On the Fundamentals of Extremal Graph Theory*

Feng Cheng

1 Introduction

Extremal graph theory emerged and initiated its development in the 1930s and 40s. It deals with maximal and minimal properties of graphs. The Hungarian mathematician Pál Turán was one of the first to study these extremal properties of graphs and became known for his Turán's theorem. Many Hungarian mathematicians such as P. Erdős, B. Bollobás, V. Sós, and E. Szemerédi have been active researchers in this field since then.

In this paper, we will focus on the earliest yet most fundamental result in extremal graph theory, Turán's theorem, proved and published in 1941. First, we will look at Mantel's theorem, a special case of Turán's theorem. Then we will shift our attention to Turán's theorem itself. The paper will produce several different proofs for both theorems. We will concentrate on some extensions alongside these two theorems as well. The properties presented and theorems proved in this paper will help the readers understand how the extremal graph theory was founded and some of its major techniques and ideas.

2 Preliminaries

In this paper, all graphs are taken as simple graphs. If the letter G is not explicitly defined in the proofs and the propositions themselves, then we recognize it as the graph G that includes all its vertices and edges. *Triangles* refer to K_3.

Let G be a graph, and v be a vertex in G. Denote the vertex set of G by $V(G)$, and the edge set of G by $E(G)$ (sometimes abbreviated as V and E). Denote the number of edges in G by $e(G) = |E(G)|$. Denote $N(v)$ as the set of vertices that are adjacent to v. Denote a graph G of order n by G_n. The *complement graph* of G, denoted by $\overline{G}$, is the graph with the same vertex set as G, but with the edge set complement to the edge set of G.

A *clique* is a subset of the vertices of a graph such that its induced subgraph is complete. A k-clique means that the clique has k vertices.

The *degree sequence* of a graph is the sequence of the degree of each vertex (usually written in nonincreasing order) in the graph.

An *independent set* of vertices is defined as a set of vertices that are pairwise nonadjacent.

3 Mantel's theorem

THEOREM 1 (Mantel). *A simple graph G with order n with no triangles has at most $\lfloor n^2/4 \rfloor$ edges.*

*This paper was produced under the supervision of Professor Carl Yerger (Davidson College) during the Pioneer Research Program that ran from early July to early October 2019. Revised in 2024.

Proof I (from maximality). Suppose v_1 is the vertex in G with maximum degree d, and let the vertices that are the neighbors of v_1 be

$$v_{n-d+1}, v_{n-d+2}, \ldots, v_n.$$

Because G doesn't contain any K_3, meaning that any two of the vertices listed above are not neighbors, the number of edges in G satisfies

$$e(G) \leqslant \sum_{i=1}^{n-d} d(v_i) \leqslant (n-d)d \leqslant \frac{(n-d+d)^2}{4} = \frac{n^2}{4}.$$

Since $e(G)$ must be an integer, G has at most $\lfloor n^2/4 \rfloor$ edges.

To satisfy the equality, v_i ($i = 1, 2, \ldots, n-d$) must not contain any pairwise neighbor, and all of these v_i must be of degree d, which satisfies the equation $n - d = d$ (indicating that d can be either $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$ as n might not be even). Since $v_{n-d+1}, v_{n-d+2}, \ldots, v_n$ do not contain any pairwise neighbor as well, G is a bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. □

As we can see from Proof I, it is possible to open up the proof if we focus on one single vertex, it is worth the attempt to focus on a pair of vertices that are adjacent to each other and form an edge.

Proof II (induction). The theorem can be verified easily when $n = 3$ and 4.

Suppose the case is true for $n < k$ ($k > 4$). Then for $n = k$, suppose $v_i v_j$ is an edge in G. Since K_3 is not contained in the graph, v_i and v_j do not have a common neighbor. Thus $d(v_i) + d(v_j) \leqslant n$, and v_i and v_j have at most $n - 1$ edges that are incident with either of them. If we remove v_i and v_j, we get a new graph with $n - 2$ vertices, which according to the inductive hypothesis should have at most $\frac{1}{4}(n-2)^2$ edges. Hence

$$e(G) \leqslant \frac{1}{4}(n-2)^2 + n - 1 = \frac{n^2}{4}.$$

The equality is satisfied only when $d(v_i) + d(v_j) = n$, which means that $N(v_i) \cup N(v_j) = V(G)$. Also note that both $N(v_i)$ and $N(v_j)$ should be independent sets since there is no K_3 in G.

Keep in mind when $n = 3$ and 4, G is a bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. For every $n \geqslant 5$, when n increases by 2, G will increase the cardinalities of the two independent subsets by 1 respectively. Hence the number of edges reaches its maximum if and only if G is a bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ (for every n). □

The next proof uses the notion $d(v_i) + d(v_j) \leqslant n$ as well, but this time we expand this inequality to the whole context and take a more algebraic approach.

Proof III (summation and transformation). As mentioned at the beginning of Proof II, $d(v_i) + d(v_j) \leqslant n$ is true for every edge $v_i v_j$. The inequality

$$\sum_{v_i v_j \in E} \big(d(v_i) + d(v_j)\big) \leqslant en \tag{1}$$

follows. Meanwhile the LHS

$$\sum_{v_i v_j \in E} \big(d(v_i) + d(v_j)\big) = \sum_{v \in V} d^2(v). \tag{2}$$

By the Cauchy-Schwarz inequality, we then have

$$\sum_{v \in V} d^2(v) \geqslant \frac{1}{n}\Big(\sum_{v \in V} d(v)\Big)^2 = \frac{1}{n} 4e^2.$$

Combining the above with (1) and (2), we get $e \leqslant n^2/4$, or more precisely.

When n is even, the equality is satisfied if and only if for every edge $v_i v_j$, $d(v_i) + d(v_j) = n$, and for all $v \in V$, $d(v)$ is the same. We need to be a bit careful when n is odd, since $\lfloor n^2/4 \rfloor < n^2/4$. We are then returned to $N(v_i) \cup N(v_j) = V(G)$ in Proof 2, and we get the same condition that G reaches its maximum only when $G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. □

Proof IV takes a more unique approach towards Mantel's theorem. It intuitively builds up a function and attempts to obtain the result through an inequality regarding the function. The core step in this proof is the use of local adjustment, a method applied to solve some complex inequalities. To put it briefly, it shifts some elements in the function once at a time, which gradually increases or decreases the value of the function.

Proof IV (weight and local adjustment). Let $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ be a vector satisfying $\sum x_i = 1$, with $0 \leqslant x_i \leqslant 1$ for each $i \in [n]$. Define

$$f(x) = \sum_{ij \in E} x_i x_j.$$

We first observe that if all x_i's are of weight $1/n$, then $f(x) = e/n^2$. This means that there must exist an x that satisfies $f(x) \geqslant e/n^2$. Note that if $ij \notin E$, shifting the weight of x_i to that of x_j (assuming the total weight of the neighbors of x_j is at least as large as the total weight assigned to the neighbors of x_i) will not decrease $f(x)$, as

$$x_i \sum_{pi \in E} x_p \leqslant x_j \sum_{qj \in E} x_q. \tag{3}$$

Repeating the "weight-shifting" procedure described above, it is clear that $f(x)$ may always increase until all the weights concentrate on a clique (by definition its induced graph is complete). Note there is no triangles, hence the weights may only concentrate on two vertices. Hence

$$\frac{e}{n^2} \leqslant f(x) \leqslant \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \tag{4}$$

which tells that $e \leqslant \lfloor n^2/4 \rfloor$. □

Remark (on Proof IV). The equality is satisfied only when the two "$\leqslant$" in (4) reaches equality. Therefore every shift of weight (3) is (almost) always an equality, which means that every vertex should have a similar number of neighbors. Yet it is still hard to give the exact equality condition. This is a major drawback of this last proof.

Remember that in Mantel's theorem, we ask about the result when there is no triangle in the graph G. It is equally important in extremal graph theory to explore the result that involves the number of existing triangles.

Here we introduce notations that will be useful. Denote by $k_r(G)$ the number of K_r's in graph G. Denote by $(d_i)_1^n$ be the degree sequence of graph G with order n. Let adjacent edges be the number of pairs of edges that share exactly one vertex.

The degree sequence of G is $(\bar{d}_i)_1^n = (n - 1 - d_i)_n^1$, and thus on one hand we get $\sum_{i=1}^n \binom{d_i}{2}$ pairs of adjacent edges in G and $\sum_{i=1}^n \binom{n-1-d_i}{2}$ pairs of adjacent edges in $\overline{G}$. On the other hand, each of the $k_3(G) + k_3(\overline{G})$ number of triangles in the two graphs G and G' contains three pairs of adjacent edges. The remaining number of three-vertices is

$$\binom{n}{3} - k_3(G) - k_3(\overline{G}),$$

where each three-vertices contains exactly one pair of adjacent edges. Therefore we have

$$\sum_{i=1}^{n}\binom{d_i}{2} + \sum_{i=1}^{n}\binom{n-1-d_i}{2} = 3k_3(G) + 3k_3(\overline{G}) + \binom{n}{3} - k_3(G) - k_3(\overline{G}).$$

Now the total number of triangles $k_3(G) + k_3(\overline{G})$

$$\begin{aligned}
&= \frac{1}{2}\Bigg[\sum_{i=1}^{n}\binom{d_i}{2} + \sum_{i=1}^{n}\binom{n-1-d_i}{2} - \binom{n}{3}\Bigg] \\
&= \frac{1}{2}\sum_{i=1}^{n}\left[\frac{1}{2}d_i(d_i-1) + \frac{1}{2}(n-1-d_i)(n-2-d_i)\right] - \frac{1}{2}\binom{n}{3} \\
&= \frac{1}{2}\sum_{i=1}^{n}\left[d_i^2 - d_i - (n-2)d_i + \frac{1}{2}(n-1)(n-2)\right] - \frac{1}{2}\binom{n}{3} \\
&= \sum_{i=1}^{n}\binom{d_i}{2} - (n-2)\cdot e(G) + \left(\frac{1}{4} - \frac{1}{12}\right)n(n-1)(n-2) \\
&= \sum_{i=1}^{n}\binom{d_i}{2} - (n-2)\cdot e(G) + \binom{n}{3}, \qquad (5)
\end{aligned}$$

which should also be equal to

$$\sum_{i=1}^{n}\binom{\bar{d}_i}{2} - (n-2)\cdot e(\overline{G}) + \binom{n}{3} \qquad (6)$$

by symmetry. We will now use e and $\bar{e}$ to denote the numbers of edges in G and $\overline{G}$, respectively.

Formula (5) above (see also [Bol78]) implies a lower bound on the number of triangles, which we will present here:

THEOREM 2 (Goodman [Goo59; Bol78]). *For a graph G of order n, we have $k_3(G) + k_3(\overline{G}) \geqslant \frac{1}{24}n(n-1)(n-5)$.*

Proof. Essentially we want to bound (5) from below. By the AM-GM inequality, we know

$$\sum_{i=1}^{n}\binom{d_i}{2} \geqslant n\binom{2e/n}{2},$$

with "=" when G is an $(2e/n)$-regular graph. Therefore

$$\begin{aligned}
k_3(G) + k_3(\overline{G}) &\geqslant e\left(\frac{2e}{n} - 1\right) - (n-2)e + \binom{n}{3} \\
&= \binom{n}{3} - \frac{2e(n^2 - n - 2e)}{2n} \\
&\geqslant \frac{1}{6}n(n-1)(n-2) - \frac{1}{8}n(n-1)^2 \\
&= \frac{1}{24}n(n-1)(n-5).
\end{aligned}$$

The lower bound is attained only when G is a $(2e/n)$-regular graph and $n^2 - n = 4e$. Note $n^2 - n = 4e$ means that

$$n - 1 = 2 \cdot \frac{2e}{n}.$$

We then know that both G and $\overline{G}$ are $(2e/n)$-regular graphs. Meanwhile n must be odd. Now given $n(n-1) = 4e$ and $\gcd(n, n-1) = 1$, we have that the lower bound is attained only when $n \equiv 1 \pmod 4$. □

There exists a sharper lower bound for the number of triangles. It was first introduced and proved in [Goo59], and was given a simpler proof (using weighting) in [Sau61].

THEOREM (Goodman refined). *For a graph* G *of order* n*,*

$$k_3(G) + k_3(\overline{G}) \geqslant \begin{cases} \frac{1}{24}n(n-2)(n-4), & \text{if } n \equiv 0 \pmod 2; \\ \frac{1}{24}n(n-1)(n-5), & \text{if } n \equiv 1 \pmod 4; \\ \frac{1}{24}(n+1)(n-3)(n-4), & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

Note that the bounds above are sharp for all three cases.

We now present a second theorem on bounding the number of triangles.

THEOREM 3 (Moon and Moser [MM62; Bol78]). *A graph* G *of order* n *and size* e *contains at least* $\frac{e}{3n}(4e - n^2)$ *triangles.*

Proof. First we have

$$3k_3(\overline{G}) \leqslant \sum_{i=1}^{n} \binom{\bar{d}_i}{2} \tag{7}$$

because each triangle contains three adjacent pairs of edges, while each adjacent pair of edges may not be in the same triangle. With (6) and the above inequality (7) we have

$$\begin{aligned}
k_3(G) &\geqslant \sum_{i=1}^{n} \binom{\bar{d}_i}{2} - (n-2)\bar{e} + \binom{n}{3} - \frac{1}{3}\sum_{i=1}^{n} \binom{\bar{d}_i}{2} \\
&= \frac{2}{3}\sum_{i=1}^{n} \binom{\bar{d}_i}{2} - (n-2)\left[\binom{n}{2} - e\right] + \binom{n}{3} \\
&\geqslant \frac{2}{3}n\binom{2\left[\binom{n}{2} - e\right]/n}{2} - (n-2)\left[\binom{n}{2} - e\right] + \binom{n}{3} \quad \text{by AM-GM inequality} \\
&= \frac{e}{3n}(4e - n^2)
\end{aligned}$$

after a significant amount of simplification in the final step. The lower bound is attained if and only if G is a regular graph. □

4 Turán's theorem

Now we turn back to Mantel's theorem. We will now generalize it from K_3 to K_n. The result is known as **Turán's theorem**, which we will focus on for the rest of our paper. Before that we again introduce some relevant definitions and notations.

Definition. Given a graph $G = (V, E)$ and a vertex subset S of V, the *subgraph* of G *induced by* S, denoted by $\langle S \rangle$, is an induced subgraph of G with vertex set S and edge set $\{uv : u, v \in S \text{ and } uv \in E\}$.

For two vertex sets X and Y, $e(X, Y)$ is defined to be the number of edges that have one vertex in X and another vertex in Y.

If two graphs G and H are isomorphic, we use the notation $G \cong H$.

The simple complete k-partite graph on n vertices in which all parts are as equal in size as possible (i.e., of sizes $\lceil n/k \rceil$ or $\lfloor n/k \rfloor$) is called a *Turán graph* and is denoted by $T_{k,n}$. Also denote the number of edges in Turán graph $T_{k,n}$ by $t_{k,n} = e(T_{k,n})$.

The origin of Turán's theorem [Sim13] is related to *Ramsey theory*, which investigates the smallest order of graph that ensures every 2-coloring of $E(G)$ using red and blue makes G contain either a red K_i or a blue K_j. Another way of interpreting Ramsey theory is that, assuming graph G of order n contains no i-clique, what is the greatest j such that a j-independent set must exist in G. In 1941 [Tur41], Turán replaced j-independent set with the number of edges $e(G)$.

THEOREM 4 (Turán). *Let G be a simple graph of order n that does not contain any k-clique ($k \geqslant 2$). Then $e(G) \leqslant e(T_{k-1,n})$, with "=" if and only if $G \cong T_{k-1,n}$.*

Turán's theorem is the first theorem of the following form:

If G_n contains no subgraph S, then G_n has at most $f(n)$ edges.

Problems of this form become known as the *Turán type problems* in graph theory (see [EG59; Dir63]). The $f(n)$ can be denoted by $\mathrm{ex}(n, S)$, where S may be a single *forbidden graph* (such as the k-clique K_k in Turán's theorem), or a family of forbidden graphs. Some famous problems of this type include *Zarankiewicz problem* [Zar51] (which considers the bipartite graph) and the *even circuit theorem* [Erd65] (which considers the cycle).

As marked by Turán in his original paper, we can restate Theorem 4 by replacing $e(T_{k-1,n})$ with

$$\frac{1}{2}\left(1 - \frac{1}{k-1}\right)(n^2 - r^2) + \binom{r}{2},$$

where r is the smallest positive integer that satisfies $n \equiv r \pmod{k-1}$. To see why this is true, consider Figure 1.

By the definition of r, there are $k-1-r$ parts of $\frac{n-r}{k-1}$ vertices and r parts of $\left(\frac{n-r}{k-1} + 1\right)$ vertices. First, we remove the additional r vertices (at the top) and see all parts to be of $\frac{n-r}{k-1}$ vertices. This means that $T_{k-1,n}$ contains $\binom{k-1}{2}\left(\frac{n-r}{k-1}\right)^2$ edges, since every time we may pick two columns of vertices and form $\left(\frac{n-r}{k-1}\right)^2$ edges, and we need to repeat this for $\binom{k-1}{2}$ number of times. Second, we look at how the additional r vertices form edges with the *existing* $n - r$ vertices. Each vertex from these additional r vertices is connected to the $(k-2)$

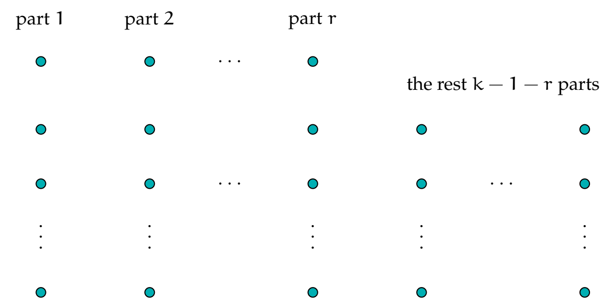

Figure 1: explicit upper bound on the number of edges in Turán's theorem

parts, whence there are an additional $r(k-2)\frac{n-r}{k-1}$ edges. Third, we know that the additional r vertices give $\binom{r}{2}$ new edges within themselves. Hence in total we have

$$\frac{1}{2}\Big(1-\frac{1}{k-1}\Big)(n^2-r^2)+\binom{r}{2} \tag{8}$$

number of edges.

We remark that the number of edges is maximized with value $\frac{(k-2)n^2}{2(k-1)}$ when $(k-1) \mid n$, i.e., $r=0$. We may replace the $e(T_{k-1,n})$ with this expression, but this is weaker than the original Turán's theorem.

The first proof of Turán's theorem is the the most known one. It looks at the vertex with the maximum degree in G (maximal properties often bring out interesting results in combinatorics). It uses induction that is simpler than the one that will be used in the second proof that comes later. Before that we present a lemma.

LEMMA 5. *The complete* k*-partite graph on* n *vertices with the largest number of edges is the Turán graph* $T_{k,n}$.

Proof. Let the number of vertices in each of the k parts be $p_1, p_2, \dots, p_k$. Then the number of edges in the graph will be

$$\sum_{1\leqslant i\leqslant j\leqslant k} p_i p_j = \frac{1}{2}\Bigg[\Big(\sum_{i=1}^{k} p_i\Big)^2 - \sum_{i=1}^{k} p_i^2\Bigg]. \tag{9}$$

Cauchy-Schwarz inequality tells us that

$$n\sum_{i=1}^{k} p_i^2 \geqslant \Big(\sum_{i=1}^{k} p_i\Big)^2 = n^2. \tag{10}$$

with "=" when $p_1 = p_2 = \cdots = p_k$.

Plugging (10) back into (9), we see the number of edges is less than or equal to $\frac{1}{2}(n^2-n)$, and achieves this maximum when all parts of the k-partite graph are of the equal size. As n might not be divisible by k, it is not hard to claim that only when the size of each part is either $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$ (as equal in size as possible) will the total number of edges be maximized. We do not write the full detail here. □

We are now ready to prove Turán's theorem.

Proof I of Turán's theorem, modified upon [Zyk49; BM08]. Obviously the theorem holds for $k=2$. Assume it holds for all integers less than k, and consider G that does not contain a k-clique. Choose the vertex x of maximum degree Δ in G, and set $X = N(x)$ and $Y = V - X$. It is easy to see that

$$e(G) = e(X) + e(X, Y) + e(Y).$$

Since G does not contain any K_k, $\langle X \rangle$ does not contain any K_{k-1}. (If not, the additional vertex x and the K_{k-1} in $\langle X \rangle$ will form a K_k in G.) Therefore by the inductive hypothesis we obtain

$$e(X) \leqslant e(T_{k-2,\Delta})$$

with equality only when $\langle X \rangle \cong T_{k-2,\Delta}$. Note that each edge incident with vertices in Y belongs to $E(X, Y)$ and $E(Y)$, which implies that

$$e(X, Y) + e(Y) \leqslant \Delta(n - \Delta),$$

with equality only when Y is an independent set and all vertices of Y is of degree Δ. Therefore $e(G) \leqslant e(H)$, where H is the combination of $T_{k-2,\Delta}$ and an independent set with $n - \Delta$ vertices, each of which is connected to all the Δ vertices in X. We see that H is a complete $(k-1)$-partite graph on n vertices. By Lemma 5, we know that $e(H) \leqslant e(T_{k-1,n})$, with equality if and only if $H \cong T_{k-1,n}$. Therefore we get $e(G) \leqslant e(T_{k-1,n})$, with equality if and only if $G \cong T_{k-1,n}$. □

The next proof is from Turán's original paper. It uses a completely different method from Zykov's proof above. It uses induction, which results from a meticulous observation that every additional $k-1$ verteics will be individually distributed into the existing $k-1$ parts of vertices, according to the definition of Turán graph.

Observe that Theorem 4 obviously holds for $n \leqslant k-1$, if we allow empty sets in the partition. Also note that if the graph G that contains no K_k has the maximal number of edges, it must have at least one K_{k-1}. (Say a new edge ℓ is added to G so that the new graph G' now contains K_k, then removing ℓ from K_k gives us two K_{k-1}.)

Proof II, modified upon Turán's proof [Tur41; Aig95]. Suppose the theorem holds for graphs of order less than n, where $n \geqslant k-1$. Let G be the graph of order n with the maximum number of edges. Let A be the vertices of a $(k-1)$-clique in G and let B be $V - A$. It follows that $|B| = n - k + 1$.

Since G has no k-clique, every vertex in B has at most $k-2$ adjacent vertices in A. If we maximize the number of edges in $\langle B \rangle$ and simultaneously let every vertex in B adjacent to exactly $k-2$ vertices in A, then the number of edges in G is maximized. Also note that all the vertices of each $(k-1)$-clique in B must not have the same neighbor in A, or there will be a K_k in G.

We can only realize the condition above with the following arrangement. By the inductive hypothesis let $B = T_{k-1,n-k+1}$ and add a distinct vertex of A to each of the $k-1$ parts of B. The resulting graph is exactly $T_{k-1,n}$.

By induction we know that if and only if $G \cong T_{k-1,n}$ will the number of edges in G be maximized. □

An immediate consequence of Turán's theorem is the upper bound on the *minimum degree* of G, denoted by $\delta(G)$.

COROLLARY 6 ([Bol78]). *If G_n does not have K_k as a subgraph, then $\delta(G_n) \leqslant \frac{k-2}{k-1}n$.*

Proof. By the weak version of Turán's theorem on page 7, we know that $e(G_n) \leqslant \frac{(k-2)n^2}{2(k-1)}$. It follows that

$$n \cdot \delta(G_n) \leqslant \sum_{v \in V} d(v) = 2e(G_n) \leqslant \frac{(k-2)n^2}{k-1},$$

which gives the desired bound. □

Turán's theorem determines a bound on the number of edges, but it also yields bounds to the clique number, the chromatic number, and the independence number, which are defined as follows for a graph G.

Definition. The *clique number*, denoted by $\omega(G)$, is the order of the largest complete graph that is a subgraph of G.

The *chromatic number*, denoted by $\chi(G)$, is the minimum integer such that G is $\chi(G)$-colorable.

The *independence number*, denoted by $\alpha(G)$, is the largest size of an independent set of vertices in G.

THEOREM 7 (lower bounds on $\omega(G), \chi(G)$, and $\alpha(G)$ [Wes01]). *For a graph G of order n and size e, it follows that $\omega(G) \geqslant \lceil \frac{n^2}{n^2-2e} \rceil$, $\chi(G) \geqslant \lceil \frac{n^2}{n^2-2e} \rceil$, and $\alpha(G) \geqslant \lceil \frac{n^2}{n+2e} \rceil$.*

Proof. From the weaker version of Turán's theorem we know that $e \leqslant \left(1 - \frac{1}{\omega(G)}\right) \cdot \frac{n^2}{2}$, which gives a lower bound on the clique number:

$$\omega(G) \geqslant \left\lceil \frac{1}{1 - \frac{2e}{n^2}} \right\rceil = \left\lceil \frac{n^2}{n^2 - 2e} \right\rceil =: p.$$

For the bound above to be sharp, we need the number of edges e in G to be at least $\binom{p}{2}$ and $T_{p,n}$ to have at least e edges. With these two conditions, we can remove edges from $T_{p,n}$ and get a graph G of order n and size e such that $\chi(G) = \omega(G) = p$.

Let $\bar{e} = e(\overline{G}) = \binom{n}{2} - e$. Then $\alpha(G) = \omega(\overline{G}) \geqslant \frac{n^2}{n^2-2\bar{e}} = \frac{n^2}{n+2e}$. □

Remark. Turán's theorem is related to vertex coloring problems. The Turán graph is the solution to finding the graph with the maximum number of edges that avoid needing k colors for vertices: every part in the Turán's graph may be colored with one distinct color. This is an opposite problem to finding the chromatic number, which means we need to find the smallest number of colors that give a proper coloring to the graph.

5 Extensions and alternative proofs of Turán's theorem

For the final section we will explore some extended properties between partite graphs, cliques, and extremal properties of a graph. These results will lead to two alternative proofs of Turán's theorem.

The third proof of Turán's theorem requires a nicely formulated result on the degree sequence on k-partite graphs, provided by Erdős.

THEOREM 8 ([Erd70; Bol78; Sim13]). *Let G be a graph of order n and vertex set $V = \{x_1, x_2, \ldots, x_n\}$ (where the vertices are indexed in ascending degrees) that contains no K_p. Then there exists a $(p-1)$-partite graph G' [which is a $(p-1)$-chromatic graph] with vertex set $V' = \{x'_1, x'_2, \ldots, x'_n\}$ again indexed in ascending degrees, such that*

$$d(x_i) \leqslant d(x'_i) \quad \text{for} i = 1, 2, \ldots, n.$$

Proof. We apply induction to p. When $p = 2$, the claim is trivially true. Suppose the case holds for all $2 \leqslant i \leqslant k$, and we consider the case when $i = k+1$. We pick the vertex with the *highest degree* v from G and use the letter N to denote the set of all its neighbors. The induced subgraph $\langle N \rangle$ must contain no K_{p-1}, or otherwise $\langle N \cup \{v\} \rangle$ would contain a K_p. Thus if we apply the inductive hypothesis to $\langle N \rangle$, we have a $(p-2)$-partite graph $\widetilde{N}$ on $|N|$ vertices that satisfies the claim. We connect each of the $n - |N|$ vertices left to all vertices of $\widetilde{N}$, which we call G'. G' is a $(p-1)$-partite graph, and by our choice of v, G' meets the claim desired. □

A simply inequality now gives Turán's theorem. However we do not readily know the exact equality condition from this approach, which we point out as a drawback.

Proof III of Turán's theorem. Combine Theorem 8 and Lemma 5, we have

$$2e(G) = \sum_{i=1}^{n} d(x_i) \leqslant \sum_{i=1}^{n} d(x_i') = 2e(G') \leqslant 2e(T_{p-1,n}).$$ □

[Dir63] mentions a few straightforward results connected to Turán's theorem. The results in the following pages are due to him, and we remind that

$$t_{k-1}(n) = \frac{1}{2}\Big(1 - \frac{1}{k-1}\Big)(n^2 - r^2) + \binom{r}{2}$$

that will be useful.

THEOREM 9. *Let k, n, α be integers such that $n \geqslant k+1 \geqslant 4$ and $\alpha \leqslant 1$. Any graph on n vertices and at least $t_{k-1}(n) + \alpha$ number of edges contains at least one subgraph with n' vertices and at least $t_{k-1}(n') + \alpha$ edges for every $n' = k, k+1, \ldots, n-1$.*

Proof. For the graph with more than $t_{k-1}(n) + \alpha$ edges, we may delete some edges and convert into a subgraph with exactly $t_{k-1}(n) + \alpha$ edges, which we denote by H.

Let us rewrite n by $(k-1)t + r$, where $1 \leqslant r \leqslant k-1$. Therefore

$$t_{k-1}(n) = \frac{1}{2}(k-1)(k-2)t^2 + (k-2)tr + \frac{1}{2}r(r-1), \tag{11}$$

which gives

$$t_{k-1}(n) - t_{k-1}(n-1) = (k-2)t + r - 1 = n - t - 1. \tag{12}$$

We may look at the problem downwards from $n-1$ to k. To accomplish this we need to prove the existence of a vertex with degree less than or equal to $n-t-1$, since this will ensure that once this vertex is removed, the induced subgraph with $n-1$ vertices will face the same situation as before.

We will now show that in our setup there exists at least one vertex in H that is of degree $\leqslant n-t-1$. Suppose all vertices are of degree $\geqslant n-t$. Then we have

$$\begin{aligned} e(H) \geqslant \frac{1}{2}n(n-t) &= \frac{1}{2}[(k-1)t + r][(k-2)t + r] \\ &= \frac{1}{2}(k-1)(k-2)t^2 + \frac{1}{2}(2k-3)tr + \frac{1}{2}r^2, \end{aligned}$$

which combined with (11) gives us

$$e(H) - t_{k-1}(n) \geqslant \frac{1}{2}tr + \frac{1}{2}r = \frac{1}{2}r(t+1).$$

When $t = 1$, $r \geqslant 2$ because $(k-1)t + r = n \geqslant k+1$; and when $t \geqslant 2$, $r \geqslant 1$. This implies that $e(H) - t_{k-1}(n) \geqslant 2$, contradicting the setup that $e(H) = t_{k-1}(n) + \alpha \leqslant t_{k-1}(n) + 1$.

Therefore there is one vertex of degree $\leqslant n - t - 1$ in H, which we call x. We may delete this x from H and obtain the induced subgraph $\langle V(H) - x\rangle$, which has at least $t_{k-1}(n) + \alpha - (n - t - 1)$ number of edges. Now by (12), we have

$$e\big(\langle V(H) - x\rangle\big) \geqslant t_{k-1}(n-1) + \alpha,$$

meaning that H, a subgraph of the original graph, contains a subgraph of order $n-1$ and size at least $t_{k-1}(n-1) + \alpha$. We have thus found the desired subgraph with $n-1$ number of vertices.

We may now repeat the same steps above and find the desired subgraph with $n-2$, $n-3$, and all the way down to k number of vertices. □

THEOREM 10. *Let k, n, q, α be integers such that $k \geqslant 3, 1 \leqslant q \leqslant k-1, n \geqslant k+q-1$, and $\alpha \leqslant 1$. Any graph G_n with n vertices and at least $t_{k-1}(n) + \alpha$ edges contains one K_{k+q-1} with $q - \alpha$ edges missing as a subgraph.*

Proof. Let G_n be a graph with n vertices and $t_{k-1}(n) + \alpha$ edges.

The theorem holds when $q = 1$ (i.e., $n = k$). Here $t_{k-1}(n) = \frac{1}{2}k(k-1) - 1$. This shows that the original graph G_n has at least $t_{k-1}(n) = \frac{1}{2}k(k-1) - (1-\alpha)$ edges, meaning exactly that it does contain a subgraph K_k with $1 - \alpha$ edges missing.

Now consider the case when $2 \leqslant q \leqslant k-1$. By Theorem 9 we know that G_n contains a subgraph H_{k+q-1} with $k+q-1$ vertices and $t_k(k+q-1) + \alpha$ edges. If we let $r = q$ and $t = 1$ in (11), we get

$$\begin{aligned} t_{k-1}(k+q-1) &= \frac{1}{2}(k-1)(k-2) + (k-2)q + \frac{1}{2}q(q-1) \\ &= \frac{1}{2}(k+q-1)(k+q-2) - q. \end{aligned}$$

Therefore H_{k+q-1} contains $\frac{1}{2}(k+q-1)(k+q-2) - (q-\alpha)$ edges, which proves that the subgraph is a K_{k+q-1} with $q - \alpha$ edges missing. □

If we replace α by 1 and every q individually by integers between 1 and $k-1$, we have an extension result of Turán's theorem.

COROLLARY 11. *Assume that $e(G_n) > t_{k-1}(n)$, where $n \geqslant k \geqslant 3$. Then for every q such that $1 \leqslant q \leqslant \min\{k-1, n-k+1\}$, G_n contains a K_k, a K_{k+1} with 1 edge missing,..., and a K_{k+q-1} with $q-1$ edges missing.* □

This gives the final proof of Turán's theorem, and concludes our expository paper on some fundamental results of extremal graph theory.

Proof IV [Dir63]. The corollary above tells us that under $n \geqslant k \geqslant 3$, if G_n does not contain a k-clique, then $e(G_n) \leqslant t_{k-1}(n)$. This is exactly Turán's theorem. (Of course the case $k = 2$ is trivially true.) □

References

[Aig95] Martin Aigner. “Turan’s Graph Theorem”. *The American Mathematical Monthly* 102 (1995), p. 808.

[BM08] J.A. Bondy and U.S.R. Murty. “Graph Theory”. Graduate Texts in Mathematics. Springer, 2008, pp. 301–305.

[Bol78] Béla Bollobás. “Extremal graph theory”. L.M.S. monographs. Academic Press, 1978. Chap. 6.

[Dir63] G. Dirac. “Extensions of Turán’s theorem on graphs”. *Acta Mathematica Academiae Scientiarum Hungaricae* 14 (1963), pp. 417–422.

[EG59] P. Erdős and T. Gallai. “On maximal paths and circuits of graphs”. *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959), pp. 337–356.

[Erd65] P. Erdős. “Extremal problems in graph theory”. *Theory of Graphs and its Applications*. Ed. by M. Fiedler. Vol. 10. New York: Academic Press, 1965, pp. 29–36.

[Erd70] P. Erdős. “On the graph theorem of Turán”. *Matematikai És Fizikai Lapok* 21 (1970), pp. 249–251.

[Goo59] A.W. Goodman. “On Sets of Acquaintances and Strangers at any Party”. *The American Mathematical Monthly* 66 (1959), pp. 778–783.

[MM62] J.W. Moon and L. Moser. “On a problem of Turán”. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 7 (1962), pp. 283–286.

[Sau61] Léopold Sauvé. “On Chromatic Graphs”. *The American Mathematical Monthly* 68 (1961), pp. 107–111.

[Sim13] Miklós Simonovits. “Paul Turán’s influence in combinatorics”. *Number Theory, Analysis, and Combinatorics*. De Gruyter, 2013, pp. 309–392.

[Tur41] P. Turán. “On an extremal problem in graph theory”. *Matematikai És Fizikai Lapok* 48 (1941), pp. 436–452.

[Wes01] Douglas B. West. “Introduction to Graph Theory”. 2nd Edition. Upper Saddle River, N.J: Prentice Hall, 2001. Chap. 5.

[Zar51] K. Zarankiewicz. “P 101”. *Colloquium Mathematicum* 2 (1951), p. 301.

[Zyk49] Alexander Aleksandrovich Zykov. “On some properties of linear complexes”. *Matematicheskii sbornik* 66 (1949), pp. 163–188.

Remark. If there are multiple references at the same time in the paper, the first one is where the original source comes from.