Lecture 18: Extremal graphs and Szemerédi's regularity lemma · 1MA020

Vilhelm Agdur[1]

[1] vilhelm.agdur@math.uu.se

11 December 2023

We discuss some basic notions of extremal graph theory, and give some nice proofs of Turán's theorem and related ideas. Then we introduce the Székmeredi regularity lemma, and use it to prove the triangle removal lemma, which we can use to prove Roth's theorem on arithmetic progressions of length three.

As usual, we start by repeating a definition from the exercise session.

Definition 1. Given any graph H, we say that a graph G is *H-free* if it has no subgraph isomorphic to H. We say that it is *maximal H-free* if adding any edge to it would create a subgraph isomorphic to H, and we say that is is *maximum H-free* (or *extremal* among H-free graphs) if additionally no other H-free graph has more edges than G.

For each integer n, we define the *extremal function for H*, denoted $\mathrm{ex}(n; H)$, to be the number of edges of a maximum H-free graph on n vertices.

In the exercises, you were asked to investigate what an extermal r-clique-free graph might look like. The answer to this problem is a famous theorem due to Turán.[2]

[2] The special case of triangle-free graphs is due to Mantel.

Definition 2. The *Turán graph $T(n, r)$* is a complete multipartite graph divided into r parts, with n vertices divided as equally as possible between the parts.[3]

[3] Concretely, if $n = qr + s$ for natural numbers q and $s < r$, it has s parts of size $q + 1$ and $r - s$ parts of size q.

This graph has

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

edges.

Theorem 3 (Turán, 1941)**.** *It holds for every r that*

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

and in particular the Turán graphs are the extremal $(r + 1)$-clique-free graphs.

We will see two different proofs of this theorem. The first will only show the upper bound, not the extremality of the Turán graphs, but the second will prove that the Turán graphs are extremal in order to show the bound.

We start with one that uses a result we have seen before, the Caro-Wei result about independent sets.[4]

[4] Which was originally invented to be used in this proof of Turán's theorem.

Proof of Theorem 3. We begin by noting that a clique in a graph is precisely an independent set in its complement graph. Now $d_{G^c}(v) = n - 1 - d_G(v)$, so by Caro-Wei we get that

$$\omega(G) = \alpha\left(G^c\right) \geq \sum_{v \in V} \frac{1}{d_{G^c}(v) + 1} = \sum_{v \in V} \frac{1}{n - d_v}.$$

Next, we are going to use the Cauchy-Schwarz inequality, which states that[5] for any $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$, we have

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right).$$

[5] Or in the more compact linear algebra formulation,

$$|\langle a, b\rangle| \leq \|a\| \, \|b\|,$$

but it's more convenient for us to just write out what it means termwise.

So let $a_v = \sqrt{n - d_v}$ and $b_v = \frac{1}{\sqrt{n - d_v}}$, so that $a_v b_v = 1$ for all v, and Cauchy-Schwarz states that

$$\begin{aligned} n^2 &\leq \left(\sum_{v \in V} n - d_v\right)\left(\sum_{v \in V} \frac{1}{n - d_v}\right) \\ &\leq \omega(G)\left(\sum_{v \in V} n - d_v\right) \\ &= \omega(G)\left(n^2 - 2|E|\right), \end{aligned}$$

where in the first step we used our bound on $\omega(G)$ we have just shown, and in the second the fact that $\sum_v d_v = 2|E|$ that we showed near the very beginning of the course.

So if we assume that $\omega(G) \leq r$, we get that

$$n^2 \leq r\left(n^2 - 2|E|\right),$$

which if you solve for $|E|$ becomes precisely the inequality of Turán's theorem. □

The second proof we present of this theorem is due to Erdős.

Proof of Theorem 3. Our proof will proceed by showing that we can always increase the edge-count of a graph, while retaining its r-clique-freeness, in a way that makes it closer to a complete multipartite graph. This will show that the extremal graph must be complete multipartite, and thus in particular must be the Turán graphs.[6]

[6] That this is so follows from the following fact:

Exercise 1. Show that the Turán graphs are the r-partite graphs on n vertices with the most edges.

Let $G = (V, E)$ be some r-clique-free graph on n vertices, and let v be a vertex of maximum degree in G. Let $S = N(v)$ and $T = V \setminus N(v)$. We note that since G is r-clique-free and v is adjacent to every vertex of S, $G[S]$ must in fact be $(r-1)$-clique-free.

Now we construct a graph H by taking G, deleting every edge inside of T, and adding every edge between S and T. This construction is illustrated in Figure 1.

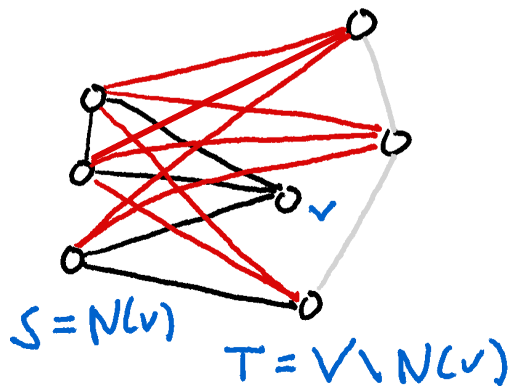

Figure 1: A figure of the construction of H from G in the Erdős proof of Theorem 3. The black edges are edges that are kept, grey edges are removed, and red edges are added. The vertex v and sets S and T are labelled in blue.

Since $G[S]$ is $(r-1)$-clique-free, adding a new neighbour to all of them cannot create an r-clique, so H is still r-clique-free. It remains to see that it has at least as many edges as G.

So consider the degree of any vertex w. If $w = v$, its degree didn't change, if $w \in S$, its degree can only have increased, and if $w \in T$ its new degree is d_v, which by assumption was maximum, so it again has not decreased. So since the degree of every vertex has increased or remained unchanged, we conclude that $|E(H)| \geq |E(G)|$.

So we conclude that the maximal r-clique-free graphs must be ones which are unchanged by this process, and a moment's thought reveals that these must be precisely the complete multipartite graphs.[7] □

[7] **Exercise 2.** Think for a moment.

Exercises