TURÁN'S GRAPH THEOREM AND EXTREMAL GRAPH THEORY

ANDREW CHANG

Abstract. Turán's graph theorem is an elegant, crucial result in graph theory. We present five unique proofs using different ideas, as well as short proofs of the simpler Mantel's theorem. At the end we introduce and prove other results, such as Ramsey's theorem, from the field of extremal graph theory of which Turán's graph theorem is a foundational discovery.

1. What is Turán's Graph Theorem

Turán's graph theorem is simple to define and is a key result in graph theory, with many beautiful proofs. Discovered in 1941, it was the starting point for the whole branch of extremal graph theory, which focuses on essentially optimizing graphs: how many edges can one have while satisfying a certain condition?

The basic question answered by Turán's graph theorem is: Suppose G is a simple graph that does not contain a p-clique. What is the largest number of edges that G can have?

Definition 1.1. A *p-clique* in G is a complete subgraph of G on p vertices, denoted by K_p.

We also use K_n to denote, in general, any complete graph with n vertices.

In other words, we want to find the maximum number of edges that can be in a graph not containing a complete subgraph with p vertices (a subgraph where every vertex is connected to every other one).

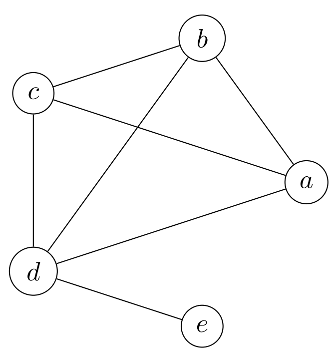

Figure 1. An example of a graph containing a 4-clique.

An easy way to create a graph that doesn't contain a p-clique is to divide the vertices into $p-1$ subsets; these should be disjoint, meaning every vertex is in exactly one subset. After partitioning the vertices like this, we draw edges between all the pairs of vertices in different subsets, but never within any subset. This means there cannot be a p-clique, because by the Pigeonhole Principle at least two of the vertices in the clique would have to be from the same subset — and then they cannot be connected!

To maximize the number of edges, we must divide the vertices as evenly as possible between the different subsets (such that none differ in size by more than 1). To prove this, suppose a graph has two subsets with sizes n_1 and n_2, where $n_1 - n_2 \geq 2$. Then we can move a vertex from n_1 to n_2. The resulting graph has $(n_1-1)(n_2+1)$ edges between the two subsets while the old graph has $n_1 n_2$, and the total number of edges with other subsets stays the same. Thus there are

$$(n_1 - 1)(n_2 + 1) - n_1 n_2 = n_1 - n_2 - 1 \geq 1$$

Date: March 20, 2022.

more edges after moving the vertex. Applying this principle in general shows that we maximize the edges in this kind of graph by having the subsets as equally sized as possible.

If n is divisible by $p-1$, we can write the total number of edges as

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

Turán's says that in general, this is indeed the maximum number of edges possible:

Theorem 1.2 (Turán). *If a graph $G=(V,E)$ on n vertices has no p-clique, $p \geq 2$, then*

$$|E| \leq \left(1-\frac{1}{p-1}\right)\frac{n^2}{2}.$$

We will give multiple proofs from [1] of the triangle-free case with Mantel's theorem, followed by several unique proofs of the full Turán's graph theorem. Last but not least we will broaden our scope with topics such as Ramsey numbers in extremal graph theory and an introduction to more complex extensions like the Erdős–Stone theorem.

2. Simple Cases

We start with a simple case, where p equals 3: in other words, we find an upper bound for the number of edges in a graph without triangles. (Note that $p=2$ is a trivial case about graphs with no edges at all, so we will skip over it.)

This result is known as Mantel's theorem:

Theorem 2.1 (Mantel's). *Suppose G is a graph on n vertices without triangles. Then G has at most $\frac{n^2}{4}$ edges, and equality holds if and only if n is even and G is the complete bipartite graph $K_{n/2,n/2}$.*

First proof — Mantel. Let the set of vertices in G be $V=\{1,\ldots,n\}$, and let the set of edges be E. Also, we define d_i as the degree of a vertex i. Since each edge in the graph will be counted in the degrees of both of its nodes, $\sum_{i\in V} d_i = 2|E|$ (summing all the degrees gives twice the number of edges). Meanwhile, for any edge ij we can say that $d_i + d_j \leq n$, because no vertex is counted in the degrees of both i and j — if a vertex was connected to both of them, the graph would have a triangle. Thus summing over all edges we have

$$\sum_{ij\in E}(d_i+d_j) \leq n|E|.$$

For any i, d_i will appear d_i times in this sum (once for every edge involving i), so we can rewrite as

$$n|E| \geq \sum_{ij\in E}(d_i+d_j) = \sum_{i\in V} d_i^2.$$

Applying the Cauchy–Schwarz inequality on $(d_1,\ldots,d_n)$ and $(1,\ldots,1)$ gives

$$n|E| \geq \sum_{i\in V} d_i^2 \geq \frac{(\sum d_i)^2}{n} = \frac{4|E|^2}{n},$$

from which the result follows. When equality holds we have $d_i = d_j$ for all i,j, and since $d_i+d_j=n$ it follows that $d_i = \frac{n}{2}$. This and the condition that G is triangle-free immediately leads to the conclusion that $G = K_{n/2,n/2}$. □

The next proof is even simpler, using only the inequality of arithmetic and geometric means, also known simply as the AM–GM inequality.

Second proof — folklore. Let a be the size of a largest independent set (a set of vertices with no edges between them) in G. Because G is triangle free, all the neighbors of a vertex i form an independent set with size d_i, so $d_i \leq a$ for all i.

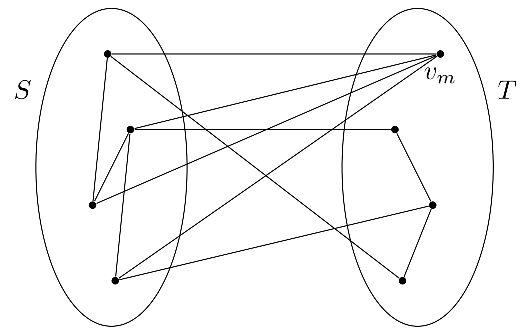

Figure 2. The original graph G.

The set $B = V \setminus A$ with size $b = n - a$ meets every edge of G, as there are no edges between two vertices in A. Therefore if we count the edges by their end vertices in B, we have $|E| \leq \sum_{i \in B} d_i$. Since $d_i \leq a$ and there are b vertices in B, $\sum_{i \in B} d_i \leq ab$. Then by the AM–GM inequality,

$$|E| \leq \sum_{i \in B} d_i \leq ab \leq \left(\frac{a+b}{2}\right)^2 = \frac{n^2}{4},$$

and the equality case is handled as last time. □

3. Proofs with Induction

First proof — Turán. We use induction on n. It is easy to show that the theorem is true when $n < p$, so we turn our attention to where $n \geq p$. Consider a graph $G = (V, E)$ with a maximum number of edges. It will contain $(p-1)$-cliques, since it has as many edges as possible, so consider one $(p-1)$-clique A and the set $B = V \setminus A$.

A has $\binom{p-1}{2}$ edges, while we find an upper bound for the number of edges within B (e_B), and the number between A and B ($e_{A,B}$). By induction we apply the theorem on B to give $e_B \leq \frac{1}{2}(1 - \frac{1}{p-1})(n-p+1)^2$. Since G has no p-clique, every $v_j \in B$ is adjacent to at most $p-2$ vertices in A (otherwise it and $p-1$ vertices in A would form a p-clique), and we obtain $e_{A,B} \leq (p-2)(n-p+1)$. Summing all this gives

$$|E| \leq \binom{p-1}{2} + \frac{1}{2}\left(1 - \frac{1}{p-1}\right)(n-p+1)^2 + (p-2)(n-p+1),$$

which after some simple algebra reduces to

$$|E| \leq \left(1 - \frac{1}{p-1}\right)\frac{n^2}{2}.$$

□

Second proof — Erdős. We use the structure of Turán graphs. Let $v_m \in V$ be a vertex with maximal degree (i.e., no vertex in G has a higher degree). Let S be the set of neighbors of v_m, and define $T = V \setminus S$. Since G does not contain a p-clique, and v_m is adjacent to all vertices of S, we see that S cannot contain a $(p-1)$-clique.

We now construct a graph H on V, as shown in figures 2 and 3, which corresponds to G on S and contains all edges between S and T, but no edges within T. In other words, T is an independent set in H, so H has again no p-cliques as there cannot be two vertices from T in a clique. Let d'_j be the degree of v_j in H. If $v_j \in S$, then we have $d'_j \geq d_j$ by the construction of H, and for $v_j \in T$, we see that $d'_j = |S| = d_m \geq d_j$ by the choice of v_m. Thus $|E(H)| \geq |E|$, and we can conclude that among all graphs with a maximal number of edges, there must be one of the form of H. By induction, the graph induced by S has no more edges than a suitable graph $K_{n_1,\ldots,n_{p-2}}$ on S. We have

$$|E| \leq |E(H)| \leq E(K_{n_1,\ldots,n_{p-1}})$$

where $n_{p-1} = |T|$, and the theorem follows from considering the edges in this type of graph. □

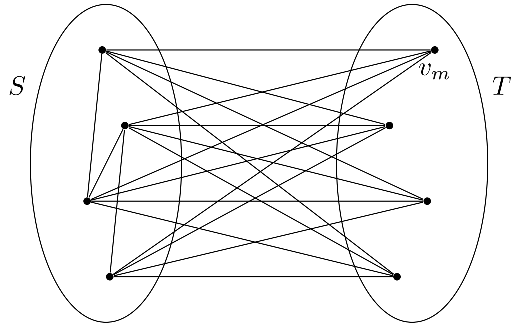

Figure 3. The new, constructed graph H.

4. Proofs with Probability

These proofs use a maximizing argument and probability theory concepts.

Third proof — Motzkin and Straus. We consider a probability distribution $\boldsymbol{w} = (w_1, \ldots, w_n)$ on the vertices, i.e. an assignment of values $w_i \geq 0$ to the vertices, so that $\sum^n w_i = 1$. We seek to maximize

$$f(\boldsymbol{w}) = \sum_{v_i v_j \in E} w_i w_j.$$

For any distribution $\boldsymbol{w}$, let v_i and v_j be a pair of nonadjacent vertices with weights w_i, w_j. Let s_i be the sum of the weights of all vertices adjacent to v_i, and similarly define s_j for v_j. We assume without loss of generality that $s_i \geq s_j$. Now we move the weight from v_j to v_i, such that the new weight of v_i is $w_i + w_j$ while the weight of v_j drops to 0. For the new distribution $\boldsymbol{w}'$ we have

$$f(\boldsymbol{w}') = f(\boldsymbol{w}) + w_j s_i - w_j s_j \geq f(\boldsymbol{w}).$$

We can repeat this (reducing the number of vertices with a positive weight by one in each step) until there are no nonadjacent vertices of positive weight anymore. At the end we conclude that there is an optimal distribution with nonzero weights all on a clique, say of size k.

Next, if for example, $w_1 > w_2 > 0$, we can choose ε with $0 < \varepsilon < w_1 - w_2$ and change w_1 to $w_1 - \varepsilon$ and w_2 to $w_2 + \varepsilon$. The new distribution $\boldsymbol{w}'$ satisfies $f(\boldsymbol{w}') = f(\boldsymbol{w}) + \varepsilon(w_1 - w_2) - \varepsilon^2 > f(\boldsymbol{w})$. Applying this argument whenever adjacent vertices are unequal in weight, we find that the maximal value of $f(\boldsymbol{w})$ is attained with $w_i = \frac{1}{k}$ on a k-clique and $w_i = 0$ otherwise. Since a k-clique contains $\frac{k(k-1)}{2}$ edges, we obtain

$$f(\boldsymbol{w}) = \left(\frac{k(k-1)}{2}\right)\frac{1}{k^2} = \frac{1}{2}\left(1 - \frac{1}{k}\right).$$

Because this expression is increasing in k, we maximize it with $k = p - 1$ (since G has no p-cliques). So we conclude

$$f(\boldsymbol{w}) \leq \frac{1}{2}\left(1 - \frac{1}{p-1}\right)$$

for any distribution $\boldsymbol{w}$. Specifically, we apply this for the uniform distribution given by $w_i = \frac{1}{n}$ for all i, giving

$$\frac{|E|}{n^2} = f\left(w_i = \frac{1}{n}\right) \leq \frac{1}{2}\left(1 - \frac{1}{p-1}\right),$$

which is the same as the inequality in the theorem. □

Fourth proof — Alon and Spencer. Let $G = (V, E)$ be any graph with vertices $v_1, \ldots, v_n$. We define d_i as the degree of v_i. Then denote $\omega(G)$ to be the number of vertices in a largest clique, which we call the *clique number* of G. We claim that

$$\omega(G) \geq \sum_{i=1}^{n} \frac{1}{n - d_i}.$$

To show this, we first choose a random permutation $\pi = v_1 v_2 \ldots v_n$ of the vertex set V, (where each permutation has equal probability $\frac{1}{n!}$ to appear), and then consider a set C_π as follows. We put a vertex

v_i into C_π if and only if v_i is adjacent to all preceding vertices $v_j (j < i)$. By definition, C_π is a clique in G since each vertex will be adjacent to the other members of the set.

Let $X = |C_\pi|$ be the random variable corresponding to the size of the set. We have $X = \sum_{i=1}^n X_i$, where X_i is the indicator random variable of the vertex v_i: $X_i = 1$ or $X_i = 0$ depending on whether or not $v_i \in C_\pi$. Note that v_i is in C_π with respect to the permutation if and only if v_i appears before all $n-1-d_i$ vertices not adjacent to v_i, i.e. if v_i is the first in the permutation among v_i and its $n-1-d_i$ non-neighbors. The probability that this happens is $\frac{1}{n-d_i}$, so $EX_i = \frac{1}{n-d_i}$. By linearity of expectation this gives

$$E(|C_\pi|) = EX = \sum_{i=1}^n EX_i = \sum_{i=1}^n \frac{1}{n-d_i}.$$

Since this is the average over many cliques formed by different permutations, there must be a clique of size at least $\sum_{i=1}^n \frac{1}{n-d_i}$ and we have proven the original claim. Then apply the Cauchy–Schwarz inequality:

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

Setting $a_i = \sqrt{n-d_i}, b_i = \frac{1}{\sqrt{n-d_i}}$, we have $a_i b_i = 1$ and thus

$$n^2 \leq \left(\sum_{i=1}^n (n-d_i)\right)\left(\sum_{i=1}^n \frac{1}{n-d_i}\right) \leq \omega(G)\sum_{i=1}^n (n-d_i)$$

By our conditions for G, $\omega(G) \leq p-1$, while we also have $\sum_{i=1}^n d_i = 2|E|$ as used earlier. Applying these gives

$$n^2 \leq (p-1)(n^2 - 2|E|),$$

which is equivalent to Turán's inequality. □

5. The Final Proof

Fifth proof — folklore. Let G be a graph on n vertices without a p-clique and with a maximal number of edges. We claim that G does not contain any three vertices u, v, w such that vw is in E but not uv or uw. Suppose this is false. Then there are two cases:

Case 1: $d_u < d_v$ or $d_u < d_w$. We assume without loss of generality that $d_u < d_v$. We duplicate v, i.e. create a new vertex v' with the same neighbors as v, and delete u. This new graph G' does not have a p-clique, but for the edge number we have

$$|E(G')| = |E(G)| + d_v - d_u > |E(G)|,$$

implying that our original graph was not maximal: a contradiction.

Case 2: $d_u \geq d_v$ and $d_u \geq d_w$. We duplicate u twice and delete v and w, once again producing a new graph G' with no p-clique and an edge number of

$$|E(G')| = |E(G)| + 2d_u - (d_v + d_w - 1) > |E(G)|,$$

a contradiction again.

It follows from our claim that two vertices u, v not being adjacent is in fact an equivalence relation:

$$u \sim v \iff uv \notin E(G)$$

is clearly reflexive and symmetric, but the fact that $uv \notin E, uw \notin E$ implies $vw \notin E$ proves transitivity. In other words G is simply a complete multipartite graph under these conditions, so we are done! □

6. Extremal Graph Theory

Extremal graph theory has since been expanded by many mathematicians, including Paul Erdős (a popular quip is that the theory was "developed and loved by Hungarians"). The common theme is the same: optimize parameters like edges & vertices and find connections between constraints on a graph and its resulting structure. We illustrate some of the other seminal results.

We first prove a well-known result about cliques and edge colorings.

Theorem 6.1 (Ramsey)**.** *For any s, t there exists $N = R(m,n)$ such that every red-blue coloring of the edges of K_N has a completely red K_m subgraph or a completely blue K_n subgraph.*

Specifically, we will show not only that $R(m,n)$ exists but also that

$$R(m,n) \leq R(m-1,n) + R(m,n-1).$$

Proof. We start with the simple case $R(m,2) = m$, where either all edges of K_m are red or a blue edge exists giving a blue K_2, and the symmetric $R(2,n) = n$.

Now, if $R(m-1,n)$ and $R(m,n-1)$ exist, then let $N = R(m-1,n) + R(m,n-1)$ and consider an arbitrary red-blue coloring of K_N. For some vertex v, let A be the set of vertices connected by a red edge and let B be the set of vertices connected by a blue edge to v.

Since $|A| + |B| = N - 1$, either $|A| \geq R(m-1,n)$ or $|B| \geq R(m,n-1)$. Suppose $|A| \geq R(m-1,n)$. Then by the definition of $R(m-1,n)$, either there is in A a subset A_R of size $m-1$ with all red edges — which together with v yields a red K_m — or there is a subset A_B of size n with all blue edges. Similar logic holds if $|B| \geq R(m,n-1)$, and it follows that K_N satisfies the (m,n)-property and our recursion relation is true.

Looking at our starting values and recursion, we find the formula

$$R(m,n) \leq \binom{m+n-2}{m-1}$$

satisfies them from the binomial relation $\binom{n}{k} = \binom{n-1}{k} + \binom{n}{k-1}$. For the case of $R(k,k)$ we have

$$R(k,k) \leq \binom{2k-2}{k-1} = \binom{2k-3}{k-1} + \binom{2k-3}{k-2} \leq 2^{2k-3}.$$

□

For over 60 years, this has remained essentially the best lower bound. In general, Ramsey numbers are very hard to find with precision with current techniques. After decades of work, the value of $R(5,5)$ is still only known to be between 43 and 48; in 2017, 49 was eliminated by computer verification, after checking about 2 trillion different cases!

We now turn our attention to finding an upper bound for the Ramsey numbers $R(k,k)$, i.e. finding a number of vertices N as large as possible where there exists a coloring without a red or blue K_k. We use probability methods to do this:

Theorem 6.2. *For all $k \geq 2$, the lower bound*

$$R(k,k) \geq 2^{\frac{k}{2}}$$

holds for the Ramsey numbers.

Proof. We have $R(2,2) = 2$. Our upper bound gives $R(3,3) \leq 6$, and 6 is indeed minimal as shown by the pentagon in 4 which does not satisfy the $(3,3)$-property.

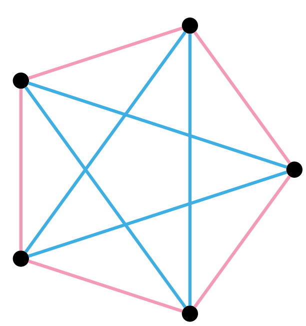

Figure 4. This coloring of K_5 does not contain either a completely red or completely blue triangle.

Now assume that $k \geq 4$. Suppose $N < 2^{\frac{k}{2}}$, and consider all red-blue colorings, where we color each edge independently red or blue with probability $\frac{1}{2}$. This way, all colorings are equally likely with probability $2^{-\binom{N}{2}}$. For some set of vertices A of size k, the probability of the event A_R that all the edges in A are colored red is $2^{-\binom{k}{2}}$. Thus, the probability p_R for at least one k-set to be colored all red is bounded by

$$p_R = P\left(\bigcup_{|A|=k} A_R\right) \leq \sum_{|A|=k} P(A_R) = \binom{N}{k} 2^{-\binom{k}{2}}.$$

Now we apply $N < 2^{\frac{k}{2}}$, $k \geq 4$, and the inequality $\binom{N}{k} \leq \frac{N^k}{2^{k-1}}$ for $k \geq 2$, the latter of which is shown by

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!} \leq \frac{n^k}{2^{k-1}}.$$

Using these we can express the probability as

$$\binom{N}{k} 2^{-\binom{k}{2}} \leq \frac{N^k}{2^{k-1}} 2^{-\binom{k}{2}} < 2^{\frac{k^2}{2} - \binom{k}{2} - k + 1} = 2^{-\frac{k}{2}+1} \leq \frac{1}{2}.$$

Thus, $p_R < \frac{1}{2}$, and by symmetry $p_B < \frac{1}{2}$ for the probability of some k vertices with all edges between them colored blue. Therefore, $p_R + p_B < 1$ for $N < 2^{\frac{k}{2}}$, so there must exist some coloring with no red or blue K_k, meaning K_N does not have the property (k, k). In other words, $R(k, k)$ must be at least $2^{\frac{k}{2}}$. □

> Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find $R(5, 5)$. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded $R(6, 6)$, however, we would have no choice but to launch a preemptive attack. (Erdős)

7. Extensions

Here are two notable ways that Turán's graph theorem has been extended. The proofs of these theorems are relatively long and will be omitted here, but are covered in detail in [2] and [3], with an overview in [4]. To state these results, we first define some common terms and notation in extremal graph theory.

Definition 7.1. The *extremal number* $\mathrm{ex}(n, H)$ is defined to be the maximum number of edges in a graph with n vertices not containing a subgraph isomorphic to H.

Definition 7.2. The *chromatic number* $\chi(H)$ of a graph H is the smallest natural number c such that the vertices of H can be colored with c colors and no two vertices of the same color are adjacent.

Now we proceed to the Erdős–Stone theorem, which bounds the number of edges in a graph that does not contain a subgraph isomorphic to some H.

Theorem 7.3 (Erdős–Stone)**.** *For any fixed graph H,*

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

The $o(1)$ here simply denotes a term that tends towards 0 as $n \to \infty$.

Another interesting extension is the hypergraph Turán problem, which concerns hypergraphs: generalizations of graphs in which edges do not have to connect only 2 vertices, but can join any number of them. These structures are far more complex to analyze, but we can obtain the following bound:

Theorem 7.4. *Let K_s^r denote the complete r-graph on s vertices. (An r-graph is one in which each edge connects r vertices.) Then*

$$\left(1 - \left(\frac{r-1}{s-1}\right)^{r-1} - o(1)\right)\binom{n}{r} \leq \mathrm{ex}\left(n, K_s^r\right) \leq \left(1 - \frac{1}{\binom{s-1}{r-1}} + o(1)\right)\binom{n}{r}.$$

References

[1] Martin Aigner and Günter M Ziegler. *Proofs from the Book*. Springer, 1999.

[2] V. Chvátal and E. Szemerédi. "Notes on the Erdös-Stone Theorem". In: *Combinatorial Mathematics*. Ed. by C. Berge et al. Vol. 75. North-Holland Mathematics Studies. North-Holland, 1983, pp. 183–190. DOI: https://doi.org/10.1016/S0304-0208(08)73386-3. URL: https://www.sciencedirect.com/science/article/pii/S0304020808733863.

[3] Peter Keevash. "Hypergraph turan problems". In: *Surveys in combinatorics* 392 (2011), pp. 83–140.

[4] Asaf Shapira. *Extremal graph theory - tau*. 2016. URL: http://www.math.tau.ac.il/~asafico/ext-graph-theory/notes.pdf.