid \bigcap_{j \in S} A_j^c\right) \leq x_i \tag{2.2}$$

for all $S \subseteq [n]$. This we will do by induction on S.

In the base case we have $S = \emptyset$ so (2.2) is true by hypothesis. Now we do the inductive step. Define

$$\mathcal{I} \coloneqq \bigcap_{j \in S : i \nsim j} A_j^c \quad \text{and} \quad \mathcal{D} \coloneqq \bigcap_{j \in S : i \sim j} A_j^c.$$

Note that A_i and $\mathcal{I}$ are independent events. Then

$$\begin{aligned}
\mathbb{P}\left(A_i \mid \bigcap_{j \in S} A_j^c\right) &= \mathbb{P}(A_i \mid \mathcal{I} \cap \mathcal{D}) \\
&= \frac{\mathbb{P}(A_i \cap \mathcal{I} \cap \mathcal{D})}{\mathcal{I} \cap \mathcal{D}} \\
&\leq \frac{\mathbb{P}(A_i \cap \mathcal{I})}{\mathcal{I} \cap \mathcal{D}} \\
&= \frac{\mathbb{P}(A_i)\mathbb{P}(\mathcal{I})}{\mathcal{I} \cap \mathcal{D}} \\
&= \frac{\mathbb{P}(A_i)}{\mathbb{P}(\mathcal{D} \mid \mathcal{I})} \\
&\leq \frac{x_i \prod_{j : j \sim i}(1 - x_j)}{\mathbb{P}(\mathcal{D} \mid \mathcal{I})} \qquad (2.3)
\end{aligned}$$

We now write $\mathcal{D} = \bigcap_{k=1}^{m} A_{i_k}^c$. Hence, by a similar argument as the one we gave at the start,

$$\mathbb{P}(\mathcal{D} \mid \mathcal{I}) = \prod_{j=1}^{m} \left(1 - \mathbb{P}(A_{i_j} \mid \mathcal{I} \cap A_{i_1}^c \cap \cdots \cap A_{i_{j-1}}^c)\right).$$

We can apply the induction hypothesis to each factor (since they are conditioned on less than $|S|$ of the events A_i^c's). It follows that

$$\mathbb{P}(\mathcal{D} \mid \mathcal{I}) \geq \prod_{j=1}^{m} (1 - x_{i_j}) \geq \prod_{j:j\sim i} (1 - x_j).$$

Applying this to (2.3) finishes the induction. □

Proof of Symmetric Local Lemma. We use the Lopsided Local Lemma with $x_i = \frac{1}{\Delta+1}$ for all i. Note that

$$x_i \prod_{j:j\sim i} (1 - x_j) \geq \frac{1}{\Delta+1}\left(1 - \frac{1}{\Delta+1}\right)^{\Delta} \geq \frac{1}{e(\Delta+1)} \geq \mathbb{P}(A_i)$$

so the hypotheses indeed hold. □

4 Upper bounds

Thus far we have gotten pretty good lower bounds on $R(3,k)$ but no upper bounds apart from the naive quadratic bound at the beginning of the chapter.

Theorem 19 (Ajtai, Kamlós, Szemerédi)**.** *For all k we have*

$$R(3,k) < c\frac{k^2}{\log k}$$

First we prove the following theorem.

Theorem 20 (Ajtai, Kamlós, Szemerédi)**.** *Let G be a triangle-free n vertex graph with maximum degree Δ. Then there is some $c > 0$ (not depending on n nor Δ) such that G contains an independent set of size at least $c\frac{n}{\Delta}\log\Delta$. In other words,*

$$\alpha(G) \geq c\frac{n}{\Delta}\log\Delta$$

Note that if we drop the assumption that G is triangle-free then we can get a bound of the form $\frac{n}{\Delta}$. Indeed, perform a random greedy algorithm: at each stage pick a vertex not adjacent to the previous ones. The assumption of triangle-freeness means that the graph is more sparse and hence we get the extra $\log\Delta$ factor.

We will present two proofs of this theorem, the first due to Shearer[1]. His approach is to first define a function $f\colon [0,\infty) \to \mathbb{R}$ so that

$$f(x) \coloneqq \frac{x \ln x - x + 1}{(x-1)^2}$$

for $0 < x < 1$, with the boundary values $f(0) = 1$ and $f(1) = 1/2$. Note that

- f is continuous and (twice) differentiable for $x \geq 0$;
- $f(x) \in [0,1]$ for all x;
- $f'(x) \leq 0$ for all x;
- $f''(x) \geq 0$ for all x;
- f always lies above its tangent line, i.e., for all x, x_0 we have

$$f(x) \geq f(x_0) + (x - x_0) f'(x_0);$$

- f satisfies the following differential equation

$$(x+1)f = 1 + (x - x^2)f'$$

These facts are not hard to prove using elementary real analysis, but, since this is not a course in real analysis, we skip the details. Shearer then proved the following.

Lemma 21 (Shearer)**.** *Let G be a triangle-free graph on n vertices with average degree $d > 0$. Then*

$$\alpha(G) \geq n f(d).$$

Proof. Induction on n. For $n = 2$ we can only have $d = 1$ and so we need to show $\alpha(G) \geq 2\frac{1}{2} = 1$ which is obviously true; this finishes the base case.

Now let $x \in V(G)$ be a vertex to be chosen later. Define $G' \coloneqq G - x - N(x)$. By induction hypothesis we have

$$\alpha(G') \geq (n - d(x) - 1) f(d')$$

where d' is the average degree of G'. It follows that

$$\begin{aligned}\alpha(G) &\geq 1 + (n - d(x) - 1) f(d') \\ &\geq 1 + (n - d(x) - 1) f(d) + (d' - d) f'(d)(n - d(x) - 1)\end{aligned}$$

since f always lies above its tangent line. Then

$$\alpha(G) \geq n f(d) + 1 - (d(x) + 1) f(d) + (n(d' - d) - d(x)d' - d' + d + d(x)d) f'(d)$$

[1] Warning: Shearer's proof reads like "***king magic"

Note that

$$\begin{aligned}
n(d-d') &= n(\frac{2e(G)}{n} - \frac{2e(G')}{n-d(x)-1}) \\
&= 2e(G) - (n-d(x)-1)\frac{2e(G')}{n-d(x)-1} + (d(x)+1)d' \\
&= 2(e(G)-e(G')) + (d(x)+1)d' \\
&= 2(\sum_{y:y\sim x} d(y)) - (d(x)+1)d',
\end{aligned}$$

where in the last equality we just noted that $e(G)-e(G')$ is the number of edges lost when deleting x and $N(x)$ from G. □

Proof 1 of Theorem 20. TODO □

Proof 2 of Theorem 20. TODO □

Now we are ready to prove the desired upper bound on $R(3,k)$.

Proof of Theorem 19. content... □

Chapter 3

Hypergraph Ramsey Numbers

Recall that an r-uniform *hypergraph* G consists of a set of vertices $V(G)$ together with a collection of (hyper)edges $E(G)$ which are subsets of $V(G)$ of size r. The graphs we have been discussing thus far are 2-uniform hypergraphs.

We define $K_n^{(r)}$ to be the complete r-uniform hypergraph, i.e.,

$$V(K_n^{(r)}) := [n] \text{ and } E(K_n^{(r)}) := [n]^{(r)}.$$

We define the r-uniform *hypergraph Ramsey numbers* by

$$R^{(r)}(l,k) := \min\{n\colon \text{all red/blue colourings of } E(K_n^{(r)}) \text{ contain a blue } K_l^{(r)} \text{ or a red } K_k^{(r)}\}.$$

As before we define the r-uniform *hypergraph diagonal Ramsey numbers* by $R^{(r)}(k) := R^{(r)}(k,k)$.

1 Off-diagonal 3-uniform hypergraph Ramsey numbers

Theorem 22 (Erdős-Rado)**.** *For all $l, k \in \mathbb{N}$ we have*

$$R^{(3)}(l,k) \leq 2^{\binom{R(l-1,k-1)}{2}}.$$

Proof. TODO □

We would like to study $R^{(3)}(4,k)$ since, given the above bound, we can use our knowledge of $R(3,k)$. We get (approximately) something of the form $R^{(3)}(4,k) \leq 2^{ck^4}$. Can we do any better?

Theorem 23 (Conlon, Fox, Sudakov, 2010)**.** *For all k we have*

$$R^{(3)}(4,k) \leq k^{2k^2}.$$

Here is the strategy. The proof of the Erdős-Rado bound suggests the following two-player game played on graphs.

Definition 24 (The Ramsey Game)**.** TODO

Lemma 25. *In the Ramsey Game, the Builder has a strategy that forces a blue K_3 or a red K_{k-1} in at most k^3 moves. Furthermore, using this strategy at most k^2 of the edges revealed will be blue.*

Proof. content... □

Proof of Theorem 23. TODO □

Now we focus on lower bounds.

Theorem 26. *For all k we have*

$$R^{(3)}(4,k) \geq 2^{\frac{k-1}{2}}$$

Proof. TODO □

Theorem 27 (Conlon, Fox, Sudakov,2010)**.** *For sufficiently large k we have*

$$R^{(3)}(4,k) \geq k^{k/5}$$

Proof. TODO □

2 Diagonal Multicolour 3-uniform hypergraph Ramsey numbers

We will show one result in this direction. By $R_4^3(k)$ we mean the minimum number n such that every 4-colouring of $E(K_n^{(3)})$ contains a monochromatic $K_k^{(3)}$.

Theorem 28 (Erdős-Hajnal-Rado, 1965)**.** *There exists some $c > 0$ such that for all k we have*

$$R_4^{(3)}(k) \geq 2^{2^{ck}}.$$

This is clearly follows from the next lemma (and the lower bound we had from diagonal Ramsey numbers in the first chapter).

Lemma 29. *For all k we have*

$$R_4^{(3)} \geq 2^{R(k-1)-1}$$

Proof. TODO □

Chapter 4

The Szemerédi Regularity Lemma

Informally, Szemerédi Regularity Lemma says that for all $\varepsilon > 0$ and $l \in \mathbb{N}$ there exists some $L(\varepsilon, l) \in \mathbb{N}$ such that every graph G can be partitioned

$$V(G) = V_1 \sqcup \cdots \sqcup V_k$$

where $k \leq L(\varepsilon, l)$ such that

- $||V_i| - |V_j|| \leq 1$ for all i, j; and
- the graph between V_i and V_j looks random-like for all but $\varepsilon\binom{k}{2}$ pairs i, j.

We will make this "random-like" notion precise.

1 The Lemma

Definition 30 (ε-uniform)**.** Let $\varepsilon > 0$ and G be a graph with some disjoint $X, Y \subseteq V(G)$. We say that the pair (X, Y) is *ε-uniform* iff for all $X' \subseteq X$ and $Y' \subseteq Y$ with $X' \geq \varepsilon|X|$ and $Y' \geq \varepsilon|Y|$ we have

$$|d(X', Y') - d(X, Y)| \leq \varepsilon$$

where we define

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}.$$

Roughly, if (X, Y) is ε-uniform, the bipartite graph they induce is almost regular (ε-regular is a synonym for ε-uniform) in the sense that the density of edges between X and Y stays roughly unchanged if we throw out some vertices from X or Y (but not too many vertices).

Proposition 31. *Let $\varepsilon > 0$ and (X, Y) a ε-uniform pair in a graph G. If $p = d(X, Y)$, then*

$$|\{x \in X : |N(x) \cap Y| < (p - \varepsilon)|Y|\}| < \varepsilon|X| \text{ and}$$
$$|\{x \in X : |N(x) \cap Y| > (p + \varepsilon)|Y|\}| < \varepsilon|X|$$

Proof. TODO □

Theorem 32 (Embedding Lemma for triangles)**.** *Let $\varepsilon > 0$, $p \in (0, 1)$, and G be a graph. Suppose $V(G)$ can be partitioned by sets V_1,V_2, and V_3 of the same size m. Furthermore, assume that these sets are pairwise ε-uniform, and that $d(V_i, V_j) \geq 2\varepsilon$ for all distinct i, j. Then there are at least $(1 - 2\varepsilon)(p - \varepsilon)^3 m^3$ triangles in G.*

Proof. TODO □

Before stating the Szemerédi Regularity Lemma we need a couple more definitions.

Definition 33 (Equipartition)**.** For G a graph, we say that a partition

$$V(G) = V_1 \sqcup \cdots \sqcup V_k$$

is an *equipartition* if $||V_i| - |V_j|| \leq 1$ for all $i, j \in [k]$.

Definition 34. For $\varepsilon > 0$ and G a graph we say that an equipartition of G of size k is ε-uniform if the elements of the partition are pairwise ε-uniform for all but $\varepsilon\binom{k}{2}$ pairs.

Theorem 35 (Szemerédi Regularity Lemma)**.** *For all $\varepsilon > 0$ and $l \in \mathbb{N}$ there is some $L(l, \varepsilon) \in \mathbb{N}$ such that every graph has a ε-uniform equipartition with k parts where $l \leq k \leq L(l, \varepsilon)$.*

Since the proof of this is technical, we will just assume the result.

Remark 36. It is really important that $L(l, \varepsilon)$ does not depend on G. The lemma basically does not say anything about graphs with $e(G) = o(n^2)$.

2 Applications

Lemma 37 (Triangle Removal Lemma)**.** *For all $\varepsilon > 0$ there exists a $\delta > 0$ such that if G is a graph with fewer than δn^3 triangles, then we can remove εn^2 edges or less from G to make it triangle free.*

Proof. TODO □

As an application, we prove the celebrated Roth's Theorem.

Theorem 38 (Roth's Theorem)**.** *For all $\varepsilon > 0$ there exists some $n_0 \in \mathbb{N}$ so that for all natural numbers $n \geq n_0$ we have that if $A \subseteq [n]$ with $|A| \geq \varepsilon n$, then A contains a non-trivial three-term arithmetic progression.*

Proof. TODO □

Recall Turán's Theorem.

Theorem 39. *Let G be a K_{r+1}-free graph on n vertices, where $r \geq 2$. Then*

$$e(G) \leq \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$$

This is sharp for complete multipartite graphs, i.e. Turán graphs.

Proof. TODO □

Definition 40. For $\alpha \in (0, 1)$ and $\varepsilon > 0$, and a partition

$$V(G) = V_1 \sqcup \cdots \sqcup V_k$$

of a graph G, we define the *reduced graph* $R_{\alpha,\varepsilon}$ by

$$V(R_{\alpha,\varepsilon}) = \{V_1, \ldots, V_k\}$$

and declaring that $V_i \sim V_j$ iff $i \neq j$ and (V_i, V_j) is ε-uniform with $d(V_i, V_j) \geq \alpha$.

Theorem 41 (Szemerédi)**.** *For all $\varepsilon > 0$ there is some $\delta > 0$ such that if G is a K_4-free graph with $\alpha(G) < \delta n$, then*

$$e(G) \leq \frac{n^2}{8} + en^2$$

Proof. TODO □