**MAT 307: Combinatorics**

# Lecture 8: Extremal combinatorics

Instructor: Jacob Fox

## 1 The Erdős-Stone theorem

We can ask more generally, what is the maximum number of edges in a graph $G$ on $n$ vertices, which does not contain a given subgraph $H$? We denote this number by $ex(n, H)$. For graphs $G$ on $n$ vertices, this question is resolved up to an additive error of $o(n^2)$ by the Erdős-Stone theorem. In order to state the theorem, we first need the notion of a *chromatic number*.

**Definition 1.** *For a graph $H$, the chromatic number $\chi(H)$ is the smallest $c$ such that the vertices of $H$ can be colored with $c$ colors with no neighboring vertices receiving the same color.*

The chromatic number is an important parameter of a graph. The graphs of chromatic number at most 2 are exactly bipartite graphs. In contrast, graphs of chromatic number 3 are already hard to decribe and hard to recognize algorithmically. Let us also mention the famous *Four Color Theorem* which states that any graph that can be drawn in the plane without crossing edges has chromatic number at most 4.

The chromatic number of $H$ turns out to be closely related to the question of how many edges are necessary for $H$ to appear as a subgraph.

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

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

In particular, for *bipartite graphs $H$*, which can be colored with 2 colors, we get that $ex(n, H) \leq \epsilon n^2$ for any $\epsilon > 0$ and sufficiently large $n$, so the theorem only says that the extremal number is very small compared to $n^2$. We denote this by $ex(H, n) = o(n^2)$. For graphs $H$ of chromatic number 3, we get $ex(n, H) = \frac{1}{4}n^2 + o(n^2)$, etc. Note that this also matches the bound we obtained for $H = K_{t+1}$ ($\chi(K_{t+1}) = t + 1$), where we got the exact answer $ex(n, K_{t+1}) = \frac{1}{2}(1 - \frac{1}{t})n^2$.

First, we prove the following technical lemma.

**Lemma 1.** *Fix $k \geq 1$, $0 < \epsilon < 1/k$ and $t \geq 1$. Then there is $n_0(k, \epsilon, t)$ such that any graph $G$ with $n \geq n_0(k, \epsilon, t)$ vertices and $m \geq \frac{1}{2}(1 - 1/k + \epsilon)n^2$ edges contains $k + 1$ disjoint sets of vertices $A_1, A_2, \ldots, A_{k+1}$ of size $t$, such that any two vertices in different sets $A_i, A_j$ are joined by an edge.*

*Proof.* First, we will find a subgraph $G' \subset G$ where *all degrees* are at least $(1 - 1/k + \epsilon/2)|V(G')|$. The procedure to find such a subgraph is very simple: as long as there is a vertex of degree smaller than $(1 - 1/k + \epsilon/2)|V(G)|$, remove the vertex from the graph. We just have to prove that this procedure terminates before the graph becomes too small.

Suppose that the procedure stops when the graph has $n_0$ vertices (potentially $n_0 = 0$, but we will prove that this is impossible). Let's count the total number of edges that we have removed from the graph. At the point when $G$ has $\ell$ vertices, we remove at most $(1 - 1/k + \epsilon/2)\ell$ edges. Therefore, the total number of removed edges is at most

$$\sum_{\ell=n_0+1}^{n} \left(1 - \frac{1}{k} + \frac{\epsilon}{2}\right) \ell = \left(1 - \frac{1}{k} + \frac{\epsilon}{2}\right)(n - n_0)\frac{n + n_0 + 1}{2} \leq \left(1 - \frac{1}{k} + \frac{\epsilon}{2}\right)\frac{n^2 - n_0^2}{2} + \frac{n - n_0}{2}.$$

At the end, $G$ has at most $\frac{1}{2}n_0^2$ edges. Therefore, the number of edges in the original graph must have been

$$|E(G)| \leq \left(1 - \frac{1}{k} + \frac{\epsilon}{2}\right)\frac{n^2 - n_0^2}{2} + \frac{n - n_0}{2} + \frac{1}{2}n_0^2 = \left(1 - \frac{1}{k} + \frac{\epsilon}{2}\right)\frac{n^2}{2} + \left(\frac{1}{k} - \frac{\epsilon}{2}\right)\frac{n_0^2}{2} + \frac{n - n_0}{2}.$$

On the other hand, we assumed that $|E(G)| \geq \left(1 - \frac{1}{k} + \epsilon\right)\frac{n^2}{2}$. Combining these two inequalities, we obtain that

$$\left(\frac{1}{k} - \frac{\epsilon}{2}\right)\frac{n_0^2}{2} - \frac{n_0}{2} \geq \frac{\epsilon n^2}{4} - \frac{n}{2}.$$

Thus if we want to get $n_0$ large enough, it's sufficient to choose $n$ appropriately larger (roughly $n \simeq n_0/\sqrt{\epsilon k}$).

From now on, we can assume that all degrees in $G$ are at least $(1 - 1/k + \epsilon/2)n$. We prove by induction on $k$ that there are $k + 1$ sets of size $t$ such that we have all edges between vertices in different sets. For $k = 0$, there is nothing to prove.

Let $k \geq 2$ and $s = \lceil t/\epsilon \rceil$. By the induction hypothesis, we can find $k$ disjoint sets of size $s$, $A_1, \ldots, A_k$ such that any two vertices in two different sets are joined by an edge. Let $U = V \setminus (A_1 \cup \ldots \cup A_k)$ and let $W$ denote the set of vertices in $U$, adjacent to at least $t$ points in each $A_i$. Let us count the edges *missing* between $U$ and $A_1 \cup \ldots \cup A_k$. Since every vertex in $U \setminus W$ is adjacent to less than $t$ vertices in some $A_i$, the number of missing edges is at least

$$\tilde{m} \geq |U \setminus W|(s - t) \geq (n - ks - |W|)(1 - \epsilon)s.$$

On the other hand, any vertex in the graph has at most $(1/k - \epsilon/2)n$ missing edges, so counting over $A_1 \cup \ldots \cup A_k$, we get

$$\tilde{m} \leq ks(1/k - \epsilon/2)n = (1 - k\epsilon/2)sn.$$

From these inequalities, we deduce

$$\begin{aligned} |W|(1-\epsilon)s &\geq (n - ks)(1 - \epsilon)s - (1 - k\epsilon/2)sn \\ &= \epsilon(k/2 - 1)sn - (1 - \epsilon)ks^2. \end{aligned}$$

Everything else being constant, we can make $n$ large enough so that $|W|$ is arbitrarily large. In particular, we make sure that

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

We know that each vertex $w \in W$ is adjacent to at least $t$ points in each $A_i$. Select $t$ specific points from each $A_i$ and denote the union of all these $kt$ points $T_w$. We have $\binom{s}{t}^k$ possible sets $T_w$; by the pigeonhole principle, at least one of them is chosen for at least $t$ vertices $w \in W$. We define these $t$ vertices to constitute our new set $A_{k+1}$, and the respective $t$-tuples of vertices connected to it $A_i' \subset A_i$. The collection of sets $A_1, \ldots, A_{k+1}$ satisfies the property that all pairs of vertices from different sets form edges. $\square$

Now we are ready to prove the Erdős-Stone theorem.

*Proof.* Let $\chi(H) = k+1$. The Turán graph $T_{n,k}$ has chromatic number $k$, hence it cannot contain $H$. This proves $ex(n,H) \geq \frac{1}{2}(1-1/k)n^2$ whenever $n$ is a multiple of $k$. Therefore,

$$ex(n,H) \geq \frac{1}{2}\left(1-\frac{1}{k}\right)(n-k)^2 \geq \frac{1}{2}\left(1-\frac{1}{k}-\frac{2k}{n}\right)n^2.$$

On the other hand, fix $t = |V(H)|$ and consider a graph $G$ with $n$ vertices and $m \geq (1-1/k+\epsilon)\frac{n^2}{2}$ edges. If $n$ is large enough, then by Lemma 1, $G$ contains sets $A_1, \ldots, A_{k+1}$ of size $t$ such that all edges between different sets are present. $H$ is a graph of chromatic number $k+1$ and therefore its vertices can be embedded in $A_1, \ldots, A_{k+1}$ based on their color. We conclude that $H$ is a subgraph of $G$ and hence

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

for any $\epsilon > 0$ and sufficiently large $n$. □