Part III — Extremal Graph Theory

Based on lectures by A. G. Thomason
Notes taken by Dexter Chua

Michaelmas 2017

Turán's theorem, giving the maximum size of a graph that contains no complete r-vertex subgraph, is an example of an extremal graph theorem. Extremal graph theory is an umbrella title for the study of how graph and hypergraph properties depend on the values of parameters. This course builds on the material introduced in the Part II Graph Theory course, which includes Turán's theorem and also the Erdös–Stone theorem.

The first few lectures will cover the Erdös–Stone theorem and stability. Then we shall treat Szemerédi's Regularity Lemma, with some applications, such as to hereditary properties. Subsequent material, depending on available time, might include: hypergraph extensions, the flag algebra method of Razborov, graph containers and applications.

Pre-requisites

A knowledge of the basic concepts, techniques and results of graph theory, as afforded by the Part II Graph Theory course, will be assumed. This includes Turán's theorem, Ramsey's theorem, Hall's theorem and so on, together with applications of elementary probability.

Contents

1   The Erdös–Stone theorem

The starting point of extremal graph theory is perhaps Turán's theorem, which you hopefully learnt from the IID Graph Theory course. To state the theory, we need the following preliminary definition:

Definition (Turán graph). The *Turán graph* $T_r(n)$ is the complete r-partite graph on n vertices with class sizes $\lfloor n/r \rfloor$ or $\lceil n/r \rceil$. We write $t_r(n)$ for the number of edges in $T_r(n)$.

The theorem then says

Theorem (Turán's theorem). If G is a graph with $|G| = n$, $e(G) \geq t_r(n)$ and $G \not\supseteq K_{r+1}$. Then $G = T_r(n)$.

This is an example of an *extremal theorem*. More generally, given a fixed graph F, we seek

$$\operatorname{ex}(n, F) = \max\{e(G) : |G| = n, G \not\supseteq F\}.$$

Turán's theorem tells us $\operatorname{ex}(n, K_{r+1}) = t_r(n)$. We cannot find a nice expression for the latter number, but we have

$$\operatorname{ex}(n, K_{r+1}) = t_r(n) \approx \left(1 - \frac{1}{r}\right)\binom{n}{2}.$$

Turán's theorem is a rather special case. First of all, we actually know the exact value of $\operatorname{ex}(n, F)$. Moreover, there is a unique *extremal graph* realizing the bound. Both of these properties do not extend to other choices of F.

By definition, if $e(G) > \operatorname{ex}(n, K_{r+1})$, then G contains a K_{r+1}. The Erdös–Stone theorem tells us that as long as $|G| = n$ is big enough, this condition implies G contains a much larger graph than K_{r+1}.

Notation. We denote by $K_r(t)$ the complete r-partite graph with t vertices in each class.

So $K_r(1) = K_r$ and $K_r(t) = T_r(rt)$.

Theorem (Erdös–Stone, 1946). Let $r \geq 1$ be an integer and $\varepsilon > 0$. Then there exists $d = d(r, \varepsilon)$ and $n_0 = n_0(r, \varepsilon)$ such that if $|G| = n \geq n_0$ and

$$e(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2},$$

then $G \supseteq K_{r+1}(t)$, where $t = \lfloor d \log n \rfloor$.

Note that we can remove n_0 from the statement simply by reducing d, since for sufficiently small d, whenever $n < n_0$, we have $\lfloor d \log n \rfloor = 0$.

One corollary of the theorem, and a good way to think about the theorem is that given numbers r, ε, t, whenever $|G| = n$ is sufficiently large, the inequality $e(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$ implies $G \subseteq K_{r+1}(t)$.

To prove the Erdös–Stone theorem, a natural strategy is to try to throw away vertices of small degree, and so that we can bound the *minimal degree* of the graph instead of the total number of edges. We will make use of the following lemma to do so:

Lemma. Let $c, \varepsilon > 0$. Then there exists $n_1 = n_1(c, \varepsilon)$ such that if $|G| = n \geq n_1$ and $e(G) \geq (c + \varepsilon)\binom{n}{2}$, then G has a subgraph H where $\delta(H) \geq c|H|$ and $|H| \geq \sqrt{\varepsilon}n$.

Proof. The idea is that we can keep removing the vertex of smallest degree and then we must eventually get the H we want. Suppose this doesn't gives us a suitable graph even after we've got $\sqrt{\varepsilon}n$ vertices left. That means we can find a sequence

$$G = G_n \supseteq G_{n-1} \supseteq G_{n-2} \supseteq \cdots G_s,$$

where $s = \lfloor \varepsilon^{1/2}n \rfloor$, $|G_j| = j$ and the vertex in $G_j \setminus G_{j-1}$ has degree $< cj$ in G_j.

We can then calculate

$$e(G_s) > (c + \varepsilon)\binom{n}{2} - c\sum_{j=s+1}^{n} j$$

$$= (c + \varepsilon)\binom{n}{2} - c\left\{\binom{n+1}{2} - \binom{s+1}{2}\right\}$$

$$\sim \frac{\varepsilon n^2}{2}$$

as n gets large (since c and s are fixed numbers). In particular, this is $> \binom{s}{2}$. But G_s only has s vertices, so this is impossible. $\qquad\square$

Using this, we can reduce the Erdös–Stone theorem to a version that talks about the minimum degree instead.

Lemma. Let $r \geq 1$ be an integer and $\varepsilon > 0$. Then there exists a $d_1 = d_1(r, \varepsilon)$ and $n_2 = n_2(r, \varepsilon)$ such that if $|G| = n \geq n_2$ and

$$\delta(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)n,$$

then $G \supseteq K_{r+1}(t)$, where $t = \lfloor d_1 \log n \rfloor$.

We first see how this implies the Erdös–Stone theorem:

Proof of Erdös–Stone theorem. Provided n_0 is large, say $n_0 > n_1\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right)$, we can apply the first lemma to G to obtain a subgraph $H \subseteq G$ where $|H| > \sqrt{\frac{\varepsilon}{2}}n$, and $\delta(H) \geq \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|H|$.

We can then apply our final lemma as long as $\sqrt{\frac{\varepsilon}{2}}n$ is big enough, and obtain $K_{r+1}(t) \subseteq H \subseteq G$, with $t > \left\lfloor d_1(r, \varepsilon/2) \log\left(\sqrt{\frac{\varepsilon}{2}}n\right)\right\rfloor$. $\qquad\square$

We can now prove the lemma.

Proof of lemma. We proceed by induction on r. If $r = 0$ or $\varepsilon \geq \frac{1}{r}$, the theorem is trivial. Otherwise, by the induction hypothesis, we may assume $G \supseteq K_r(T)$ for

$$T = \left\lfloor \frac{2t}{\varepsilon r} \right\rfloor.$$

Call it $K = K_r(T)$. This is always possible, as long as

$$d_1(r, \varepsilon) < \frac{\varepsilon r}{2} d_1\left(r - 1, \frac{1}{r(r-1)}\right).$$

The exact form is not important. The crucial part is that $\frac{1}{r(r-1)} = \frac{1}{r-1} - \frac{1}{r}$, which is how we chose the ε to put into $d_1(r-1, \varepsilon)$.

Let U be the set of vertices in $G - K$ having at least $\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|$ neighbours in K. Calculating

$$\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K| = \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)rT = (r-1)T + \frac{\varepsilon r}{2}T \geq (r-1)T + t,$$

we see that every element in U is joined to at least t vertices in each class, and hence is joined to some $K_r(t) \subseteq K$.

So we have to show that $|U|$ is large. If so, then by a pigeonhole principle argument, it follows that we can always find enough vertices in U so that adding them to K gives a $K_{r+1}(t)$.

Claim. There is some $c > 0$ (depending on r and ε) such that for n sufficiently large, we have
$$|U| \geq cn.$$

The argument to establish these kinds of bounds is standard, and will be used repeatedly. We write $e(K, G - K)$ for the number of edges between K and $G - K$. By the minimum degree condition, each vertex of K sends at least $\left(1 - \frac{1}{r} + \varepsilon\right)n - |K|$ edges to $G - K$. Then we have two inequalities

$$|K|\left\{\left(1 - \frac{1}{r} + \varepsilon\right)n - |K|\right\} \leq e(K, G - K)$$

$$\leq |U||K| + (n - |U|)\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|.$$

Rearranging, this tells us

$$\frac{\varepsilon n}{2} - |K| \leq |U|\left(\frac{1}{r} - \frac{\varepsilon}{2}\right).$$

Now note that $|K| \sim \log n$, so for large n, the second term on the left is negligible, and it follows that $U \sim n$.

We now want to calculate the number of $K_r(t)$'s in K. To do so, we use the simple inequality
$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$
Then we have

$$\#K_r(t) = \binom{T}{t}^r \leq \left(\frac{eT}{t}\right)^{rt} \leq \left(\frac{3e}{\varepsilon r}\right)^{rt} \leq \left(\frac{3e}{\varepsilon r}\right)^{rd\log n} = n^{rd\log(3e/\varepsilon r)}$$

Now if we pick d sufficiently small, then this tells us $\#K_r(t)$ grows sublinearly with n. Since $|U|$ grows linearly, and t grows logarithmically, it follows that for n large enough, we have
$$|U| \geq t \cdot \#K_r(t).$$

Then by the pigeonhole principle, there must be a set $W \subseteq U$ of size t joined to the same $K_r(t)$, giving a $K_{r+1}(t)$. $\qquad\square$

Erdös and Simonovits noticed that Erdös–Stone allows us to find $\mathrm{ex}(n, F)$ asymptotically for all F.

Theorem (Erdös–Simonovits)**.** Let F be a fixed graph with chromatic number $\chi(F) = r + 1$. Then

$$\lim_{n \to \infty} \frac{\mathrm{ex}(n, F)}{\binom{n}{2}} = 1 - \frac{1}{r}.$$

Proof. Since $\chi(F) = r + 1$, we know F cannot be embedded in an r-partite graph. So in particular, $F \not\subseteq T_r(n)$. So

$$\mathrm{ex}(n, F) \geq t_r(n) \geq \left(1 - \frac{1}{r}\right)\binom{n}{2}.$$

On the other hand, given any $\varepsilon > 0$, if $|G| = n$ and

$$e(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2},$$

then by the Erdös–Stone theorem, we have $G \supseteq K_{r+1}(|F|) \supseteq F$ provided n is large. So we know that for every $\varepsilon > 0$, we have

$$\limsup \frac{\mathrm{ex}(n, F)}{\binom{n}{2}} \leq 1 - \frac{1}{r} + \varepsilon.$$

So we are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

If $r > 1$, then this gives us a genuine asymptotic expression for $\mathrm{ex}(n, F)$. However, if $r = 1$, i.e. F is a bipartite graph, then this only tells us $\mathrm{ex}(n, F) = o\left(\binom{n}{2}\right)$, but doesn't tell us about the true asymptotic behaviour.

To end the chapter, we show that $t \sim \log n$ is indeed the best we can do in the Erdös–Stone theorem, by actually constructing some graphs.

Theorem. Given $r \in \mathbb{N}$, there exists $\varepsilon_r > 0$ such that if $0 < \varepsilon < \varepsilon_r$, then there exists $n_3(r, \varepsilon)$ so that if $n > n_3$, there exists a graph G of order n such that

$$e(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$$

but $K_{r+1}(t) \not\subseteq G$, where

$$t = \left\lceil \frac{3 \log n}{\log 1/\varepsilon} \right\rceil.$$

So this tells us we cannot get better than $t \sim \log n$, and this gives us some bound on $d(r, \varepsilon)$.

Proof. Start with the Turán graph $T_r(n)$. Let $m = \left\lceil \frac{n}{r} \right\rceil$. Then there is a class W of $T_r(n)$ of size m. The strategy is to add $\varepsilon\binom{n}{2}$ edges inside W to obtain G such that $G[W]$ (the subgraph of G formed by W) does not contain a $K_2(t)$. It then follows that $G \not\supseteq K_{r+1}(t)$, but also

$$e(G) \geq t_r(n) + \varepsilon\binom{n}{2} \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2},$$

as desired.

To see that such an addition is possible, we choose edges inside W independently with probability p, to be determined later. Let X be the number of edges chosen, and Y be the number of $K_2(t)$ created. If $\mathbb{E}[X - Y] > \varepsilon\binom{n}{2}$, then this means there is an actual choice of edges with $X - Y > \varepsilon\binom{n}{2}$. We then remove an edge from each $K_2(t)$ to leave a $K_2(t)$-free graph with at least $X - Y > \varepsilon\binom{n}{2}$ edges.

So we want to actually compute $\mathbb{E}[X - Y]$. Seeing that asymptotically, m is much larger than t, we have

$$
\begin{aligned}
\mathbb{E}[X - Y] &= \mathbb{E}[X] - \mathbb{E}[Y] \\
&= p\binom{m}{2} - \frac{1}{2}\binom{m}{t}\binom{m-t}{t}p^{t^2} \\
&\sim \frac{1}{2}pm^2 - \frac{1}{2}m^{2t}p^{t^2} \\
&= \frac{1}{2}pm^2(1 - m^{2t-2}p^{t^2-1}) \\
&= \frac{1}{2}pm^2(1 - (m^2 p^{t+1})^{t-1}).
\end{aligned}
$$

We pick $p = 3\varepsilon r^2$ and $\varepsilon_r = (3r^2)^{-6}$. Then $p < \varepsilon^{5/6}$, and

$$
m^2 p^{t+1} < m^2 \varepsilon^{5/6(t+1)} < (m^2 m^{-5/2})^{t-1} < \frac{1}{2}.
$$

Hence, we find that

$$
\mathbb{E}[x - y] \geq \frac{1}{4}pm^2 > \varepsilon\binom{n}{2}. \qquad \square
$$

Let's summarize what we have got so far. Let $t(n, r, \varepsilon)$ be the largest value of t such that a graph of order n with $\left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$ edges is guaranteed to contain $K_{r+1}(t)$. Then we know $t(n, r, \varepsilon)$ grows with n like $\log n$. If we are keen, we can ask how t depends on the other terms r and ε. We just saw that

$$
t(n, r, \varepsilon) \leq \frac{3\log n}{\log 1/\varepsilon}.
$$

So we see that the dependence on ε is at most logarithmic, and in 1976, Bollobás, Erdös and Simonovits showed that

$$
t(n, r, \varepsilon) \geq c\frac{\log n}{r\log 1/\varepsilon}
$$

for some c. Thus, $t(n, r, \varepsilon)$ also grows (inverse) logarithmically with ε.

In our original bound, there is no dependence on r, which is curious. While Bollobás, Erdös and Simonovits's bound does, in 1978, Chvátal and Szemerédi showed that

$$
t \geq \frac{1}{500}\frac{\log n}{\log 1/\varepsilon}
$$

if n is large. So we know there actually is no dependence on r.

We can also ask about the containment of less regular graphs than $K_r(t)$. In 1994, Bollobás and Kohayakawa adapted the proof of the Erdös–Stone theorem

to show that there is a constant c such that for any $0 < \gamma < 1$, if $e(G) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$ and n is large, then we can find a complete $(r+1)$-partite subgraph with class sizes

$$\left\lfloor c\gamma \frac{\log n}{r \log 1/\varepsilon} \right\rfloor, \left\lfloor c\gamma \frac{\log n}{\log r} \right\rfloor, r - 1 \text{ times}, \left\lfloor c\varepsilon^{\frac{3}{2} - \frac{\gamma}{2}} n^{1-\gamma} \right\rfloor.$$

We might wonder if we can make similar statements for hypergraphs. It turns out all the analogous questions are open. Somehow graphs are much easier.

2 Stability

An extremal problem is *stable* if every near-optimal extremal example looks close to some optimal (unique) example. Stability pertain for $\mathrm{ex}(n, F)$.

Theorem. Let $t, r \geq 2$ be fixed, and suppose $|G| = n$ and $G \not\supseteq K_{r+1}(t)$. If

$$e(G) = \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}.$$

Then

  (i) There exists $T_r(n)$ on $V(G)$ with $|E(G)\Delta E(T_r(n))| = o(n^2)$.

 (ii) G contains an r-partite subgraph with $\left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}$ edges.

(iii) G contains an r-partite subgraph with minimum degree $\left(1 - \frac{1}{r} + o(1)\right) n$.

The outline of the proof is as follows: we first show that the three statements are equivalent, so that we only have to prove (ii). To prove (ii), we first use Erdös–Stone to find some $K_r(s)$ living inside G (for some s), which is a good r-partite subgraph of G, but with not quite the right number of vertices. We now look at the remaining vertices. For each vertex v, find the C_i such that v is connected to as few vertices in C_i as possible, and enlarge C_i to include that as well. After doing this, C_i will have quite a few edges between its vertices, and we throw those away. The hope is that there are only $o(n^2)$ edges to throw away, so that we still have $\left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}$ edges left. It turns out as long as we throw out some "bad" vertices, we will be fine.

Proof. We first prove (ii), which looks the weakest. We will then argue that the others follow from this (and in fact they are equivalent).

By Erdös–Stone, we can always find some $K_r(s) = K \subseteq G$ for some $s = s(n) \to \infty$. We can, of course, always choose $s(n)$ so that $s(n) \leq \log n$, which will be helpful for our future analysis. We let C_i be the ith class of K.

 – By throwing away $o(n)$ vertices, we may assume that

$$\delta(G) \geq \left(1 - \frac{1}{r} + o(1)\right) n$$

 – Let X be the set of all vertices joined to at least t vertices in each C_i. Note that there are $\binom{s}{t}^r$ many copies of $K_r(t)$'s in K. So by the pigeonhole principle, we must have $|X| < t\binom{s}{t}^r = o(n)$, or else we can construct a $K_{r+1}(t)$ inside G.

 – Let Y be the set of vertices joined to fewer than $(r-1)s - \frac{s}{2t} + t$ other vertices. By our bound on $\delta(G)$, we certainly have

$$e(K, G - K) \geq s(r - 1 + o(1))n.$$

On the other hand, every element of G is certainly joined to at most $(r-1)s + t$ edges, since X was thrown away. So we have

$$((r-1)s + t)(n - |Y|) + |Y|\left((r-1)s - \frac{s}{2t} + t\right) \geq e(K, G - K).$$

So we deduce that $|Y| = o(n)$. Throw Y away.

9

Let V_i be the set of vertices in $G \setminus K$ joined to fewer than t of C_i. It is now enough to show that $e(G[V_i]) = o(n^2)$. We can then throw away the edges in each V_i to obtain an r-partite graph.

Suppose on the contrary that $e(G[V_j]) \geq \varepsilon n^2$, say, for some j. Then Erdös–Stone with $r = 1$ says we have $G[V_j] \supseteq K_2(t)$. Each vertex of $K_2(t)$ has at least $s - \frac{s}{2t} + 1$ in each other C_i, for $i \neq j$. So we see that $K_2(t)$ has at least $s - 2t \left(\frac{s}{2t} - 1 \right) > t$ common neighbours in each C_i, giving $K_{r+1}(t) \subseteq G$.

It now remains to show that (i) and (iii) follows from (ii). The details are left for the example sheet, but we sketch out the rough argument. (ii) $\Rightarrow$ (i) is straightforward, since an r-partite graph with that many edges is already pretty close to being a Turán graph.

To deduce (iii), since we can assume $\delta(G) \geq \left(1 - \frac{1}{r} + o(1)\right) n$, we note that if (iii) did not hold, then we have εn vertices of degree $\leq \left(1 - \frac{1}{r} - \varepsilon\right) n$, and their removal leaves a graph of order $(1 - \varepsilon)n$ with at least

$$\left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2} - \varepsilon n \left(1 - \frac{1}{r} - \varepsilon\right) n > \left(1 - \frac{1}{r} + \varepsilon^2\right) \binom{(1-\varepsilon)n}{2},$$

which by Erdös–Stone would contain $K_{r+1}(t)$. So we are done. $\qquad \square$

Corollary. Let $\chi(F) = r + 1$, and let G be extremal for F, i.e. $G \not\supseteq F$ and $e(G) = \mathrm{ex}(|G|, F)$. Then $\delta(G) = \left(1 - \frac{1}{r} + o(1)\right) n$.

Proof. By our asymptotic bound on $\mathrm{ex}(n, F)$, it cannot be that $\delta(G)$ is greater than $\left(1 - \frac{1}{r} + o(1)\right) n$. So there must be a vertex $v \in G$ with $d(v) \leq \left(1 - \frac{1}{r} - \varepsilon\right) n$.

We now apply version (iii) of the stability theorem to obtain an r-partite subgraph H of G. We can certainly pick $|F|$ vertices in the same part of H, which are joined to $m = \left(1 - \frac{1}{r} + o(1)\right)$ common neighbours. Form G^* from $G - v$ by adding a new vertex u joined to these m vertices. Then $e(G^*) > e(G)$, and so the maximality of G entails $G^* \supseteq F$.

Pick a copy of F in G^*. This must involve u, since G, and hence $G - v$ does not contain a copy of F. Thus, there must be some x amongst the $|F|$ vertices mentioned. But then we can replace u with x and we are done. $\qquad \square$

Sometimes, stability and bootstrapping can give exact results.

Example. $\mathrm{ex}(n, C_5) = t_2(n)$. In fact, $\mathrm{ex}(n, C_{2k+1}) = t_2(n)$ if n is large enough.

Theorem (Simonovits)**.** Let F be $(r + 1)$-edge-critical, i.e. $\chi(F) = r + 1$ but $\chi(F) \setminus e = r$ for every edge e of F. Then for large n,

$$\mathrm{ex}(n, F) = t_r(n).$$

and the only extremal graph is $T_r(n)$.

So we get a theorem like Turán's theorem for large n.

Proof. Let G be an extremal graph for F and let H be an r-partite subgraph with

$$\delta(H) = \left(1 - \frac{1}{r} + o(1)\right) n.$$

Note that H necessarily have r parts of size $\left(\frac{1}{r} + o(1)\right) n$ each. Assign each of the $o(n)$ vertices in $V(G) \setminus V(H)$ to a class where it has fewest neighbours.

Suppose some vertex v has εn neighbours in its own class. Then it has $\geq \varepsilon n$ in each class. Pick εn neighbours of v in each class of H. These neighbours span a graph with at least $\left(1 - \frac{1}{r} + o(1)\right) \binom{r\varepsilon n}{2}$ edges. So by Erdös–Stone (or arguing directly), they span $F - w$ for some vertex w (contained in $K_r(|F|)$). Hence $G \supseteq F$, contradiction.

Thus, each vertex of G has only $o(n)$ vertices in its own class. So it is joined to all but $o(n)$ vertices in every other class.

Suppose some class of G contains an edge xy. Pick a set Z of $|F|$ vertices in that class with $\{x, y\} \in Z$. Now Z has $\left(\frac{1}{r} + o(1)\right) n$ common neighbours in each class, so (by Erdös–Stone or directly) these common neighbours span a $K_{r-1}(|F|)$. But together with Z, we have a $K_r(|F|)$ but with an edge added inside a class. But by our condition that $F \backslash e$ is r-partite for any e, this subgraph contains an F. This is a contradiction.

So we conclude that no class of G contains an edge. So G itself is r-partite, but the r-partite graph with most edges is $T_r(n)$, which does not contain F. So $G = T_r(n)$. $\qquad\square$

3 Supersaturation

Suppose we have a graph G with $e(G) > \mathrm{ex}(n, F)$. Then by definition, there is at least one copy of F in G. But can we tell how many copies we have? It turns out this is not too difficult to answer, and in fact we can answer the question for any hypergraph.

Recall that an *ℓ-uniform hypergraph* is a pair (V, E), where $E \subseteq V^{(\ell)}$. We can define the extremal function of a class of hypergraphs $\mathcal{F}$ by

$$\mathrm{ex}(n, \mathcal{F}) = \max\{e(G) : |G| = n, G \text{ contains no } f \in \mathcal{F}\}.$$

Again we are interested in the limiting density

$$\pi(\mathcal{F}) = \lim_{n \to \infty} \frac{\mathrm{ex}(n, \mathcal{F})}{\binom{n}{\ell}}.$$

It is an easy exercise to show that this limit always exists. We solved it explicitly for graphs explicitly previously, but we don't really need the Erdös–Stone theorem just to show that the limit exists.

The basic theorem in supersaturation is

Theorem (Erdös–Simonovits)**.** Let H be some ℓ-uniform hypergraph. Then for all $\varepsilon > 0$, there exists $\delta(H, \varepsilon)$ such that every ℓ-uniform hypergraph G with $|G| = n$ and

$$e(G) > (\pi(H) + \varepsilon) \binom{n}{\ell}$$

contains $\lfloor \delta n^{|H|} \rfloor$ copies of H.

Note that $n^{|H|}$ is approximately the number of ways to choose $|H|$ vertices from n, so it's the number of possible candidates for subgraphs H.

Proof. For each m-set $M \subseteq V(G)$, we let $G[M]$ be the sub-hypergraph induced by these vertices. Let the number of subsets M with $e(G[M]) > \left(\pi(H) + \frac{\varepsilon}{2}\right) \binom{m}{\ell}$ be $\eta \binom{n}{m}$. Then we can estimate

$$(\pi(H) + \varepsilon) \binom{n}{\ell} \leq e(G) = \frac{\sum_M e(G[M])}{\binom{n-\ell}{m-\ell}}$$
$$\leq \frac{\eta \binom{n}{m} \binom{m}{\ell} + (1 - \eta) \binom{n}{m} \left(\pi(H) + \frac{\varepsilon}{2}\right) \binom{m}{\ell}}{\binom{n-\ell}{m-\ell}}.$$

So if $n > m$, then

$$\pi(H) + \varepsilon \leq \eta + (1 + \eta) \left(\pi(H) + \frac{\varepsilon}{2}\right).$$

So

$$\eta \geq \frac{\frac{\varepsilon}{2}}{1 - \pi(H) - \frac{\varepsilon}{2}} > 0.$$

The point is that it is positive, and that's all we care about.

We pick m large enough so that

$$\mathrm{ex}(m, H) < \left(\pi(H) + \frac{\varepsilon}{2}\right) \binom{m}{\ell}.$$

Then $H \subseteq G[M]$ for at least $\eta\binom{n}{m}$ choices of M. Hence G contains at least

$$\frac{\eta\binom{n}{m}}{\binom{n-|H|}{m-|H|}} = \frac{\eta\binom{n}{|H|}}{\binom{m}{|H|}}$$

copies of H, and we are done. (Our results hold when n is large enough, but we can choose δ small enough so that $\delta n^{|H|} < 1$ when n is small) $\qquad\square$

Let $k_P(G)$ be the number of copies of K_p in G. Ramsey's theorem tells us $k_p(G) + k_p(\bar{G}) > 0$ if $|G|$ is large (where $\bar{G}$ is the complement of G). In the simplest case where $p = 3$, it turns out we can count these monochromatic triangles exactly.

Theorem (Lorden, 1961)**.** Let G have degree sequence $d_1, \ldots, d_n$. Then

$$k_3(G) + k_3(\bar{G}) = \binom{n}{3} - (n-2)e(G) + \sum_{i=1}^{n}\binom{d_i}{2}.$$

Proof. The number of paths of length 2 in G and $\bar{G}$ is precisely

$$\sum_{i=1}^{n}\left(\binom{d_i}{2} + \binom{n-1-d_i}{2}\right) = 2\sum_{i=1}^{n}\binom{d_i}{2} - 2(n-2)e(G) + 3\binom{n}{3},$$

since to find a path of length 2, we pick the middle vertex and then pick the two edges. A complete or empty K_3 contains 3 such paths; Other sets of three vertices contain exactly 1 such path. Hence

$$\binom{n}{3} + 2(k_3(G) + k_3(\bar{G})) = \text{number of paths of length 2}. \qquad\square$$

Corollary (Goodman, 1959)**.** We have

$$k_3(G) + k_3(\bar{G}) \geq \frac{1}{24}n(n-1)(n-5).$$

In particular, the Ramsey number of a triangle is at most 6.

Proof. Let $m = e(G)$. Then

$$k_3(G) + k_3(\bar{G}) \geq \binom{n}{3} - (n-2)m + n\binom{2m/n}{2}.$$

Then minimize over m. $\qquad\square$

This shows the minimum density of a monochromatic K_3 in a red/blue colouring of K_n (for n large) is $\sim \frac{1}{4}$. But if we colour edges monochromatic, then $\frac{1}{4}$ is the probability of having a triangle being monochromatic. So the density is achieved by a "random colouring". Recall also that the best bounds on the Ramsey numbers we have are obtained by random colourings. So we might think the best way of colouring if we want to minimize the number of monochromatic cliques is to do so randomly.

However, this is not true. While we do not know what the minimum density of monochromatic K_4 in a red/blue colouring of K_n, it is already known to be $< \frac{1}{33}$ (while $\frac{1}{32}$ is what we expect from a random colouring). It is also known by flag algebras to be $> \frac{1}{35}$. So we are not very far.

Corollary. For $m = e(G)$ and $n = |G|$, we have

$$k_3(G) \geq \frac{m}{3n}(4m - n^2).$$

Proof. By Lorden's theorem, we know

$$k_3(G) + k_3(\bar{G}) = \binom{n}{3} - (n-2)e(\bar{G}) + \sum \binom{\bar{d}_i}{2},$$

where $\bar{d}_i$ is the degree sequence in $\bar{G}$. But

$$3k_3(\bar{G}) \leq \sum \binom{\bar{d}_i}{2},$$

since the sum is counting the number of paths of length 2. So we find that

$$k_3(G) \geq \binom{n}{3} - (n-2)\bar{m} - \frac{2}{3}n\binom{2\bar{m}/n}{2},$$

and $\bar{m} = \binom{n}{2} - m$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Observe that equality is almost never attained. It is attained only for regular graphs with no subgraphs looking like

$$\bullet$$

$$\bullet\!\!-\!\!-\!\!-\!\!\bullet$$

So non-adjacency is an equivalence relation, so the graph is complete multi-partite and regular. Thus it is $T_r(n)$ with $r \mid n$.

Theorem. Let G be a graph. For any graph F, let $i_F(G)$ be the number of induced copies of F in G, i.e. the number of subsets $M \subseteq V(G)$ such that $G[M] \cong F$. So, for example, $i_{K_p}(G) = k_p(G)$.

Define

$$f(G) = \sum_F \alpha_F i_F(G),$$

with the sum being over a finite collection of graphs F, each being complete multipartite, with $\alpha_F \in \mathbb{R}$ and $\alpha_F \geq 0$ if F is not complete. Then amongst graphs of given order, $f(G)$ is maximized on a complete multi-partite graph. Moreover, if $\alpha_{\bar{K}_3} > 0$, then there are no other maxima.

Proof. We may suppose $\alpha_{\bar{K}_3} > 0$, because the case of $\alpha_{\bar{K}_3} = 0$ follows from a limit argument. Choose a graph G maximizing f and suppose G is not complete multipartite. Then there exist non-adjacent vertices x, y whose neighbourhoods X, Y differ.

There are four contributions to $i_F(G)$, coming from

 (i) F's that contain both x and y;

 (ii) F's that contain y but not x;

 (iii) F's that contain x but not y;

 (iv) F's that contain neither x nor y.

We may assume that the contribution from (iii) $\geq$ (ii), and if they are equal, then $|X| \leq |Y|$.

Consider what happens when we remove all edges between y and Y, and add edges from y to everything in X. Clearly (iii) and (iv) are unaffected.

- If (iii) > (ii), then after this move, the contribution of (ii) becomes equal to the contribution of (iii) (which didn't change), and hence strictly increased.

  The graphs F that contribute to (i) are not complete, so $\alpha_F \geq 0$. Moreover, since F is complete multi-partite, it cannot contain a vertex in $X \Delta Y$. So making the move can only increase the number of graphs contributing to (i), and each contribution is non-negative.

- If (iii) = (ii) and $|X| \leq |Y|$, then we make similar arguments. The contribution to (ii) is unchanged this time, and we know the contribution of (i) strictly increased, because the number of $\bar{K}_3$'s contributing to (i) is the number of points not connected to x and y.

In both cases, the total sum increased. $\square$

Perhaps that theorem seemed like a rather peculiar one to prove. However, it has some nice consequences. The following theorem relates $k_p(G)$ with $k_r(G)$ for different p and r:

Theorem (Bollobás, 1976)**.** Let $1 \leq p \leq r$, and for $0 \leq x \leq \binom{n}{p}$, let $\psi(x)$ be a maximal convex function lying below the points

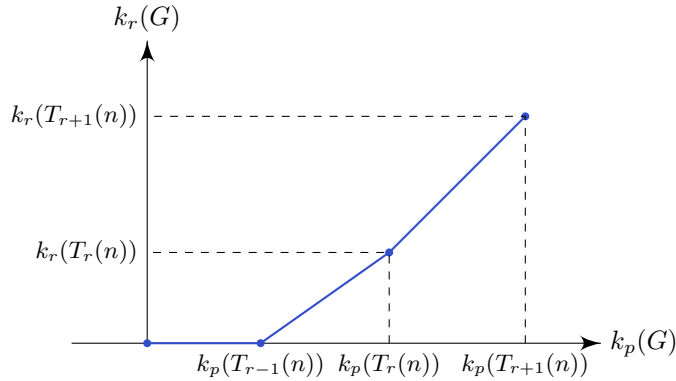$$\{(k_p(T_q(n)), k_r(T_q(n))) : q = r - 1, r, \ldots\} \cup \{(0, 0)\}.$$

Let G be a graph of order n. Then

$$k_r(G) \geq \psi(k_p(G)).$$

Proof. Let $f(G) = k_p(G) - ck_r(G)$, where $c > 0$.

Claim. It is enough to show that f is maximized on a Turán graph for any c.

Indeed, suppose we plot out the values of $(k_p(T_q(n)), k_r(T_q(n)))$:

If the theorem doesn't hold, then we can pick a G such that $(k_p(G), k_r(G))$ lies below ψ. Draw a straight line through the point with slope parallel to ψ. This has slope $\frac{1}{c} > 0$ for some c. The intercept on the x-axis is then $k_p(G) - ck_r(G)$, which would be greater than f(any Turán graph) by convexity, a contradiction.

Now the previous theorem immediately tells us f is maximized on some complete multi-partite graph. Suppose this has q class, say of sizes $a_1 \leq a_2 \leq \cdots \leq a_q$. It is easy to verify $q \geq r - 1$. In fact, we may assume $q \geq r$, else the maximum is on a Turán graph $T_{r-1}(n)$.

Then we can write

$$f(G) = a_1 a_q A - c a_q a_q B + C,$$

where A, B, C are rationals depending only on $a_2, \ldots, a_{q-1}$ and $a_1 + a_q$ (A and B count the number of ways to pick a K_p and K_r respectively in a way that involves terms in both the first and last classes).

We wlog assume c is irrational. Hence $a_q a_q A - c a_1 a_q = a_1 a_q (A - cB) \neq 0$.

- If $A - cB < 0$, replace a_1 and a_q by 0 and $a_1 + a_q$. This would then increase f, which is impossible.

- If $A - cB > 0$ and $a_1 \leq a_q - 2$, then we can replace a_1, a_q by $a_1 + 1, a_q - 1$ to increase f.

Hence $a_1 \geq a_q - 1$. So $G = T_q(n)$. $\qquad\square$

4 Szemerédi's regularity lemma

Szmerédi's regularity lemma tells us given a very large graph, we can always equipartition it into pieces that are "uniform" in some sense. The lemma is arguably "trivial", but it also has many interesting consequences. To state the lemma, we need to know what we mean by "uniform".

Definition (Density)**.** Let U, W be disjoint subsets of the vertex set of some graph. The number of edges between U and W is denoted by $e(U, W)$, and the *density* is

$$d(U, W) = \frac{e(U, W)}{|U||W|}.$$

Definition (ε-uniform pair)**.** Let $0 < \varepsilon < 1$. We say a pair (U, W) is *ε-uniform* if

$$|d(U', W') - d(U, W)| < \varepsilon$$

whenever $U' \subseteq U$, $W' \subseteq W$, and $|U'| \geq \varepsilon|U|$, $|W'| \geq \varepsilon|W|$.

Note that it is necessary to impose some conditions on how small U' and W' can be. For example, if $|U'| = |W'| = 1$, then $d(U', W')$ is either 0 or 1. So we cannot have a sensible definition if we want to require the inequality to hold for arbitrary U', W'.

But we might be worried that it is unnatural to use the same ε for two different purposes. This is not something one should worry about. The Szemerédi regularity lemma is a fairly robust result, and everything goes through if we use different ε's for the two different purposes. However, it is annoying to have to have many different ε's floating around.

Before we state and prove Szemerédi's regularity lemma, let's first try to understand why uniformity is good. The following is an elementary observation.

Lemma. Let (U, W) be an ε-uniform pair with $d(U, W) = d$. Then

$$|\{u \in U : |\Gamma(u) \cap W| > (d - \varepsilon)|W|\}| \geq (1 - \varepsilon)|U|$$
$$|\{u \in U : |\Gamma(u) \cap W| < (d + \varepsilon)|W|\}| \geq (1 - \varepsilon)|U|,$$

where $\Gamma(u)$ is the set of neighbours of u.

Proof. Let

$$X = \{u \in U : |\Gamma(u) \cap W| \leq (d - \varepsilon)|W|\}.$$

Then $e(X, W) \leq (d - \varepsilon)|X||W|$. So

$$d(X, W) \leq d - \varepsilon = d(U, W) - \varepsilon.$$

So it fails the uniformity condition. Since W is definitely not small, we must have $|X| < \varepsilon|U|$.

The other case is similar, or observe that the complementary bipartite graph between U and W has density $1 - d$ and is ε-uniform. $\qquad\square$

What is good about ε-uniform pairs is that if we have enough of them, then we can construct essentially any subgraph we like. Later, Szemerédi's regularity lemma says any graph large enough has ε-uniform equipartitions, and together, they can give us some pretty neat results.

Lemma (Graph building lemma)**.** Let G be a graph containing distinct vertex subsets $V_1, \ldots, V_r$ with $|V_i| = u$, such that (V_i, V_j) is ε-uniform and $d(V_i, V_j) \geq \lambda$ for all $1 \leq i \leq j \leq r$.

Let H be a graph with maximum degree $\Delta(H) \leq \Delta$. Suppose H has an r-colouring in which no colour is used more than s times, i.e. $H \subseteq K_r(s)$, and suppose $(\Delta + 1)\varepsilon \leq \lambda^\Delta$ and $s \leq \lfloor \varepsilon u \rfloor$. Then $H \subseteq G$.

To prove this, we just do what we did in the previous lemma, and find lots of vertices connected to lots of other vertices, and then we are done.

Proof. We wlog assume $V(H) = \{1, \ldots, k\}$, and let $c : V(H) \to \{1, \ldots, r\}$ be a colouring of $V(H)$ using no colour more than s times. We want to pick vertices $x_1, \ldots, x_k$ in G so that $x_i x_j \in E(G)$ if $ij \in E(H)$.

We claim that, for $0 \leq \ell \leq k$, we can choose distinct vertices $x_1, \ldots, x_\ell$ so that $x_j \in C_{c(j)}$, and for $\ell < j \leq k$, a set X_j^ℓ of *candidates* for x_j such that

(i) $X_j^\ell \subseteq V_{c(j)}$;

(ii) $x_i y_j \in E(G)$ for all $y_j \in X_j^\ell$ and $i \leq \ell$ such that $ij \in E(H)$.

(iii) $|X_j^\ell| \geq (\lambda - \varepsilon)^{|N(j,\ell)|}|V_{c(j)}|$, where

$$N(j, \ell) = \{x_i : 1 \leq i \leq \ell \text{ and } ij \in E(H)\}.$$

The claim holds for $\ell = 0$ by taking $X_j^0 = V_{c(j)}$.

By induction, suppose it holds for ℓ. To pick $x_{\ell+1}$, of course we should pick it from our candidate set $X_{\ell+1}^\ell$. Then the first condition is automatically satisfied. Define the set

$$T = \{j > \ell + 1 : (\ell + 1)j \in E(H)\}.$$

Then each $t \in T$ presents an obstruction to (ii) and (iii) being satisfied. To satisfy (ii), for each $t \in T$, we should set

$$X_t^{\ell+1} = X_t^\ell \cap \Gamma(x_{\ell+1}).$$

Thus, to satisfy (iii), we want to exclude those $x_{\ell+1}$ that make this set too small. We define

$$Y_t = \left\{ y \in X_{\ell+1}^\ell : |\Gamma(y) \cap X_t^\ell| \leq (\lambda - \varepsilon)|X_t^\ell| \right\}.$$

So we want to find something in $X_{\ell+1}^\ell \setminus \bigcup_{t \in T} Y_t$. We also cannot choose one of the x_i already used. So our goal is to show that

$$\left| X_{\ell+1}^\ell - \bigcup_{t \in T} Y_t \right| > s - 1.$$

This will follow simply from counting the sizes of $|X_{\ell+1}^\ell|$ and $|Y_t|$. We already have a bound on the size of $|X_{\ell+1}^\ell|$, and we shall show that if $|Y_t|$ is too large, then it violates ε-uniformity.

Indeed, by definition of Y_t, we have

$$d(Y_t, X_{\ell+1}^\ell) \leq \lambda - \varepsilon \leq d(V_{c(t)}, V_{c(\ell+1)}) - \varepsilon.$$

So either $|X_{\ell+1}^\ell| < \varepsilon|V_{c(t)}|$ or $|Y_t| < \varepsilon|V_{c(\ell+1)}|$. But the first cannot occur.

Indeed, write $m = |N(\ell+1, \ell)|$. Then $m + |T| \leq \Delta$. In particular, since the $|T| = 0$ case is trivial, we may assume $m \leq \Delta - 1$. So we can easily bound

$$|X_{\ell+1}| \geq (\lambda - \varepsilon)^{\Delta-1}|V_{c(t)}| \geq (\lambda^{\Delta-1} - (\Delta-1)\varepsilon)|V_{c(t)}| > \varepsilon|V_{c(t)}|.$$

Thus, by ε-uniformity, it must be the case that

$$|Y_t| \leq \varepsilon|V_{c(\ell+1)}|.$$

Therefore, we can bound

$$\left| X_{\ell+1}^{\ell} - \bigcup_{t \in T} Y_t \right| \geq (\lambda - \varepsilon)^m|V_{c(\ell+1)}| - (\Delta - m)\varepsilon|V_{c(\ell+1)}|$$

$$\geq (\lambda^m - m\varepsilon - (\Delta - m)\varepsilon)u \geq \varepsilon u > s - 1.$$

So we are done. $\qquad\square$

Corollary. Let H be a graph with vertex set $\{v_1, \ldots, v_r\}$. Let $0 < \lambda, \varepsilon < 1$ satisfy $r\varepsilon \leq \lambda^{r-1}$.

Let G be a graph with disjoint vertex subsets $V_1, \ldots, V_r$, each of size $u \geq 1$. Suppose each pair (V_i, V_j) is ε uniform, and $d(V_i, V_j) \geq \lambda$ if $v_iv_j \in E(H)$, and $d(V_i, V_j) \leq 1 - \lambda$ if $v_iv_j \notin E(H)$. Then there exists $x_i \in V_i$ so that the map $v_i \to x_i$ is an isomorphism $H \to G[\{x_1, \ldots, x_r\}]$.

Proof. By replacing the V_i-V_j edges by the complementary set whenever $v_iv_j \notin E(H)$, we may assume $d(V_i, V_j) \geq \lambda$ for all i, j, and H is a complete graph.

We then apply the previous lemma with $\Delta = r - 1$ and $s = 1$. $\qquad\square$

Szemerédi showed that every graph that is sufficiently large can be partitioned into finitely many classes, with most pairs being ε-uniform. The idea is simple — whenever we see something that is not uniform, we partition it further into subsets that are more uniform. The "hard part" of the proof is to come up with a measure of how far we are from being uniform.

Definition (Equipartition)**.** An *equipartition* of $V(G)$ into k parts is a partition into sets $V_1, \ldots, V_k$, where $\lfloor \frac{n}{k} \rfloor \leq V_i \leq \lceil \frac{n}{k} \rceil$, where $n = |G|$.

We say that the partition is ε-uniform if (V_i, V_j) is ε-uniform for all but $\varepsilon\binom{k}{2}$ pairs.

Theorem (Szemerédi's regularity lemma)**.** Let $0 < \varepsilon < 1$ and let ℓ be some natural number. Then there exists some $L = L(\ell, \varepsilon)$ such that every graph has an ε-uniform equipartition into m parts for some $\ell \leq m \leq L$, depending on the graph.

This lemma was proved by Szemerédi in order to prove his theorem on arithmetic progressions in dense subsets of integers.

When we want to apply this, we usually want at least ℓ many parts. For example, having 1 part is usually not very helpful. The upper bound on m is helpful for us to ensure the parts are large enough, by picking graphs with sufficiently many vertices.

We first need a couple of trivial lemmas.

Lemma. Let $U' \subseteq U$ and $W' \subseteq W$, where $|U'| \geq (1 - \delta)|U|$ and $|W'| \geq (1 - \delta)|W|$. Then

$$|d(U', W') - d(U, W)| \leq 2\delta.$$

Proof. Let $d = d(U, W)$ and $d' = d(U', W')$. Then

$$d = \frac{e(U, W)}{|U||W|} \geq \frac{e(U', W')}{|U||W|} = d' \frac{|U'||W'|}{|U||W|} \geq d'(1 - \delta)^2.$$

Thus,

$$d' - d \leq d'(1 - (1 - \delta)^2) \leq 2\delta d' \leq 2\delta.$$

The other inequality follows from considering the complementary graph, which tells us

$$(1 - d') - (1 - d) \leq 2\delta. \qquad \square$$

Lemma. Let $x_1, \ldots, x_n$ be real numbers with

$$X = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

and let

$$x = \frac{1}{m} \sum_{i=1}^{m} x_i.$$

Then

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 \geq X^2 + \frac{m}{n - m}(x - X)^2 \geq X^2 + \frac{m}{n}(x - X)^2.$$

If we ignore the second term on the right, then this is just Cauchy–Schwarz.

Proof. We have

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 = \frac{1}{n} \sum_{i=1}^{m} x_i^2 + \frac{1}{n} \sum_{i=m+1}^{n} x_i^2$$

$$\geq \frac{m}{n} x^2 + \frac{n - m}{n} \left(\frac{nX - mx}{n - m} \right)^2$$

$$\geq X^2 + \frac{m}{n - m}(x - X)^2$$

by two applications of Cauchy–Schwarz. $\qquad \square$

We can now prove Szemerédi's regularity lemma.

Proof. Define the index $\mathrm{ind}(\mathcal{P})$ of an equipartition $\mathcal{P}$ into k parts V_i to be

$$\mathrm{ind}(P) = \frac{1}{k^2} \sum_{i < j} d^2(V_i, V_j).$$

We show that if P is not ε-uniform, then there is a refinement equipartition $\mathcal{Q}$ into $k4^k$ parts, with $\mathrm{ind}(\mathcal{Q}) \geq \mathrm{ind}(\mathcal{P}) + \frac{\varepsilon^5}{8}$.

This is enough to prove the theorem. For choose $t \geq \ell$ with $4^t \varepsilon^5 \geq 100$. Define recursively a function f by

$$f(0) = t, \quad f(j+1) = f(j)4^{f(j)}.$$

Let

$$N = f(\lceil 4\varepsilon^{-5} \rceil),$$

and pick $L = N16^N$.

Then, if $n \leq L$, then just take an equipartition into single vertices. Otherwise, begin with a partition into t parts. As long as the current partition into k parts is not ε uniform, replace it a refinement into $4k^4$ parts. The point is that $\text{ind}(\mathcal{P}) \leq \frac{1}{2}$ for any partition. So we can't do this more than $4\varepsilon^{-5}$ times, at which point we have partitioned into $N \leq L$ parts.

Note that the reason we had to set $L = N16^N$ is that in our proof, we want to assume we have many vertices lying around.

The proof really is just one line, but students tend to complain about such short proofs, so let's try to explain it in a bit more detail. If the partition is not ε-uniform, this means we can further partition each part into uneven pieces. Then our previous lemma tells us this discrepancy allows us to push up $\frac{1}{n}\sum x_i^2$.

So given an equipartition $\mathcal{P}$ that is not ε-uniform, for each non-uniform pair (V_i, V_j) of P, we pick witness sets

$$X_{ij} \subseteq V_i, \quad X_{ji} \subseteq V_j$$

with $|X_{ij}| \geq \varepsilon|V_i|$, $|X_{ji}| \geq |V_j|$ and $|d(X_{ij}, X_{ji}) - D(V_i, V_j)| \geq \varepsilon$.

Fix i. Then the sets X_{ij} partition V_i into at most 2^{k-1} *atoms*. Let $m = \lfloor \frac{n}{k4^k} \rfloor$, and let $n = k4^k m + ak + b$, where $0 \leq a \leq 4^k$ and $b \leq k$. Then we see that

$$\lfloor n/k \rfloor = 4^k m + a$$

and the parts of $\mathcal{P}$ have size $4^k m + a$ or $4^k m + a + 1$, with b of the larger size.

Partition each part of $\mathcal{P}$ into 4^k sets, of size m or $m + 1$. The smaller V_i having a parts of size $m + 1$, and thelargrer having $a + 1$ such pairs.

We see that any such partition is an equipartition into $k4^k$ parts of size m or $m + 1$, with $ak + b$ parts of larger size $m + 1$.

Let's choose such an equipartition $\mathcal{Q}$ with parts as nearly as possible inside atoms, so each atom is a union of parts of $\mathcal{Q}$ with at most m extra vertices.

All that remains is to check that $\text{ind}(\mathcal{Q}) \geq \text{ind}(\mathcal{P}) + \frac{\varepsilon^5}{8}$.

Let the sets of $\mathcal{Q}$ within V_i be $V_i(s)$, where $1 \leq s \leq 4^k \equiv q$. So

$$V_i = \bigcup_{s=1}^{q} V_i(s).$$

Now

$$\sum_{1 \leq s,t \leq q} e(V_i(s), V_j(t)) = e(V_i, V_j).$$

We'd like to divide by some numbers and convert these to densities, but this is where we have to watch out. But this is still quite easy to handle. We have

$$\frac{m}{m+1} \leq q|V_i(s)| \leq |V_i| \leq \frac{m+1}{m}q|V_i(s)|$$

for all s. So we want m to be large for this to not hurt us too much.

So

$$\left(\frac{m}{m+1}\right)d(V_i, V_j) \leq \frac{1}{q^2}\sum_{s,t}d(V_i(s), V_j(t)) \leq \left(\frac{m+1}{m}\right)^2 d(V_i, V_j).$$

Using $n \geq k16^k$, and hence

$$\left(\frac{m}{m+1}\right)^2 \geq 1 - \frac{2}{m} \geq 1 - \frac{2}{4^k} \geq 1 - \frac{\varepsilon^5}{50},$$

we have

$$\left|\frac{1}{q^2}\sum_{s,t}d(V_i(s), V_j(t)) - d(V_i, V_j)\right| \leq \frac{\varepsilon^5}{49} + d(V_i, V_j).$$

In particular,

$$\frac{1}{q^2}\sum d^2(V_i(s), V_j(t)) \geq d^2(V_i, V_j) - \frac{\varepsilon^5}{25}.$$

The lower bound can be improved if (V_i, V_j) is not ε-uniform.

Let X_{ij}^* be the largest subset of X_{ij} that is the union of parts of $\mathcal{Q}$. We may assume

$$X_{ij}^* = \bigcup_{1 \leq s \leq r_i} V_i(s).$$

By an argument similar to the above, we have

$$\frac{1}{r_i r_j}\sum_{\substack{1 \leq s \leq r_i \\ 1 \leq t \leq r_j}} d(V_i(s), d_j(t)) \leq d(X_{ij}^*, X_{ji}^*) + \frac{\varepsilon^5}{49}.$$

By the choice of parts of $\mathcal{Q}$ within atoms, and because $|V_i| \geq qm = 4^k m$, we have

$$|X_{ij}^*| \geq |X_{ij}| - 2^{k-1}m$$
$$\geq |X_{ij}|\left(1 - \frac{2^k m}{\varepsilon|V_i|}\right)$$
$$\geq |X_{ij}|\left(1 - \frac{1}{2^k \varepsilon}\right)$$
$$\geq |X_{ij}|\left(1 - \frac{\varepsilon}{10}\right).$$

So by the lemma, we know

$$|d(X_{ij}^*, X_{ji}^*) - d(X_{ij}, X_{ji})| < \frac{\varepsilon}{5}.$$

Recalling that

$$|d(X_{ij}, X_{ji}) - d(V_i, V_j)| > \varepsilon,$$

we have

$$\left|\frac{1}{r_i r_j}\sum_{\substack{1 \leq s \leq r_i \\ 1 \leq t \leq r_j}} d(V_i(s), V_j(t)) - d(V_i, V_j)\right| > \frac{3}{4}\varepsilon.$$

We can now allow our Cauchy–Schwarz inequality with $n = q^2$ and $m = r_i r_j$ gives

$$\frac{1}{q^2} \sum_{1 \le s,t \le q} d^2(V_i(s), V_j(t)) \ge d^2(V_i, V_j) - \frac{\varepsilon^5}{25} + \frac{r_i r_j}{q^2} \cdot \frac{9\varepsilon^2}{16}$$

$$\ge d^2(V_i, V_j) - \frac{\varepsilon^5}{25} + \frac{\varepsilon^4}{3},$$

using the fact that

$$\frac{r_i}{q} \ge \left(1 - \frac{1}{m}\right) \frac{r_i}{q} \frac{m+1}{m} \ge \left(1 - \frac{1}{m}\right)$$

$$\ge \frac{|X_{ij}^*|}{|V_i|} \ge \left(1 - \frac{1}{m}\right)\left(1 - \frac{\varepsilon}{10}\right) \frac{|X_{ij}|}{|V_i|} > \frac{4\varepsilon}{5}.$$

Therefore

$$\mathrm{ind}(\mathcal{Q}) = \frac{1}{k^2 q^2} \sum_{\substack{1 \le i < j < k \\ 1 \le s,t \le q}} d^2(V_i(s), V_j(t))$$

$$\ge \frac{1}{k^2} \sum_{1 \le i < j \le k} d^2(V_i, V_j) - \frac{\varepsilon^5}{25} + \varepsilon \binom{k}{2} \frac{\varepsilon^4}{3}$$

$$\ge \mathrm{ind}(P) + \frac{\varepsilon^5}{8}. \qquad \square$$

The proof gives something like

$$L \sim 2^{2^{\cdot^{\cdot^{\cdot^2}}}},$$

where the tower is ε^{-5} tall. Can we do better than that?

In 1997, Gowers showed that the tower of height at least $\varepsilon^{-1/16}$. More generally, we can define $V_1, \ldots, V_k$ to be $(\varepsilon, \delta, \eta)$-uniform if all but $\eta \binom{k}{2}$ pairs (V_i, V_j) satisfy $|d(V_i, V_j) - d(V_i', V_j')| \le \varepsilon$ whenever $|V_i'| \ge \delta |V_i|$. Then there is a graph for which every $(1 - \delta^{1/16}, \delta, 1 - 20\delta^{1/16})$-uniform partition needs a tower height $\delta^{-1/16}$ parts.

More recently, Moskowitz and Shapira (2012) improved these bounds. Most recently, a reformulation of the lemma due to Lovász and Szegedy (2007) for which the upper bound is $\mathrm{tower}(\varepsilon^{-2})$ was shown to have lower bound $\mathrm{tower}(\varepsilon^{-2})$ by Fox and Lovász (2014) (note that these are different Lovász's!).

Let's now turn to some applications of the Szemerédi's regularity lemma. Recall that Ramsey's theorem says there exists $R(k)$ so every red-blue colouring of the edges of K_n yeidls a monochromatic K_k provided $n \ge R(k)$. There are known bounds

$$2^{k/2} \le R(k) \le 4^k.$$

The existence of $R(k)$ implies that for every graph G, there exists a number $r(G)$ minimal so if $n \ge r(G)$ and we colour K_n, we obtain a monochromatic G. Clearly, we have

$$r(G) \le R(|G|).$$

How much smaller can $r(G)$ be compared to $R(|G|)$?

Theorem. Given an integer d, there exists $c(d)$ such that

$$r(G) \leq c|G|$$

for every graph G with $\Delta(G) \leq d$.

Proof. Let $t = R(d+1)$. Pick $\varepsilon \leq \min\left\{\frac{1}{t}, \frac{1}{2^d(d+1)}\right\}$. Let $\ell \geq t^2$, and let $L = L(\ell, \varepsilon)$. We show that $c = \frac{L}{\varepsilon}$ works.

Indeed, let G be a graph. Colour the edges of K_n by red and blue, where $n \geq c|G|$. Apply Szemirédi to the red graph with ℓ, ε as above. Let H be the graph whose vertices are $\{V_1, \ldots, V_m\}$, where $V_1, \ldots, V_m$ is the partition of the red graph. Let $V_i, V_j \in E(H)$ if (V_i, V_j) is ε-uniform. Notice that $m = |H| \geq \ell \geq t^2$, and $e(\bar{H}) \leq \varepsilon\binom{m}{2}$. So $H \supseteq K_t$, or else by Turán's theroem, there are integers $d_1, \ldots, d_{t-1}$ and $\sum d_i m$ and

$$e(\bar{H}) \geq \binom{d_i}{2} \geq (t-1)\binom{m/(t-1)}{2} \geq \varepsilon\binom{m}{2}$$

by our choice of ε and m.

We may as well assume all pairs V_i, V_j for $1 \leq i < j \leq t$ are ε-uniform. We colour the edge of K_t green if $d(V_i, V_j) \geq \frac{1}{2}$ (in the red graph), or white if $< \frac{1}{2}$ (i.e. density $> \frac{1}{2}$ in the blue graph).

By Ramsey's theorem, we may assume all pairs $V_i V_j$ for $1 \leq i < j \leq d+1$ are the same colour. We may wlog assume the colour is green, and we shall find a red G (a similar argument gives a blue G if the colour is white).

Indeed, take a vertex colouring of G with at most $d+1$ colours (using $\Delta(G) \leq d$), with no colour used more than $|G|$ times. By the building lemma with H (in the lemma) being G (in this proof), and G (in the lemma) equal to the subgrpah of the red graph spanned by $V_1, \ldots, V_{d+1}$ (here),

$$u = |V_i| \geq \frac{n}{L} \geq \frac{c|G|}{L} \geq \frac{|G|}{\varepsilon},$$

$r = d+1$, $\lambda = \frac{1}{2}$, and we are done. $\qquad\square$

This proof is due to Chvátal, Rödl, Szeméredi, Trotter (1983). It was extended to more general graphs by Chen and Schelp (1993) including planar graphs. It was conjectured by Burr and Erdös (1978) to be true for d-degenerate graphs ($e(H) \leq d|H|$ for all $H \subseteq G$).

Kostochka–Rödl (2004) introduced "dependent random choice", used by Fox–Sudakov (2009) and finally Lee (2015) proved the full conjecture.

An important application of the Szeméredi regularity lemma is the *triangle removal lemma*.

Theorem (Triangle removal lemma)**.** Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $|G| = n$ and G contains at most δn^3 triangles, then there exists a set of at most εn^2 edges whose removal leaves no triangles.

Proof. Exercise. (See example sheet) $\qquad\square$

Appropriate modifications hold for general graphs, not just triangles.

Corollary (Roth, 1950's)**.** Let $\varepsilon > 0$. Then if n is large enough, and $A \subseteq [n] = \{1, 2, \ldots, n\}$ with $|A| \geq \varepsilon n$, then A contains a 3-term arithmetic progression.

Roth originally proved this by some sort of Fourier analysis arguments, while Szmerédi came and prove this for all lengths later.

Proof. Define
$$B = \{(x, y) \in [2n]^2 : x - y \in A\}.$$
Then certainly $|B| \geq \varepsilon n^2$. We form a 3-partite graph with disjoint vertex classes $X = [2n]$ and $Y = [2n]$, $Z = [4n]$. If we have $x \in X, y \in Y$ and $z \in Z$, we join x to y if $(x, y) \in B$; join x to z if $(x, z - x) \in B$ and join y to z if $(z - y, y) \in B$.

A triangle in G is a triple (x, y, z) where $(x, y), (x, y + w), (x + w, y) \in B$ where $w = z - x - y$. Note that $w < 0$ is okay. A 0-triangle is one with $w = 0$. There are at least εn^2 of these, one for each $(x, y) \in B$, and these are edge disjoint, because $z = x + y$.

Hence the triangles cannot be killed by removing $\leq \varepsilon n^2 / 2$ edges. By the triangle removal lemma, we must have $\geq \delta n^3$ triangles for some δ. In particular, for n large enough, there is some triangle that is not a 0-triangle.

But then we are done, since
$$x - y - w, x - y, x - y + w \in A$$
where $w \neq 0$, and this is a 3-term arithemtic progression. $\qquad\square$

There is a simple generalization of this argument which yields k-term arithmetic progressions provided we have a suitable removal lemma. This needs a Szemerédi regularity lemma for $(k - 1)$-uniform hypergraphs, instead of just graphs, and a corresponding version of the building lemma.

The natural generalizations of Szemerédi's lemma to hypergraphs is easily shown to be true (exercise). The catch is that what the Szemerédi's lemma does not give us a strong enough result to apply the building lemma.

What we need is a stronger version of the regularity that allows us to build graphs, but not too strong that we can't prove the Szemerédi regularity lemma. A workable hypergraph regularity lemma along these lines was proved by Nagle, Rödl, Skokan, and by Gowers (those weren't quite the same lemma, though).

5 Subcontraction, subdivision and linking

We begin with some definitions.

Definition. Let G be a graph, $e = xy$ an edge. Then the graph G/e is the graph formed from $G \setminus \{x, y\}$ by adding a new vertex joined to all neighbours of x and y. We say this is the graph formed by *contracting* the edge e.

Definition ((Sub)contraction). A *contraction* of G is a graph obtained by a sequence of edge contractions. A *subcontraction* of G is a contraction of a subgraph. We write $G \succ H$ if H is a subcontraction of G.

If $G \succ H$, then G has disjoint vertex subsets W_v for each $v \in V(H)$ such that $G[W_v]$ is connected, and there is an edge of G between W_u and W_v if $w \in E(H)$.

Definition (Subdivision). If H is a graph, TH stands for any graph obtained from H by replacing its edges by vertex disjoint paths (i.e. we subdivide edges).

The T stands for "topological", since the resulting graph has the same topology.

Clearly, if $G \supseteq TH$, then $G \succ H$.
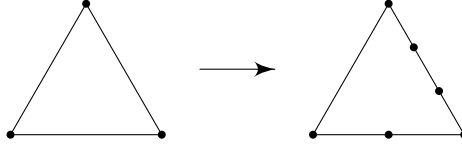
Recall the following theorem:

Theorem (Menger's theorem). Let G be a graph and $s_1, \ldots, s_k, t_1, \ldots, t_k$ be distinct vertices. If $\kappa(G) \geq k$, then there exists k vertex disjoint paths from $\{s_1, \ldots, s_k\}$ to $\{t_1, \ldots, t_k\}$.

This is good, but not good enoguh. It would be nice if we can have paths that join s_i to t_i, as opposed to joining s_i to any element in $\{t_1, \ldots, t_k\}$.

Definition (k-linked graph). We say G is *k-linked* if there exists vertex disjoint s_i-t_i paths for $1 \leq i \leq k$ for any choice of these $2k$ vertices.

We want to understand how these three notions interact. There are some obvious ways in which they interact. For example, it is not hard to see that $G \supseteq TK_t$ if G is k-linked for some large k. Here is a kind of converse.

Lemma. If $\kappa(G) \geq 2k$ and $G \supseteq TK_{5k}$, then G is k-linked.

Proof. Let B be the set of "branch vertices" of TK_{5k}, i.e. the $5k$ vertices joined by paths. By Menger's theorem, there exists $2k$ vertex disjoint paths joining $\{s_1, \ldots, s_k, t_1, \ldots, t_k\}$ to B (note that these sets might intersect). We say one of our $2k$ paths *impinges* on a path in TK_{5k} if it meets that path and subsequently leaves it. Choose a set of $2k$ paths impinging minimally (counting 1 per impingement).

Let these join $\{s_1, \ldots, t_k\}$ to $\{v_1, \ldots, v_{2k}\}$, where $B = \{v_1, \ldots, v_{5k}\}$. Then no path impinges on a path in TK_{5k} from $\{v_1, \ldots, v_{2k}\}$ to $\{v_{2k+1}, \ldots, v_{5k}\}$.

Otherwise, pick the path impinging closest to v_{2k+j}, and reroute it to v_{2k+j} rather than to something in $\{v_1, \ldots, v_{2k}\}$, reducing the number of impingements.

Hence each path (of any $2k$) meet at most one path in $\{v_1, \ldots, v_{2k}\}$ to $\{v_{2k+1}, \ldots, v_{5k}\}$ (once we hit it, we must stay there).

Thus, we may assume the paths from $\{v_1, \ldots, v_{2k}\}$ to $\{v_{4k+1}, \ldots, v_{5k}\}$ are not met at all. Then we are done, since we can link up v_j to v_{k+j} via v_{4k+j} to join s_j to t_k.   □

The argument can be improved to use TK_{3k}.

Since we are doing extremal graph theory, we want to ask the following question — how many edges do we need to guarantee $G \succ K_t$ or $G \supseteq TK_t$?

For two graphs G and H, we write $G + H$ for the graph formed by taking the disjoint union of G and H and then joining everything in G to everything in H.

Lemma. If $e(G) \geq k|G|$, then there exists some H with $|H| \leq 2k$ and $\delta(H) \geq k$ such that $G \succ K_1 + H$.

Proof. Consider the minimal subcontraction G' of G among those that satisfy $e(G') \geq k|G'|$. Then we must in fact have $e(G') = k|G'|$, or else we can just throw away an edge.

Since $e(G') = k|G'|$, it must be the case that $\delta(G') \leq 2k$. Let v be of minimum degree in G'. and set $H = G'[\Gamma(v)]$. Then $G \succ K_1 + H$ and $|H| \leq 2v$.

To see that $\delta(H) \geq k$, suppose $u \in V(H) = \Gamma(v)$. Then by minimality of G', we have

$$e(G'/uv) \leq k|G'| - k - 1.$$

But the number of edges we kill when performing this contraction is exactly 1 plus the number of triangles containing uv. So uv lies in at least k triangles of G'. In other words, $\delta(H) \geq k$.   □

By iterating this, we can find some subcontractions into K_t.

Theorem. If $t \geq 3$ and $e(G) \geq 2^{t-3}|G|$, then $G \succ K_t$.

Proof. If $t = 3$, then G contains a cycle. So $G \succ K_3$. If $t > 3$, then $G \succ K_1 + H$ where $\delta(H) \geq 2^{t-3}$. So $e(H) \geq 2^{t-4}|H|$ and (by induction) $H \succ K_{t-1}$.   □

We can prove similar results for the topological things.

Lemma. If $\delta(G) \geq 2k$, then G contains vertex disjoint subgraphs H, J with $\delta(H) \geq k$, J connected, and every vertex in H has a neighbour in J.

If we contract J to a single vertex, then we get an $H + K_1$.

Proof. We may assume G is connected, or else we can replace G by a component.

Now pick a subgraph J maximal such that J is connected and $e(G/J) \geq k(|G| - |J| + 1)$. Note that any vertex could be used for J. So such a J exist.

Let H be the subgraph spanned by the vertices having a neighbour in J. Note that if $v \in V(H)$, then v has at least k neighbours in H. Indeed, when we contract $J \cup \{v\}$, maximality tells us

$$e(G/(J \cup \{v\})) \leq k(|G| - |J| + 1) - k - 1,$$

and the number of edges we lose in the contraction is 1 plus the number of neighbours of v (it is easier to think of this as a two-step process — first contract

J, so that we get an $H + K_1$, then contract v with the vertex coming from K_1). $\qquad\square$

Again, we iterate this result.

Theorem. Let F be a graph with n edges and no isolated vertices. If $\delta(G) \geq 2^n$, then $G \supseteq TF$.

Proof. If $n = 1$, then this is immediate. If F consists of n isolated edges, then $F \subseteq G$ (in fact, $\delta(G) \geq 2n - 1$ is enough for this). Otherwise, pick an edge e which is not isolated. Then $F - e$ has at most 1 isolated vertex. Apply the previous lemma to G to obtain H with $\delta(H) \geq 2^{n-1}$. Find a copy of $T(F - e)$ in H (apart from the isolated vertex, if exists).

If e was just a leaf, then just add an edge (say to J). If not, just construct a path going through J to act as e, since J is connected. $\qquad\square$

It is convenient to make the definition

$$c(t) = \inf\{c : e(G) \geq c|G| \Rightarrow G \succ K_t\}$$
$$t(t) = \inf\{c : e(G) \geq c|G| \Rightarrow G \supseteq TK_t\}.$$

We can interpret the first result as saying $c(t) \leq 2^{t-3}$. Moreover, note that if $e(G) \geq k|G|$, then G must contain a subgraph with $\delta \geq k$. Otherwise, we can keep removing vertices of degree $< k$ till we have none left, but still have a positive number of edges, which is clearly nonsense. Hence, the second result says $t(t) \leq 2^{\binom{t}{2}}$.

Likewise, we can define

$$f(k) = \min\{c : \kappa(G) \geq c \Rightarrow G \text{ is } k\text{-linked}\}.$$

Since $\kappa(G) \geq c$ implies $\delta(G) \geq c$, the existence of $t(t)$ and the beginning lemma implies that $f(k)$ exists.

In 1967, Mader showed $c(t) = t - 2$ for $t \leq 7$. Indeed, for $t \leq 7$, we have the exact result

$$\text{ex}(n; \succ K_t) = (t - 2)n - \binom{t - 1}{2}.$$

So for some time, people though c is linear. But in fact, $c(t)$ is superlinear.

Lemma. We have
$$c(t) \geq (\alpha + o(1))t\sqrt{\log t}.$$

To determine α, let $\lambda < 1$ be the root to $1 - \lambda + 2\lambda \log \lambda = 0$. In fact $\lambda \approx 0.284$. Then

$$\alpha = \frac{1 - \lambda}{2\sqrt{\log 1/\lambda}} \approx 0.319.$$

Proof. Consider a random graph $G(n, p)$, where p is a constant to be chosen later. So $G(n, p)$ has n vertices and edges are chosen independently at random, with probability p. Here we *fix* a t, and then try to find the best combination of n and p to give the desired result.

A partition $V_1, \ldots, V_t$ of $V(G)$ is said to be *complete* if there is an edge between V_i and V_j for all i and j. Note that having a complete partition is a

necessary, but not sufficient condition for having a K_t minor. So it suffices to show that with probability > 0, there is no complete partition.

Writing $q = 1-p$, a given partition with $|V_i| = n_i$ is complete with probability

$$\prod_{i,j}(1 - q^{n_i n_j}) \leq \exp\left(-\sum_{i<j} q^{n_i n_j}\right)$$

$$\leq \exp\left(-\binom{t}{2}\prod q^{n_i n_j/\binom{t}{2}}\right) \qquad \text{(AM-GM)}$$

$$\leq \exp\left(-\binom{t}{2}q^{n^2/t^2}\right).$$

The expected number of complete partitions is then

$$\leq t^n \exp\left(-\binom{t}{2}q^{n^2/t^2}\right),$$

As long as we restrict to the choices of n and q such that

$$t > n\sqrt{\frac{\log(1/q)}{\log n}},$$

we can bound this by

$$\leq \exp\left(n\log t - \binom{t}{2}\frac{1}{n}\right) = o(1)$$

in the limit $t \to \infty$. We set

$$q = \lambda, \quad n = \frac{t\sqrt{\log t}}{\sqrt{\log 1/\lambda}}.$$

Then with probability > 0, we can find a graph with only $o(1)$ many complete partitions, and by removing $o(n)$ many edges, we have a graph with

$$p\binom{n}{2} - o(n) = (\alpha + o(1))t\sqrt{\log t} \cdot n$$

many edges with no K_t minor. $\qquad\square$

At this point, the obvious question to ask is — is $t\sqrt{\log t}$ the correct growth rate? The answer is yes, and perhaps surprisingly, the proof is also probabilistic. Before we prove that, we need the following lemma:

Lemma. Let $k \in \mathbb{N}$ and G be a graph with $e(G) \geq 11k|G|$. Then there exists some H with
$$|H| \leq 11k + 2, \quad 2\delta(H) \geq |H| + 4k - 1$$
such that $G \succ H$.

Proof. We use our previous lemma as a starting point. Letting $\ell = 11k$, we know that we can find H_1 such that

$$G \succ K_1 + H_1,$$

with $|H_1| \leq 2\ell$ and $\delta(H_1) \geq \ell$. We shall improve the connectivity of this H_1 at the cost of throwing away some elements.

By divine inspiration, consider the constant $\beta \approx 0.37$ defined by the equation

$$1 = \beta \left(1 + \frac{\log 2}{\beta} \right),$$

and the function ϕ defined by

$$\phi(F) = \beta\ell \frac{|F|}{2} \left(\log \frac{|F|}{\beta\ell} + 1 \right).$$

Now consider the set of graphs

$$\mathcal{C} = \{ F : |F| \geq \beta\ell, e(F) \geq \phi(F) \}.$$

Observe that $H_1 \in \mathcal{C}$, since $\delta(H_1) \geq \ell$ and $|H_1| \leq 2\ell$.

Let H_2 be a subcontraction of H_1, minimal with respect to $\mathcal{C}$.

Since the complete graph of order $\lceil \beta\ell \rceil$ is not in $\mathcal{C}$, as $\phi > \binom{\beta\ell}{2}$, the only reason we are minimal is that we have hit the bound on the number of edges. Thus, we must have

$$|H_2| \geq \beta\ell + 1, \quad e(H_2) = \lceil \phi(H_2) \rceil,$$

and

$$e(H_2/uv) < \phi(H_2/uv)$$

for all edges uv of H_2.

Choose a vertex $u \in H_2$ of minimum degree, and put $H_3 = H_2[\Gamma(u)]$. Then we have

$$|H_3| = \delta(H_2) \leq \frac{2\lceil \phi(H_2) \rceil}{|H_2|} \leq \left\lfloor \beta\ell \left(\log \left(\frac{|H_2|}{\beta\ell} \right) + 1 \right) + \frac{2}{|H_2|} \right\rfloor \leq \ell,$$

since $\beta\ell \leq |H_2| \leq 2\ell$.

Let's write $b = \beta\ell$ and $h = |H_2|$. Then by the usual argument, we have

$$2\delta(H_3) - |H_3|$$
$$\geq 2(\phi(H_2) - \phi(H_2/uv) - 1) - |H_3|$$
$$\geq bh \left(\log \frac{h}{\beta} + 1 \right) - b(h-1) \left(\log \frac{h-1}{b} + 1 \right) - b \left(\log \frac{h}{b} + 1 \right) - 3$$
$$= b(h-1) \log \frac{h}{h-1} - 3$$
$$> b - 4,$$

because $h \geq b + 1$ and $x \log \left(1 + \frac{1}{x} \right) > 1 - \frac{1}{x}$ for real $x > 1$. So

$$2\delta(H_3) - |H_3| \geq \beta\ell - 4 > 4k - 4.$$

If we put $H = K_2 + H_3$, then $G \succ H$, $|H| \leq 11k + 2$ and

$$2\delta(H) - |H| \geq 2\delta(H_3) - |H_3| + 2 \geq 4k - 1. \qquad \square$$

Theorem. We have
$$c(t) \le 7t\sqrt{\log t}$$
if t is large.

The idea of the proof is as follows: First, we pick some H with $G \succ H$ such that H has high minimum degree, as given by the previous lemma. We now randomly pick disjoint subsets $U_1, \ldots, U_{2t}$, and hope that they give us a subcontraction to K_t. There are more sets that we need, because some of them might be "bad", in the sense that, for example, there might be no edges from U_1 to U_{57}. However, since H has high minimum degree, the probability of being bad is quite low, and with positive probability, we can find t subsets that have an edge between any pair. While the U_i are not necessarily connected, this is not a huge problem since H has so many edges that we can easily find paths that connect the different vertices in U_i.

Proof. For large t, we can choose $k \in \mathbb{N}$ so that
$$5.8t\sqrt{\log_2 t} \le 11k \le 7t\sqrt{\log t}.$$
Let $\ell = \lceil 1.01\sqrt{\log_2 t} \rceil$. Then $k \ge t\ell/2$. We want to show that $G \succ K_t$ if $e(G) \ge 11k|G|$.

We already know that $G \succ H$ where $h = |H| \le 11k + 2$, and $2\delta(H) \ge |H| + 4k - 1$. We show that $H \succ K_t$.

From now on, we work entirely in H. Note that any two adjacent vertices have $\ge 4k + 1$ common neighbours. Randomly select $2t$ disjoint ℓ sets $U_1, \ldots, U_{2t}$ in $V(H)$. This is possible since we need $2t\ell$ vertices, and $2t\ell \le 4k$.

Of course, we only need t many of those ℓ sets, and we want to select the ones that have some chance of being useful to us.

Fix a part U. For any vertex v (not necessarily in U), its degree is at least $h/2$. So the probability that v has no neighbour in U is at most $\binom{h/2}{\ell}/\binom{h}{\ell} \le 2^{-\ell}$. Write $X(U)$ for the vertices having no neighbour in U, then $\mathbb{E}|X(U)| \le 2^{-\ell}h$.

We say U is *bad* if $|X(U)| > 8 \cdot 2^{-\ell}h$. Then by Markov's inequality, the probability that U is bad is $< \frac{1}{8}$. Hence the expected number of bad sets amongst $U_1, \ldots, U_{2t}$ is $\le \frac{t}{4}$. By Markov again, the probability that there are more than $\frac{t}{2}$ bad sets is $< \frac{1}{2}$.

We say a pair of sets (U, U') is *bad* if there is no $U - U'$ edge. Now the probability that U, U' is bad is $\mathbb{P}(U' \subseteq X(U))$. If we condition on the event that U is good (i.e. not bad), then this probability is bounded by

$$\binom{8 \cdot 2^{-\ell}h}{\ell}/\binom{h-\ell}{\ell} \le 8^\ell 2^{-\ell^2}\left(1 + \frac{\ell}{h-\ell}\right)^\ell \le 8^\ell 2^{-\ell^2}e^{\ell^2/(h-\ell)} \le 9^\ell 2^{-\ell^2}$$

if t is large (hence ℓ is slightly large and h is very large). We can then bound this by $\frac{1}{8t}$.

Hence, the expected number of bad pairs (where one of them is good) is at most
$$\frac{1}{8t}\binom{2t}{2} \le \frac{t}{4}.$$

So the probability that there are more than $\frac{t}{2}$ such bad pairs is $< \frac{1}{2}$.

Now with positive probability, $U_1, \ldots, U_{2t}$ has at most $\frac{t}{2}$ bad sets and at most $\frac{t}{2}$ bad pairs amongst the good sets. Then we can find t many sets that are pairwise good. We may wlog assume they are $U_1, \ldots, U_t$.

Fixing such a choice $U_1, \ldots, U_t$, we now work deterministically. We would be done if each U_i were connected, but they are not in general. However, we can find $W_1, \ldots, W_t$ in $V(H) \setminus (U_1 \cup \cdots U_t)$ such that $U_i \cup W_i$ is connected for $1 \leq i \leq t$.

Indeed, if $U_i = \{u_1, \ldots, u_\ell\}$, we pick a common neighbour of u_{j-1} and u_j if $u_{j-1} u_j \notin E(H)$, and put it in W_i. In total, this requires us to pick $\leq t\ell$ distinct vertices in $V(H) - (U_1 \cup \cdots \cup U_t)$, and we can do this because u_{j-1} and u_j have at least $4k - t\ell \geq t\ell$ common neighbours in this set. $\qquad\square$

It has been shown (2001) that equality holds in the previous lemma.

So we now understand $c(t)$ quite well. How about subdivision and linking? We decide to work on $f(k)$ now, because Robertson and Seymour (1995) showed that our first lemma that holds if $G \succ TK_{3k}$ instead of $G \supseteq K_{3k}$. We also know how many edges we need to get a minor. Combining with our previous theorem, we know

$$f(k) = O(k\sqrt{\log k}).$$

We know we can't do better than k, but we can get it down to order k by proving our first lemma under the weaker hypothesis that $G \succ H$ where H is dense.

So far, we have proven theorems that say, roughly, "if G is a graph with enough edges, then it subcontracts to some highly connected thing". Now saying that G subcontracts to some highly connected thing is equivalent to saying we can find disjoint non-empty connected subgraphs $D_1, \ldots, D_m$ such that each D_i is connected, and there are many edges between the D_i. The lemma we are going to prove next says under certain conditions, we can choose the D_i in a way that they contain certain prescribed points.

Definition (*S*-cut). Given $S \subseteq V(G)$, an *S*-cut is a pair (A, B) of subsets of the vertices such that $A \cup B = V(G)$, $S \subseteq A$ and $e(A \setminus B, B \setminus A) = 0$. The *order* of the cut is $|A \cap B|$.

We say (A, B) *avoids* C if $A \cap V(C) = \emptyset$.

Example. For any $S \subseteq A$, the pair $(A, V(G))$ is an *S*-cut.

Lemma. Let $d \geq 0$, $k \geq 2$ and $h \geq d + \lfloor 3k/2 \rfloor$ be integers.

Let G be a graph, $S = \{s_1, \ldots, s_k\} \subseteq V(G)$. Suppose there exists disjoint subgraphs $C_1, \ldots, C_h$ of G such that

($*$) Each C_i is either connected, or each of its component meets S. Moreover, each C_i is adjacent to all but at most d of the C_j, $j \neq i$ not meeting S.

($\dagger$) Moreover, no *S*-cut of order $< k$ avoids $d + 1$ of $C_1, \ldots, C_h$.

Then G contains disjoint non-empty connected subgraphs $D_1, \ldots, D_m$, where

$$m = h - \lfloor k/2 \rfloor,$$

such that for $1 \leq i \leq k$, $s_i \in D_i$, and D_i is adjacent to all but at most d of $D_{k+1}, \ldots, D_m$.

Proof. Suppose the theorem is not true. Then there is a minimal counterexample G (with minimality defined with respect to proper subgraphs).

We first show that we may assume G has no isolated vertices. If v is an isolated vertex, and $v \notin S$, then we can simply apply the result to $G - v$. If $v \in S$, then the S-cut $(S, V(G) \setminus \{v\})$ of order $k - 1$ avoids $h - k \geq d - 1$ of the C_i's.

Claim. If (A, B) is an S-cut of order k avoiding $d + 1$ many of the C_i's, then $B = V(G)$ and A is discrete.

Indeed, given such an S-cut, we define

$$S' = A \cap B$$
$$G' = G[B] - E(S')$$
$$C_i' = C_i \cap G'.$$

We now make a further claim:

Claim. $(G', S', \{C_i'\})$ satisfies the hypothesis of the lemma.

We shall assume this claim, and then come back to establishing the claim.

Assuming these claims, we show that G isn't a countrexample after all. Since $(G', S', \{C_i'\})$ satisfies the hypothesis of the lemma, by minimality, we can find subgraphs $D_1', \ldots, D_m'$ in G' that satisfy the conclusion of the lemma for G' and S'. Of course, these do not necessarily contain our s_iAssuming these claims, we can construct the desired D_i. However, we can just add paths from S' to the s_i.

Indeed, $G[A]$ has no S-cut (A'', B'') of order less than k with $S' \subseteq B''$, else $(A'', B'' \cup B)$ is an S-cut of G of the forbidden kind, since A, and hence A'' avoids too many of the C_i. Hence by Menger's theorem, there are vertex disjoint paths $P_1, \ldots, P_k$ in $G[A]$ joining s_i to s_i' (for some labelling of S'). Then simply take

$$D_i = D_i' \cup P_i.$$

To check that $(G', S', \{C_i'\})$ satisfy the hypothesis of the lemma, we first need to show that the C_i' are non-empty.

If (A, B) avoids C_j, then by definition $C_j \cap A = \emptyset$. So $C_j = C_j'$. By assumption, there are $d + 1$ many C_j's for which this holds. Also, since these don't meet S, we know each C_i is adjacent to at least one of these C_j's. So in particular, C_i' is non-empty since $C_i' \supseteq C_i \cap C_j' = C_i \cap C_j$.

Let's now check the conditions:

$(*)$ For each C_i', consider its components. If there is some component that does not meet S', then this component is contained in $B \setminus A$. So this must also be a component of C_i. But C_i is connected. So this component is C_i. So $C_i' = C_i$. Thus, either C_i' is connected, or all components meet S'.

Moreover, since any C_j' not meeting S' equals C_j, it follows that each C_i and so each C_i' is adjacent to all but at most d of these C_j's. Therefore $(*)$ holds.

(†) If (A', B') is an S'-cut in G' avoiding some C_i', then $(A \cup A', B')$ is an S-cut of G of the same order, and it avoids C_i (since C_i' doesn't meet S', and we have argued this implies $C_i' = C_i$, and so $C_i \cap A = \emptyset$). In particular, no S'-cut (A', B') of G' has order less than k and still avoids $d + 1$ of $C_1', \ldots, C_h'$.

––––––––––––

We can now use our original claim. We know what S-cuts of order k loop like, so we see that if there is an edge that does not join two C_i's, then we can contract the eedge to get a smaller counterexample. Recalling that G has no isolated vertices, we see that in fact $V(G) = \cup V(C_i)$, and $|C_i| = 1$ unless $C_i \subseteq S$.

Let

$$C = \bigcup \{V(C_i) : |C_i| \geq 2\}.$$

We claim that there is a set I of $|C|$ independent edges meeting C. If not, by Hall's theorem, we can find $X \subseteq C$ whose neighbourhood Y (in $V(G) - S$) such that $|Y| < |X|$. Then $(S \cup Y, V(G) - X)$ is an S-cut of order $|S| - |X| + |Y| < |S| = k$ avoiding $\geq |G| - |S| - |Y|$ many C_i's, and this is

$$\geq |G| - |S| - |X| + 1 \geq |G| - |C| - k + 1.$$

But $h \leq |G| - |C| + \frac{|C|}{2}$, since $|G| - |C|$ is the number of C_i's of size 1, so we can bound the above by

$$\geq h - \frac{|C|}{2} - k + 1 \geq h - \frac{3k}{2} + 1 \geq d + 1.$$

This contradiction shows that I exists.

Now we can just write down the D_i's. For $1 \leq i \leq k$, set D_i to be s_i if s_i is a C_ℓ with $|C_\ell| = 1$, i.e. $s_i \notin C$. If $s_i \in C$, let D_i be the edge of I meeting S_i.

Let $D_{k+1}, \ldots, D_m$ each be single vertices of $G - S - $ (ends of I). Note those exist because $\geq |G| - k - |C| \geq h - \lfloor \frac{3k}{2} \rfloor = m - k$ vertices are avaiable. Note that each D_i contains a C_ℓ with $|C_\ell| = 1$. So each D_i is joined to all but $\leq d$ many D_j's. $\qquad\square$

The point of this result is to prove the following theorem:

Theorem. Let G be a graph with $\kappa(G) \geq 2k$ and $e(G) \geq 11k|G|$. Then G is k-linked. In particular, $f(k) \leq 22k$.

Note that if $\kappa(G) \geq 22k$ then $\delta(G) \geq 22k$, so $e(G) \geq 11k|G|$.

Proof. We know that under these conditions, $G \succ H$ with $|H| \leq 11k + 2$ and $2\delta(H) \geq |H| + 4k - 1$. Let $h = |H|$, and $d = h - 1 - \delta(H)$. Note that

$$h = 2d - 2 + 2\delta(H) - h \geq 2d + 4k.$$

Let $C_1, \ldots, C_k$ be connected subgraphs of G which contract to form H. Then clearly each C_i is joined to all but at most $h - 1 - \delta(H) = d$ other C_j's. Now let $s_1, \ldots, s_k, t_1, \ldots, t_k$ be the vertices we want to link up. Let $S = \{s_1, \ldots, s_k, t_1, \ldots, t_k\}$. Observe that G has no S-cut of order $< 2k$ at all since $\kappa(G) \geq 2k$. So the conditions are satisfied, but with $2k$ instead of k.

So there are subgraphs $D_1, \ldots, D_m$, where $m = h - k \geq 2d + 3k$ as described. We may assume $s_i \in D_i$ and $t_i \in D_{k+i}$. Note that for each pair (D_i, D_j), we can find $\geq m - 2d \geq 3k$ other D_ℓ such that D_ℓ is joined to both D_i and D_j. So for each $i = 1, \ldots, k$, we can find an unused D_ℓ not among $D_1, \ldots, D_{2k}$ that we can use to connect up D_i and D_{k+i}, hence s_i and t_i. $\qquad\square$

In 2004, the coefficient was reduced from 11 to 5. This shows $f(k) \leq 10k$. It is known that $f(1) = 3, f(2) = 6$ and $f(k) \geq 3k - 2$ for $k \geq 3$.

Finally, we consider the function $t(t)$. We begin by the following trivial lemma:

Lemma. If $\delta(a) \geq t^2$ and G is $\binom{t+2}{3}$-linked, then $G \supseteq TK_t$.

Proof. Since $\delta(G) \geq t^2$, we may pick vertices $v_1, \ldots, v_t$ and sets $U_1, \ldots U_t$ all disjoint so that $|U_i| = t - 1$ and v_i is joined to U_i. Then $G - \{v_1, \ldots, v_k\}$ is still $\binom{t+2}{2} - t = \binom{t}{2}$-linked. So $U_1, \ldots, U_{t-1}$ may be joined with paths to form a TK_t in G. $\qquad\square$

To apply our previous theorem together with this lemma, we need the following lemma:

Lemma. Let $k, d \in \mathbb{N}$ with $k \leq \frac{d+1}{2}$, and suppose $e(G) \geq d|G|$. Then G contains a subgraph H with

$$e(H) = d|H| - kd + 1, \quad \delta(H) \geq d + 1, \quad \kappa(H) \geq k.$$

Proof. Define

$$\mathcal{E}_{d,k} = \{F \subseteq G : |F| \geq d, e(F) > d|F| - kd\}.$$

We observe that $D \in \mathcal{E}_{d,k}$, but $K_d \notin \mathcal{E}_{d,k}$. Hence $|F| > d$ for all $F \in \mathcal{E}_{d,k}$. let H be a subgraph of G minimal with respect to $H \in \mathcal{E}_{d,k}$. Then $e(H) = d|H| - kd + 1$, and $\delta(H) \geq d + 1$, else $H - v \in \mathcal{E}_{d,k}$ for some $v \in H$.

We only have to worry about the connectivity condition. Suppose S is a cutset of H. Let C be a component of $G - S$. Since $\delta(h) \geq d + 1$, we have $|C \cup S| \geq d + 2$ and $|H - C| \geq d + 2$.

Since neither $C \sup S$ nor $H - C$ lies in that class, we have

$$e(C \cup S) \leq d|C \cup S| - kd \leq d|C| + d|S| - kd$$

and

$$e(H - C) \leq d|H| - d|C| - kd.$$

So we find that

$$e(H) \leq d|H| + d|S| - 2kd.$$

But we know this is equal to $d|H| - kd + 1$. So we must have $|S| \geq k$. $\qquad\square$

Theorem.
$$\frac{t^2}{16} \leq t(t) \leq 13\binom{t+1}{2}.$$

Recall our previous upper bound was something like 2^{t^2}. So this is an improvement.

Proof. The lower bound comes from (disjoint copies of) $K_{t^2/8, t^2/8}$. For the uppper bound, write

$$d = 13\binom{t+1}{2}, \quad k = t \cdot (t+1).$$

Then we can find a subgraph H with $\delta(H) \geq d+1$ and $\kappa(H) \geq k+1$.

Certainly $|H| \geq d+2$, and we know

$$e(H) \geq d|H| - kd > (d-k)|H| \geq 11\binom{t+1}{2}|H|.$$

So we know H is $\binom{t+1}{2}$-linked, and so $H \supseteq TK_t$. $\square$

So $t(t)$ is, more or less, t^2!

6 Extremal hypergraphs

Define $K_\ell^\ell(t)$ be the complete ℓ-partite ℓ-uniform hypergraph with ℓ classes of size t, containing all t^ℓ edges with one vertex in each class.

Theorem (Erdös)**.** Let G be ℓ-uniform of order n with $p\binom{n}{\ell}$ edges, where $p \geq 2n^{-1/t^{\ell-1}}$. Then G contains $K_\ell^\ell(t)$ (provided n is large).

If we take $\ell = 2$, then this agrees with what we know about the exact value of the extremal function.

Proof. Assume that $t \geq 2$. We proceed by induction on $\ell \geq 2$. For each $(\ell-1)$-set $\sigma \subseteq V(G)$, let
$$N(\sigma) = \{v : \sigma \cup \{v\} \in E(G)\}.$$

Then the average degree is
$$\binom{n}{\ell-1}^{-1} \sum |N(\sigma)| = \binom{n}{\ell-1}^{-1} \ell|E(G)| = p(n - \ell + 1).$$

For each of the $\binom{n}{t}$ many t-sets $T \subseteq V(G)$, let
$$D(T) = \{\sigma : T \subseteq N(\sigma)\}.$$

Then
$$\sum_T |D(T)| = \sum_\sigma \binom{|N(\sigma)|}{t} \geq \binom{n}{\ell-1}\binom{p(n-\ell+1)}{t},$$

where the inequality follows from convexity, since we can assume that $p(n - \ell + 1) \geq t - 1$ if n is large.

In particular, there exists a T with
$$|D(T)| \geq \binom{n}{t}^{-1}\binom{n}{\ell-1}\binom{p(n-\ell+1)}{t} \geq \frac{1}{2}p^t\binom{n}{\ell-1}.$$

when n is large. By our choice of t, we can write this as
$$|D(T)| \geq 2^{t-1}n^{-1/t^{\ell-2}}\binom{n}{\ell-1}.$$

If $\ell = 2$, then this is $\geq 2^{t-1} \geq t$. So T is joined to a T-set giving $K_{t,t}$.

If $\ell \geq 3$, then $|D(T)| \geq 2n^{-1/t^{\ell-2}}\binom{n}{\ell-1}$. So, by induction, the $(\ell-1)$-uniform hypergraph induced by the σ with $T \subseteq N(\sigma)$ contains $K_{\ell-1}^{\ell-1}(t)$, giving $K_\ell^\ell(t)$ with T. $\qquad\square$

We can do a simple random argument to show that this is a right-order of magnitude.

Index