Lecture 11: Extremal graph theory II

October 13, 2020

Review

Recall that if H is a (finite, simple) graph, we denote by $\mathrm{ex}(n, H)$ the maximum number of edges of a (simple) graph G on n vertices that does not contain any subgraph isomorphic to H. Note: we always assume that $E(H) \neq \emptyset$.

The Erdös-Stone theorem implies that:

- If $\chi(H) \geq 3$, then $\lim_{n\to\infty} \frac{\mathrm{ex}(n,H)}{n^2} = 1 - \frac{1}{\chi(H)-1} > 0$.
- If $\chi(H) = 2$ (equivalently, H is bipartite), then for every $\epsilon > 0$, we have $\mathrm{ex}(n, H) \leq \epsilon n^2$ for $n \gg 0$.

In this lecture we begin by discussing the behavior of $\mathrm{ex}(n, H)$ for $n \to \infty$ for certain examples of bipartite graphs H.

Trivial example. If H is the graph below and if G does not contain H, then $\deg(v) \leq 1$ for every $v \in V(G)$. Hence $\#E(G) \leq \frac{1}{2}\#V(G)$.

The case of P_4

The first interesting case of a bipartite graph is that of $H = P_4$.

Theorem 1. If G is a simple graph on n vertices that does not contain any P_4, then

$$\#E(G) \leq \frac{n}{4}(1 + \sqrt{4n-3}).$$

Proof. For every $v \in G$, let's put $d_v = \deg(v)$. We consider

$$F = \{(u, v, w) \mid \{u, v\} \in E(G), \{u, w\} \in E(G), v \neq w\}.$$

We count $\#F$ in two ways:

- Every $u \in V(G)$ contributes $d_u(d_u - 1)$ to F, hence $\#F = \sum_{u \in V(G)} d_u(d_u - 1)$.
- For every pair of distinct vertices (v, w), there is at most one $(u, v, w) \in F$ since G contains no P_4, hence $\#F \leq n(n-1)$.

The case of P_4, cont'd

We thus conclude that

$$n(n-1) \geq \sum_{u \in V(G)} d_u(d_u - 1) = \sum_{u \in V(G)} d_u^2 - \sum_{u \in V(G)} d_u. \qquad (1)$$

Recall now that if $m = \#E(G)$, then $\sum_u d_u = 2m$. We thus have

$$n\big(n(n-1) + 2m\big) \geq n \cdot \sum_{u \in V(G)} d_u^2 \geq \left(\sum_{u \in V(G)} d_u \right)^2 = 4m^2,$$

where the first inequality follows from (1) and the second one from Cauchy-Schwarz. This gives

$$4m^2 - 2mn - n^2(n-1) \leq 0$$

and by computing the roots of this quadratic expression, we get $m \leq \frac{n}{4}(1 + \sqrt{4n-3})$.

Sharpness of the estimate for $\mathrm{ex}(n, P_4)$

Remark. The result in Theorem 1 is sharp in the sense that there is $c > 0$ such that $\mathrm{ex}(n, P_4) > cn^{3/2}$ for infinitely many n.

In order to see this, consider $n = p^2$, where p is a prime number and

$$V = \mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z},$$

hence $n := \#V = p^2$. Two distinct (x, y) and (x', y') in V are joined by an edge if $x + x' = yy'$.

Note that for every $(x, y) \in V$ there are precisely p pairs $(x', y') \in V$ with $x + x' = yy'$. This is clear if $y \neq 0$ (we can take x' arbitrary and then $y' = y^{-1}(x + x')$); on the other hand, if $y = 0$, then $x' = -x$ and y' can be arbitrary. However, it might happen that $(x', y') = (x, y)$. In any case, it follows that for every $v \in V$, we have $\deg(v) \in \{p - 1, p\}$, hence

$$\#E(G) = \frac{1}{2}\sum_v \deg(v) \geq \frac{p^2(p-1)}{2} > cn^{3/2}$$

for some $c > 0$ and all $p \gg 0$.

Sharpness of the estimate for $\mathrm{ex}(n, P_4)$, cont'd

Claim: for every p, the graph we have described contains no subgraph isomorphic to P_4.

Indeed, suppose that two distinct (x_1, y_1) and $(x_2, y_2) \in V$ have two distinct neighbors (x, y) and (x', y'). Since

$$x + x_1 = yy_1 \quad \text{and} \quad x + x_2 = yy_2,$$

it follows that

$$x_1 - x_2 = y(y_1 - y_2).$$

If $y_1 = y_2$, this implies $x_1 = x_2$, a contradiction.

On the other hand, if $y_1 \neq y_2$, we see that y is determined uniquely by $y = (x_1 - x_2)(y_1 - y_2)^{-1}$, and then also x is determined uniquely by $x = yy_1 - x_1$. We thus again obtain a contradiction.

Conclusion: we have $\mathrm{ex}(n, P_4) > cn^{3/2}$ for $n = p^2 \gg 0$.

The complete bipartite graphs

Recall that the complete bipartite graph $K_{m,n}$ is the bipartite graph on two sets A and B with m and n elements, respectively, and such that every element of A is a neighbor of every element of B.

Note that P_4 is isomorphic to the complete bipartite graph $K_{2,2}$. Our next goal is to extend (partially) the result of Theorem 1 to all graphs $K_{n,n}$.

Theorem 2. For every $t \geq 2$, there is $c > 0$ such that

$$\mathrm{ex}(n, K_{t,t}) < c \cdot n^{2-\frac{1}{t}} \quad \text{for all} \quad n \gg 0.$$

Remark. Of course, if this is the case, then after possibly replacing c by a larger constant, we may assume that the inequality holds for all n.

The complete bipartite graphs, cont'd

Proof of Theorem 2. Suppose that G is a graph on $\{1, \ldots, n\}$ that does not contain any subgraph isomorphic to $K_{t,t}$. Put $d_i = \deg(i)$. Note that the set consisting of the neighbors of i has precisely $\binom{d_i}{t}$ subsets with t elements. Since G contains no $K_{t,t}$, when we vary i, every such subset appears $\leq (t-1)$ times. We thus have

$$\sum_{i=1}^{n} \binom{d_i}{t} \leq (t-1)\binom{n}{t} \leq (t-1)\frac{n^t}{t!}. \tag{2}$$

Note that $d_1 + \ldots + d_n = 2e$, where $e = \#E(G)$. Since the function $f(x) = \binom{x}{t} = \frac{x(x-1)\ldots(x-t+1)}{t!}$ is convex on $[t, \infty)$, it follows that we have the following lower bound for the left-hand side of (2):

$$\sum_{i=1}^{n} \binom{d_i}{t} \geq n\binom{2e/n}{t} = \frac{n \cdot \frac{2e}{n} \ldots \left(\frac{2e}{n} - t + 1\right)}{t!} \geq \frac{n \cdot \left(\frac{2e}{n} - t + 1\right)^t}{t!}.$$

The complete bipartite graphs, cont'd

By combining the inequalities on the previous slide, we obtain

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

Taking t-roots, we get

$$\frac{2e}{n} \leq (t-1)^{1/t} n^{1-\frac{1}{t}} + t - 1,$$

and thus

$$e < c \cdot n^{2-\frac{1}{t}}$$

for a suitable $c > 0$ and all $n \gg 0$.

Polygons with an even number of vertices

One can extend the result in Theorem 1 in another direction, by considering polygons with an even number of vertices. We state the following result, without proof:

Theorem 3. There is $c > 0$ such that

$$\mathrm{ex}(P_{2k}, n) \leq c \cdot n^{1+\frac{1}{k}} \quad \text{for all} \quad n.$$

Edge density

We end our introduction to extremal combinatorics with the presentation of a powerful result, due to Szemerédi, which over the years found many striking applications. We will only discuss the statement and an outline of the proof, and describe briefly some applications.

In what follows we only consider finite simple graphs. Szmerédi's Regularity Lemma roughly says that for a graph with many vertices, the vertex set admits a partition into a bounded number of subsets such that the edges between most different parts behave "random-like".

Definition. Given a graph G with vertex set V and two (nonempty) subsets $X, Y \subseteq E(G)$, we put

$$E_G(X, Y) = \{(x, y) \in X \times Y \mid \{x, y\} \in E(G)\}$$

and the edge density between X and Y is given by

$$d_G(X, Y) = \frac{|E_G(X, Y)|}{|X| \cdot |Y|} \in [0, 1].$$

ϵ-regular pairs and ϵ-regular partitions

We fix a finite simple graph G with vertex set V and $\epsilon > 0$. We now introduce a condition on two subsets X, Y of V, that roughly says that for two large subsets of X and Y, the edge density between these subsets approximates the edge density between X and Y.

Definition. A pair (X, Y) of nonempty subsets of V is ϵ-regular if for every subsets $A \subseteq X$ and $B \subseteq Y$, with $|A| \geq \epsilon|X|$ and $|B| \geq \epsilon|Y|$, we have

$$|d_G(A, B) - d_G(X, Y)| \leq \epsilon.$$

We next consider partitions of V. Ideally, we would want partitions such that any pair of sets is ϵ-regular. However, we have to settle for less: most of such pairs will be regular, with the precise meaning of "most" depending on the size of the sets.

More precisely, we make the following definition.

ϵ-regular pairs and ϵ-regular partitions

Definition. A partition $\mathcal{P}$ of V given by $\{V_1, \ldots, V_r\}$ is an ϵ-regular partition if

$$\sum_{\substack{1 \le i,j \le r \\ (V_i, V_j) \neq \epsilon - \text{regular}}} |V_i| \cdot |V_j| \le \epsilon |V|^2.$$

With this definition, we can now state

Theorem 4 (Szemerédy's Regularity Lemma). For every $\epsilon > 0$ and every positive integer m, there is an integer $M \ge m$ such that for every graph G on a set of vertices V, with $|V| \ge M$, there is an ϵ-regular partition of V into k sets, with $m \le k \le M$.

Endre Szemerédi



Endre Szemerédi, Hungarian mathematician, Professor at Rutgers University

He is famous (among other things) for proving a conjecture of Erdös and Turán on sequences of positive integers with positive density. In order to prove this he proved the Regularity Lemma, that has become an important tool in combinatorics.

He was awarded the Abel prize in 2012 “for his fundamental contributions to discrete mathematics and theoretical computer science”.

Szemerédi's theorem

Theorem (Szemerédi 1975, conjectured by Erdös-Turán 1936). If $A \subset \mathbf{Z}_{>0}$ is a set of positive density, then A contains arbitrarily long arithmetic progressions.

Recall that A has positive density if

$$\limsup_{n\to\infty} \frac{A \cap \{1, \ldots, n\}}{n} > 0.$$

Several other proofs of this theorem have been given, especially by Furstenberg (using ergodic theory) and Gowers (using Fourier analysis and combinatorics).
One should compare this with the following more recent result (building on Szemerédi's theorem):
Theorem (Green-Tao, 2004). The set of prime integers contains arbitrarily long arithmetic progressions.

Szemerédi's Regularity Lemma: outline of the proof

The proof of the theorem proceeds following an algorithm:

- Begin with the trivial partition.

- Suppose that at some step in the algorithm, we have a partition $\{V_1, \ldots, V_r\}$ of V. If this is ϵ-regular, then we are done. If it is not ϵ-regular, then for every pair (V_i, V_j) that is not ϵ-regular, we choose subsets $A^{i,j} \subseteq V_i$ and $A^{j,i} \subseteq V_j$ witnessing this fact: this means that

$$|A^{i,j}| \geq \epsilon |V_i|, |A^{j,i}| \geq \epsilon |V_j|, \quad \text{and} \quad |d_G(A^{i,j}, A^{j,i}) - d_G(V_i, V_j)| > \epsilon.$$

- Using the sets $A^{i,j}$, one refines the given partition. The goal is to show that if M is large enough, the algorithm stops after finitely many steps.

This is done by introducing a notion of energy.

Szemerédi's Regularity Lemma: outline of the proof

Definition. Given $U_1, U_2 \subseteq V(G)$, we put

$$q(U_1, U_2) := \frac{|U_1| \cdot |U_2|}{|V(G)|^2} d_G(U_1, U_2).$$

Definition. For a partition $\mathcal{P} = \{U_1, \ldots, U_r\}$ of $V(G)$, the energy of $\mathcal{P}$ is

$$q(\mathcal{P}) = \sum_{i,j=1}^{r} q(U_i, U_j).$$

It is easy to see that $q(\mathcal{P}) \leq 1$ for all $\mathcal{P}$.

Another fact that is easy to see is that if $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}') \geq q(\mathcal{P})$.

Szemerédi's Regularity Lemma: outline of the proof

The key ingredient is the following

Lemma. If $|V(G)|$ is large enough, then if a partition $\mathcal{P} = \{U_1, \ldots, U_r\}$ of $V(G)$ is not ϵ-regular, then there is a partition $\mathcal{P}'$ refining $\mathcal{P}$, with each U_i partitioned into at most 2^k parts, such that

$$q(\mathcal{P}') \geq q(\mathcal{P}) + \epsilon^5.$$

Note that since the energy of every partition is bounded above by 1, we see that after applying the above lemma at most ϵ^{-5} steps, we obtain the conclusion in the Regularity Lemma.

We end here our discussion of Szemerédi's Regularity Lemma.