Turán's Theorem

Theorem 1 (weak Turán's Theorem, 1941)**.** *The maximum number of edges in a graph on n vertices with no $(k+1)$-clique subgraph is at most*

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

Let $T_k(n)$ be a complete k-partite graph on n vertices with parts of as equal sizes as possible, i.e., sizes are $\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$. Another way of defining it would be a balanced blow-up of K_k. Such graph is called the *Turán graph*.

1: Show that $T_k(n)$ gives asymptotically tight lower bound for Turán's theorem.

Solution: It is easy to see that $T_k(n)$ cannot contain a $(k+1)$-clique as any set of $k+1$ vertices in $T_k(n)$ will have two vertices in the same class and therefore not connected by an edge. Furthermore, $e(T_k(n)) \geq \binom{k}{2}(\lfloor \frac{n}{k} \rfloor)^2 > \binom{k}{2}(\frac{n}{k}-1)^2 = (1-\frac{1}{k})\frac{n^2}{2} - O(n)$.

2: Prove Turán's theorem by induction. Idea: Find a clique A of size k and remove it for induction.

Solution: [First proof.] (Turán, 1941) Induction on n. The theorem is trivially true for $n \leq k$, so let $n > k$ and assume the theorem holds for smaller graphs. Let G be a graph on n vertices with no $(k+1)$-clique and the maximum number of edges. Therefore, G must contain a k-clique as otherwise we could add edges to G contradicting maximality. Let A be a clique of size k and let B be the remaining $n-k$ vertices. The graph B has no $(k+1)$-clique so by induction $e(B) \leq (1-\frac{1}{k})\frac{(n-k)^2}{2}$. Furthermore, each vertex in B can have at most $k-1$ neighbors in A, so we have

$$e(G) \leq \binom{k}{2} + \left(1-\frac{1}{k}\right)\frac{(n-k)^2}{2} + (n-k)(k-1) = \left(1-\frac{1}{k}\right)\frac{n^2}{2}.$$

Actually, Turán proved more.

Theorem 2 (Turán's Theorem, 1941)**.** *The maximum number of edges in an n-vertex graph with no $(k+1)$-clique is exactly $e(T_k(n))$. Furthermore, $T_k(n)$ is the unique graph attaining this maximum.*

We show several more proofs that prove one version or the other. Note that $T_k(n)$ is a complete multipartite graph and among complete multipartite graphs with no $(k+1)$-clique is it the largest. This leads to a ingenious approach: if we can show that a graph G with no $(k+1)$-clique and the maximum number of edges is complete multipartite then we are done.

3: Prove $T_k(n)$ is the unique extremal graph by using induction on k. Idea: Take a vertex x of maximum degree, use induction on neighbors of x, and maximality arguments on non-neighbors of x.

Solution: [Second proof.] (Erdős, 1970) Induction on k. The theorem is trivially true for $k=1$, so let $k>1$ and assume the theorem holds for $k-1$. We will prove that if G has no $(k+1)$-clique and the maximum number of edges, then $G = T_k(n)$. Let x be a vertex of maximum degree, let S be the neighbors of x and let $T = G - S$ be the remaining vertices. The graph S has no k-clique (as otherwise we could build a $(k+1)$-clique with x). Let us construct a new graph H on the vertex set of G as follows. The graph H is the same as G on S, it contains all edges between S and T, and it has no edges in T. Observe that all degrees in H are at least as large as in G. Thus H has at least as many edges as G. However, if the graph G has an edge in T, then H would have more edges. Therefore, as G was maximal, T contains no edges in G and thus $G = H$. Because S does not contain a k-clique, H does not contain a $(k+1)$-clique and therefore H is a graph with the maximum number of edges with this property. Furthermore, S must be an edge-maximal graph with no k-clique. Thus, by induction on k, we have $S = T_{k-1}(|S|)$, so H is a complete multipartite graph. The largest (most edges) n-vertex complete multipartite graph with no $(k+1)$-clique is $T_k(n)$.

Another proof is based on the characterization of complete multipartite-graphs by a forbidden subgraph on 3 vertices.

4: Let G be a graph with no three vertices that induce exactly one edge. That is, G does not contain $\overline{P_3}$, so called co-cherry. Show that G is a complete multipartite graph.

Solution: Suppose G does not contain $\overline{P_3}$. If G is a complete graph, we are done. Let x, y be two non-adjacent vertices. Since G does not contain $\overline{P_3}$, all other vertices are partitioned to sets A and B, where A are neighbors of both x and y, while B are non-neighbors of both x and y. Let $a \in A$ and $b \in B$. By considering triple a, b, x, we conclude ab is an edge. Let $b_1, b_2 \in B$ By considering triple b_1, b_2, x, we conclude b_1, b_2 is not an edge. Hence $B \cup \{x, y\}$ is an independent set and there is a complete bipartite graph between $B \cup \{x, y\}$ and A. Hence $B \cup \{x, y\}$ is one part of the multi-partite graph and other parts can be obtained by repeating the same argument on A.

5: Prove $T_k(n)$ is the unique extremal graph by showing the extremal graph must be a complete multilartite graph. Idea: Take a co-cherry and consider degrees of its vertices. Duplicate/erase vertices to get more edges.

Solution: [Third proof.] (Zykov, 1949)[1] Let G be an n-vertex graph with no $(k+1)$-clique and the maximum number of edges. We will show that G is a complete multipartite graph. If G is not multipartite, then there is a pair of non-adjacent vertices x and y and a vertex z such that xz is an edge, but yz is not an edge. If $d(x) > d(y)$, then remove all edges incident to y and connect y to all neighbors of x (but not x itself). The resulting graph has no $(k+1)$-clique but has more edges than G which contradicting maximality. Thus $d(y) \geq d(x)$ and $d(y) \geq d(z)$ (by the same argument for z). Now remove all edges incident to x and z and connect both vertices to the neighbors of y. The resulting graph has more edges than G and has no $(k+1)$-clique again contradicting maximality. Therefore G must be a complete multipartite graph. The largest complete multipartite graph with no $(k+1)$-clique is $T_k(n)$.

[1] Due to WWII, Zykov's work was done without the knowledge of Turán's.

[Fourth proof on Turán's Theorem](Katona-Nemetz-Simonovits, 1965)[2]

6: Denote the number of edges of the Turán graph on n vertices with k classes by

$$t_k(n) = |E(T_k(n))|.$$

Show a technical lemma that

$$(n+1)t_k(n) < (n-1)t_k(n+1) + (n-1)$$

Idea: Write $n = qk + r$ where $r < k$ and precisely calculate $2t_k(n)$ and $t_k(n) = t_k(n+1) - \textit{something}$. Then prove the inequality.

Solution: Put $n = qk + r$ where $r < k$. Because $n > k$ we have either $r > 0$ or $q > 1$. Then $T_k(n)$ has $k - r$ classes of size q and r classes of size $q + 1$. The maximum degree in $T_k(n)$ is $n - q$ and there are $r(q + 1)$ many vertices of degree 1 less, so summing degrees gives

$$2t_k(n) = n(n-q) - r(q+1).$$

Adding a vertex to a class of size q gives $T_k(n+1)$, thus

$$t_k(n) = t_k(n+1) - (n-q).$$

Now we use the above two equalities to get

$$(n+1)t_k(n) = (n-1)t_k(n) + 2t_k(n) = (n-1)t_k(n+1) + (n-q) - r(q+1).$$

Observe that $(n-q)-t(q+1) < n-1$ as long as either $r > 0$ or $q > 1$ which completes the proof.

7: Now for the main proof, let $g_k(n)$ be the maximum number of edges possible in an n-vertex graph with no K_{k+1} subgraph. Obviously $t_k(n) \leq g_k(n)$ we we would like to show $t_k(n) \geq g_k(n)$.
Idea: Proceed by induction on n. Double count pairs (v, e) where v is a vertex and e is an edge NOT incident to v. Use the previous inequality.

Solution: We proceed by induction on n. The assertion is true for $n \leq k + 1$ so let us assume the inequality holds for $n \geq k + 1$ and now show it holds for $n + 1$.

Let G be an $(n+1)$-vertex K_{k+1}-free graph with the maximum number of edges. We will double count the pair (v, e) where v is a vertex and e is an edge NOT incident to v. We can fix v first in $n + 1$ ways and then fix e in at most $g_k(n)$ ways. Alternatively, we can fix e in $g_k(n+1)$ ways and then fix v in exactly $n - 1$ ways. Combining these estimates and applying induction to $g_k(n)$ gives

$$g_k(n+1)(n-1) \leq (n+1)g_k(n) \leq (n+1)t_k(n). \tag{1}$$

To complete the proof it is enough to recall that $(n+1)t_k(n) < (n-1)t_k(n+1) + (n-1)$.

[2] This proof is from the first paper of Nemetz and Simonovits and the second of Katona.