Lecture 10: Extremal graph theory I

October 8, 2020

What is extremal combinatorics

Extremal combinatorics deals with questions of the following form: given a graph G on n vertices, how big certain global invariants can be when G does not exhibit certain local substructures? What are the graphs without such local substructures that have maximal global invariants?

The questions we will discuss: how many edges can G have if G does not have certain types of subgraphs.

We start with the easiest such result. All graphs in these two lectures will be simple.

Proposition 1. If G is a bipartite graph on n vertices, then G can have $\leq \frac{n^2}{4}$ edges.

Remark. This falls under the previous paradigm since G is bipartite if and only if G does not contain any polygon of odd length.

The maximal number of edges of a bipartite graph

Proof of the proposition. If G is a bipartite graph on n vertices, then we have disjoint subsets A and B of $V(G)$, with k and $n-k$ elements, such that every edge of G joins a vertex in A with a vertex in B. It follows that

$$\#E(G) \leq k(n-k) \leq \left(\frac{k+(n-k)}{2}\right)^2 = \frac{n^2}{4},$$

where the inequality is given by the inequality between the arithmetic and geometric means.

Proof of the remark. The fact that a bipartite graph contains no polygon of odd length is clear. Suppose now that G contains no such polygon and let's show that G is bipartite. By treating each connected component of G separately, we may assume G is connected. Fix a vertex v_0 of G and let A and B be the subsets of $V(G)$ that can be connected to v_0 via a path of even (respectively, odd) length. We clearly have $A \cup B = V(G)$ and the hypothesis implies $A \cap B = \emptyset$. Moreover, it implies that no two vertices in A (or B) can be neighbors in G. Hence $G = G(A, B)$ is bipartite.

Graphs with no triangles

With a bit more work, we can prove a stronger version of the previous result.

Proposition 2 (Mantel, 1907). If G is a graph on n vertices that does not contain any triangle, then $\#E(G) \leq \frac{n^2}{4}$.

Proof. Let $m = \#E(G)$. For every $x \in V(G)$, let $d_x = \deg(x)$. Given any two adjacent vertices $x, y \in V(G)$, the hypothesis implies that x and y have no common neighbors. We thus have $d_x + d_y \leq n$. Therefore

$$\sum_{\{x,y\} \in E(G)} (d_x + d_y) \leq mn.$$

On the other hand, we have

$$\sum_{\{x,y\} \in E(G)} (d_x + d_y) = \sum_{x \in V(G)} d_x^2.$$

Graphs with no triangles, cont'd

Recall now that the Cauchy-Schwarz inequality says that for every $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, we have

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

By taking $b_1 = \ldots = b_n = 1$, we obtain

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

Applying this and our inequalities, we obtain

$$mn^2 \geq n \cdot \sum_{x \in V(G)} d_x^2 \geq \left(\sum_{x \in V(G)} d_x\right)^2 = (2m)^2 = 4m^2,$$

hence $m \leq n^2/4$.

Graphs with no triangles, cont'd

Remark. It follows from the proof of Proposition 2 that if G is as in the proposition and $m = \frac{n^2}{4}$, then for every edge $\{x, y\}$ in $E(G)$, every vertex of G is a neighbor of precisely one of x and y. If A consists of the neighbors of x and B consists of the neighbors of y, since G contains no triangles, we see that G is the complete bipartite graph on A and B. If $a = \#A$ and $b = \#B$, then $n = a + b$ and $\#E(G) = ab = (a + b)^2/4$, hence $a = b$.

Remark. If n is odd and G is a graph on n vertices that does not contain any triangles, then it follows from Proposition 2 that $\#E(G) \leq \frac{n^2-1}{4}$. This is sharp, as can be seen by taking the complete bipartite graph $K_{k+1,k}$, where $k = \frac{n-1}{2}$.

Exercise. Show that in this case, too, if $\#E(G) = \frac{n^2-1}{4}$, then G is isomorphic to $K_{k+1,k}$.

Graphs with no K_{t+1}

Next: we want to extend the result in Proposition 2, by asking what is the maximum number of edges for a graph on n vertices that does not contain any K_{t+1}, for some $t \geq 2$ (recall: K_{t+1} is the complete graph on $t+1$ vertices).

How to construct graphs without any K_{t+1}: let $V = V_1 \sqcup \ldots \sqcup V_t$. If we consider the graph G with $V(G) = V$ and where the edges in G are given by all edges that connect vertices in different V_i, then G does not contain any K_{t+1} (for any $t+1$ vertices of G, two of them must lie in the same V_i by the pigeonhole principle).

Note: in this example, if $d_i = \#V_i$, then $\#V = d_1 + \ldots + d_t$ and $\#E(G) = \sum_{i<j} d_i d_j$.

Easy fact: if $n = d_1 + \ldots + d_t$ is fixed, then $\sum_{i<j} d_i d_j$ is maximal precisely when

$$d_i - d_j \in \{0, 1, -1\} \quad \text{for all} \quad i, j.$$

Graphs with no K_{t+1}, cont'd

Let's prove this fact when t divides n. In this case, we have

$$\sum_{i<j} d_i d_j = \frac{(d_1 + \ldots + d_t)^2 - (d_1^2 + \ldots + d_t^2)}{2}.$$

The Cauchy-Schwarz inequality gives

$$t \cdot (d_1^2 + \ldots + d_t^2) \geq (d_1 + \ldots + d_t)^2$$

(it also says that equality holds iff $d_1 = \ldots = d_t$). We thus conclude that

$$\sum_{i<j} d_i d_j \leq \frac{n^2 - \frac{n^2}{t}}{2} = \frac{n^2}{2}\left(1 - \frac{1}{t}\right)$$

(with equality iff $d_i = n/t$ for all i).

Graphs with no K_{t+1}, cont'd

The proof of the fact in the general setting (when t might not divide n is similar). Note that the condition $d_i - d_j \in \{0, 1, -1\}$ for all i and j uniquely determines the graph we constructed, up to isomorphism. Indeed, if $n = tk + r$, with $r \in \{0, 1, \ldots, t-1\}$, then after reordering the V_i, we may assume that $d_i = k + 1$ for $1 \leq i \leq r$ and $d_i = k$ for $r + 1 \leq i \leq t$. The corresponding graph is the Turán graph $T_{n,t}$.

Remark. The number of edges of the Turán graph $T_{n,t}$ is equal to

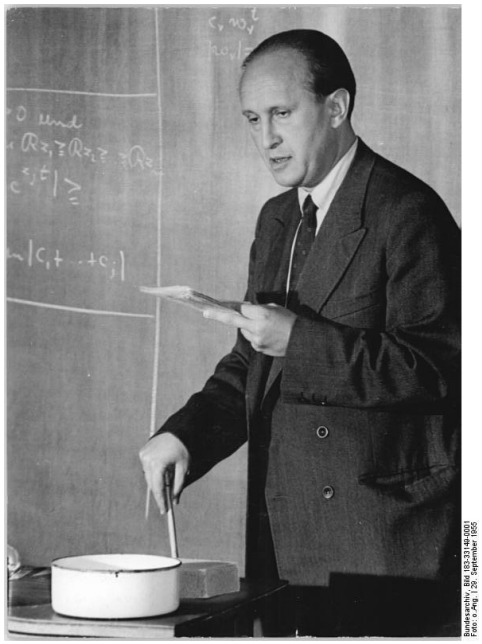
$$\frac{n^2}{2}\left(1 - \frac{1}{t}\right) + \text{lower degree terms in } n. \tag{1}$$

This is immediate if t divides n: in this case every V_i has $\frac{n}{t}$ elements, hence the number of edges is

$$\frac{1}{2}t \cdot \frac{n}{t} \cdot \left(n - \frac{n}{t}\right) = \frac{n^2}{2}\left(1 - \frac{1}{t}\right).$$

Exercise. Prove the formula in (1) in general.

Pál Turán



Pál Turán is another famous Hungarian mathematician, who lived 1910-1976.

Worked mostly in number theory, but also in analysis and graph theory. He collaborated extensively with Paul Erdös.

As a Jew, he could not get a university job for several years and was sent to labour service at various times between 1940-44.

Peter Franks: "Mathematicians have only paper and pen, he doesn't have anything in camp. So he created extremal combinatorics for which he did not need either."

Turán's theorem

Theorem (Turán). For every $t \geq 2$, among all graphs on n vertices that contain no K_{t+1}, the graph $T_{n,t}$ has the most edges; moreover, it is unique with this property.

Proof. We argue by induction on $t \geq 2$. For $t = 2$, we have seen that among the graphs on n vertices that contain no triangles, the unique one with the maximal number of edges is $K_{k,k}$ (if $n = 2k$ is even) or $K_{k+1,k}$ if $n = 2k + 1$ is odd. We now consider $t \geq 3$ and assume the theorem known for $t - 1$.

Let G be a graph on n vertices with no K_{t+1} subgraphs. Choose $v \in V(G)$ such that its degree $d_v := \deg_G(v)$ is maximal. Let $S_v \subseteq V(G)$ be the subset consisting of the neighbors of v (hence $\#S_v = d_v$) and $T_v = V(G) \smallsetminus S_v$.

Clear: since v is a common neighbor of all vertices in S_v, the assumption on G implies that the subgraph of G spanned by S_v contains no K_t subgraph.

Turán's theorem

We now modify G to get a new graph G' with $V(G') = V(G)$, as follows:
- We keep all edges in $E(G)$ between the vertices in S_v.
- We add edges between the vertices in S_v and the vertices in T_v.
- We remove all edges in $E(G)$ between the vertices in T_v.

Claim. For every vertex w in $V(G)$, we have $\deg_{G'}(w) \geq \deg_G(w)$.
- This is clear if $w \in S_v$ (we have only added some new edges incident to w).
- If $w \in T_v$, then

$$\deg_{G'}(w) = \#S_v = d_v \geq \deg_G(w),$$

by the maximality in our choice of v.
The claim, together with the formula relating the sum of the degrees with the number of edges gives

$$\#E(G') \geq \#E(G).$$

Turán's theorem

By induction hypothesis for the subgraph $\langle S_v \rangle$ of G spanned by S_v:

$$\#E(\langle S_v \rangle) \leq \#E(T_{d_v,t-1}).$$

Since G' contains no edges between vertices in T_v and contains all edges between vertices in S_v and vertices in T_v, we conclude that

$$\#E(G) \leq \#E(G') = (\#T_v) \cdot d_v + \#E(\langle S_v \rangle)$$

$$\leq (\#T_v) \cdot d_v + \#E(T_{d_v,t-1}) \leq \#E(T_{n,t}).$$

Furthermore, if $\#E(G) = \#E(T_{n,t})$, then we see that in G already every vertex in S_v was adjacent to every vertex in T_v (otherwise, by the previous computation $\#E(G') > \#E(G)$). Moreover, S_v has to be isomorphic to $T_{d_v,t-1}$. Since G contains no K_{t+1}, it follows that no vertices of T_v are adjacent in G. Hence G is constructed out of subsets $V_1, \ldots, V_t$. Maximality in the number of edges implies that G is isomorphic to $T_{n,t}$.

The Erdös-Stone theorem

We now consider the following general problem. Given a fixed graph H, estimate for large n the number $\mathrm{ex}(n, H)$ consisting of the maximum number of edges of a graph on n vertices that does not contain any graphs isomorphic to H. The answer is provided by the Erdös-Stone theorem. Of course, this is not interesting if $E(H) = \emptyset$ (in which case every graph with $n \geq \#V(H)$ vertices contains a subgraph isomorphic to H).

Recall that the chromatic number $\chi(H)$ is the smallest number c such that the vertices of H can be colored with c colors such that no two adjacent vertices get the same color. Note that $\chi(H) = 1$ if and only if $\#E(G) = \emptyset$.

Theorem (Erdös-Stone). For every (finite simple) graph H with $E(G) \neq \emptyset$ and for every $\epsilon > 0$, for $n \gg 0$, we have

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

The Erdös-Stone theorem

Example. If $H = K_{t+1}$ (when $\chi(H) = t+1$), this matches the assertion given by Turán's theorem, which gives

$$\mathrm{ex}(n, K_{t+1}) = \frac{1}{2}\left(1 - \frac{1}{t}\right) n^2 + \text{lower order terms in } n$$

(depending on the residue of n when divided by t).
Remark. The formula in the Erdös-Stone theorem implies that if $\chi(H) \geq 3$, then

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n, H)}{n^2} = \frac{1}{2}\left(1 - \frac{1}{\chi(H) - 1}\right) > 0.$$

Hence we understand the behavior of $\mathrm{ex}(n, H)$ for $n \to \infty$. The situation is more delicate when $\chi(H) = 2$ (that is, for bipartite graphs), when we can only conclude that for every $\epsilon > 0$, we have $\mathrm{ex}(n, H) \leq \epsilon n^2$ for $n \gg 0$. We will discuss some examples in the next lecture.

Proof of the lower bound in Erdös-Stone

This part is easy: say $\chi(H) = k + 1$. Since the Turán graph $T_{n,k}$ has vertices in k disjoint sets, with neighbors being precisely vertices in different such subsets, it follows that $\chi(T_{n,k}) = k$. Since $\chi(H) = k + 1$, we deduce that $T_{n,k}$ has no subgraph isomorphic to H. We thus conclude that $\mathrm{ex}(n, H) \geq \#E(T_{n,k})$. We have seen that

$$\#E(T_{n,k}) = \frac{1}{2}\left(1 - \frac{1}{k}\right) n^2 + \text{lower order terms in } n$$

(with the precise formula depending on the residue of n divided by t).

It thus follows that given $\epsilon > 0$, we have

$$\mathrm{ex}(n, H) \geq \#E(T_{n,k}) > \frac{1}{2}\left(1 - \frac{1}{k} - \epsilon\right) n^2 \quad \text{for} \quad n \gg 0.$$

A lemma for the upper bound in Erdös-Stone

Lemma. Given positive integers k and t, and $0 < \epsilon < 1/k$, for every graph G on n vertices, with n large enough (depending on k, t, and ϵ) and with $m \geq \frac{1}{2}\left(1 - \frac{1}{k} + \epsilon\right) n^2$ edges, there are disjoint subsets $A_1, \ldots, A_{k+1}$ of $V(G)$, all of size t, such that any two vertices in two different A_i and A_j are neighbors in G.

Proof. We first prove the following **Claim**:

Given any p and every ϵ', with $0 < \epsilon' < \epsilon$, if n is large enough, then for every graph G on n vertices, with $\#E(G) \geq \frac{1}{2}\left(1 - \frac{1}{k} + \epsilon\right) n^2$, we can find a subgraph G' of G on $\geq p$ vertices, such that for every $v \in V(G')$, we have

$$\deg_{G'}(v) \geq \left(1 - \frac{1}{k} + \epsilon'\right) \cdot \#V(G').$$

We construct G' by successively removing the "bad" vertices.

A lemma for the upper bound in Erdös-Stone, cont'd

More precisely, we construct a sequence of subgraphs $G_1, G_2, \ldots$ of $G = G_0$ as follows: if G_i has been constructed and it contains a vertex v_i with $\deg_{G_i}(v_i) < \left(1 - \frac{1}{k} + \epsilon'\right) \cdot \#V(G_i)$, then we take G_{i+1} to be the subgraph of G_i spanned by $V(G_i) \setminus \{v_i\}$; if there is no such v_i, then we stop.

We need to show that if $n \gg 0$, we can't reach a subgraph G_ℓ such that $\#V(G_\ell) < n'$. If this is the case, note that $q := \#V(G_\ell) = n - \ell < n'$. In this case, we have $\#E(G_\ell) \leq \frac{q(q-1)}{2} < \frac{{n'}^2}{2}$.
On the other hand, for every $i \geq 0$, we have $\#E(G_i) \leq$

$$\#E(G_{i+1}) + \left(1 - \frac{1}{k} + \epsilon'\right) \cdot \#V(G_i) = \#E(G_{i+1}) + \left(1 - \frac{1}{k} + \epsilon'\right) \cdot (n - i).$$

By combining all these, we get $\#E(G) \leq$

$$\frac{{n'}^2}{2} + \sum_{i=0}^{\ell-1} \left(1 - \frac{1}{k} + \epsilon'\right) \cdot (n-i) = \frac{{n'}^2}{2} + \left(1 - \frac{1}{k} + \epsilon'\right) \cdot \left(\frac{n(n+1)}{2} - \frac{q(q+1)}{2}\right)$$

A lemma for the upper bound in Erdös-Stone, cont'd

Since q is bounded above and $\epsilon' < \epsilon$, this implies that for n large enough, we get

$$\#E(G) < \frac{1}{2}\left(1 - \frac{1}{k} + \epsilon\right) n^2,$$

a contradiction. This completes the proof of the claim.

The claim implies that from now on (after possibly replacing ϵ by a smaller value), we may assume that we have a graph on n vertices, such that every vertex has degree $\geq \left(1 - \frac{1}{k} + \epsilon\right) n$. We show by induction on q, with $1 \leq q \leq k+1$, that for all t, if $n \gg 0$, then we can find disjoint subsets $A_1, \ldots, A_q$ of $V(G)$, all of size t, such that any two vertices in two different A_i and A_j are neighbors in G. For $q = k+1$, we obtain the assertion in the lemma.

A lemma for the upper bound in Erdös-Stone, cont'd

The assertion to prove is clear for $q = 1$, hence we may assume $q \geq 2$. We apply the induction hypothesis to get disjoint subsets $A'_1, \ldots, A'_q$ of $V(G)$, all of size $s = \lceil t/\epsilon \rceil$, such that any two vertices in two different A'_i and A'_j are neighbors in G.

Let $U = V(G) \setminus (A'_1 \cup \ldots \cup A'_q)$ and consider

$$W = \{v \in U \mid v \text{ has } \geq t \text{ neighbors in each } A'_i\}.$$

We first show that we can make $\#W$ arbitrarily large by taking $n \gg 0$.

We bound in two ways the number N of missing edges in G between $U \setminus W$ and $A'_1 \cup \ldots \cup A'_q$. Since every vertex in $U \setminus W$ has $< t$ neighbors in some A'_i, it follows that

$$N \geq \#(U \setminus W) \cdot (qs - t) = (n - qs - \#W) \cdot (qs - t) \geq (n - qs - \#W) s (q - \epsilon).$$

A lemma for the upper bound in Erdös-Stone, cont'd

On the other hand, since every vertex in G has degree $\geq \left(1 - \frac{1}{q} + \epsilon\right) n$ neighbors, we have

$$N \leq \#(A_1' \cup \ldots \cup A_q') \cdot \left(\frac{1}{q} - \epsilon\right) n = qsn \left(\frac{1}{q} - \epsilon\right) = sn(1 - q\epsilon).$$

We now combine the two inequalities involving N to get a lower bound for $\#W$:

$$(q - \epsilon)\#W \geq (n - qs)(q - \epsilon) - n(1 - q\epsilon) = n(\epsilon + 1)(q - 1) - qs(q - \epsilon).$$

Since q, ϵ, and s are fixed, we see that when we make n large enough, we may assume that $\#W$ is as large as we want.

In particular, we may and will assume that $\#W > \binom{s}{t}^q (t - 1)$.

Proof of the upper bound in Erdös-Stone

We now make the following construction: for every $w \in W$, we choose t neighbors of w in each of the A'_i and let Γ_w be their union.

Clearly, there are at most $\binom{s}{t}^q$ such subsets Γ_w. By our condition on $\#W$ and the pigeonhole principle, we can find distinct $w_1, \ldots, w_t \in W$ such that $\Gamma_{w_1} = \ldots = \Gamma_{w_t} = \Gamma$.

If we put $A_{q+1} = \{w_1, \ldots, w_t\}$ and let $A_i = \Gamma \cap A'_i$ for $1 \leq i \leq q$, we see that $A_1, \ldots, A_{q+1}$ all have size t, and any two vertices in two different A_i and A_j are neighbors in G.
This completes the proof of the induction step for the assertion depending of q and thus the proof of the lemma.

Proof of the upper bound in Erdös-Stone

We can now prove the upper bound in the Erdös-Stone theorem. Let $k = \chi(H) - 1$ and $t = \#V(H)$. It follows the lemma that if $n \gg 0$ and G is a graph on n vertices and with $m \geq \frac{1}{2}\left(1 - \frac{1}{k} + \epsilon\right) n^2$ edges, we have disjoint subsets $A_1, \ldots, A_{k+1}$ of $V(G)$ such that any two vertices in two different A_i, A_j are neighbors in G.

Since $\chi(H) = k + 1$, if we consider a coloring of H with colors $1, \ldots, k+1$ and if we map the vertices colored by i to distinct vertices in A_i, then we see that G has a subgraph isomorphic to H. We thus have

$$\mathrm{ex}(n, H) < \frac{1}{2}\left(1 - \frac{1}{k} + \epsilon\right) n^2 \quad \text{for} \quad n \gg 0.$$

This completes the proof of the Erdös-Stone theorem.

Reference

In the next lecture, we will consider in some examples the asymptotic behavior of $\mathrm{ex}(n, H)$ for certain bipartite graphs H.

For more on extremal graph theory, see Jacob Fox's MIT course, available at *http://math.mit.edu/~fox/MAT307*