These notes are from https://mathematicaster.org/teaching/graphs2022/extra_05-03.pdf

Recall that for graphs G, H, the graph G is said to be H-free if G does not contain a copy of H.

Today we explore the following natural question: How many edges can an H-free graph have?

Definition 1. *Let H be a graph. The* extremal number *(or* Turán number*) of H is defined to be*

$$\text{ex}(n, H) = \max\big\{|E(G)| : G \textit{ an } n\textit{-vertex, } H\textit{-free graph}\big\}.$$

A bit of jargon as well: G is said to be an extremal example for $\text{ex}(n, H)$ if G is an n-vertex, H-free graph and has $|E(G)| = \text{ex}(n, H)$, i.e. G is an example of such a graph with the maximum possible number of edges. We will also say that G is the unique extremal example if any extremal example is isomorphic to G.

Let's start with two silly examples:

- $\text{ex}(n, K_2) = 0$ and the unique extremal example is $\overline{K_n}$.

  Indeed, any edge is a copy of K_2.

- $\text{ex}(n, P_3) = \lfloor n/2 \rfloor$ and the unique extremal example is a matching on $\lfloor n/2 \rfloor$ many edges.

  Indeed, any pair of incident edges create a copy of P_3.

Let's now meet our first really interesting example: triangle-free graphs. To begin, note that $K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$ is an n-vertex, triangle-free graph, so

$$\text{ex}(n, K_3) \geq |E(K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor})| = \left\lceil \frac{n}{2} \right\rceil \left\lfloor \frac{n}{2} \right\rfloor = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

It turns out that you can't do any better!

Theorem 2 (Mantel)**.** $\text{ex}(n, K_3) = \lfloor n^2/4 \rfloor$ *and the unique extremal example is* $K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$.

Proof. Let $G = (V, E)$ be an n-vertex, triangle-free graph. We need to show that $|E| \leq \lfloor n^2/4 \rfloor$ with equality if and only if $G \cong K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$.

Set $\alpha = \alpha(G)$ and let A be a maximum independent set of G, so $|A| = \alpha$. Recall HW2.5:

$$\sum_{v \in A} \deg v \leq |E|,$$

with equality if and only if $V \setminus A$ is also an independent set.

Next, since G is triangle-free, $N(v)$ is an independent set for all $v \in V$ and so $\deg v \leq \alpha$. We therefore bound

$$2|E| = \sum_{v \in V} \deg v = \sum_{v \in A} \deg v + \sum_{v \in V \setminus A} \deg v \leq |E| + \sum_{v \in V \setminus A} \alpha = |E| + (n - \alpha)\alpha \implies |E| \leq (n - \alpha)\alpha,$$

with equality if and only if $V \setminus A$ is also an independent set and $\deg v = \alpha$ for all $v \in V \setminus A$. In other words, equality above holds iff $G \cong K_{\alpha, n-\alpha}$.

Now, we apply the AM–GM inequality to bound

$$(n-\alpha)\alpha \le \frac{((n-\alpha)+\alpha)^2}{4} = \frac{n^2}{4},$$

with equality iff $\alpha = n/2$. Of course, n and α are both integers, so $\alpha = n/2$ is impossible if n is odd. However, a very slight modification of the proof of the AM–GM inequality yields

$$(n-\alpha)\alpha \le \left\lfloor \frac{n^2}{4} \right\rfloor,$$

with equality if and only if $\alpha \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$.

Putting everything together, we have shown that $|E| \le \lfloor n^2/4 \rfloor$ with equality if and only if $G \cong K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$. □

What's next? Well, we now understand $\mathrm{ex}(n, K_2)$ and $\mathrm{ex}(n, K_3)$, so let's figure out $\mathrm{ex}(n, K_t)$ for larger values of t.

We first try to find a good construction for $\mathrm{ex}(n, K_t)$. First consider a complete multipartite with $r \le t-1$ many parts; since K_t is t-partite, this graph cannot contain a copy of K_t. Furthermore, if this complete multipartite graph has parts $A_1, \dots, A_r$ ($r \le t-1$), then it has

$$\binom{n}{2} - \sum_{i=1}^{r} \binom{|A_i|}{2}$$

many edges. How should we pick $r \le t-1$ and the A_i's so as to maximize the above quantity?

Let $T_k(n)$ denote the balanced complete k-partite graph on n vertices (often known as "the Turán graph"). Here, "balanced" means that the sizes of any two parts differ by at most one, i.e. the part-sizes are as close to equal as possible. This means that each part has size either $\lfloor n/k \rfloor$ or $\lfloor n/k \rfloor + 1$. Note that $T_2(n) = K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$.

If $n \mod k = r$, then $T_k(n)$ has r parts of size $\lfloor n/k \rfloor + 1$ and $k - r$ parts of size $\lfloor n/k \rfloor$, so

$$|E(T_k(n))| = \binom{n}{2} - r\binom{\lfloor n/k \rfloor + 1}{2} - (k-r)\binom{\lfloor n/k \rfloor}{2}.$$

Now, when n is large, this is approximately

$$|E(T_k(n))| \approx \frac{n^2}{2} - k\frac{(n/k)^2}{2} = \left(1 - \frac{1}{k}\right)\frac{n^2}{2} \approx \left(1 - \frac{1}{k}\right)\binom{n}{2},$$

so $T_k(n)$ has approximately a $1 - \frac{1}{k}$ proportion of all possible edges, which is quite a bit!

We claim that $T_{t-1}(n)$ is the "best" complete multipartite graph for avoiding K_t's.

Lemma 3. *If G is an n-vertex, complete multipartite, K_t-free graph, then $|E(G)| \le |E(T_{t-1}(n))|$ with equality if and only if $G \cong T_{t-1}(n)$.*

Proof. Let $G = (V, E)$ be such a graph with the maximum number of edges; we claim that $G \cong T_{t-1}(n)$ which implies the claim.

Since G is complete multipartite, we may denote the parts by $A_1, \dots, A_r$ for some r where each A_i is non-empty. We begin by noting that $r \le t-1$ or else we could pick one vertex from the first t parts to find a copy of K_t (since G has all edges between every pair of parts).

Now, if $r < t-1$, then introduce $t-1-r$ many independent sets, so that we can assume that G is a complete multipartite graph with parts $A_1, \ldots, A_{t-1}$, where some parts may be empty. If we can show that

$$\big||A_i| - |A_j|\big| \leq 1,$$

for all $i \neq j \in [t-1]$, then we will have shown that $G \cong T_{t-1}(n)$.

Suppose for the sake of contradiction that there is some $i \neq j \in [t-1]$ for which $\big||A_i|-|A_j|\big| \geq 2$. By relabeling the A_ℓ's if necessary, we may suppose that $|A_1| \geq |A_2|+2$. Then let G' be the complete multipartite graph with parts $A_1', A_2', A_3, \ldots, A_{t-1}$ where $|A_1'| = |A_1| - 1$ and $|A_2'| = |A_2| + 1$ (i.e. move one vertex from A_1 to A_2). We find that the number of edges from $A_1' \cup A_2'$ to $A_3 \cup \cdots \cup A_{t-1}$ in G' is precisely the same as the number of edges from $A_1 \cup A_2$ to $A_3 \cup \cdots \cup A_{t-1}$ in G. Therefore, setting $x = |A_1|$ and $y = |A_2|$ (so $x \geq y+2$), we have

$$|E(G')| - |E(G)| = |A_1'| \cdot |A_2'| - |A_1| \cdot |A_2| = (x-1)(y+1) - xy = x - y - 1 \geq 1;$$

a contradiction to the maximality of G. ∎

Theorem 4 (Turán)**.** *$\mathrm{ex}(n, K_t) = |E(T_{t-1}(n))| \approx \left(1 - \frac{1}{t-1}\right)\binom{n}{2}$ and the unique extremal graph is $T_{t-1}(n)$.*

Note that Turán's theorem implies Mantel's theorem. Also, there are tons and tons of proofs of Turán's theorem. Since this is our last day and I'm not expecting you to solve questions along these lines, I want to share the most clever proof that I've ever seen of it.

Proof. Let $G = (V, E)$ be an n-vertex, K_t-free graph with the maximum number of edges possible. We claim that $G \cong T_{t-1}(n)$ which will prove the claim. By Lemma 3, it suffices to show that G is a complete multipartite graph.

Claim 5. *For any three distinct vertices $x, y, z \in V$, if $xy, yz \notin E$, then also $xz \notin E$.*

Proof. Suppose for the sake of contradiction that $xz \in E$. Without loss of generality, we may suppose that $\deg z \geq \deg x$. We break into cases depending on the degree of y.

Case 1: $\deg y < \deg z$. Create a new graph G' by deleting y and "duplicating" z. By "duplicating" z, we mean introduce a new vertex z' which has the same neighborhood as z (note that z and z' are not adjacent). Certainly G' also has n vertices (we deleted one and added another); we claim that G' additionally has no copy of K_t.

Indeed, if G' contains a copy of K_t, then such a copy *must* use the vertex z' since otherwise this would be a copy of K_t in G. But if it uses z', then it cannot use z since $zz' \notin E(G')$. However, since z' is a duplicate of z, we may replace z' by z to find a copy of K_t in G, which we know is impossible.

Thus, G' is an n-vertex, K_t-free graph. However, since $\deg y < \deg z$ and $yz \notin E$, we have

$$|E(G')| = |E| - \deg y + \deg z > |E|;$$

a contradiction to the maximality of G.

Case 2: $\deg y \geq \deg z$. Create a new graph G' by deleting both x and z and duplicating y twice. This new graph G' still has n vertices and, by the same logic as above (duplicating a vertex cannot suddenly introduce a copy of K_t), G' is K_t-free. However, since $\deg y \geq \deg z \geq \deg x$, $xy, yz \notin E$ and $xz \in E$, we have

$$|E(G')| = |E| - (\deg x + \deg z - 1) + 2\deg y \geq |E| + 1 > |E|;$$

a contradiction to the maximality of G. $\square$

Now, define the relation R on V by $x\,R\,y \iff xy \notin E$. Certainly R is reflexive and symmetric, and the above claim tells us that R is also transitive. So R is actually an equivalence relation! Therefore, let label the equivalence classes of R as $A_1, \ldots, A_r$ for some r. By definition G has no edges within any A_i and G contains every edge between A_i and A_j for all $i \neq j$. In other words, G is a complete multipartite graph and so the claim follows from Lemma 3. $\square$

Now that we understand the extremal numbers of cliques in general, let's work on another small graph: the four-cycle.

Theorem 6 (Kővári–Sós–Turán (special case))**.** $\mathrm{ex}(n, C_4) \leq \frac{n}{4}\big(1 + \sqrt{4n-3}\big) \approx \frac{1}{2}n^{3/2}$.

The full Kővári–Sós–Turán theorem gives an upper bound on $\mathrm{ex}(n, K_{s,t})$ (note that $C_4 \cong K_{2,2}$).

In order to prove this, we will require a special case of the Cauchy–Schwarz inequality, which is arguably the most important inequality in all of mathematics.

Lemma 7 (Cauchy–Schwarz inquality)**.** *For any real numbers* $a_1, \ldots, a_n, b_1, \ldots, b_n$.

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

The full Cachy–Schwarz inequality is an inequality for general inner products.

Proof. We expand

$$0 \leq \frac{1}{2}\sum_{i,j}(a_i b_j - a_j b_i)^2 = \frac{1}{2}\sum_{i,j}(a_i^2 b_j^2 + a_j^2 b_i^2 - 2a_i b_i a_j b_j)$$
$$= \sum_{i,j} a_i^2 b_j^2 - \sum_{i,j} a_i b_i a_j b_j = \sum_i a_i^2 \cdot \sum_i b_i^2 - \left(\sum_i a_i b_i\right)^2. \qquad \square$$

Explicitly, we will use the following corollary:

Corollary 8. *For any real numbers* $a_1, \ldots, a_n$*,*

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

Proof. We apply Cauchy–Schwarz to bound

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

Proof of Kővári–Sós–Turán. Let $G = (V, E)$ be an n-vertex, C_4-free graph. We need to show that $|E| \leq \frac{n}{4}\big(1 + \sqrt{4n-3}\big)$.

Fix any $x \neq y \in V$; we begin by observing that $|N(x) \cap N(y)| \leq 1$. Indeed, if there were $u \neq v \in N(x) \cap N(y)$, then (x, u, y, v) would create a 4-cycle.

In particular, $\binom{N(x)}{2} \cap \binom{N(y)}{2} = \varnothing$ for every $x \neq y \in V$. Of course, $\binom{N(x)}{2} \subseteq \binom{V}{2}$ for each x, so, using these observations, we bound

$$\binom{n}{2} = \left|\binom{V}{2}\right| \geq \left|\bigsqcup_{x \in V}\binom{N(x)}{2}\right| = \sum_{x \in V}\left|\binom{N(x)}{2}\right| = \sum_{x \in V}\binom{\deg x}{2} = \frac{1}{2}\sum_{x \in V}(\deg^2 x - \deg x).$$

From here, we apply Cauchy–Schwarz and the handshaking lemma to bound

$$\begin{aligned}\binom{n}{2} &\geq \frac{1}{2}\sum_{x\in V}\left(\deg^2 x - \deg x\right) = \frac{1}{2}\left(\sum_{x\in V}\deg^2 x\right) - |E| \\ &\geq \frac{1}{2}\cdot\frac{1}{n}\left(\sum_{x\in V}\deg x\right)^2 - |E| = \frac{2|E|^2}{n} - |E|.\end{aligned}$$

Solving for $|E|$ in this inequality, we get

$$|E| \leq \frac{n}{4}\left(1+\sqrt{4n-3}\right). \qquad \square$$