Lecture 3

For general graphs H, we are interested in the function $ex(n, H)$, defined as follows.

$$ex(n, H) = \max\{e(G) : |G| = n, H \not\subset G\}.$$

Turán's theorem itself tells us that

$$ex(n, K_{r+1}) \leq \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

We are now going to deal with the general case. We will show that the behaviour of the extremal function $ex(n, H)$ is tied intimately to the chromatic number of the graph H.

Definition 1 *The chromatic number $\chi(H)$ of a graph H is the smallest natural number c such that the vertices of H can be coloured with c colours and no two vertices of the same colour are adjacent.*

The fundamental result which we shall prove, known as the Erdős-Stone-Simonovits theorem, is the following.

Theorem 1 (Erdős-Stone-Simonovits) *For any fixed graph H and any fixed $\epsilon > 0$, there is n_0 such that, for any $n \geq n_0$,*

$$\left(1 - \frac{1}{\chi(H) - 1} - \epsilon\right)\frac{n^2}{2} \leq ex(n, H) \leq \left(1 - \frac{1}{\chi(H) - 1} + \epsilon\right)\frac{n^2}{2}.$$

For the complete graph K_{r+1}, the chromatic number is $r+1$, so in this case the Erdős-Stone-Simonovits theorem reduces to an approximate version of Turán's theorem. For bipartite H, it gives $ex(n, H) \leq \epsilon n^2$ for all $\epsilon > 0$. This is an important theme, one we will return to later in the course.

To prove the Erdős-Stone-Simonovits theorem, we will first prove the following lemma, which already contains most of the content.

Lemma 1 *For any natural numbers r and t and any positive ϵ with $\epsilon < 1/r$, there exists an n_0 such that the following holds. Any graph G with $n \geq n_0$ vertices and $\left(1 - \frac{1}{r} + \epsilon\right)\frac{n^2}{2}$ edges contains $r + 1$ disjoint sets of vertices $A_1, \ldots, A_{r+1}$ of size t such that the graph between A_i and A_j, for every $1 \leq i < j \leq r + 1$, is complete.*

Proof To begin, we find a subgraph G' of G such that every degree is at least $\left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)|V(G')|$. To find such a graph, we remove one vertex at a time. If, in this process, we reach a graph with ℓ vertices and there is some vertex which has fewer than $\left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\ell$ neighbors, we remove it.

Suppose that this process terminates when we have reached a graph G' with n' vertices. To show that n' is not too small, consider the total number of edges that have been removed from the graph. When the graph has ℓ vertices, we remove at most $\left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\ell$ edges. Therefore, the total number of edges removed is at most

$$\sum_{\ell=n'+1}^{n}\left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\ell = \left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\frac{(n - n')(n + n' + 1)}{2} \leq \left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\frac{(n^2 - n'^2)}{2} + \frac{(n - n')}{2}.$$

Also, since G' has at most $\frac{n'^2}{2}$ edges, we have

$$|e(G)| \leq \left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\frac{(n^2 - n'^2)}{2} + \frac{(n-n')}{2} + \frac{n'^2}{2} = \left(1 - \frac{1}{r} + \frac{\epsilon}{2}\right)\frac{n^2}{2} + \left(\frac{1}{r} - \frac{\epsilon}{2}\right)\frac{n'^2}{2} + \frac{(n-n')}{2}.$$

But we also have $|e(G)| \geq \left(1 - \frac{1}{r} + \epsilon\right)\frac{n^2}{2}$. Therefore, the process stops once

$$\left(\frac{1}{r} - \frac{\epsilon}{2}\right)\frac{n'^2}{2} - \frac{n'}{2} < \epsilon\frac{n^2}{4} - \frac{n}{2},$$

that is, when $n' \approx \sqrt{\epsilon r}n$. From now on, we will assume that we are working within this large well-behaved subgraph G'.

We will show, by induction on r, that there are $r+1$ sets $A_1, A_2, \ldots, A_{r+1}$ of size t such that every edge between A_i and A_j, with $1 \leq i < j \leq r+1$, is in G'. For $r = 0$, there is nothing to prove.

Given $r > 0$ and $s = \lceil 3t/\epsilon \rceil$, we apply the induction hypothesis to find r disjoint sets $B_1, B_2, \ldots, B_r$ of size s such that the graph between every two disjoint sets is complete. Let $U = V(G') \backslash \{B_1 \cup \cdots \cup B_r\}$ and let W be the set of vertices in U which are adjacent to at least t vertices in each B_i.

We are going to estimate the number of edges missing between U and $B_1 \cup \cdots \cup B_r$. Since every vertex in $U \backslash W$ is adjacent to fewer than t vertices in some B_i, we have that the number of missing edges is at least

$$\tilde{m} \geq |U \backslash W|(s-t) \geq (n' - rs - |W|)\left(1 - \frac{\epsilon}{3}\right)s.$$

On the other hand, every vertex in G' has at most $\left(\frac{1}{r} - \frac{\epsilon}{2}\right)n'$ missing edges. Therefore, counting over all vertices in $B_1 \cup \cdots \cup B_r$, we have

$$\tilde{m} \leq rs\left(\frac{1}{r} - \frac{\epsilon}{2}\right)n' = \left(1 - \frac{r\epsilon}{2}\right)sn'.$$

Therefore,

$$|W|\left(1 - \frac{\epsilon}{3}\right)s \geq (n' - rs)\left(1 - \frac{\epsilon}{3}\right)s - \left(1 - \frac{r\epsilon}{2}\right)sn' = \epsilon\left(\frac{r}{2} - \frac{1}{3}\right)sn' - \left(1 - \frac{\epsilon}{3}\right)rs^2.$$

Since ϵ, r and s are constants, we can make $|W|$ large by making n' large. In particular, we may make $|W|$ such that

$$|W| > \binom{s}{t}^r (t-1).$$

Every element in W has at least t neighbours in each B_i. There are at most $\binom{s}{t}^r$ ways to choose a t-element subset from each of $B_1 \cup \cdots \cup B_r$. By the pigeonhole principle and the size of $|W|$, there must therefore be some subsets $A_1, \ldots, A_r$ and a set A_{r+1} of size t from W such that every vertex in A_{r+1} is connected to every vertex in $A_1 \cup \cdots \cup A_r$. Since $A_1, \ldots, A_r$ are already joined in the appropriate manner, this completes the proof. □

Note that a careful analysis of the proof shows that one may take $t = c(r, \epsilon)\log n$. It turns out that this is also best possible (see example sheet).

It remains to prove the Erdős-Stone-Simonovits theorem itself.

Proof of Erdős-Stone-Simonovits For the lower bound, we consider the Turán graph given by $r = \chi(H) - 1$ sets of almost equal size $\lfloor n/r \rfloor$ and $\lceil n/r \rceil$. This has roughly the required number of vertices and it is clear that every subgraph of this graph has chromatic number at most $\chi(H) - 1$.

For the upper bound, note that if H has chromatic number $\chi(H)$, then, provided t is large enough, it can be embedded in a graph G consisting of $\chi(H)$ sets $A_1, A_2, \ldots, A_{\chi(H)}$ of size t such that the graph between any two disjoint A_i and A_j is complete. We simply embed any given colour class into a different A_i. The theorem now follows from an application of the previous lemma. $\square$