Extremal Graph Theory

Lectured by A. Thomason

Lent Term 2013

Please let me know of corrections: glt1000@cam.ac.uk

Last updated: Sat 7th May, 2016

Course description

Extremal graph theory is an umbrella title for the study of graph properties and their dependence on the values of graph parameters. This course builds on the material introduced in the Part II Graph Theory course, in particular Turán's theorem and the Erdős-Stone theorem, as well as developing the use of randomness in combinatorial proofs. Further techniques and extensions to hypergraphs will be discussed. It is intended to cover some reasonably large subset of the following.

The Erdős-Stone theorem and stability. Supersaturation. Szemerédi's Regularity Lemma, with applications. The number of complete subgraphs.

Hypergraphs. Erdős's r-partite theorem. Instability. The Fano plane. Razborov's flag algebras. Hereditary properties and their sizes.

Probabilistic tools: the Local Lemma and concentration inequalities. The chromatic number of a random graph. The semi-random method, large independent sets and the Erdős-Hanani problem. Dependent random choice.

Pre-requisite Mathematics

A knowledge of the basic concepts, techniques and results of graph theory, such as that afforded by the Part II Graph Theory course.

Literature

No book covers the course but the following can be helpful.

B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics **184**, Springer-Verlag, New York (1998), xiv+394 pp.

N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, 3rd ed. (2008)

Lecture 1

1. The Erdős-Stone Theorem

Recall:

Turán's theorem. If $|G| = n$, $e(G) \geqslant t_r(n)$ and $G \not\supset K_{r+1}$, then $G = T_r(n)$, the r-partite Turán graph of order n.

Note. $T_r(n)$ is the complete r-partite graph with class sizes $\lceil n/r \rceil$ and $\lfloor n/r \rfloor$, and $t_r(n)$ is $e(T_r(n))$.

This answers the *extremal* problem for K_{r+1}, and there is a unique extremal graph. The structure of $T_r(n)$ invites many proofs by induction.

In general, we are interested in $ex(n, F)$ for some fixed graph F, where

$$ex(n, F) = \max \left\{ e(G) : |G| = n, F \not\subset G \right\}$$

So Turán's theorem states that $ex(n, K_{r+1}) = t_r(n) \approx \left(1 - \frac{1}{r}\right)\binom{n}{2}$.

Comment. It is true that $t_r(n) \geqslant \left(1 - \frac{1}{r}\right)\binom{n}{2}$. This is equivalent to the average degree being $\geqslant \left(1 - \frac{1}{r}\right)(n-1)$. But in fact the minimum degree is $\geqslant \left(1 - \frac{1}{r}\right)(n-1)$, as can be seen by looking at a vertex in the largest class and noting it misses at most $\frac{1}{r}(n-1)$ vertices as neighbours. Equality holds only if there is precisely one largest class, in which case the average degree is greater than the minimum anyway, so in fact we always have $t_r(n) > \left(1 - \frac{1}{r}\right)\binom{n}{2}$.

For general F there might be several extremal graphs and the extremal function might be hard to evaluate exactly.

Denote by $K_r(t)$ the complete r-partite graph with t vertices per class.

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

Lemma 1.1. Let $r \geqslant 0$ be an integer and $\varepsilon > 0$. Then there exist $d = d'(r, \varepsilon)$ and $n_1 = n_1(r, \varepsilon)$ such that, if $|G| = n \geqslant n_1$ and, if $r \geqslant 1$,

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

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

Proof. If $r = 0$ or $\varepsilon \geqslant 1/r$ then the assertion is trivial. We proceed by induction on r.

By the induction hypothesis, we may assume that G has a subgraph $K = K_r(T)$, where $T = \lceil 2t/\varepsilon r \rceil$. (This requires only that $d'(r, \varepsilon) < \frac{\varepsilon r}{3} d'\left(r-1, \frac{1}{r(r-1)}\right)$.)

Now, each vertex of K sends at least $\left(1 - \frac{1}{r} + \varepsilon\right)n - |K|$ edges to $G - K$. Let U be the set of vertices of $G - K$ having at least $\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|$ neighbours in K.

Writing $e(G - K, K)$ for the number of edges between $G - K$ and K, we have

$$|K| \left[\left(1 - \frac{1}{r} + \varepsilon\right) n - |K|\right] \ \leqslant\ e(G-K, K) \ \leqslant\ |U||K| + \big(n - |U|\big)\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|K|$$

or

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

which, if $n_1(r,\varepsilon)$ is large enough, implies $\dfrac{|U|}{r} \geqslant \dfrac{\varepsilon n}{3}$.

Se we may assume that $|U| \geqslant \dfrac{\varepsilon r n}{3}$.

Now, each vertex in U is joined to at least

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

vertices in each class of K, and so is joined to some $K_r(t)$ in K. But there are only $\binom{T}{t}^r$ many $K_r(t)$ in K, and, recalling that $\binom{n}{k} \leqslant (\frac{en}{k})^k$, we have

$$\binom{T}{t}^r \leqslant \left(\frac{eT}{t}\right)^{tr} \leqslant \left(\frac{3e}{\varepsilon r}\right)^{rd\log n} \leqslant \frac{\varepsilon r n}{3t} \leqslant \frac{|U|}{t}$$

if $d'(r,\varepsilon)$ is small and $n_1(r,\varepsilon)$ is large.

Hence there exists $W \subset U$, with $|W| \geqslant t$, joined to the same $K_r(t)$ in K.

Hence $K_{r+1}(t) \subset G$. □

Lemma 1.2. Let $c, \varepsilon > 0$. Then there exists $n_2 = n_2(c,\varepsilon)$ with the following property.

Suppose that $|G| = n \geqslant n_2$ and $e(G) \geqslant (c+\varepsilon)\binom{n}{2}$. Then G has a subgraph H such that $\delta(H) \geqslant c|H|$ and $|H| \geqslant \varepsilon^{1/2} n$.

Proof. If not, there is a sequence $G = G_n \supset G_{n-1} \supset G_{n-2} \supset \cdots \supset G_s$, where $s = \lfloor \varepsilon^{1/2} n \rfloor$, such that $|G_j| = j$ and the only vertex in G_j not in G_{j-1} has degree less than cj. Then

$$e(G_s) \geqslant (c+\varepsilon)\binom{n}{2} - \sum_{j=s+1}^{n} cj = (c+\varepsilon)\binom{n}{2} - c\left[\binom{n+1}{2} - \binom{s+1}{2}\right] > \frac{\varepsilon n^2}{2} > \binom{s}{2}$$

provided n_2 is large enough. Contradiction. □

Lecture 2 **Theorem 1.3 (Erdős-Stone, 1946).** Let $r \geqslant 0$ be an integer and $\varepsilon > 0$. Then there exist $d = d(r,\varepsilon)$ and $n_0 = n_0(r,\varepsilon)$ such that, if $|G| = n \geqslant n_0$ and if for $r \geqslant 1$ we have $e(G) \geqslant \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$, then $G \supset K_{r+1}(t)$ where $t = \lfloor d \log n \rfloor$.

Proof. Provided $n_0 \geqslant n_2\left(1 - \frac{1}{r} + \frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right)$, we may apply Lemma 1.2 to G to obtain a subgraph H with $\delta(H) \geqslant \left(1 - \frac{1}{r} + \frac{\varepsilon}{2}\right)|H|$ and $|H| \geqslant \varepsilon^{1/2} n$.

Provided $n_0 \geqslant \varepsilon^{-1/2} n_1(r, \frac{\varepsilon}{2})$, we may apply Lemma 1.1 to H to obtain $K_{r+1}(t)$ with $t \geqslant \left\lfloor d_1(r, \frac{\varepsilon}{2}) \log \varepsilon^{1/2} n \right\rfloor$.

Provided $n_0 > \frac{1}{\varepsilon}$, if we take $d(r,\varepsilon) = \frac{1}{2} d_1(r, \frac{\varepsilon}{2})$, we are done. □

As observed by Erdős and Simonovits in the mid-1960s, we can determine $ex(n,F)$ asymptotically for *every* F.

Theorem 1.4. Let F be a fixed graph with chromatic number $r = \chi(F)$. Then

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

Proof. Since $\chi(T_{r-1}(n)) = r-1$, we have $F \not\subset T_{r-1}(n)$, hence

$$ex(n,F) \;\geqslant\; e(T_{r-1}(n)) \;\geqslant\; \left(1 - \frac{1}{r-1}\right)\binom{n}{2}.$$

On the other hand, given $\varepsilon > 0$, if $|G| = n$ and $e(G) \geqslant \left(1 - \frac{1}{r-1} + \varepsilon\right)\binom{n}{2}$, then $G \supset K_r(|F|) \supset F$ if n is large enough, by the Erdős-Stone theorem.

Thus for every $\varepsilon > 0$, we have $\limsup \dfrac{ex(n,F)}{\binom{n}{2}} \;\leqslant\; \left(1 - \dfrac{1}{r-1} + \varepsilon\right)$. □

Here's a pretty consequence of Erdős-Stone.

Define the **upper density** of an infinite graph to be the supremum of densities of large finite subgraphs:

$$ud(G) = \lim_{n\to\infty} \sup\left\{x : \exists F \subset G, |F| \geqslant n, e(F) \geqslant x\binom{|F|}{2}\right\}$$

Corollary 1.5. $ud(G) \in \{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\} \cup \{1\}$.

Can we strengthen Erdős-Stone to obtain larger t?

Theorem 1.6 Given $r \in \mathbb{N}$, there exists $\varepsilon_r > 0$ such that, if $\varepsilon < \varepsilon_r$, there exists $n(r,\varepsilon)$ such that for all $n \geqslant n(r,\varepsilon)$ there is a graph G of order n with $e(G) \geqslant \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$ and $G \not\supset K_{r+1}(t)$ where $t = \left\lfloor \frac{3\log n}{\log(1/\varepsilon)} \right\rfloor$.

Proof. Let W be a largest vertex class of $T_r(n)$ with $|W| = w = \lceil n/r \rceil$. Form G by adding $\varepsilon\binom{n}{2}$ edges within W so that $G[W] \not\supset K_2(t)$ and hence $G \not\supset K_{r+1}(t)$.

To see that this addition is possible, choose edges inside W independently with probability $p = 3\varepsilon r^2$. Take $\varepsilon_r = (3r^2)^{-6}$. So $p < 1$.

Let X be the number of edges chosen and Y the number of $K_2(t)$ formed by them. Then

$$\mathbb{E}(X - Y) = \mathbb{E}X - \mathbb{E}Y = p\binom{w}{2} - \frac{1}{2}\binom{w}{t}\binom{w-t}{t}p^{t^2}$$

Now,

$$w^{2t-2}p^{t^2-1} = \left[w^2 p^{t+1}\right]^{t-1} < \left[w^2 \varepsilon^{\frac{5}{6}(t+1)}\right]^{t+1} < \left[w^2 n^{-5/2}\right]^{t-1} < \frac{1}{2}.$$

Hence $\mathbb{E}(X-Y) > \dfrac{p}{2}\dbinom{w}{2}$. So there is a choice with $X - Y > \dfrac{p}{2}\dbinom{w}{2}$.

Remove an edge from each $K_2(t)$ to leave at least $X - Y > \dfrac{p}{2}\dbinom{w}{2} > \varepsilon\dbinom{n}{2}$ edges with no $K_2(t)$. □

This shows dependence on n is correct. Our argument gives $d(r,\varepsilon) \geqslant \dfrac{\varepsilon}{2^r(r-1)!}$.

Bollobás, Erdős, Simonovits (1976) proved that $t > c_r \dfrac{\log n}{\log(1/\varepsilon)}$.

Chvatal, Szemerédi (1978) proved $t > \dfrac{\log n}{\log(1/\varepsilon)}$.

Lecture 3

2. Stability

An extremal problem is **stable** if all nearly-optimal examples have the same structure.

Theorem 2.1. Let $t, r \geqslant 2$ be fixed and suppose that $G \not\supset K_{r+1}(t)$. If $e(G) = \big(1 - \frac{1}{r} + o(1)\big)\binom{n}{2}$ then

(a) there exists $T_r(n)$ on $V(G)$ with $|E(G) \triangle E(T_r(n))| = o(n^2)$
(b) G contains an r-partite subgraph of size $\big(1 - \frac{1}{r} + o(1)\big)\binom{n}{2}$
(c) G contains an r-partite subgraph of minimum degree $\big(1 - \frac{1}{r} + o(1)\big)n$

Global statement. Throughout, unless otherwise stated, $|G| = n$.

Proof. Note that (a) $\iff$ (b) and (c) $\implies$ (b). Note too that we may jettison $o(n)$ vertices should we wish to. In particular, we may assume that $\delta(G) \geqslant \big(1-\frac{1}{r}+o(1)\big)n$, so (b) $\implies$ (c): for otherwise, there exist εn vertices of degree at most $\big(1 - \frac{1}{r} + \eta\big)n$ for some small $\varepsilon < \frac{1}{4}\eta$, and their removal leaves a graph of order $(1-\varepsilon)n$ and size $\geqslant \big(1 - \frac{1}{r} + \frac{\varepsilon^2}{r}\big)\binom{(1-\varepsilon)n}{2}$, so by Erdős-Stone we would have a $K_{r+1}(t)$.

Now, $G \supset K = K_r(s)$ where $\log n \geqslant s = s(n) \to \infty$. Let C_i be the i^{th} class of K, and let X be the set of vertices joined to at least t in every class of K. THere are $\binom{s}{t}^r$ many $K_r(t)$ in J, and since $K_{r+1}(t) \subset G$, we have $|X| \leqslant t\binom{s}{t}^r = o(n)$. Jettison X.

Let Y be the set of vertices joined to fewer than $(r-1)s - \frac{s}{2t} + t$ vertices of K. Since $\delta(G) \geqslant \big(1 - \frac{1}{r} + o(1)\big)n$, we have $e(K, G-K) \geqslant s(r - 1 + o(1))n$. Because X has been jettisoned, we have

$$(n - |Y|)\big((r-1)s + t\big) + |Y|\left((r-1)s - \frac{s}{2t} + t\right) \geqslant s\big(r - 1 + o(1)\big)n$$

So $|Y| = o(n)$. Jettison Y.

The remaining vertices are partitioned by V_i, $1 \leqslant i \leqslant r$, where V_i is the set of vertices joined to fewer than t of C_i. Since $\delta(G) \geqslant \big(1 - \frac{1}{r} + o(1)\big)n$, it is enough to show that $e\big(G[V_i]\big) = o(n^2)$ for each i, for then $V_1, \ldots, V_r$ are the parts of the r-partite graph we seek.

Suppose instead that $e\big(G[V_i]\big) \geqslant \varepsilon n^2$, say. Then $G[V_1] \supset K_2(t)$ by Erdős-Stone with $r = 1$. Each vertex of $K_s(t)$ is joined to at least $s - \frac{s}{2t} + 1$ vertices in each class C_i. Hence C_i contains a set of $s - 2t\big(\frac{s}{2t} - 1\big) > t$ vertices joined to all of $K_2(t)$. But then $G \supset K_{r+1}(t)$, a contradiction. □

Corollary 2.2. Let $\chi(F) = r + 1$ and let G be extremal for F – i.e., $F \not\subset G$, $e(G) = ex(n, F)$. Then $\delta(G) = \big(1 - \frac{1}{r} + o(1)\big)n$.

Proof. If not, by Theorem 1.4 we have $\delta(G) \leqslant \big(1 - \frac{1}{r} - \varepsilon\big)n$. By Theorem 2.1(c) with $t = |F|$, G has $|F|$ vertices $x_1, \ldots, x_{|F|}$ joined to $m = \big(1 - \frac{1}{r} + o(1)\big)n$ common neighbours $y_1, \ldots, y_m$.

Form G^* from $G - v$ (where v is a vertex of lowest degree) by adding a new vertex u joined to $y_1, \ldots, y_m$. Then $e(G^*) > e(G)$, so $G^* \supset F$. This copy of F in G^* contains u but not x_i, say. But then $(F - u) \cup x_i$ contains F in G. Contradiction. □

Stability can sometimes be used as a bootstrapping device to obtain exact results. For example, $ex(n, C_5) = t_2(n)$ for $n \geqslant 6$. In fact, $ex(n, C_{2k+1}) = t_2(n)$ if n is large. This is a special case of the following theorem.

Theorem 2.3 (Simonovits). Let F be $(r+1)$-**edge-critical**, i.e. $\chi(F) = r+1$ but $\chi(F-e) = r$ for all edges $e \in E(F)$. Then $ex(n, F) = t_r(n)$ for large n, and $T_r(n)$ is the unique extremal graph.

Proof. Let G be an extremal graph of order n. Select, by Theorem 2.1(c) using $t = |F|$, an r-partite subgraph H of minimum degree $\big(1 - \frac{1}{r} + o(1)\big)n$. Necessarily, each part of H is of order $\big(\frac{1}{r} + o(1)\big)n$. Assign the $o(n)$ vertices of $G - H$ to the parts of H with the fewest neighbours.

Lecture 4

Suppose a vertex x is joined to εn vertices in its own class. Then it has εn neighbours in each part of H. These r sets of εn vertices span $\big(1 - \frac{1}{r} + o(1)\big)\binom{r\varepsilon n}{2}$ edges because $\delta(H) = \big(1 - \frac{1}{r} + o(1)\big)n$.

By Erdős-Stone the neighbours of x span $K_r(|F|)$ which contains $F - v$ for any $v \in V(F)$. Hence $G \supset F$, a contradiction.

So each vertex has only $o(n)$ neighbours in its own part. By Corollary 2.2, we know $\delta(G) = \big(1 - \frac{1}{r} + o(1)\big)n$. So each vertex of G is joined to all but $o(n)$ vertices in each other class. Suppose xy is an edge inside some class. Pick a set Z of $|F|$ vertices in this class, including $\{x, y\}$.

Then all but $o(n)$ vertices of the other classes are common neighbours of Z. Hence these common neighbours span $K_{r-1}(|F|)$, either by Erdős-Stone or directly. But Z together with $K_{r-1}(|F|)$ contains F, a contradiction.

Hence G is r-partite, and since $T_r(n)$ is the unique r-partite graph of maximum size, we have $G = T_r(n)$. □

Here is another example.

Theorem 2.4. Let r, s be fixed. For large n, the unique extremal graph for sK_{r+1} (i.e., s disjoint copies of K_{r+1}) is $K_{s-1} + T_r(n - s + 1)$ (where '+' means 'join all to all').

Proof. The proof is similar to that of Theorem 2.3. Proceed by induction on s. The case $s = 1$ is Turán's theorem.

As before, choose H and assign $G - H$ to the classes of H. Once again, if some vertex x has εn neighbours in its own class, then the neighbours of x span some $K_r(s(r + 1))$. But $G - x$ cannot contain $(s - 1)K_{r+1}$, hence $e(G - x) \leqslant e\big(K_{s-2} + T_r(n - s + 1)\big)$, so $e(G) \leqslant e\big(K_{s-1} + T_r(n - s + 1)\big)$.

So equality holds in both cases, so $G - x = K_{s-2} + T_r(n - s + 1)$ by the induction hypothesis, so $G = K_{s-1} + T_r(s - n + 1)$.

Hence we may suppose that each vertex is joined to $o(n)$ in its own class. Suppose some class contains s independent edges. Let Z be the $2s$ endvertices of these edges. As in the previous proof, the common neighbours of Z span $K_{r-1}(s(r-1))$, in which case $G \supset sK_{r+1}$.

Thus no class contains s independent edges, so the j^{th} class contains a set A_j of $2(s-1)$ vertices such that every edge in the j^{th} class meets A_j. But each vertex in A_j has $o(n)$ neighbours in the j^{th} class.

Hence $e(G) \leqslant 2r(s-1)o(n) + t_r(n) < e\big(K_{s-1} + T_r(n-s+1)\big)$. □

3. Supersaturation

Supersaturation is the study of how many copies of F must exist in G if $e(G) > ex(n, F)$. The basic theorem holds in a general context.

Recall that an **ℓ-uniform hypergraph** is a pair $G = (V, E)$ where

$$E \subset V^{(\ell)} = \big\{Y \subset V : |Y| = \ell\big\}.$$

For a class $\mathcal{F}$ of ℓ-uniform hypergraphs, define

$$ex(n, \mathcal{F}) = \max\big\{e(G) : |G| = n, G \text{ is } \ell\text{-uniform}, G \text{ contains no } F \in \mathcal{F}\big\}$$

and

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

Exercise. This limit exists.

Theorem 3.1. Let H be an ℓ-uniform hypergraph. Then for all $\varepsilon > 0$ there exists $\delta = \delta(H, \varepsilon) > 0$ such that every ℓ-uniform hypergraph G with $|G| = n$ and $e(G) \geqslant \big(\pi(H) + \varepsilon\big)\binom{n}{\ell}$ contains $\big\lfloor \delta n^{|H|} \big\rfloor$ copies of H.

Lecture 5 **Proof.** For each m-set $M \in V^{(m)}$, let $G[M]$ be the hypergraph induced by G on M. If there are $\eta\binom{n}{m}$ sets M with $e(G[M]) \geqslant \big(\pi(H) + \frac{\varepsilon}{2}\big)\binom{m}{\ell}$ then

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

So, assuming $n \geqslant m \geqslant \ell$, we have

$$\pi(H) + \varepsilon \leqslant \eta + (1-\eta)\big(\pi(H) + \tfrac{\varepsilon}{2}\big) \quad \text{or} \quad \eta \geqslant \frac{\frac{\varepsilon}{2}}{1 - \pi(H) - \frac{\varepsilon}{2}} > 0\ .$$

Pick m so that $ex(m, H) < \big(\pi(H) + \frac{\varepsilon}{2}\big)\binom{m}{\ell}$. Then $H \subset G[M]$ for each of $\eta\binom{n}{m}$ subsets M.

So G contains $\geqslant \eta\binom{n}{m}\binom{n-|H|}{m-|H|}^{-1}$ distinct copies of H, i.e., $\geqslant \eta\binom{n}{|H|}\binom{m}{|H|}^{-1}$ copies.

Pick $n_0 \geqslant m$ so that $\displaystyle\binom{n}{|H|} \geqslant \frac{1}{2}\frac{n^{|H|}}{|H|!}$.

If $n \geqslant n_0$ then pick $\displaystyle\delta \leqslant \frac{1}{2}\frac{1}{|H|!}\binom{m}{|H|}^{-1}$ small enough that $\big\lfloor \delta n^{|H|} \big\rfloor$ works for $n \leqslant n_0$. □

Back to graphs.

Let $k_p(G)$ denote the number of copies of K_p in G. Ramsey's theorem says $k_p(G) + k_p(\overline{G}) > 0$, if G is large (where $\overline{G}$ is the complement of G).

Theorem 3.2 (Lorden, 1962). Let G have degree sequence $d_1, \ldots, d_n$. Then

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

Proof. The number of paths of length 2 in G and $\overline{G}$ is $\sum \binom{d_i}{2} + \sum \binom{n-1-d_i}{2}$.

A complete or empty vertex triple contains 3 such paths, and every other triple contains exactly 1. So

$$\begin{aligned}\binom{n}{3} + 2\big(k_3(G) + k_3(\overline{G})\big) &= \sum \binom{d_i}{2} + \sum \binom{n-1-d_i}{2} \\ &= 2\sum \binom{d_i}{2} - 2(n-2)e(G) + 3\binom{n}{3} .\end{aligned}$$

□

Corollary 3.3 (Goodman, 1959). $k_3(G) + k_3(\overline{G}) \geqslant \frac{1}{24}n(n-1)(n-5)$.

Proof. Let $m = e(G)$. Then $k_3(G) + k_3(\overline{G}) \geqslant \binom{n}{3} - (n-2)m + n\binom{2m/n}{2}$. □

These results show that $k_3(G) + k_3(\overline{G})$ depends only on the degree sequence, and the minimum density of monochromatic triangles is at least 1/4, as attained by a random colouring. No such result holds for K_4 – it is known that the density can be $< 1/33$.

Corollary 3.4. $k_3(G) \geqslant \dfrac{m}{3n}(4m - n^2)$ where $n = |G|$, $m = e(G)$.

Proof. Theorem 3.2 implies that $k_3(G)+k_3(\overline{G}) = \binom{n}{3} - (n-2)\overline{m} + \sum \binom{\overline{d_i}}{2}$, where $\overline{m} = e(\overline{G})$ and $(\overline{d_i})$ is the degree sequence of $\overline{G}$.

But $3k_3(\overline{G}) \leqslant \sum \binom{\overline{d_i}}{2}$. So using $\overline{m} = \binom{n}{2} - m$, we have

$$k_3(G) \geqslant \binom{n}{3} - (n-2)\overline{m} + \frac{2}{3}n\binom{2\overline{m}/n}{2} = \frac{m}{3n}(4m - n^2) .$$

□

This bound is tight only for regular graphs containing no triples with just one edge. This means that $\overline{G}$ is a union of complete graphs, so G is complete multipartite. Such graphs are rare: they are $T_r(n)$ where $r \mid n$.

Given F, let $i_F(G)$ be the number of **induced** subgraphs of G isomorphic to F.

So, e.g., $i_{\overline{K_p}}(G) = k_p(\overline{G})$.

Theorem 3.5. Let $f(G) = \sum_F \alpha_F i_F(G)$, the sum being over a finite collection of F, each of which is complete multipartite, with $\alpha_F \in \mathbb{R}$, and $\alpha_F \geqslant 0$ unless F is complete.

Then, amongst graphs G of given order, $f(G)$ is maximized on a complete multipartite graph. Moreover, if $\alpha_{\overline{K_3}} > 0$, there are no other maxima.

Proof. We may suppose that $\alpha_{\overline{K_3}} > 0$, the case $\alpha_{\overline{K_3}} = 0$ following by a limiting argument.

Choose a graph G of order n that maximizes $f(G)$. Suppose G is not complete multipartite. Then G contains two non-adjacent vertices x, y whose neighbourhoods X, Y differ.

The number $i_F(\alpha)$ contains contributions from four kinds of F: those containing, respectively, x but not y, y but not x, both x and y, neither x nor y. The first contribution depends only on X, and the second only on Y. Moreover the third depends only on $X \cap Y$ and $V - (X \cup Y)$ where $V = V(G)$, since F is complete multipartite.

This fourfold partition of $i_F(G)$ means we can write

$$f(G) = g(X) + g(Y) + h(X \cap Y, V - (X \cup Y)) + C,$$

where C is independent of X and Y.

Lecture 6

Note that $h(A, B) \leqslant h(A', B')$ if $A \subset A'$, $B \subset B'$, because F makes no contribution to h if F is complete, and otherwise $\alpha_F \geqslant 0$. Moreover, if also $B \neq B'$, then $h(A, B) < h(A', B')$, but $\alpha_{\overline{K_3}} > 0$ and $i_{\overline{K_3}}(G)$ contributes exactly $\alpha_{\overline{K_3}}|B|$ to $h(A, B)$.

We may suppose that $g(X) \geqslant g(Y)$ and, if $g(X) \neq g(Y)$, that $|X| \leqslant |Y|$ and so $X \neq X \cup Y$. Therefore

$$g(X) + h(X, V - X) > g(Y) + h(X \cap Y, V - (X \cup Y))$$

Form H from G by removing all edges y to Y and inserting all edges y to X. Then

$$f(H) = 2g(X) + h(X, V - X) + C > f(G).$$

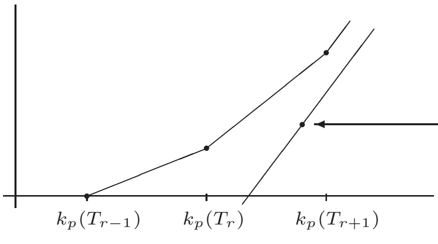
Contradiction. □

Note that this transformation does not increase the clique size or the chromatic number.

Let $1 \leqslant p \leqslant r$. For $0 \leqslant x \leqslant \binom{n}{p}$, let $\psi(x)$ be the maximal convex function defined by $\psi(0) = 0$, $\psi\big(k_p(T_q(n))\big) = k_r(T_q(n))$. (Where $q = r-1, r, r+1, \ldots$)

Theorem 3.6 (Bollobás, 1976). Let G be a graph of order n. Then $k_r(G) \geqslant \psi(k_p(G))$

Proof. Let $f(G) = k_p(G) - ck_r(G)$ for some real $c > 0$. Since ψ is convex, it suffices to prove that f is maximized on a Turán graph. So let G be a graph on which f is maximized.



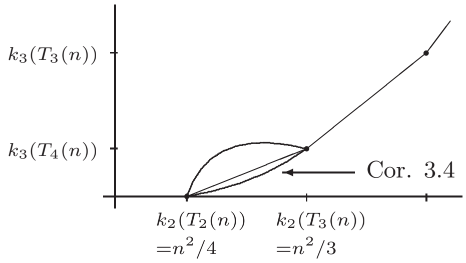
Suppose there exists G below ψ. Consider the straight line of gradient $c > 0$. The intercept on the axis is $k_p(G) - ck_r(G)$.

By Theorem 3.5, we know that f is maximized on a complete-partite graph, with q classes of sizes $0 < a_1 \leqslant a_2 \leqslant \cdots \leqslant a_q$. We may assume that $q \geqslant r$, else $T_{r-1}(n)$ maximizes f.

Now, $f(G) = a_1 a_q A - c a_1 a_q B + C$, where A, B, C are non-zero rationals depending on $a_2, a_3, \ldots, a_{q-1}$ and $a_1 + a_q$.

We may assume without loss of generality that c is irrational, so $A - cB \neq 0$. If $A - cB < 0$, then replace a_1 by 0 and a_q by $a_1 + a_q$ to increase f. If $A - cB > 0$ and $a_1 \leqslant a_q - 2$, then replace a_1 by $a_1 + 1$ and a_q by $a_q - 1$ to increase f. Thus $a_1 \geqslant a_q - 1$ and so G is a Turán graph. □

The exact value of $\min\{k_3(G) : |G| = n, k_2(G) = m\}$ is unknown. It is conjectured to be given by r-partite graphs where r is minimal.



I think. I foolishly tried to draw this during the lecture, and so it might not be right…

The continuous envelope here for $n^2/4 \leqslant m \leqslant n^2/3$ was posed as a lower bound by Fisher (1989). The whole range $m \leqslant \binom{n}{2}$ was proved (again in the limit) by Razborov, who introduced the method of flag algebras (2002), a lowbrow view of this being a massive generalisation of Cauchy-Schwarz, using CSP to find optimal quadratic forms.

Nikiforov (2008) did likewise for $n = 4$. Reiher (2012) did likewise for all r, for $p = 2$.

Open problem. On inducibility: what is $\max\{i_{\text{length 3 paths}}(G)\}$?

Lecture 7

4. Szemerédi's Regularity Lemma

A graph having the (large scale) property that its (induced) subgraphs all have roughly the same density, the graph itself can be regarded as "pseudo-random" in a sense that can be made precise. Consider the following bipartite version.

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

The **density** $d(U, W)$ is $\dfrac{e(U, W)}{|U||W|}$.

Definition 4.1. Let $0 < \varepsilon < 1$. The pair (U, W) is said to be ε**-uniform** (or ε-regular) if for $U' \subset U$, $W' \subset W$,

$$\big|d(U', W') - d(U, W)\big| < \varepsilon \quad \text{whenever} \quad |U'| > \varepsilon|U|, \ |W'| > \varepsilon|W|$$

(Clearly a lower bound on $|U'|$ is necessary – e.g., $|U'| = 1$ is hopeless.)

An ε-uniform pair is roughly regular.

Lemma 4.2. Let (U, W) be ε-uniform and $d(U, W) = d$. Then

$$\big|\{u \in U : |\Gamma(u) \cap W| > (d - \varepsilon)|W|\}\big| > (1 - \varepsilon)|U|$$

and

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

(Recall that $\Gamma(u)$ is the neighbourhood of u.)

Proof. Let $X = \{u \in U : |\Gamma(u) \cap W| \leqslant (d-\varepsilon)|W|\}$. Then $e(X,W) \leqslant (d-\varepsilon)|X||W|$, so $d(X,W) \leqslant d-\varepsilon$. By the definition of density, $|X| \leqslant \varepsilon|U|$, which proves the first inequality.

The proof of the other is similar. Alternatively, observe that in the complementary graph, (U,W) is ε-uniform with density $1-d$, and apply the first inequality. $\square$

Repeated applications give lots of information about graph structure. For example, given a graph H with vertices $v_1,\ldots,v_k$ and subsets $V_1,\ldots,V_k$ of $V(G)$ such that (V_i,V_j) is ε-uniform, we may find a copy of H in G. More is true.

Lemma 4.3. Let H be a graph with $\Delta(H) = d$, and suppose that H has an r-colouring in which no colour is used more than s times. Let G be a graph containing disjoint vertex subsets $V_1,\ldots,V_r$ with $|V_i| = u$ such that for all i,j, (V_i,V_j) is ε-uniform and $d(V_i,V_j) \geqslant \lambda$.

Suppose that $(d+1)\varepsilon \leqslant \lambda^d$ and $s \leqslant \lceil \varepsilon u \rceil$. Then G contains a copy of H.

Proof. Let $c : V(H) \to \{1,\ldots,r\}$ be an r-colouring of H in which no colour is used more than s times. Let $V(H) = \{v_1,\ldots,v_k\}$. We will select vertices $x_1,\ldots,x_k$ in G so that $x_ix_j \in E(G)$ if $v_iv_j \in E(H)$.

We claim that for $0 \leqslant \ell \leqslant k$, vertices $x_1,\ldots,x_\ell$ can be chosen so that $x_i \in V_{c(v_i)}$ and, for $\ell < j \leqslant k$ there is a set $x_j^\ell \subset V_{c(v_j)}$ of **candidates** for x_j at stage ℓ, meaning that $x_iy_j \in E(G)$ for every $y_j \in X_j^\ell$ and every $x_i \in N(j,\ell) = \{x_i : 1 \leqslant i \leqslant \ell \text{ and } v_iv_j \in E(H)\}$.

Moreover, $|X_j^\ell| \geqslant (\lambda-\varepsilon)^{|N(j,\ell)|}|V_{c(v_j)}|$.

The claim clearly holds for $\ell = 0$ – just take $X_j^0 = V_{c(v_j)}$. Proceed by induction on ℓ. In general, for each $t \in T = \{j > \ell+1 : v_{\ell+1}v_j \in E(H)\}$, let

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

Let $m = N(\ell+1,\ell)$, and note that $m + |T| \leqslant d$.

Since $d(Y_t, X_t^\ell) \leqslant \lambda - \varepsilon \leqslant d(V_{c(v_{\ell+1})}, V_{c(v_t)}) - \varepsilon$ and

$$|X_t^\ell| \geqslant (\lambda-\varepsilon)^{d-1}|V_{c(v_t)} \;>\; (\lambda^{d-1} - (d-1)\varepsilon)|V_{c(v_t)}| \;>\; \varepsilon|V_{c(v_t)}|$$

it follows that $|Y_t| \leqslant \varepsilon|V_{c(v_{\ell+1})}|$.

Therefore,

$$\begin{aligned}
\left|X_{\ell+1}^\ell - \bigcup_{t\in T} Y_t\right| &\geqslant (\lambda-\varepsilon)^m|V_{c(v_{\ell+1})} - (d-m)|V_{c(v_{\ell+1})}| \\
&\geqslant \big(\lambda^m - m\varepsilon - (d-m)\varepsilon\big)|V_{c(v_{\ell+1})}| \\
&\geqslant \lceil \varepsilon u \rceil \geqslant s
\end{aligned}$$

At most $s-1$ vertices of $X_{\ell+1}^\ell - \bigcup_{t\in T} Y_t$ have been used for $x_1,\ldots,x_\ell$, so we may select $x_{\ell+1}$ from this set. Take $X_t^{\ell+1} - X_t^\ell \cap \Gamma(x_{\ell+1})$ for $t \in T$, and $X_t^{\ell+1} = X_t^\ell$ for $t \notin T$. $\square$

Lecture 8

Corollary 4.4. Let H be a graph with vertex set $v_1,\ldots,v_k$. Let $0 < \lambda, \eta < 1$ satisfy $k\eta \leqslant \lambda^{k-1}$. Let G be a graph with vertex set $V_1 \cup \cdots \cup V_k$, where the V_i are disjoint sets of size $y \geqslant 1$.

Suppose that each pair (V_i,V_j) is η-uniform, that $d(V_i,V_j) \geqslant \lambda$ if $v_iv_j \in E(H)$, and that $d(V_i,V_j) \leqslant 1-\lambda$ if $v_iv_j \notin E(H)$.

Then there exist vertices $x_1, \ldots, x_k$ ($x_i \in V_i$) such that the map $v_i \mapsto x_i$ gives an isomorphism between H and $G[\{x_1, \ldots, x_k\}]$.

Proof. Note that, by replacing the set of $V_i - V_j$ edges by the complementary set if $v_iv_j \notin E(H)$, we may assume that H is complete and $d(V_i, V_j) \geqslant \lambda$ for all i, j.

The result is then immediate from Lemma 4.3 on taking $\varepsilon = \eta$, $r = k$, $d = k - 1$, and $s = 1$. □

It is a remarkable fact that *every* graph can be partitioned into a *bounded* number of pieces, almost all pairs of which are ε-uniform. This result is due to Szemerédi and is used in his proof that sets of integers with positive density contain arbitrarily long arithmetic progressions.

An **equipartition** of $V(G)$ into k parts is a partition $V_1, \ldots, V_k$ such that $\lfloor n/k \rfloor \leqslant |V_i| \leqslant \lceil n/k \rceil$ for all $1 \leqslant i \leqslant k$, where $n = |V(G)|$. The partition is **ε-uniform** if (V_i, V_j) is ε-uniform for all but at most $\varepsilon\binom{k}{2}$ pairs, for $1 \leqslant i < j \leqslant k$.

Theorem 4.5 (Szemerédi's Regularity Lemma). Let $0 < \varepsilon < 1$ and let ℓ be a natural number. Then there exists $L = L(\ell, \varepsilon)$ such that every graph has an ε-uniform equipartition into m parts, for some $\ell \leqslant m \leqslant L$.

Before proving the lemma, we establish two useful facts.

Lemma 4.6. Let $U' \subset U$ and $W' \subset W$ satisfy $|U'| \geqslant (1 - \delta)|U|$ and $|W'| \geqslant (1 - \delta)|W|$.

Then $\left|d(U', W') - d(U, W)\right| \leqslant 2\delta$.

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

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

So $d' - d \leqslant d'\big(1 - (1 - \delta)^2\big) \leqslant 2\delta d' \leqslant 2\delta$.

By considering the complementary graph, $(1 - d') - (1 - d) \leqslant 2\delta$, i.e., $d - d' \leqslant 2\delta$. □

Lemma 4.7. Let $x_1, \ldots, x_n$ be real numbers such that $X = \frac{1}{n}\sum_{i=1}^{n} x_i$. Let $x = \frac{1}{m}\sum_{i=1}^{m} x_i$. Then

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

Proof. We have

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} x_i^2 &= \frac{1}{n}\sum_{i=1}^{m} x_i^2 + \frac{1}{n}\sum_{i=m+1}^{n} x_i^2 \\ &\geqslant \frac{m}{n}x^2 + \frac{n - m}{n}\left(\frac{nX - mx}{n - m}\right)^2 \\ &= X^2 + \frac{m}{n - m}(x - X)^2 \end{aligned}$$

by two applications of Cauchy-Schwarz. □

Proof of Theorem 4.5. Define the **index** ind(P) of an equipartition P with k parts $V_1, \ldots, V_k$ to be

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

We shall show that, if $n \geqslant k16^k$ and $4^k\varepsilon^5 > 100$, and P is not ε-uniform, then there exists an equipartition Q into $k4^k$ parts with $\mathrm{ind}(Q) \geqslant \mathrm{ind}(P) + \varepsilon^5/8$.

This will be enough to prove the theorem. For, choose $t \geqslant \ell$ with $4^t\varepsilon^5 > 100$. Define $f(0) = t$ and $f(j+1) = f(j)4^{f(j)}$. Let $N = f(\lceil 4\varepsilon^{-5}\rceil)$. Let $L = N16^N$. Then if $n \leqslant L$, take an equipartition into n single vertices. Otherwise, begin with an equipartition P into t parts.

So long as the current partition into k parts is not ε-uniform, replace it with one into $k4^k$ parts and larger index. Since $\mathrm{ind}(P) \leqslant \frac{1}{2}$, replacement can occur at most $\lceil 4\varepsilon^{-5}\rceil$ times, so an ε-uniform partition is found with $\leqslant L$ parts.

For each (V_i, V_j) that is not ε-uniform, select witness sets $X_{ij} \subset V_i$ and $X_{ji} \subset V_j$ such that $|X_{ij}| \geqslant \varepsilon|X_i|$ and $|X_{ji}| \geqslant \varepsilon|X_j|$, and $\big|d(X_{ij}, X_{ji}) - d(V_i, V_j)\big| \geqslant \varepsilon$. For each i, the sets X_{ij} partition V_i into at most 2^{k-1} **atoms**.

Let $m = \lfloor n/k4^k\rfloor$ and let $n = k4^k m + ak + b$ where $0 \leqslant a < 4^k$ and $0 \leqslant b \leqslant k$. Then $\lfloor n/k\rfloor = 4^k m + a$, so the parts of P have size $4^k m + a$ or $4^k m + a + 1$, with B parts have the larger.

Lecture 9 Partition each part of P into 4^k sets of size m or $m+1$. Any such partition is a partition of G into $k4^k$ sets of sizes m or $m+1$, i.e. an equipartition.

Choose such a partition Q whose parts are as much as possible inside atoms: that is, every atom is a union of part of Q together with at most m extra vertices. All that remains is to check that $\mathrm{ind}(Q) \geqslant \mathrm{ind}(P) + \varepsilon^5/8$.

Let the setes of Q within V_i be $V_i(s)$, for $1 \leqslant s \leqslant q = 4^k$. That is, $V_i = \bigcup_{s=1}^q V_i(s)$. Note that $\sum_{1\leqslant s,t\leqslant q} e\big(V_i(s), V_j(t)\big) = e(V_i, V_j)$ and $|V_i| \geqslant q|V_i(s)|\frac{m}{m+1}$ for each s, and so

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

Since $n \geqslant k16^k$, we have

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

Therefore

$$\frac{1}{q^2}\sum_{1\leqslant s,t\leqslant q} d^2\big(V_i(s), V_j(t)\big) \overset{\text{C.S.}}{\geqslant} \left(\frac{1}{q^2}\sum_{1\leqslant s,t\leqslant q} d\big(V_i(s), V_j(t)\big)\right)^2 \geqslant d^2(V_i, V_j) - \frac{\varepsilon^5}{25}$$

The main point is that we can improve this bound if (V_i, V_j) is not uniform. Let X_{ij}^* be the largest subset of X_{ij} that is a union of parts of Q. Because we choose parts of Q as much as possible within atoms, and because $|V_i| \leqslant \lfloor nl\rfloor 4^k m + a \geqslant 4^k m$, we have

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

so by Lemma 4.6, we have

$$\left|d(X^*_{ij}, X^*_{ji}) - d(X_{ij}, X_{ji})\right| \leqslant \frac{\varepsilon}{5}\ .$$

We may assume that $X^*_{ij} = \bigcup_{1\leqslant s\leqslant r_i} V_i(s)$ for some number r_i.

By a similar argument to that above,

$$\left|\frac{1}{r_i r_j}\sum_{1\leqslant s\leqslant r_i}\sum_{1\leqslant t\leqslant r_j} d\big(V_i(s), V_j(t)\big) - d(X^*_{ij}, X^*_{ji})\right| \leqslant \frac{\varepsilon^5}{50}$$

Recalling that $|d(X_i, X_j) - d(V_i, V_j)| \geqslant \varepsilon$, we see that

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

Applying Lemma 4.7 in the above, use of Cauchy-Schwarz with $n = q^2$, $m = r_i r_j$ gives

$$\frac{1}{q^2}\sum_{1\leqslant s,t\leqslant q} d^2\big(V_i(s), V_j(t)\big) \ \geqslant\ d^2(V_i, V_j) - \frac{\varepsilon^5}{25} + \frac{r_i r_j}{q^2}\frac{9\varepsilon^2}{16} \ \geqslant\ d^2(V_i, V_j) - \frac{\varepsilon^5}{25} + \frac{\varepsilon^4}{3}$$

because

$$\frac{r_i}{q} \geqslant \left(1 - \frac{1}{m}\right) r_i \left(\frac{m+1}{m}\right)\frac{1}{q} \geqslant \left(1 - \frac{1}{m}\right)\frac{|X^*_{ij}|}{|V_i|} \geqslant \left(1 - \frac{1}{m}\right)\left(1 - \frac{\varepsilon}{10}\right)\frac{|X_{ij}|}{|V_i|} \geqslant \frac{4\varepsilon}{5}\ .$$

Therefore,

$$\begin{aligned}
\mathrm{ind}(Q) &= \frac{1}{k^2 q^2}\sum_{1\leqslant i,j\leqslant k}\sum_{1\leqslant s,t\leqslant q} d^2\big(V_i(s), V_j(t)\big) \\
&\geqslant \frac{1}{k^2}\sum_{1\leqslant i,j\leqslant k}\left(d^2(V_i, V_j) - \frac{\varepsilon^5}{25}\right) + \frac{1}{k^2}\varepsilon\binom{k}{2}\frac{\varepsilon^4}{3} \\
&\geqslant \mathrm{ind}(P) + \frac{\varepsilon^5}{8}
\end{aligned}$$

completing the proof. □

The proof gives something like $L = 2^{2^{2^{2}}}$, with ε^{-5} many 2s in the tower.

Gowers (1997) showed a tower of height $\varepsilon^{-1/16}$ is necessary. Indeed, define an equipartition $V_1, \ldots, V_k$ to be $(\varepsilon, \delta, \eta)$-**uniform** if for all but $\eta\binom{k}{2}$ of the pairs (V_i, V_j), it holds that $\big|d(V_i', V_j') - d(V_i, V_j)\big| \leqslant \varepsilon$ whenever $|V_i'| \geqslant \delta|V_i|$.

Gowers proved that for small δ there is a graph such that every $(1-\delta^{1/16}, \delta, 1-20\delta^{1/16})$-uniform partition has at least $2^{2^{2^{2}}}$ (with $\delta^{-1/16}$ many $2s$) parts.

Lecture 10

5. A couple of applications

The simplest non-trivial case of Ramsey's theorem asserts that there is a smallest integer $R(k)$ such that if the edges of $K_{R(k)}$ are coloured red or blue then we get a monochromatic K_k. It is known that $2^{k/2} \leqslant R(k) \leqslant 4^k$. The existence of $R(k)$ implies that for any graph G there is a minimal integer $r(G)$ such that any red/blue colouring of of $K_{r(G)}$ contains a monochromatic G. Clearly $r(G) \leqslant R(|G|)$.

Theorem 5.1. Given an integer d, there exists a number $c(d)$ such that $r(G) \leqslant c(d)|G|$ if $\Delta(G) \leqslant d$.

Remark. The same is conjectured to be true for every graph with $e(H) \leqslant d|H|$ for all $H \subset G$.

It is known for a wider class than in the theorem, including planar graphs.

Proof. Let $t = R(d+1)$. Choose $\varepsilon \leqslant \min\left\{\frac{1}{t}, \frac{1}{2^d(d+1)}\right\}$.

Let $\ell \geqslant t^2$ and let $L = L(\ell, \varepsilon)$ be the number given by Szemerédi's Regularity Lemma (SRL). Finally, let $c = L/\varepsilon$.

Let G be a gaph a maximum degree $\leqslant d$ and let the edges of K_n, where $n \geqslant c|G|$, be coloured red/blue. Apply Szemerédi's Lemma to R, the red subgraph of K_n, with ℓ, ε as above.

Let H be the graph with vertex set $\{V_1, \dots, V_m\}$, where $V_1, \dots, V_m$ is the partition of R given by SRL. Let $V_iV_j \in E(H)$ if (V_i, V_j) is ε-uniform. Then $|H| = m \geqslant t^2$ and $e(\overline{H}) \leqslant \varepsilon\binom{m}{2}$. So $H \supset K_t$. [Or else, by Turán's theorem, there exist integers $m_1, \dots, m_{t-1}$ with $\sum m_i = m$ and $e(\overline{H}) \geqslant \sum \binom{m_i}{2} \geqslant (t-1)\binom{m/(t-1)}{2} > \varepsilon\binom{m}{2}$ where $\varepsilon \leqslant \frac{1}{t}$ and $m \geqslant t^2$.]

Thus we may assume that every pair (V_i, V_j), $1 \leqslant i < j \leqslant t$ is ε-uniform. Colour the edges of K_t green if $d(V_i, V_j) \geqslant \frac{1}{2}$, and yellow otherwise.

Since $t = R(d+1)$, there exists a monochromatic K_{d+1}, and we may assume that it is spanned by $V_1, \dots, V_{d+1}$. Therefore, $d(V_i, V_j) \geqslant \frac{1}{2}$ for $1 \leqslant i < j \leqslant d+1$, or $d(V_i, V_j) < \frac{1}{2}$ for $1 \leqslant i < j \leqslant d+1$.

We will show that in the first case we have $R \supset G$, and that in the other case we obtain similarly a blue G in K_n.

Take a vertex colouring of G with $d+1$ colours in which no colour is used more than $|G|$ times. Lemma 4.3 applied with H(there) $= G$(here), G(there) $=$ (subgraph of R spanned by $V_1, \dots, V_{d+1}$), $u = |V_i| \geqslant n/L \geqslant c|G|/L \geqslant |G|/\varepsilon$, $s = |G|$, $r = d+1$, $\lambda = \frac{1}{2}$, $(d+1)\varepsilon \leqslant 1/2^d$.

This completes the proof. $\square$

Theorem 5.1 was proved by Chvatál, Rödl, Szemerédi, Trotter in 1983. It was extended to a larger class (using a modification of Lemma 4.3) by Chen and Schelp in 1993. Graham, Rödl, Rucinski showed in 1999 that $c(d) \leqslant 2^{Cd\log^2 d}$ (obviously without SRL).

Observe that the Erdős-Stone theorem follows from SRL and Turán's theorem. Indeed, if $e(G) \geqslant \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$, form the "reduced" graph H whose edges correspond to pairs of positive density. Then $e(H) \geqslant t_r(|H|)$ so $H \supset K_{r+1}$, and so $G \supset K_{r+1}$ by Lemma 4.3.

We make this precise, but, more interestingly, we recover stability too. We use an argument of Erdős.

Theorem 5.2 (Erdős, 1970). Let G be a K_{r+1}-free graph. Then there is a (complete) r-partite graph H with $V(H) = V(G)$ and $d_H(v) \geqslant d_G(v)$ for all $v \in V(G)$.

Proof. By induction on r. The case $r = 1$ is trivial.

In general, let x be a vertex of maximum degree. Then $G' = G[\Gamma(x)]$ is K_r-free, so by induction there is an $(r-1)$-partite graph H' with $V(H') = V(G') = \Gamma(x)$, and $d_{H'}(v) \geqslant d_{G'}(v)$ for all $v \in \Gamma(x)$.

Form H by joining every vertex of $V(G) \setminus \Gamma(x)$ to every vertex in $\Gamma(x)$. Then H is r-partite and if $v \notin \Gamma(x)$ then $d_H(x) = |\Gamma(x)| = d_G(x) \geqslant d_G(v)$, while if $v \in \Gamma(x)$ then $d_H(v) = d_{H'}(v) + |V(G) \setminus \Gamma(x)| \geqslant d_{G'}(v) + |V(G) \setminus \Gamma(x)| \geqslant d_G(v)$. □

Lecture 11

Theorem 5.3 (Füredi, 2010). Let G be a K_{r+1}-free graph. Then there is a complete r-partite graph H with $V(H) = V(G)$ and $|E(G) \setminus E(H)| \leqslant \frac{1}{2}|E(H) \setminus E(G)|$.

In particular, if $e(G) \geqslant t_r(|G|) - k$ then $|E(H) \triangle E(G)| \leqslant 3k$.

Proof. Construct H as in the previous proof. The proof is via induction, the case $r = 1$ being trivial.

Now $|E(G') \setminus E(H')| \leqslant \frac{1}{2}|E(H') \setminus E(G')|$. Let there be e edges from G inside $V(G) \setminus \Gamma(x)$, and f edges *missing* from G between $\Gamma(x)$ and $V(G) \setminus \Gamma(x)$. Then $|E(G) \setminus E(H)| = |E(G') \setminus E(H')| + e$, and $|E(H) \setminus E(G)| = |E(H') \setminus E(G')| + f$, so it suffices to show that $e \leqslant \frac{1}{2} f$.

For each $v \in V(G) \setminus \Gamma(x)$, let $e(v)$ be the number of edges of G inside $V(G) \setminus \Gamma(x)$ meeting v, and let $f(v)$ be the number of edges missing from G between $\Gamma(x)$ and $V(G) \setminus \Gamma(x)$ meeting v. Then $e = \frac{1}{2} \sum_v e(v)$ and $f = \frac{1}{2} \sum_v f(v)$.

But for each v, $d_G(x) \geqslant d_G(v) = d_G(x) - f(v) + e(v)$, so $e(v) \leqslant f(v)$, and we are done.

Finally, if $e(G) \geqslant t_r(|G|) - k$, then $t_r(|G|) \geqslant e(H) = e(G) - |E(G) \setminus E(H)| + |E(H) \setminus E(G)| \geqslant t_r(G) - k + \frac{1}{2}|E(H) \setminus E(G)|$, so $|E(H) \setminus E(G)| \leqslant 2k$, hence $|E(G) \setminus E(H)| \leqslant k$, and so $|E(H) \triangle E(G)| \leqslant 3k$. □

Definition 5.4. Let G be a graph with an ε-uniform partition P into m parts $V_1, \ldots, V_m$. Let $\lambda \in \mathbb{R}$. The **reduced graph** $G(P, \lambda)$ is the graph of order M with vertex set $\{V_1, \ldots, V_m\}$ and edge set $\{V_i V_j : (V_i, V_j)$ is ε-uniform and $d(V_i, V_j) \leqslant \lambda\}$.

Lemma 5.5. Let G, P be as in this definition. If $e(G) \geqslant c\binom{|G|}{2}$ then

$$e(G(P, \lambda)) \leqslant \left(c - \varepsilon - \lambda - \frac{1}{m-1}\right)\binom{m}{2},$$

provided $|G| \geqslant 2m^2$.

Proof. Let $e(G(P, \lambda)) = d\binom{m}{2}$ and $n = |G|$. Then

$$\begin{aligned}
e(G) &\leqslant d\binom{m}{2}\left(\frac{n}{m}+1\right)^2 + \binom{m}{2}\varepsilon\left(\frac{n}{m}+1\right)^2 + \binom{m}{2}\lambda\left(\frac{n}{m}+1\right)^2 + m\binom{\frac{n}{m}+1}{2} \\
&\leqslant \left(d + \varepsilon + \lambda + \frac{1}{m-1}\right)\binom{m}{2}\left(\frac{n}{m}+1\right)^2 \\
&\leqslant \left(d + \varepsilon + \lambda + \frac{1}{m-1}\right)\binom{n}{2}
\end{aligned}$$

□

Theorem 5.6. Let $r, t \in \mathbb{N}$ and let $\varepsilon > 0$. Then there exists $n_0(r, t, \varepsilon)$ such that, if $|G| = n \geqslant n_0$ and $G \not\supset K_{r+1}(t)$, then G has a subgraph G' with $e(G') \geqslant e(G) - \varepsilon\binom{n}{2}$ and $G' \not\supset K_{r+1}$.

Proof. Let $\lambda = \varepsilon/2$. Pick η so that $(rt+1)\eta \leqslant \lambda^{rt}$ and $\eta \leqslant \varepsilon/4$. Let $n_0 = 2L(\ell, n)^2$, where ℓ is picked so that $\frac{1}{\ell-1} < \frac{\varepsilon}{4}$. Then $t \leqslant \lceil \varepsilon \lfloor n/L \rfloor \rceil$ (if not, make n_0 bigger).

Take an η-uniform partition P of G into $\ell \leqslant m \leqslant L$ parts, as given by Theorem 4.5. If $K_{r+1} \subset G(P, \lambda)$, then by Lemma 4.3, $G \supset K_{r+1}(t)$. Hence $K_{r+1} \not\subset G(P, \lambda)$.

Let G' be the subgraph of G consisting of all edges between pairs V_i and V_j where $V_iV_j \in E(G(P, \lambda))$. Then $G' \not\supset K_{r+1}$, for such a subgraph would have vertices in distinct parts of P.

The edges of G not in G' are those in non-uniform pairs, pairs of density $< \lambda$, and inside parts, so

$$\begin{aligned} e(G) - e(G') &\leqslant \eta \binom{m}{2} \left(\frac{n}{m} + 1\right)^2 + \binom{m}{2} \lambda \left(\frac{n}{m} + 1\right)^2 + m \binom{\frac{n}{m} + 1}{2} \\ &\leqslant \left(\eta + \lambda + \frac{1}{m-1}\right) \binom{n}{2} \; < \; \varepsilon \binom{n}{2} \end{aligned}$$

□

Theorem 5.8. Let $r, t \in \mathbb{N}$ and let $\varepsilon > 0$. Then there exists $n_0(r, t, \varepsilon)$ such that, if $|G| = n \geqslant n_0$, $G \not\supset K_{r+1}(t)$ and $e(G) \geqslant \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}$, then there is a complete r-partite graph T with $V(T) = V(G)$ and $|E(T) \triangle E(G)| \leqslant 11\varepsilon\binom{n}{2}$.

Remarks.

1. This is essentially Theorem 2.1.
2. Hence $e(G) \leqslant e(T) + 11\varepsilon\binom{n}{2} \leqslant \left(1 - \frac{1}{r} + 12\varepsilon\right)\binom{n}{2}$, which is the Erdős-Stone theorem (without the $\log n$).

Proof. By Lemma 5.6, choose $G' \subset G$ with $e(G') \geqslant \left(1 - \frac{1}{r} - 2\varepsilon\right)\binom{n}{2}$ and $G' \supset K_{r+1}$. Then $e(G') \geqslant t_r(n) - 3\varepsilon\binom{n}{2}$.

Apply Theorem 5.3 to G' to obtain T with $|E(G') \triangle E(T)| \leqslant 9\varepsilon\binom{n}{2}$.

Then $|E(G) \triangle E(T)| \leqslant 11\varepsilon\binom{n}{2}$. □

Lecture 12

6. Hypergraphs

Let F be an ℓ-uniform hypergraph. Define

$$ex(n, F) = \max\{e(G) : G \text{ is } \ell\text{-uniform}, \; |G| = n, \; G \not\supset F\}$$

and

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

Denote by K_r^ℓ the complete ℓ-uniform hypergraph of order r with $\binom{r}{\ell}$ edges. The value of $\pi(K_r^\ell)$ is unknown for $r > \ell > 3$. Turán conjectured that $\pi(K_4^3) = 5/9$. Erdős offered \$5000 for a solution in his memory.

Best bounds are $\frac{1}{6}(-1 + \sqrt{21}) = 0.5971$ (Giraud, 1989), $\frac{1}{12}(3 + \sqrt{17}) = 0.5935$ (Chung & Lu, 1999), 0.561555 (Razborov, using flag algebras, 2008).

Turán further conjectured $ex(n, K_5^3) = \binom{n}{3} - \binom{\lfloor n/2 \rfloor}{3} - \binom{\lceil n/2 \rceil}{3}$, but this is false for odd $n > 9$ (Sidorenko, 1990s).

The situation appears *unstable.*

$\pi(F)$ is known only in one or two simple cases; more recently, in a couple of less trivial examples, e.g. the Fano plane. But stability holds here. (Füredi, Simonovits, Pikhurko, Sudakov, Keevadi.)

Clearly $\pi(K_r^\ell) \leqslant 1 - \frac{1}{\binom{r}{\ell}}$. The best known general bound is as follows.

Lemma 6.1. Let G be an ℓ-uniform hypergraph with n_r copies of K_r^l (for $r \geqslant \ell$), where $n_\ell = e(G)$ and $n_{\ell-1} = \binom{n}{\ell-1}$, and $n = |G|$. Then provided $n_{r-1} > 0$,

$$\frac{n_{r+1}}{n_r} \;\geqslant\; \frac{r^2}{(r-\ell+1)(r+1)} \times \frac{n_r}{n_{r-1}} \;-\; \frac{(\ell-1)(n-r)+r}{(r-\ell+1)(r+1)}$$

Proof. Let $A_1, \ldots, A_{n_{r-1}}$ be an enumeration of the K_{r-1}'s, and $B_1, \ldots, B_{n_r}$ enumerate the K_r's. Let a_i be the number of K_r's containing A_i and b_j be the number of K_{r+1}'s containing B_j. Then $\sum a_i = rn_r$ and $\sum b_j = (r+1)n_{r+1}$.

Let N be the number of pairs (S,T) where S is the vertex set of a K_r, T is an r-set not K_r, and $|S \cap T| = r-1$. Clearly

$$N \;=\; \sum a_i(n-r+1-a_i) = (n-r+1)rn_r - \sum a_i^2 \;\leqslant\; (n-r+1)rn_r - \frac{r^2 n_r}{n_{r-1}}$$

On the other hand, for each B_j we can find $n-r-b_j$ vertices x such that $B \cup \{x\}$ doesn't span K_{r+1}. So there is an $(\ell-1)$-set $Y \subset B_j$ such that $Y \cup \{x\}$ is not an edge.

Then each $z \in B_j \setminus Y$ gives a pair $(S,T) = (B_j, B_i \setminus \{z\} \cup \{x\})$. Thus

$$N \;\geqslant\; \sum_j (n-r-b_j)(r-\ell+1) \;=\; (r-\ell+1)\big((n-r)n_r - (r+1)n_{r+1}\big)$$

Comparing bounds on N gives the result. □

Corollary 6.2 (de Caen, 1983). $ex(n, K_r^\ell) \leqslant \binom{n}{l}\left(1 - \frac{1}{\binom{r-1}{\ell-1}} \times \frac{n-r+1}{n-\ell+1}\right)$.

In particular, $\pi(K_r^\ell) \leqslant 1 - \frac{1}{\binom{r-1}{\ell}}$.

Proof. Let G be maximal with no K_r^ℓ. Then $n_{\ell-1} = \binom{n}{\ell-1}$ and $n_s > 0$ for $\ell \leqslant s < r$.

We show by induction that

$$\frac{n_s}{n_{s-1}} \;\geqslant\; \frac{n-\ell+1}{s}\binom{s-1}{\ell-1}\left(\frac{n_\ell}{\binom{n}{\ell}} - 1 + \frac{n-s+1}{n-\ell+1} \times \frac{1}{\binom{s-1}{\ell-1}}\right)$$

holds for $\ell \leqslant s < r$, which proves the Corollary since $n_r = 0$.

The case $s = \ell$ holds automatically since $n_{s-1} = \binom{n}{s-1}$.

Writing $q_s = n_s/n_{s-1}$ the desired inequality is

$$q_s \;\geqslant\; \frac{n-\ell+1}{s}\binom{s-1}{\ell-1}\left(\frac{n_\ell}{\binom{n}{\ell}}-1\right)+\frac{n-s+1}{s}$$

which follows from

$$q_s \;\geqslant\; \frac{(s-1)^2}{(s-\ell)^2}q_{s-1}-\frac{(\ell-1)(n-s+1)+s-1}{(s-\ell)s}$$

(from Lemma 6.1) and induction. □

The next theorem (Erdős, 1964) shows that ℓ-partite, ℓ-uniform hypergraphs have extremal function $o(n^\ell)$, i.e. $\pi = 0$.

An ℓ-uniform hypergraph H is **ℓ-partite** if $V(H) = V_1 \cup \ldots \cup V_\ell$ disjoint V_i, and each edge has one vertex in each class. The complete ℓ-partite ℓ-uniform $K^\ell(t_1, \ldots, t_\ell)$ has $|V_i| = t_i$ and all possible edges.

Lecture 13 **Theorem 6.3.** Let G be an ℓ-uniform hypergraph of order n and size $pn^\ell/\ell!$. Let $t_1 \leqslant t_2 \leqslant \ldots \leqslant t_\ell$ be positive integers and suppose that $p^{t_1 \cdots t_{\ell-1}} > T^2 n^{-1}$, where $T = \sum t_i$.

Then G contains at least $\frac{1}{2T!}p^{t_1 \cdots t_\ell} n^T$ copies of $K^\ell(t_1, \ldots, t_\ell)$.

Remarks.

1. Note that this quantity is similar to the expected number if the edges were chosen randomly.
2. This implies that $ex(n, K^\ell(t_1, \ldots, t_\ell)) \leqslant c_T n^{\ell - 1/t_1 \cdots t_\ell}$.

Proof. Let $\chi : V^\ell \to \{0, 1\}$ be the characteristic (indicator) function of edges.

Then $\sum_{x_1,\ldots,x_\ell} \chi(x_1, \ldots, x_l) = pn^l$.

Let $f(t_1, \ldots, t_\ell) = \dfrac{1}{n^T} \sum_{x_1^1,\ldots,x_1^{t_1}} \cdots \sum_{x_\ell^1,\ldots,x_\ell^{t_\ell}} \prod_{x_1^{i_1}} \cdots \prod_{x_\ell^{i_\ell}} \chi(x_1^{i_1} \cdots x_\ell^{i_\ell})$.

Note that $n^T f(t_1, \ldots, t_\ell)$ is the number of labelled homomorphic copies of $K^\ell(t_1, \ldots, t_\ell)$ in G (copies where vertices in the same class might coincide). I.e., $f()$ is the probability that a randomly chosen collection of T vertices spans a homomorphic copy of $K^\ell(t_1, \ldots, t_\ell)$. Now

$$\begin{aligned}
f(t_1, \ldots, t_\ell) &= \frac{1}{n^{T-t_\ell}} \sum_{x_1^1,\ldots,x_1^{t_1}} \cdots \sum_{x_{\ell-1},\ldots,x_{\ell-1}^{t_{\ell-1}}} \prod_{x_1^{i_1}} \cdots \prod_{x_{\ell-1}^{i_{\ell-1}}} \sum_{x_\ell^1,\ldots,x_\ell^{t_\ell}} \prod_{x_\ell} \chi(x_1^{i_1}, \ldots, x_{\ell-1}^{i_{\ell-1}}, x_\ell^{i_\ell}) \\
&= \frac{1}{n^{T-t_\ell}} \sum_{x_1^1,\ldots,x_1^{t_1}} \cdots \sum_{x_{\ell-1},\ldots,x_{\ell-1}^{t_{\ell-1}}} \prod_{x_1^{i_1}} \cdots \prod_{x_{\ell-1}^{i_{\ell-1}}} \left(\frac{1}{n}\sum_{x_\ell} \chi(x_1^{i_1}, \ldots, x_{\ell-1}^{i_{\ell-1}}, x_\ell)\right)^{t_\ell} \\
\text{use Jensen's inequality} &\geqslant \left(\frac{1}{n^{T-t_\ell}} \sum_{x_1^1,\ldots,x_1^{t_1}} \cdots \sum_{x_{\ell-1},\ldots,x_{\ell-1}^{t_{\ell-1}}} \prod_{x_1^{i_1}} \cdots \prod_{x_{\ell-1}^{i_{\ell-1}}} \frac{1}{n}\sum_{x_\ell} \chi(x_1^{i_1}, \ldots, x_{\ell-1}^{i_{\ell-1}}, x_\ell)\right)^{t_\ell} \\
&= f(t_1, \ldots, t_{\ell-1}, 1)^{t_\ell} \\
&\geqslant f(t_1, \ldots, t_{\ell-2}, 1, 1)^{t_{\ell-1}t_\ell} \;\geqslant\; \cdots \;\geqslant\; f(1, \ldots, 1)^{t_1 t_2 \cdots t_\ell} \;=\; p^{t_1 \cdots t_\ell}
\end{aligned}$$

The contribution to $n^T f(t_1,\ldots,t_\ell)$ from terms where, say, $x_1^i = x_1^j$ is $n^{T-1} f(t_1-1, t_2, .., t_\ell)$, so the contribution from terms where all variables are distinct, i.e. the number of labelled copies of $K^\ell(t_1,\ldots,t_\ell)$ is at least

$$\begin{aligned} p^{t_1\ldots t_\ell} n^T - \sum_{i=1}^{\ell} \binom{t_i}{2} p^{t_1\ldots(t_i-1)\ldots t_\ell} n^{T-1} &\geqslant p^{t_1\ldots t_\ell} n^T \left(1 - \binom{T}{2} p^{-t_1\ldots t_{\ell-1}} n^{-1}\right) \\ &\geqslant \frac{1}{2} p^{t_1\ldots t_\ell} n^T \end{aligned}$$

□

7. The size of a hereditary property

A(n **ℓ-uniform hyper**)**graph property** $\mathcal{P}$ is a class of (ℓ-uniform hyper)graphs closed under isomorphism.

It is **non-trivial** if it is not the class of all graphs.

It is **hereditary** if it is closed under taking induced subgraphs: if $G \in \mathcal{P}$ and H is an induced subgraph of G, then $H \in \mathcal{P}$. Thus $\mathcal{P}$ is hereditary if and only if it is closed under the removal of vertices.

$\mathcal{P}$ is **monotone** if it is closed under taking *any* subgraph – i.e., it is closed under the removal of edges and vertices. Thus monotone $\implies$ hereditary.

Examples.

- "3-colourable" is monotone
- "planar" is monotone
- "K_4-free" is monotone
- "no induced C_4" is hereditary

Observe that a hereditary property $\mathcal{P}$ can be written $\mathcal{P} = \mathrm{Forb}(\mathcal{F})$, where $\mathcal{F}$ is a class of graphs and $\mathrm{Forb}(\mathcal{F})$ is the class of graphs containing no member of $\mathcal{F}$ as an induced subgraph. To see this, we could take $\mathcal{F}$ to be $\mathcal{F}_1$ = graphs not in $\mathcal{P}$. More usefully, we have $\mathcal{P} = \mathrm{Forb}(\mathcal{F}_0)$, where $\mathcal{F}_0$ = minimal elements of $\mathcal{F}_1$ with respect to inclusion, i.e. $\mathcal{F}_0 = \{F \notin \mathcal{P} :$ every proper induced subgraph of F is in $\mathcal{P}\}$.

For monotone properties, it is more natural to write $\mathcal{P} = \mathrm{Forb}_*(\mathcal{F})$ where "induced" is removed from the above.

Examples.

- "2-colourable" $= \mathrm{Forb}_*(\text{odd cycles})$
- "no induced C_4" $= \mathrm{Forb}(C_4)$

Observe that, for any $\mathcal{F}$, $\mathrm{Forb}(\mathcal{F})$ is hereditary, and $\mathrm{Forb}_*(\mathcal{F})$ is monotone.

How "large" is $\mathcal{P}$?

Let $\mathcal{P}_n$ be the graphs in $\mathcal{P}$ with labelled vertex set $[n]$.

Example: let $\mathcal{P} = \text{Forb}(K_4) = \text{Forb}_*(K_4)$. Then $|\mathcal{P}_n| \geqslant 2^{t_3(n)} = 2^{\left(1-\frac{1}{3}+o(1)\right)\binom{n}{2}}$, because we can take a Turán graph $T_3(n)$ and all its subgraphs.

Lecture 14 We extend Definition 5.4 slightly.

Definition 7.1. Let G be a graph with an ε-uniform partition P into $m = |P|$ parts $V_1, \ldots, V_m$. Let $\lambda, \mu \in \mathbb{R}$. The **reduced graph** $G(P, \lambda, \mu)$ is the graph of order m with vertex set $\{V_1, \ldots, V_m\}$ and edge set $\{V_iV_j : (V_i, V_j) \text{ is } \varepsilon\text{-uniform}, \lambda \leqslant d(V_i, V_j) \leqslant 1 - \mu\}$.

Hence $G(P, \lambda) = G(P, \lambda, 0)$.

Lemma 7.2. Let $V_1, \ldots, V_m$ be a partition of $[n]$, where $n \geqslant 2m^2$. Let $0 < \varepsilon, \lambda < \frac{1}{4}$. Let H be a graph with vertex set $\{V_1, \ldots, V_m\}$, where $e(H) = d\binom{m}{2}$.

Then the number of graphs having vertex set $[n]$ and having an ε-uniform partition P into parts $V_1, \ldots, V_m$ such that $G(P, \lambda, \lambda) = H$ is at most

$$2^{\left(d+\varepsilon+8\lambda\log\frac{1}{\lambda}+\frac{1}{m-1}\right)\binom{n}{2}}$$

Proof. Consider a graph G of this kind. The number of possible edges of $[n]^{(2)}$ between parts corresponding to edges of $G(P, \lambda, \lambda)$, or to non-uniform pairs, or within parts, is (as in the proof of Lemma 5.5), at most $\left(d + \varepsilon + \frac{1}{m-1}\right)\binom{n}{2}$.

Hence there are at most $2^{\left(d+\varepsilon+\frac{1}{m-1}\right)\binom{n}{2}}$ ways to choose edges of G from these. In how many ways can we choose edges of G between pairs where we must have $d(V_i, V_j) \leqslant \lambda$ or $d(V_i, V_j) \geqslant 1 - \lambda$?

Let $E = \left(\frac{n}{m} + 1\right)^2$. The number of choices between some fixed pair (V_i, V_j) is at most

$$2\sum_{i=0}^{\lambda E}\binom{E}{i} \leqslant 2(\lambda E + 1)\binom{E}{\lambda E} \leqslant \binom{E}{\lambda E}^2 \leqslant \left(\frac{e}{\lambda}\right)^{2\lambda E} \leqslant 2^{\left(8\lambda\log\frac{1}{\lambda}\right)E}$$

(using $\binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k$).

Since $E\binom{m}{2} \leqslant \binom{n}{2}$, the total number of choices for G is at most

$$2^{\left(d+\varepsilon+\frac{1}{m-1}\right)\binom{n}{2}} \times 2^{\left(8\lambda\log\frac{1}{\lambda}\right)\binom{n}{2}}$$

□

Theorem 7.3 (Erdős, Frankl, Rödl, 1936). Let $\mathcal{P} = \text{Forb}_*(\mathcal{F})$ and let $r + 1 = \min\{\chi(F) : F \in \mathcal{F}\}$.

Then $|\mathcal{P}_n| = 2^{\left(1-\frac{1}{r}+o(1)\right)\binom{n}{2}}$.

Proof. Since every subgraph of $T_r(n)$ is in $\mathcal{P}$, we have $|\mathcal{P}_n| \geqslant 2^{\left(1-\frac{1}{r}+o(1)\right)\binom{n}{2}}$.

To prove the converse inequality, let $\varepsilon > 0$. We will show that, if n is large, then $|\mathcal{P}_n| \leqslant 2^{\left(1-\frac{1}{r}+\varepsilon\right)\binom{n}{2}}$.

Pick $F \in \mathcal{F}$ with $\chi(F) = r + 1$. Let $t = |F|$. Thus if $G \in \mathcal{P}$ then $G \not\supset K_{r+1}(t)$. Pick λ so that $8\lambda\log\frac{1}{\lambda} < \frac{\varepsilon}{5}$. Pick $0 < \eta < \frac{\varepsilon}{5}$ so that $(rt + 1)\eta < \lambda^{rt}$. Pick ℓ so that $\frac{1}{\ell-1} < \frac{\varepsilon}{5}$ and $t_r(\ell) < \left(1 - \frac{1}{r} + \frac{\varepsilon}{5}\right)\binom{\ell}{2}$.

By Theorem 4.5, each $G \in \mathcal{P}_n$ has an η-uniform partition P into $\ell \leqslant m \leqslant L$ parts, where $L = L(\ell, \eta)$. Since $K_{r+1}(t) \not\subset G$, then $K_{r+1} \not\subset G(P, \lambda, \lambda)$ if n is large, by Lemma 4.3. So $e(G(P, \lambda, \lambda)) \leqslant t_r(m) \leqslant \left(1 - \frac{1}{r} + \frac{\varepsilon}{5}\right)\binom{m}{2}$.

There are at most L^n choices for $V_1, \ldots, V_m$ and at most $2^{\binom{L}{2}}$ choices for H such that $G(P, \lambda, \lambda) = H$, and at most $2^{\binom{L}{2}}$ ways to specify which are the non-uniform pairs, and at most $2^{\binom{L}{2}}$ ways to specify whether a non-edge has density $< \lambda$ or $> 1 - \lambda$.

By Lemma 7.2, the number of $G \in \mathcal{P}_n$ is at most

$$2^{3\binom{L}{2}} L^n 2^{\left(1 - \frac{1}{r} + \frac{\varepsilon}{5} + \frac{\varepsilon}{5} + \frac{\varepsilon}{5} + \frac{\varepsilon}{5}\right)\binom{n}{2}} < 2^{\left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}}$$

□

How about Forb($\mathcal{F}$)? Can we find some simple hereditary property to play the role of r-colourable graphs in the previous proof?

Definition 7.4. A graph is (r, s)**-colourable** if its vertex set can be partitioned into r classes, s of which span complete graphs, and the others of which contains no edges.

Denote by $\mathcal{C}(r, s)$ the class of (r, s)-colourable graphs.

Note. $\mathcal{C}(r, 0)$ is the class of r-colourable graphs. $\mathcal{C}(r, r)$ is their complements. And $\mathcal{C}(r, s)$ is hereditary.

Lecture 15 **Definition.** For a non-trivial property $\mathcal{P}$, let $r(\mathcal{P}) = \max\{r : \exists s \mathcal{C}(r, s) \subset \mathcal{P}\}$.

Note that $r(\mathcal{P})$ is well-defined. For there exists some graph $H \notin \mathcal{P}$ and $H \in \mathcal{C}(r, s)$ for all $r \geqslant |H|$ and all s. Thus $r(\mathcal{P}) < |H|$. On the other hand, if $\mathcal{P}$ contains arbitrarily large graphs and is hereditary, then by Ramsey's theorem either $\mathcal{C}(1, 0) \subset \mathcal{P}$ or $\mathcal{C}(1, 1) \subset \mathcal{P}$. Thus $r(\mathcal{P}) \geqslant 1$.

Lemma 7.6. Let $t \geqslant 3$ and let $R = R(t)$ be the Ramsey number of t. Let K_n, $n \geqslant R$, be edge-coloured red, blue and grey, with fewer than $\frac{1}{R^t}\binom{n}{2}$ grey edges. Then there exists a red or blue K_t.

Proof. Suppose we recolour the grey edges red. Every set of R vertices contains a red or blue K_t, so we have at least $\binom{n}{R}\binom{n-R}{R-t}^{-1} = \binom{n}{t}\binom{R}{t}^{-1}$ of these. The number of these containing an originally grey edge is at most $\frac{1}{R^t}\binom{n}{2}\binom{n-2}{t-2} = \frac{1}{R^t}\binom{t}{2}\binom{n}{t} < \binom{n}{t}\binom{R}{t}^{-1}$. □

Lemma 7.7. Let $r, t \in \mathbb{N}$ and $0 \leqslant \lambda \leqslant \frac{1}{4}$. Then there exists $n_0 \in \mathbb{N}$ and $\eta_0 > 0$ such that the following holds.

Let G be a graph with vertex set $\bigcup_{i=1}^{r+1} V_i$, there V_i being disjoint and $|V_i| \geqslant n_0$. Suppose, for $1 \leqslant i < j \leqslant r + 1$, then $\lambda \leqslant d(V_i, V_j) \leqslant 1 - \lambda$, and (V_i, V_j) is η_0-uniform. Then there exists $0 \leqslant s \leqslant r + 1$ such that G contains every graph in $\mathcal{C}(r + 1, s)_t$ as an induced subgraph.

Proof. Let $R = R(t)$. Pick $\varepsilon < \left\{\frac{1}{R^t}, \frac{1}{t+1}\left(\frac{\lambda}{2}\right)^t\right\}$. Let $L = L(R, \varepsilon)$. Take an ε-uniform partition Q of V_1 into $R \leqslant m \leqslant L$ parts.

Colour the edges of the reduced graph $G[V_1](Q, 0, 0)$ red or blue according to whether the density of the pair is $\leqslant \frac{1}{2}$ or $< \frac{1}{2}$. Then, by Lemma 7.6, there is a monochromatic K_t. So we may pick subsets V_{1p}, $1 \leqslant p \leqslant t$, in V_1, so that (V_{1p}, V_{1q}) is ε-uniform, $1 \leqslant p, q \leqslant t$, and either $d(V_{1p}, V_{1q}) \leqslant \frac{1}{2}$ for $1 \leqslant p, q \leqslant t$ or $d(V_{1p}, V_{1q}) > \frac{1}{2}$ for $1 \leqslant p, q \leqslant t$.

We call V_1 a "red" class in the first case, and a "blue" class in the second case. Repeat this on the classes $V_2, \ldots, V_{r+1}$. Let there be s blue classes, wlog $V_1, \ldots, V_s$.

Note that $|V_{ip}| \geqslant \dfrac{|V_i|}{L+1}$ for all i, p. Let $\eta_0 = \min\left\{\dfrac{\varepsilon}{L+1}, \dfrac{\lambda}{2}\right\}$.

Then for all $U \subset V_{ip}$, $W \subset V_{jq}$ with $|U| \geqslant \varepsilon|V_{ip}| \geqslant \eta_0|V_i|$, $|W| \geqslant \varepsilon|V_{jq}| \geqslant \eta_0|V_j|$, we have $|d(U,W) - d(V_i, V_j)| < \eta_0$. In particular, $|d(V_{ip}, V_{jq}) - d(V_i, V_j)| < \eta_0$. Since $\eta_0 < \frac{\varepsilon}{2}$, we have that (V_{ip}, V_{jq}) is ε-uniform.

Moreover, $\frac{\lambda}{2} \leqslant \lambda - \eta_0 \leqslant d(V_{ip}, V_{jq}) \leqslant 1 - \lambda + \eta_0 \leqslant 1 - \frac{\lambda}{2}$ for all $1 \leqslant i < j \leqslant r+1$, $1 \leqslant p, q \leqslant t$.

Let $H \in \mathcal{C}(r+1, s)_t$. Then, provided that n_0 is large enough, Corollary 4.4 shows that we can find H between some t of the classes V_{ip}. Indeed, H has a partition into s classes of sizes $t_1, \ldots, t_s$ spanning complete graphs and $r - s$ classes of sizes $t_{s+1}, \ldots, t_r$ spanning empty graphs, $\sum t_i = t$. Apply Corollary 4.4 to V_{ip}, $1 \leqslant i \leqslant r+1$, $1 \leqslant p \leqslant t_i$. All pairs are ε-uniform, with densities $\geqslant \frac{\lambda}{2}$ where an edge is needed, and densities $\leqslant 1 - \frac{\lambda}{2}$ where a non-edge is needed. □

The following theorem was proved by Prőmel & Steger (1991) for $\mathcal{P} = \mathrm{Forb}(F)$, by Bollobás & Thomason and by Alekseev (1992) in general. The proof here is due to Montgomery (2011).

Theorem 7.8. Let $\mathcal{P}$ be a hereditary property. Then $|\mathcal{P}_n| = 2^{\left(1-\frac{1}{r}+o(1)\right)\binom{n}{2}}$, where $r = r(\mathcal{P})$.

Proof. By definition of $r(\mathcal{P})$ there exists s such that $\mathcal{C}(r,s) \subset \mathcal{P}$. Let H be any subgroup of $T_r(n)$. Fill in s of the classes of H so that they are complete. This graph is in $\mathcal{C}(r,s)$, so in $\mathcal{P}$. There are $2^{t_r(n)}$ choices of H, so $|\mathcal{P}_n| \geqslant 2^{\left(1-\frac{1}{r}+o(1)\right)\binom{n}{2}}$.

To prove the converse inequality, let $\varepsilon > 0$. We show $|\mathcal{P}_n| \leqslant 2^{\left(1-\frac{1}{r}+\varepsilon\right)\binom{n}{2}}$ for all large n.

Suppose not. Then there are infinitely many n with $|\mathcal{P}_n| \geqslant 2^{\left(1-\frac{1}{r}+\varepsilon\right)\binom{n}{2}}$. Pick $t \in \mathbb{N}$. By applying the argument of Theorem 7.8, where η is picked with $\eta < \eta_0$, we see that there is an infinite number of n for which some $G \in \mathcal{P}_n$ has $K_{r+1} \subset G(P, \lambda, \lambda)$. But then, by Lemma 7.7, $\mathcal{C}(r+1, s)_t \subset \mathcal{P}_n \subset \mathcal{P}$.

Hence for each $t \in \mathbb{N}$, ther exists s with $\mathcal{C}(r+1, s)_t \subset \mathcal{P}$. But there are only finitely many possible s, so for some s, $\mathcal{C}(r+1, s)_t \subset \mathcal{P}$ for infinitely many t. But $\mathcal{P}$ is hereditary, so for this s, $\mathcal{C}(r+1, s) \subset \mathcal{P}$. This contradicts the definition of $r(\mathcal{P})$. □

Lecture 16

8. Containers

Let H be an ℓ-uniform hypergraph. Let $r = e(H)$. The r-uniform hypergraph $G(N, H)$ has vertex set $[N]^{(\ell)}$ where $B = \{v_1, \ldots, v_r\}$ is an edge whenever B, regarded as an ℓ-graph on vertex set $[N]$, is isomorphic to H.

Notice that H-free ℓ-uniform hypergraphs on vertex set $[N]$ correspond exactly to independent sets in $G(N, H)$ (i.e., sets $I \subset V(G)$ containing no edge of G). Hence Theorem 7.3 amounts to counting independent sets in $G(N, H)$.

How many independent sets can there be in a graph? If it's d-regular? Can be more than $2^{n/2}$ (beautiful entropy argument of Kahn, 2001, shows that the maximum is found in $\frac{n}{2d}K_{d,d}$).

How many maximal independent sets? Still many (add 1 factor to each class in previous example gives $> 2^{n/4}$).

For many applications we would need fewer. Sometimes it is enough to have instead a small set of containers.

Definition 8.1. A set of **containers** $\mathcal{C}$ for a hypergraph G is a collection of subsets $C \subset V(G)$ such that, for every independent set I, there exists $C \in \mathcal{C}$ with $I \subset C$.

We would like a small collection $\mathcal{C}$ of containers so each $C \in \mathcal{C}$ is "not too big".

We do this for graphs. It is not too optimistic to require $|C|$ not too big – e.g., $K_{d,n-d}$ needs a container with $|C| \geqslant n - d$.

Definition 8.2. Let G be a graph with average degree d and of order n. The **degree measure** of a subset $S \subset V(G)$ is

$$\mu(S) = \frac{1}{nd} \sum_{v \in S} d(v) \ .$$

Theorem 8.3. Let G be a graph of order n and average degree d. Let $\zeta > 0$. Then there is a function $C : \mathcal{P}[n] \to \mathcal{P}[n]$ such that, for every independent set $I \subset [n]$, there exists $T \subset I$ with

(a) $I \subset C(T)$
(b) $\mu(T) \leqslant 2/\zeta d$
(c) $|T| \leqslant 2n/\zeta^2 d$
(d) $\mu(C(T)) \leqslant \frac{1}{2} + 2\zeta + \mu(T)$ for all $T \subset \mathcal{P}[n]$

Remark. This gives a collection of containers with $|\mathcal{C}| \leqslant \binom{n}{2n/\zeta^2 d} \leqslant e^{\frac{n}{d}\log d}$ and each container is not big.

Proof. Order the vertices $v_1, \ldots, v_n$ by decreasing degree. Let $m = \max\{j : d(v_j) \geqslant \zeta d\}$. Given I, run the following algorithm to find T. Begin with sets $T = \emptyset$, $\Gamma = \emptyset$.

For $j = 1$ to m : do
  if $v_j \in I$
    if $\left|\{i > j : v_i v_j \in E(G) \text{ but } v_i \notin \Gamma\}\right| \geqslant \zeta d(v_j)$, add v_j to T
    for $i = j+1$ to n : do
      if $\left|\Gamma(v_i) \cap T\right| > \frac{d(v_i)}{d}$, add v_i to Γ

Suppose instead we are given some set T. Run the next algorithm, beginning with sets $D = [n]$, $\Gamma = \emptyset$.

For $j = 1$ to m : do
  if $\left|\{i > j : v_i v_j \in E(G) \text{ but } v_i \notin \Gamma\}\right| \geqslant \zeta d(v_j)$, remove v_j from D
  for $i = j+1$ to n : do
    if $\left|\Gamma(v_i) \cap T \cap \{v_1, \ldots, v_j\}\right| > \frac{d(v_i)}{d}$, add v_i to Γ

Observe that the set Γ constructed at each stage is the same. Observe that if T came from a run of the first algorithm, then $I \subset D \cup T$ where D is produced from the second algorithm. Also, in this case, $T \subset I$, and each vertex of Γ has a neighbour in T, so $I \cap \Gamma = \emptyset$.

Define $C(T) = (D \cup T) \setminus \Gamma$ via the second algorithm. The first algorithm shows (a) is true.

Note that when a vertex v_i enters Γ, it has at most $\frac{d(v_i)}{d} + 1$ neighbours in T, since it had fewer than $\frac{d(v_i)}{d} + 1$ neighbours before. So, counting edges between T and Γ,

$$\zeta n d \mu(T) = \zeta \sum_{v_j \in T} d(v_j) \leqslant \sum_{v_i \in [n]} 1 + \frac{d(v_i)}{d} = n + n = 2n$$

so (b) follows.

Since $v_j \in T$ means $j \leqslant m$, so $d(v_j) \geqslant \zeta d$ and

$$|T| \zeta d \leqslant \sum_{v_j \in T} d(v) = n d \mu(T)$$

and (c) follows.

Lecture 17 Let Y be the set of edges inside $D \setminus \Gamma$.

$$\begin{aligned} \sum_{v \in D \setminus \Gamma} d(v) &\leqslant 2|Y| + \sum_{v \notin D \setminus \Gamma} d(v) \\ \mu(D \setminus \Gamma) &\leqslant \frac{2|Y|}{nd} + 1 - \mu(D \setminus \Gamma) \\ \mu(D \setminus \Gamma) &\leqslant \frac{1}{2} + \frac{|Y|}{nd} \end{aligned}$$

Now

$$\begin{aligned} |Y| &= \big|\{v_j v_i \in E(G) : j < i \text{ and } v_j, v_i \in D \setminus \Gamma\}\big| \\ &\leqslant \sum_{j=1}^{m} \zeta d(v_j) + \sum_{j=m+1}^{n} \zeta d \ \leqslant \ \zeta n d + \zeta n d \end{aligned}$$

Therefore $\mu(D \setminus \Gamma) \leqslant \frac{1}{2} + 2\zeta$, so $\mu(C) \leqslant \mu(D \setminus \Gamma) + \mu(T)$ giving (d). □

This can be made to work for hypergraphs. Applied to $G(N, H)$ it gives the following, where

$$m(H) = \max_{H' \subset H, e(H) > 1} \frac{e(H') - 1}{v(H') - \ell}$$

Theorem 8.4. Let $\varepsilon > 0$. Then there exists $c > 0$ and n_0 such that, for $N > n_0$, there is a collection of ℓ-uniform hypergraphs on vertex set $[N]$ such that

(a) every H-free ℓ-uniform hypergraph is a subgraph of some $C \in \mathcal{C}$
(b) for every $C \in \mathcal{C}$, $e(C) \leqslant \big(\pi(H) + \varepsilon\big)\binom{N}{\ell}$
(c) $\log |\mathcal{C}| \leqslant c N^{\ell - 1/m(H)} \log N$

Corollary 8.5. The number of H-free ℓ-graphs is $2^{(\pi(H)+\varepsilon)\binom{N}{\ell}}$.

9. The Local Lemma

Suppose we have n events $A_1, \ldots, A_n$ and we want $\mathbb{P}\left(\bigcap_{j=1}^n \overline{A_j}\right) > 0$, i.e. it is possible that no A_j occurs.

Suppose $\mathbb{P}(A_j) = p$, $1 \leqslant j \leqslant n$, $0 < p < 1$.

If the A_i are independent, then $\mathbb{P}(\bigcap \overline{A_i}) = (1-p)^n > 0$.

Otherwise, if $p < 1/n$, then $\mathbb{P}(\bigcup A_i) < np < 1$, so $\mathbb{P}(\bigcap \overline{A_i}) > 0$.

What if $p > 1/n$ but the A_i are "somewhat independent"? For each A_i, choose $J_i \subset [n]$ such that A_i is independent of the system $\{A_j : j \notin J_i\}$. Note there is no unique minimal choice for J_i.

Theorem 9.1 (Local Lemma: Erdős, Lovász, 1975). Let $A_1, \ldots, A_n$ and $J_1, \ldots, J_n$ be as above. Suppose there exist $0 < \gamma_i < 1$ such that $\mathbb{P}(A_i) \leqslant \gamma_i \prod_{j \in J_i}(1-\gamma_j)$.

Then $\mathbb{P}\left(\bigcap_{i=1}^n \overline{A_i}\right) \geqslant \prod_{j=1}^n (1-\gamma_j) > 0$.

Corollary 9.2. Suppose $|J_i| \leqslant \Delta$ for all i and that $\mathbb{P}(A_i) \leqslant \dfrac{1}{e(\Delta+1)}$. Then $\mathbb{P}\left(\bigcap_{i=1}^n \overline{A_i}\right) > 0$.

Proof of 9.1. By induction on n, we may assume that $\mathbb{P}\left(\bigcap_{i=1}^{n-1} \overline{A_i}\right) > 0$, so we may condition on this event. We shall prove (by induction) that $\mathbb{P}\left(A_n \middle| \bigcap_{i=1}^{n-1} \overline{A_i}\right) \leqslant \gamma_n$.

This implies, for any $S \subset [n-1]$ and $i \notin S$, that $\mathbb{P}\left(A_i \middle| \bigcap_{j \in S} \overline{A_j}\right) \leqslant \gamma_i$.

The theorem follows because $\mathbb{P}\left(\bigcap_{i=1}^n \overline{A_i}\right) = \prod_{i=1}^n \mathbb{P}\left(\overline{A_i} \middle| \bigcap_{j<i} \overline{A_j}\right) \geqslant \prod_{i=1}^n (1-\gamma_i)$. Now

$$\begin{aligned}
\mathbb{P}\left(A_n \middle| \bigcap_{j<n} \overline{A_j}\right) &= \frac{\mathbb{P}\left(A_n \cap \bigcap_{j<n} \overline{A_j}\right)}{\mathbb{P}\left(\bigcap_{j \notin J_n \cup \{n\}} \overline{A_j}\right)} \times \frac{\mathbb{P}\left(\bigcap_{j \notin J_n \cup \{n\}} \overline{A_j}\right)}{\mathbb{P}\left(\bigcap_{j<n} \overline{A_j}\right)} \\
&\leqslant \frac{\mathbb{P}\left(A_n \cap \bigcap_{j \notin J_n} \overline{A_j}\right)}{\text{as above}} \times \frac{\text{as above}}{\text{as above}} \\
&= \frac{\mathbb{P}\left(A_n \middle| \bigcap_{j \notin J_n} \overline{A_j}\right)}{\mathbb{P}\left(\bigcap_{j<n} \overline{A_j} \middle| \bigcap_{j \notin J_n \cup \{n\}} \overline{A_j}\right)}
\end{aligned}$$

By definition of J_n, the top is $\mathbb{P}(A_n)$. Relabelling so that $J_n = \{1, 2, \ldots, d\}$, the bottom is

$$\mathbb{P}\left(\bigcap_{j=1}^{d} \overline{A_j} \middle| \bigcap_{j=d+1}^{n-1} \overline{A_j}\right) = \prod_{i=1}^{d} \mathbb{P}\left(\overline{A_i} \middle| \bigcap_{j=i+1}^{n-1} \overline{A_j}\right) \geqslant \prod_{i=1}^{d} (1-\gamma_i) = \prod_{j \in J_n} (1-\gamma_j)$$

□

Eek. Fiddly. I hope that at least some of that was transcribed correctly... Let me know if not.

Lecture 18 **Theorem 9.3.** Every r-regular r-uniform hypergraph is 2-colourable if $r \geqslant 9$ (i.e., the vertices can be coloured red/blue so that no edge is monochromatic).

Proof. Let the hypergraph have order n. Take a random 2-colouring. Let A_i, $1 \leqslant i \leqslant \frac{nr}{r} = n$, be the event that edge i is monochromatic. This is independent of $\{A_j : \text{edge } j \text{ is disjoint from edge } i\}$.

So we can take $J_i = \{A_j : \text{edge } j \text{ meets edge } i\}$ and $|J_i| \leqslant r(r-1)$, since $\frac{1}{2^i} \leqslant \frac{1}{e(r(r-1)+1)}$ if $r \geqslant 9$. □

We can use the Local Lemma to improve lower bounds for Ramsey. A simple argument gives $R(3,t) = \Omega\Big(\big(\frac{t}{\log t}\big)^{3/2}\Big)$.

Theorem 9.4. $R(3,t) \geqslant \big(1+o(1)\big)\Big(\dfrac{t}{\log t}\Big)^2$.

Proof. Fix $\varepsilon > 0$. Let $n = (1-\varepsilon)\big(\frac{t}{\log t}\big)^2$. Colour the edges of K_n red with probability $p = \frac{1-\varepsilon}{\sqrt{n}}$ and blue otherwise.

For $\alpha \in [n]^{(3)}$, let A_α be "$G[\alpha]$ is red", and for $\beta \in [n]^{(t)}$, let B_β be "$G[\beta]$ is blue".

We can take $J_\alpha = \{A_{\alpha'}, B_{\beta'} : |\alpha \cap \alpha'| \geqslant 2, |\alpha \cap \beta'| \ \geqslant 2\}$, and J_β likewise. So

$$|J_\alpha| \ \leqslant 3(n-3) + \binom{n}{t} \quad \text{and} \quad |J_\beta| \leqslant \binom{t}{3} + \binom{t}{2}(n-t) + \binom{n}{t}$$

By the Local Lemma, it suffices to find γ and δ such that

$$\begin{aligned} p^3 &\leqslant \gamma(1-\gamma)^{3(n-3)}(1-\delta)^{\binom{n}{t}} \\ (1-p)^{\binom{t}{2}} &\leqslant \delta(1-\gamma)^{\binom{t}{3}+\binom{t}{2}(n-t)}(1-\delta)^{\binom{n}{t}} \end{aligned}$$

Erdős-Szekeres showed $R(3,t) \leqslant \dbinom{s+t-2}{s-1} = \dbinom{t+1}{2}$.

Ajtai-Komlós-Szemerédi (1980) and Shearer (1986) showed $R(3,t) = O\Big(\dfrac{t^2}{\log t}\Big)$.

Kim (1995) showed $R(3,t) = \Theta\Big(\dfrac{t^2}{\log t}\Big)$.

10. Tail Estimation

Let X be a random variable with mean μ. Chebychev's bound $\mathbb{P}(|X-\mu| > t) \leqslant \frac{\operatorname{Var} X}{t^2}$ gives a simple bound that X is far from μ, but it is weak.

E.g., if $X \sim \operatorname{Bin}(n, \frac{1}{2})$, then $\frac{X-n/2}{\sqrt{n/4}} \sim N(0,1)$, so $\mathbb{P}\big(|X-\frac{n}{2}| > \varepsilon n\big) \leqslant e^{-2\varepsilon^2 n}$.

$\mathbb{P}\big(|X-\frac{n}{2}| > x\sqrt{n/4}\big) \sim \frac{1}{\sqrt{2\pi}} e^{-x^2}$.

Chebychev gives only $\mathbb{P}\big(|X-\frac{n}{2}| > \varepsilon n\big) < \frac{1}{4\varepsilon^2 n}$.

Theorem 10.1 (Chernoff, 1952). Let $I_1, \ldots, I_n$ be independent indicators with $\mathbb{P}(I_i = 1) = p_i$. Let $p = \frac{1}{n} \sum p_i$. Let $S = \sum I_i$. Then, for $h \geqslant 0$,

$$\mathbb{P}\big(S \geqslant (p+h)n\big) \leqslant e^{-nh^2/2a} \quad \text{and} \quad \mathbb{P}\big(S \leqslant (p-h)n\big) \leqslant e^{-nh^2/2b}$$

where

$$a = \max\{\alpha(1-\alpha) : p \leqslant \alpha \leqslant p+h\} \quad \text{and} \quad b = \max\{\beta(1-\beta) : p-h \leqslant \beta \leqslant p\}$$

Corollary 10.2. Let $X \sim \mathrm{Bin}(n,p)$. Then

$$\mathbb{P}\big(X \leqslant (1-\varepsilon)pn\big) \leqslant e^{\varepsilon^2 pn/2} \text{ if } p \leqslant \tfrac{1}{2}$$

$$\mathbb{P}\big(X \geqslant (1+\varepsilon)pn\big) \leqslant e^{\varepsilon^2 pn/4} \text{ if } \varepsilon \leqslant 1$$

Proof of 10.2. Let $h = \varepsilon p$. By Theorem 10.1, we obtain the first inequality. Likewise the second. □

Proof of 10.1. Let $z \in \mathbb{R}$, $z \geqslant 1$. Then

$$\begin{aligned} \mathbb{P}\big(S \geqslant \gamma n\big) &= \mathbb{P}\big(z^S \geqslant z^{\gamma n}\big) \\ &\leqslant \frac{\mathbb{E}(z^S)}{z^{\gamma n}} \quad \text{by Markov's inequality} \\ &= z^{-\gamma n} \prod_{i=1}^{n} \big((1-p_i) + zp_i\big) \\ &\leqslant z^{-\gamma n}\big((1-p) + zp\big)^n \quad \text{by AM-GM} \\ &= \big((1-p)z^{-\gamma} + pz^{(1-\gamma)}\big)^n \\ &= e^{-Hn} \end{aligned}$$

Put $z = \dfrac{\gamma(1-p)}{p(1-\gamma)}$. Note $\gamma \geqslant p$, so $z \geqslant 1$. Write $\gamma = p + h$.

$H = (p+h)\log\big(1 + \frac{h}{p}\big) + (q-h)\log\big(1 - \frac{h}{q}\big)$, where $q = 1-p$.

By Taylor's theorem, $H = \dfrac{h^2}{2\alpha(1-\alpha)}$, for some α with $p \leqslant \alpha \leqslant p+h$.

The second inequality comes from the complementary $\mathbb{P}\big(\overline{S} \leqslant (q+h)n\big)$. □

Lecture 19

Theorem 10.3. Let X be hypergeometrically distributed with parameters (n, N, pN). Then $X = \sum_{i=1}^n I_i$ for some independent indicators I_i with some probabilities p_i, such that $\frac{1}{n}\sum_{i=1}^n p_i = p$.

In particular, the bounds in Theorem 10.1 and Corollary 10.2 hold for X.

Proof. $\mathbb{P}\{X = k\}$ is the probability that, if n balls are chosen without replacement from an urn with N balls, pN of which are red, then k balls are red. Write $R = pN$.

So $\mathbb{P}\{X = k\} = \dbinom{R}{k}\dbinom{N-R}{n-k}\dbinom{N}{n}^{-1}$, whence

$$
\begin{aligned}
\mathbb{E}\binom{X}{\ell} &= \sum_k \binom{R}{k}\binom{N-R}{n-k}\binom{N}{n}^{-1}\binom{k}{\ell} \\
&= \binom{R}{\ell}\binom{N}{n}^{-1}\sum_k \binom{R-\ell}{k-\ell}\binom{N-R}{n-k} \\
&= \binom{R}{\ell}\binom{N}{n}^{-1}\binom{N-\ell}{n-\ell} \\
&= \binom{R}{\ell}\binom{n}{\ell}\binom{N}{\ell}^{-1}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\mathbb{E}(z^X) &= \mathbb{E}\sum_\ell \binom{X}{\ell}(z-1)^\ell \\
&= \sum_\ell \binom{n}{\ell}\binom{R}{\ell}\binom{N}{\ell}^{-1}(z-1)^\ell
\end{aligned}
$$

Noting that $\binom{R}{\ell}\binom{N}{\ell}^{-1} \leqslant p^\ell$, we see that

$$
\mathbb{E}(z^X) \leqslant \sum_\ell \binom{n}{\ell} p^\ell (z-1)^\ell = [q+pz]^n, \quad \text{for } z \geqslant 1,
$$

from which we could jump in to the proof of Theorem 10.1. To get Theorem 10.3 in all its fullness,

$$
\begin{aligned}
\mathbb{E}(z^X) &= \frac{R!}{N!}\sum_\ell \binom{n}{\ell}(N-\ell)(N-\ell-1)\cdots(R-\ell+1)(z-1)^\ell \\
&= \frac{R!}{N!}\frac{1}{x^R}\frac{d^{N-R}}{dx^{N-R}}\sum_\ell \binom{n}{\ell}x^{N-\ell}, \quad \text{where } x = \frac{1}{z-1} \\
&= \frac{R!}{N!}\frac{1}{x^R}\frac{d^{N-R}}{dx^{N-R}}x^N[1+x^{-1}]^n
\end{aligned}
$$

which has real roots. That is, $\mathbb{E}(z^X)$, which is a polynomial of degree n, has real roots.

Hence $\mathbb{E}(z^X) = \prod_{i=1}^n(\alpha_i + \beta_i z)$ with $\alpha_i, \beta_i \in \mathbb{R}$. Since $\mathbb{E}(1^X) = 1$, $\alpha_i + \beta_i \neq 0$ for all i, we may write $\mathbb{E}(z^X) = c\prod_{i=1}^n(q_i + p_i z)$ with $q_i + p_i = 1$. Again, $\mathbb{E}(1^X) = 1$ means $c = 1$.

Moreover, $q_i, p_i \geqslant 0$ or else there exists $z > 0$ with $\mathbb{E}(z^X) = 0$, contradicting $\mathbb{E}(z^X) > 0$ for all $z > 0$.

Therefore $\mathbb{E}(z^X) = \prod_{i=1}^n(q_i+p_i z)$ with $q_i+p_i = 1$ and $q_i, p_i \geqslant 0$. But then $X = I_1+\ldots+I_n$, where there I_i are independent indicators, $\mathbb{P}(I_i = 1) = p_i$.

Also, $\sum p_i = \mathbb{E}X = pn$. □

11. Martingales Inequalities

We would like bounds akin to Theorem 10.1 for variables not the sum of indicators. There are nowadays several inequalities available (Janson, Hoeffding, Talagrand) that give different information under different conditions. We look at Hoeffding.

A **filter** on a space Ω is a sequence $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n$ of successively finer partitions of Ω.

A sequence of random variables $X_0, X_1, \ldots, X_n$ (with X_i defined on $(\Omega, \mathcal{F}_i)$, meaning X_i constant on parts of $\mathcal{F}_i$) is a **martingale** if $\mathbb{E}(X_{i+1} \mid \mathcal{F}_i) = X_i$. Typically X is some random variable and $X_i = \mathbb{E}(X \mid \mathcal{F}_i)$.

The classical example is that of a gambler in a fair (zero expectation) game. If $\mathcal{F}_i$ is the game and X_i = current winnings, then $\mathbb{E}(X_{i+1} \mid \mathcal{F}_i) = X_i$.

Let $f(G)$ be a graph-theoretic function. Label the edges of $[n]^{(2)}$ as $1, 2, \ldots, N = \binom{n}{2}$. Insert the edges at random with probability p one by one to obtain a random graph $G \in \mathcal{G}(n, p)$.

Let $X_i = \mathbb{E}\big(f(G) \mid \text{the first } i \text{ edges are determined}\big)$.

Then $X_0 = \mathbb{E}(f(G))$, $X_N = f(G)$ and $(X_i)_{i=0}^N$ is a martingale. This is called the **edge exposure** martingale. (This is mildly ambiguous – depends on the edge labelling.)

We could instead, at stage i, add all edges between vertices $\{1, \ldots, i-1\}$ and vertex i. $X_i = \mathbb{E}\big(f(G) \mid \text{all edges inside vertices } \{1, \ldots, i\} \text{ are determined}\big)$. This is the **vertex exposure** martingale and is a subsequence of (some) edge martingale.

Lecture 20 Given a martingale, its **difference sequence** is $Y_i = X_i - X_{i-1}$. Thus $\mathbb{E}(Y_i \mid \mathcal{F}_{i-1}) = 0$.

We aim to show that X is concentrated near its mean if the differences $|X_i - X_{i-1}|$ are bounded (e.g., when $X = \chi(G)$).

Lemma 11.1. Let Y be a random variable with $-r \leqslant Y \leqslant 1 - r$, where $r \in \mathbb{R}^+$ and $\mathbb{E}Y = 0$. Then, for $s \geqslant 0$, we have

$$\mathbb{E}(e^{sY}) \leqslant (1-r)e^{-sr} + r^{s(1-r)} \leqslant e^{s^2/8}$$

Proof. Since e^{sy} is a convex function of y, we have $e^{sy} \leqslant (1-r-y)e^{-sr} + (y+r)e^{s(1-r)}$. Take expectations.

For the second inequality, apply Taylor's theorem to $f(s) = \log\big(e^{-sr}(1 - r + re^s)\big)$. □

Theorem 11.2 (Hoeffding, 1963). Let $Y_1, \ldots, Y_n$ be a martingale difference sequence ($X_i = \mathbb{E}(X \mid \mathcal{F}_i)$) with $-r_i \leqslant Y_i \leqslant 1 - r_i$, $1 \leqslant i \leqslant n$. Let $r = \frac{1}{n}\sum_{i=1}^n r_i$, with $0 \leqslant r \leqslant 1$.

Then $\mathbb{P}\big(X \geqslant \mathbb{E}X + hn\big) = \mathbb{P}\big(\sum_{i=1}^n Y_i \geqslant hn\big) \leqslant e^{-nh^2/2a}$

and $\mathbb{P}\big(X \leqslant \mathbb{E}X - hn\big) = \mathbb{P}\big(\sum_{i=1}^n Y_i \leqslant -hn\big) \leqslant e^{-nh^2/2b}$

where $a = \max\{x(1-x) : r \leqslant x \leqslant r + h\}$ and $b = \max\{x(1-x) : r - h \leqslant x \leqslant r\}$.

For the equalities, recall: $\mathcal{F}_0$ = trivial, $X_0 = \mathbb{E}X$; $\mathcal{F}_n$ = complete partition, $X_n = X$. So $X - \mathbb{E}X = X_n - X_0 = \sum Y_i$.

Proof. We proceed in a similar way to Theorem 10.1. Thus, for $s \geqslant 0$,

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} Y_i \geqslant hn\right) &= \mathbb{P}\left(e^{s\sum_{i=1}^{n} Y_i} \geqslant e^{shn}\right) \\
&\leqslant \mathbb{E}\left(e^{s\sum_{i=1}^{n} Y_i} e^{-shn}\right) \\
&= e^{-shn}\mathbb{E}\left(e^{s\sum_{i=1}^{n} Y_i} \mid \mathcal{F}_{n-1}\right) \\
&= e^{-shn}\mathbb{E}\left(e^{s\sum_{i=1}^{n-1} Y_i}\mathbb{E}\left(e^{sY_n} \mid \mathcal{F}_{n-1}\right)\right) \quad \text{using } \mathbb{E}(A) = \mathbb{E}(\mathbb{E}(A \mid B)) \\
&\leqslant e^{-shn}\mathbb{E}\left(e^{s\sum_{i=1}^{n-1} Y_i} e^{-sr_n}\left(1 - r_n + r_n e^s\right)\right) \\
&\leqslant e^{-shn}\prod_{i=1}^{n} e^{-sr_i}\left(1 - r_i + r_i e^s\right) \quad \text{by induction} \\
&\leqslant e^{-shn} e^{-srn}\left(1 - r + re^s\right)^n \quad \text{by AM-GM} \\
&= \left((1-r)e^{-s(h+r)} + re^{s(1-h-r)}\right)^n \\
&\leqslant e^{-nh^2/2a} \quad \text{as before}
\end{aligned}$$

The second inequality comes from considering the martingale $(-X_i)$, with difference sequence $(-Y_i)$ and using the first inequality. □

We could estimate $e^{-nh^2/2a}$ as in Corollary 10.2, but the following is more useful.

Theorem 11.3. Let $X_i = \mathbb{E}(X \mid \mathcal{F}_i)$ be a martingale whose difference sequence satisfies $a_i \leqslant Y_i \leqslant a_i + c_i$, where a_i is a function on $(\Omega, \mathcal{F}_{i-1})$ and $c_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$.

Then, for $t > 0$,

$$\mathbb{P}\big(X \geqslant \mathbb{E}X + t\big) \leqslant e^{-2t^2/\sum_{i=1}^{n} c_i^2}$$
$$\mathbb{P}\big(X \leqslant \mathbb{E}X - t\big) \leqslant e^{-2t^2/\sum_{i=1}^{n} c_i^2}$$

Proof. Applying Lemma 11.1 to $Y = \mathbb{E}\left(\frac{Y_n}{c_n} \mid \mathcal{F}_{n-1}\right)$ we obtain

$$\mathbb{E}\left(e^{sY_n} \mid \mathcal{F}_{n-1}\right) = \mathbb{E}\left(e^{c_n sY}\right) \leqslant e^{s^2 c_n^2/8}$$

Thus, following the previous proof,

$$\mathbb{P}\left(X \geqslant \mathbb{E}X + t\right) = \mathbb{P}\left(\sum Y_i \geqslant t\right) \leqslant e^{-ts} e^{\frac{s^2}{8}\sum_{i=1}^{n} c_i^2}$$

Now take $s = 4t/\sum_{i=1}^{n} c_i^2$. □

We can think of the c_i as the maximum change in X_i at stage i.

A weaker form of Theorem 11.3, known as Azuma's inequality, gives the bound $e^{-t^2/2\sum c_i^2}$, based on the assumption that $|Y_i| \leqslant c_i$. But in most applications, Y_i is not symmetric about zero and so 11.3 is stronger.

A useful corollary follows.

Corollary 11.4. Let $Z_1, \ldots, Z_n$ be independent random variables with Z_i taking values in space A_i, $1 \leqslant i \leqslant n$. Suppose that $f : \prod_{i=1}^{n} A_i \to \mathbb{R}$ satisfies $|f(z) - f(z')| \leqslant c_i$ whenever z, z' differ only in the i^{th} coordinate. Let $Z = f(Z_1, \ldots, Z_n)$.

Then $\mathbb{P}\left(|Z - \mathbb{E}Z| \geqslant r\right) \leqslant 2e^{-2t^2/\sum c_i}$.

Proof. Define a filter on $\Omega = \prod_{i=1}^n A_i$ where $\mathcal{F}_i$ partitions Ω into $\prod_{j=1}^i |A_j|$ parts according to the first i coordinates. Then $X_i = \mathbb{E}(Z \mid \mathcal{F}_i)$ is a martingale where $X_0 = \mathbb{E}Z$, $X_n = Z$. The conditions on f imply that

$$\min_z X_i \leqslant X_{i-1} = \mathbb{E}(X_i) \leqslant \max_z X_i \leqslant \min_z X_i + c_i$$

where z runs over A_i. Hence (Y_i) satisfy Theorem 11.3 with $a_i = \min_z X_i - X_{i-1}$. □

Lecture 21

12. The Chromatic Number of a Random Graph

As usual, $\mathcal{G}(n,p)$ is the space of random graphs on vertex set $[n]$, with edges chosen independently at random with probability p (we take p constant).

We say that an event holds **with high probability** (whp) if $\mathbb{P}(A) \to 1$ as $n \to \infty$. (Strictly speaking, we have a sequence of events $A_n \subset \mathcal{G}(n,p)$.)

Martingale inequalities show that $\chi(G)$ is highly concentrated around its mean. But what is the mean?

Lemma 12.1. Let $0 < p < 1$ be constant. Let $G \in \mathcal{G}(n,p)$.

Then $\chi(G) \geqslant \big(1 + o(1)\big)\dfrac{n}{2\log_{1/q} n}$ whp, where $q = 1 - p$.

Proof. Let $\varepsilon > 0$. Let $d = \lceil (2+\varepsilon)\log)1/qn\rceil + 1$. Let X be the number of independent sets of size d in G. Then

$$\mathbb{P}\left(\chi(G) \leqslant \frac{n}{d}\right) \leqslant \mathbb{P}(X \geqslant 1) \leqslant \mathbb{E}X = \binom{n}{d}(1-p)^{\binom{d}{2}} \leqslant \left[nq^{(d-1)/2}\right]^d \to 0 \quad \text{as} \quad n \to \infty$$

This is true for all $\varepsilon > 0$. □

It is not hard to show that the greedy algorithm uses $\big(1 + o(1)\big)n/\log_{1/q} n$ colours.

Chebychev's bound shows that there exist independent sets of size $2\log_{1/q} n$ whp, but not with very high probability.

Theorem 12.2 (Bollobás, 1988). Let $0 < p < 1$ be constant. Let $G \in \mathcal{G}(n,p)$.

Then $\chi(G) = \big(1 + o(1)\big)\dfrac{n}{2\log_{1/q} n}$ whp, where $q = 1 - p$.

Proof. Lemma 12.1 gives the lower bound.

Let $0 < \varepsilon < \frac{1}{20}$, and let $d = \lceil (2-\varepsilon)\log_{1/q} n\rceil$. We shall find m with $n^{1-\varepsilon} < m < n^{1-\varepsilon/2}$ such that, whp, the event

$$A = \text{"every subset of } m \text{ vertices contains an independent set of size } d\text{"}$$

holds. Note that, if A holds, then $\chi(G) \leqslant \dfrac{n}{d} + m$ since we use colour 1 on d vertices, then colour 2 on d vertices, and so on until fewer than m vertices remain, at which point we finish off with m more colours. Since $m = o(n/d)$, and since the statement holds for all $\varepsilon > 0$, the theorem follows.

Let $P_m = \mathbb{P}\big(H \in \mathcal{G}(n,p)$ has no independent d-set$\big)$. Then the probability of A failing is at most

$$\binom{n}{m} P_m \leqslant n^m P_m = e^{m \log n} P_m \leqslant e^{\frac{1}{1+\varepsilon} m \log m} P_m$$

So it suffices to show that $P_m < e^{-m^{1+\delta}}$ for some constant $\delta > 0$.

Let X be the number of independent d-sets in H. Then $\mathbb{E}X = \binom{m}{d} q^{\binom{d}{2}}$.

If $m < n^{1-\varepsilon}$ then $\mathbb{E}X \leqslant \left(\frac{em}{d} \frac{q^{2/d}}{\sqrt{q}}\right)^d < 1$.

If $m > n^{1-\varepsilon/2}$ then $\mathbb{E}X \geqslant \left(\frac{m}{d} \frac{q^{q/2}}{\sqrt{d}}\right)^d \geqslant n^{d\varepsilon/8} \geqslant n^2$.

If n is large. Clearly $\mathbb{E}X$ increases as m increases, and moving from $m-1$ to m the value increases by a factor $\frac{m}{m-d} < 2$ for large n. It follows that there is some value of m for which $2m^{7/4} \leqslant \mathbb{E}X \leqslant 4m^{7/4}$.

Define the random variables

$$\begin{aligned} Y &= \text{max number of independent } d\text{-sets with } \leqslant 1 \text{ vertex in common.} \\ Z &= \text{number of independent } d\text{-sets meeting every other in } \leqslant 1 \text{ vertex.} \end{aligned}$$

Clearly $Z \leqslant Y \leqslant X$. Consider H as an element of the product space $\{0,1\}^{\binom{m}{2}}$. The value of Y changes by at most 1 if we change one edge. So we can apply Corollary 11.4 to Y with $c_i = 1$, $1 \leqslant i \leqslant \binom{m}{2}$, so $\mathbb{P}(Y \leqslant \mathbb{E}Y - t) \leqslant e^{-2t^2/\binom{m}{2}}$.

If we can show that $\mathbb{E}Y \geqslant m^{7/4}$ then we are done, because then

$$P_m = \mathbb{P}(X=0) \leqslant \mathbb{P}(Y=0) \leqslant \mathbb{P}(Y \leqslant \mathbb{E}Y - m^{7/4}) \leqslant e^{-2m^{7/2}/\binom{m}{2}} \leqslant e^{-4m^{3/2}}$$

Lecture 22

Now $\mathbb{E}Y \geqslant \mathbb{E}Z$, so it's enough to show that $\mathbb{E}Z \geqslant m^{7/4}$. For $2 \leqslant r \leqslant d-1$, let

$$W_r = \text{number of independent } d\text{-sets meeting another in } r \text{ vertices}$$

Then $X \leqslant Z + W_2 + W_3 + \ldots + W_{d-1}$. So $\mathbb{E}Z \geqslant \mathbb{E}X - \sum_{r=2}^{d-1} \mathbb{E}W_r$. Then

$$\mathbb{E}W_r \leqslant \binom{m}{d} q^{\binom{d}{2}} \binom{d}{r} \binom{m-d}{d-r} q^{\binom{d}{2}-\binom{r}{2}} = \mathbb{E}X f_r$$

where $f_r = \binom{d}{r}\binom{m-d}{d-r} q^{\binom{d}{2}-\binom{r}{2}}$.

So it suffices to show that $\sum_{r=2}^{d-1} f_r \leqslant \frac{1}{2}$, since $\mathbb{E}X \geqslant 2m^{7/4}$.

First,

$$\begin{aligned} f_r &\leqslant \binom{d}{r}^2 \binom{m}{d} \binom{m}{r}^{-1} q^{\binom{d}{2}-\binom{r}{2}} = \mathbb{E}X \binom{d}{r}^2 \binom{m}{r}^{-1} q^{-\binom{r}{2}} \\ &\leqslant \mathbb{E}X \left(\frac{d^2\sqrt{q}}{mq^{r/2}}\right)^r \leqslant \mathbb{E}X m^{-9r/10} \end{aligned}$$

for $r \leqslant d/12$. Since $\mathbb{E}X \leqslant m^{7/4}$, we have $\sum_{r=2}^{d/12} f_r \leqslant 1/4$.

Secondly, writing $s = d - r$, we have

$$f_r = \binom{d}{s}\binom{m-d}{s} q^{sd-\binom{s+1}{2}} \leqslant \left(dmq^{d-\frac{s+1}{2}}\right)^s \leqslant \left(dmq^{-1/2}q^{13d/24}\right)^s$$

for $s \leqslant 11d/12$, i.e. for $r \geqslant d/12$.

Now $\varepsilon < \frac{1}{20}$, so $\frac{13d}{24} \geqslant \frac{13}{12}(1-\varepsilon)\log n \geqslant \left(1+\frac{1}{40}\right)\log n \geqslant \left(1+\frac{1}{40}\right)\log m$.

So $\sum_{r=d/12}^{d-1} f_r \leqslant dm^{-1/50} \leqslant 1/4$. □

13. The Semi-Random Method

In 1980, Ajtai-Komlós-Szemerédi proved that a triangle-free graph of order n and average degree d has an independent set of size $\frac{n}{d}\log n$ (as opposed to $\geqslant \frac{n}{d}$ by Turán).

They achieved this by choosing a small random subset of size $\frac{1}{2}|X|$, such that the conditions in $G - X - \Gamma(X)$ could be described, not much worse than in G. Repeat. This is better than

(a) removing randomly one-by-one

(b) removing a large set once

The method was used by Rödl (then Frankl-Rödl) to prove the Erdős-Hanani conjecture.

We use Chebychev's inequality:

$$\mathbb{P}\big(|X - \mathbb{E}X| \geqslant t\big) = \mathbb{P}\big((X-\mathbb{E}X)^2 \geqslant t^2\big) \leqslant \frac{\mathbb{E}\big((X-\mathbb{E}X)^2\big)}{t^2} = \frac{\operatorname{Var} X}{t^2}$$

Note that $\mathbb{E}\big((X-\mathbb{E}X)^2\big) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.

Note that if $X = \sum I_\alpha$, then $\mathbb{E}X^2 = \mathbb{E}\sum_{\alpha,\beta} I_\alpha I_\beta = \sum_{\alpha,\beta} \mathbb{P}(I_\alpha \text{ and } I_\beta)$.

Call an r-uniform hypergraph δ**-semiregular** (of degree d) if

$$(1-\delta)d \leqslant d(v) \leqslant (1+\delta)d \ \text{ for all but } \delta|G| \text{ vertices}$$

$$d(v) \leqslant e^r d \ \text{ for all } v$$

$$d(u,v) \leqslant \delta d \ \text{ for all } u, v \text{ with } u \neq v$$

where $d(u,v)$ is the **codegree**, the number of edges containing $\{u,v\}$.

Lemma 13.1. For all $r \geqslant 2$ and $0 < \varepsilon < 1$, there exists δ_0 such that, if $\delta < \delta_0$ and G is r-uniform δ-semiregular, then G has a subhypergraph G' of order at most $(1-\varepsilon+\varepsilon^2)|G|$ that is $\delta^{1/4}$-semiregular, and $V(G) - V(G')$ is covered by at most $(\varepsilon+\varepsilon^2)|G|/r$ edges.

Lecture 23 **Proof.** Let G be δ-semiregular of degree d. Call the vertices not satisfying $(1-\delta)d \leqslant d(v) \leqslant (1+\delta)d$ **bad**. Let B be the set of bad vertices. Choose a set X of edges independently at random with probability ε/d. Let G' be the subgraph on vertex set $V(G) - V(X) - B$.

We show that, with positive probability, G' is $\delta^{1/4}$-semiregular and $|X|+|B| \leqslant (\varepsilon+\varepsilon^2)n/r$, where $n = |G|$. This will prove the lemma, because our edge cover for $V(G) - V(G')$ ... *(something I missed, sorry - anyone help?)*

Let $\mathcal{A}$ be the event "$|X| + |B| \leqslant (\varepsilon + \varepsilon^2)n/r$"
$\mathcal{B}$ be the event "$n/4 \leqslant |G'| \leqslant (1 - \varepsilon + \varepsilon^2)n$"
$\mathcal{C}$ be the event "G' is $\delta^{1/4}$-semiregular"

We show that $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C} \neq \emptyset$.

Note $\delta d \geqslant 1$, so we can make d, and hence also n, as large as we like, by making δ_0 small.

(a) $e(G) \leqslant \frac{1}{r}\big((1-\delta)n(1+\delta)d + \delta n e^r d\big) \leqslant \big(1 + \frac{\varepsilon}{2}\big)\frac{nd}{r}$ is δ_0 is small.
By Corollary 10.2,

$$\mathbb{P}\left(|X| \geqslant \mathbb{E}X + \frac{\varepsilon^2}{4}\frac{n}{r}\right) \;\leqslant\; \mathbb{P}\left(|X| \geqslant \left(\varepsilon + \frac{3\varepsilon^2}{4}\right)\frac{n}{r}\right) \;\leqslant\; e^{-\frac{\varepsilon^2}{16}\frac{n\varepsilon}{4r}} \;<\; \frac{1}{3}$$

if δ_0 is small. Hence $\mathbb{P}(\mathcal{A}) \geqslant 2/3$. ($|B| \leqslant \delta n \leqslant \varepsilon^2 n/4r$.)

(b) Let v be a good (i.e, not bad) vertex, and I_v be the event $v \in V(G')$.
So $|G'| = \sum_v I_v$. Now

$$\frac{1}{3} \;<\; \left(1 - \frac{\varepsilon}{d}\right)^{(1+\delta)d} \;\leqslant\; \mathbb{P}(v \in G') \;=\; \left(1 - \frac{\varepsilon}{d}\right)^{d(v)}$$
$$\leqslant\; \left(1 - \frac{\varepsilon}{d}\right)^{(1-\delta)d} \;\leqslant\; e^{-\varepsilon(1-\delta)} \;\leqslant\; 1 - \varepsilon + \frac{2}{3}\varepsilon^2$$

So by Chebychev, $\mathbb{P}(\overline{\mathcal{B}}) \leqslant \dfrac{\operatorname{Var}|G'|}{(\varepsilon^2 n/12)^2}$, since $\mathbb{E}|G'| \leqslant \big(1 - \varepsilon + \frac{2}{3}\varepsilon^2\big)n$.
Let $u, v \in V(G) - B$, $u \neq v$. Then

$$\mathbb{P}(I_u \text{ and } I_v) \;=\; \left(1 - \frac{\varepsilon}{d}\right)^{d(v)+d(v)-d(u,v)} \;\leqslant\; \left(1 - \frac{\varepsilon}{d}\right)^{-\delta d}\mathbb{P}(I_u)\mathbb{P}(I_v)$$

So

$$\mathbb{E}(|G'|^2) \;=\; \sum_{u,v}\mathbb{P}(I_u \text{ and } I_v) \;=\; \sum_{u=v} + \sum_{u\neq v} \;\leqslant\; \mathbb{E}(|G'|) + \left(1 - \frac{\varepsilon}{d}\right)^2 \mathbb{E}(|G'|)^2$$

since $\left(1 - \dfrac{\varepsilon}{d}\right)^{-\delta d} \leqslant e^{\varepsilon\delta} \leqslant 1 + 2\delta\varepsilon$. So

$$\operatorname{Var}(|G'|) \;=\; \mathbb{E}|G'|^2 - (\mathbb{E}|G'|)^2 \;\leqslant\; \mathbb{E}|G'| + 2\delta\varepsilon\mathbb{E}|G'|^2 \;\leqslant\; n + 2\delta\varepsilon n^2$$

Hence $\mathbb{P}(\overline{\mathcal{B}}) \;\leqslant\; \dfrac{n + 2\delta\varepsilon n^2}{(\varepsilon^2 n/12)^2} \;<\; \dfrac{1}{3}$ if δ_0 is small.

(c) See handout.

Theorem 13.2 (Frankl-Rödl, 1985). Given $r, \varepsilon > 0$, there exists $\delta > 0$ such that if G is d-regular r-uniform of maximum codegree $< \delta d$, then G can be covered by $(1+\varepsilon)|G|/r$ edges.

Remark. Hence there is a set of $(1 - r\varepsilon)|G|/r$ independent edges, i.e. pairwise disjoint. For let W be the set of vertices in more than one cover edge. Then $|G| + |W| \leqslant (1+\varepsilon)|G|$, so $|W| \leqslant \varepsilon|G|$. Remove edges meeting $|W|$.

Proof. Choose $\eta > 0$ and $k \in \mathbb{N}$ so that $\dfrac{1+\eta}{1-\eta} + r(1-\eta+\eta^2) < 1+\varepsilon$.

Then $\delta < \delta_0(r,\eta)^{4^k}$ will work. For Lemma 13.1 can be applied k times with r and η. The edges selected amount to at most

$$(\eta+\eta^2)\Big(1 + (1-\eta+\eta^2) + (1-\eta+\eta^2)^2 + \cdots + (1-\eta+\eta^2)^{k-1}\Big)|G| \leqslant \frac{1+\eta}{1-\eta}\frac{|G|}{r}$$

and there remain $\leqslant (1-\eta+\eta^2)^k|G|$ vertices, which can be covered by an edge apiece. □

Define:

$m(n,k,\ell)$ = maximum number of k-sets in $[n]$ such that each ℓ-set is in at most one of them.

$M(n,k,\ell)$ = minimum number of k-sets in $[n]$ such that each ℓ-set is in at least one of them.

Erdős-Hanani (1963) conjectured

$$m(n,k,\ell) \sim \frac{\binom{n}{\ell}}{\binom{k}{\ell}} \sim M(n,k,\ell)$$