ELEMENTS OF EXTREMAL GRAPH THEORY

ANDREI ILIAŞENCO[1)] and MARCEL TELEUCĂ[2)]

Abstract. In this paper we present and prove several properties from extremal graph theory. We also discuss several illustrative applications.
Keywords: Graph, Extremal Graph Theory
MSC : Primary 05C90

1. Introduction and Notations

A graph G on n vertices is *complete* if for any two vertices of G there exists an edge of G connecting these vertices. The complete graph on n vertices is denoted by K_n and has $\dfrac{n(n-1)}{2}$ edges. A graph G is *bipartite* if there exists a partition of its set of vertices $V(G) = A \cup B$, $A \cap B = \varnothing$, such that every edge connects some vertex in A to some vertex in B. A bipartite graph is *complete bipartite* if it contains all edges $\{a, b\}$ such that $a \in A$ and $b \in B$. If $|A| = p$ and $|B| = q$, the complete bipartite graph is denoted by $K_{p,q}$ and has pq edges.

Definition 1. *A complete bipartite graph $K_{1,n}$ is called a star graph.*

Definition 2.*Let G be a graph on n vertices $F_1, F_2, \ldots, F_n$. A graph $D(G)$ is called the doubling of the graph G if the following conditions hold*

i) *The graph $D(G)$ is on $2n$ vertices $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_n$.*

ii) *If there exists an edge F_iF_j in the graph G, there exist edges A_iB_j and A_jB_i in the graph $D(G)$.*

Definition 3. *A graph that does not contain complete subgraphs on k vertices is called a k-free graph.*

2. The Applications

2.1. Some Starting Problems

Let us begin with a simple problem.

Problem 1. Find the maximal number of edges of a graph G on m vertices, which does not contain star subgraphs on n vertices (not necessarily induced subgraphs).

Solution. If $n > m$, there can be no subgraph on n vertices. The maximal number of edges of G is thus $\dfrac{m(m-1)}{2}$.

If $m \geq n$, suppose the degree of a vertex A is at least $n-1$. Then the vertex A and any $n-1$ of the vertices adjacent to A form a star subgraph on n vertices. Suppose then the degree of every vertex is at most $n-2$. Then the graph G obviously does not contain any star subgraphs on n vertices. Thus a graph contains a star subgraph on n vertices if and only if there exists a vertex of degree at least $n-1$.

Is it always possible to find a graph with degree of any vertex at most $n-2$? Arrange the vertices in a circle. Let us consider 3 cases.

[1)] Lyceum „Orizont“, Chişinău, Republic of Moldova, andrei.iliasenco@gmail.com
[2)] State University of Tiraspol, Republic of Moldova, mteleuca@gmail.com

Case 1. n is even. Join any vertex A to $\frac{n-2}{2}$ vertices left to A and $\frac{n-2}{2}$ vertices right to A. The degree of any vertex will be equal to $n-2$. Therefore the maximal number of edges will be equal to $\frac{(n-2)m}{2}$.

Case 2. n is odd and m is even. Join any vertex A to $\frac{n-3}{2}$ vertices left to A and $\frac{n-3}{2}$ vertices right to A. Because m is even, for every vertex A there is an opposite vertex on the circle. So join any vertex A to its opposite vertex.

Case 3. n is odd and m is odd. If the degree of every vertex would be equal to $n-2$, the sum of degrees of all vertices would be equal to $(n-2)m$. This number is odd, but it must be equal $2\times$ number of vertices. Thus all vertices can't have degree $n-2$. Case 1 implies that we can add edges to the graph in such a way that every vertex will have degree $n-3$. Recall that $n-3 \le m-3 < m-1$. Let's assign numbers to all vertices on the circle. Then because of procedure of joining vertices in case 1, no vertex with number k will be connected to the vertex with number $k+\frac{m-1}{2}$. If we connect the vertices with numbers $1, 2, \ldots, \frac{n-1}{2}$ to the vertices with the numbers $1+\frac{n+1}{2}, \ldots, n$, we will get the graph with all vertices of degree $n-2$ and one vertex of degree $n-3$. This means that the maximal number is equal to $\frac{(n-2)m-1}{2}$.

Therefore the answer is equal to $\max\left(\frac{m(m-1)}{2}, \left[\frac{(n-2)m}{2}\right)\right]$. ■

We continue with a more difficult problem.

Problem 2. Find the maximal number of edges of a graph G on n vertices, which does not contain triangles (K_3).

Solution. We claim the maximal number of edges is equal to $\left[\frac{n^2}{4}\right]$. The proof is by induction on the number n of vertices. For $n=1$ and $n=2$ the result is trivial.

Suppose $n > 2$. Assume the graph has an edge AB. Since the subgraph determined by the other $n-2$ vertices does not contain triangles, using the induction hypothesis we find that the number of its edges is at most $\left[\frac{(n-2)^2}{4}\right]$. Let C be a vertex different from A and B. Because the graph G does not contain triangles, it contains at most one of the edges AB and AC. Thus the number of edges with one end in $\{A, B\}$ and with the other different from A and B is at most $n-2$. Note that we must also count the edge AB. Thus the number of edges of the graph G is at most

$$\left[\frac{(n-2)^2}{4}\right] + (n-2) + 1 = \left[\frac{(n-2)^2 + 4(n-2) + 4}{4}\right] = \left[\frac{n^2}{4}\right].$$

The graph $K_{\left[\frac{n}{2}\right], \left[\frac{n+1}{2}\right]}$ is an example of a graph with $\left[\frac{n^2}{4}\right]$ edges. ■

2.2. Theoretical Results

The above problem is a particular case of *Turán*'s theorem. To prove this useful theorem we need *Zarankiewicz*'s lemma.

Lemma (Zarankiewicz). *If G is a k-free graph, then there exists a vertex having degree at most* $\left[\dfrac{k-2}{k-1}n\right]$.

Proof. Assume the converse. To prove the statement we need some notations. Consider an arbitrary vertex V_1. Denote by A_i the set of vertices adjacent to V_i. Let x be equal to $\left[\dfrac{k-2}{k-1}n\right]$. By the assumption, we get:

$$|A_1| > \left[\frac{k-2}{k-1}n\right] > 0.$$

Hence there exists $V_2 \in A_1$. By the assumption $|A_i| > 1 + x$ for any i. Furthermore we have:

$$|A_1 \cap A_2| = |A_1| + |A_2| - |A_1 \cup A_2| \geq 2(1+x) - n > 0.$$

Hence there exists $V_3 \in A_1 \cap A_2$. Then we get:

$$|A_1 \cap A_2 \cap A_3| = |A_1 \cap A_2| + |A_3| - |(A_1 \cap A_2) \cup A_3| \geq$$
$$\geq (2(1+x) + (1+x)) - n = 3(1+x) - n > 0.$$

Continuing in the same way we see that the for any i

$$\left|\bigcap_{i=1}^{j} A_i\right| \geq j\left(1 + \left[\frac{k-2}{k-1}n\right]\right) - (j-1)n.$$

For $i = k-1$:

$$\left|\bigcap_{i=1}^{k-1} A_i\right| \geq (k-1)\left(1 + \left[\frac{k-2}{k-1}n\right]\right) - (k-2)n > 0.$$

Therefore, there exists a vertex $V_k \in \bigcap\limits_{i=1}^{k-1} A_i$.

In this case the vertices $V_1, V_2, .., V_k$ form a complete graph with k vertices. Hence the graph G is not k-free. The contradiction completes the proof.

Now let's prove *Turán*'s theorem.

Theorem (Turán). *The maximal number of edges of a k-free graph with n vertices is*

$$\left[\frac{k-2}{k-1} \cdot \frac{n^2 - r^2}{2} + \binom{r}{2}\right],$$

where r is the remainder of n when divided by $k-1$.

Proof. The proof is by induction on n. Assume the result true for all k-free graphs having at most $n-1$ vertices.

Using *Zarankiewicz*'s lemma we can pick a vertex with at most $\left[\dfrac{k-2}{k-1}n\right]$ adjacent vertices. The subgraph determined by another $n-1$ vertices is k-free by the inductive assumption. Hence the number of edges is at most

$$\left[\frac{k-2}{k-1}n\right] + \frac{k-2}{k-1} \cdot \frac{(n-1)^2 - s^2}{2} + \binom{s}{2},$$

where s is the remainder of $n-1$ when divided by $k-1$. Let us consider two cases: $s = k-2$, $r = 0$; and $r = s+1$.

For both cases we check that:

$$\left[\frac{k-2}{k-1}n\right] + \frac{k-2}{k-1} \cdot \frac{(n-1)^2 - s^2}{2} + \frac{s^2 - s}{2} = \frac{k-2}{k-1} \cdot \frac{n^2 - r^2}{2} + \frac{r^2 - r}{2}.$$

This completes the proof. ■

2.3. Some More Problems We continue with a similar problem.

Problem 3. Find the maximal number of edges of a graph that does not contain two triangles with a common edge.

Ssolution The complete bipartite graph $K_{[\frac{n}{2}],[\frac{n+1}{2}]}$ satisfies the conditions of the problem. Hence we want to prove that the maximal number is equal to $\left[\frac{(n-1)^2}{4}\right]$. We can rewrite the condition of the problem in the following way: for any 4 points the number of edges connecting these points is at most 4. It is easy to see that this condition is equivalent to the condition of problem.

The proof is by induction on n. For $n = 1, 2, 3, 4$ there is nothing to prove. Suppose that for any graphs with a number of edges smaller than n the inductive assumption was proved. Let us consider two cases.

Case 1. There are no 4 points in the graph having exactly 4 edges connecting them. In this case we can add at least one more edge to the graph while the condition remains satisfied. Hence the number of edges is not maximal.

Case 2. There are 4 points A, B, C, D such that the edges AB, BC, CD, DA are connecting them. Therefore the number of edges connecting these points is equal to 4. By inductive assumption we have that the number of edges connecting the other $n-4$ points is equal to $\left[\frac{(n-4)^2}{2}\right]$. Let's find the maximal number of edges connecting one of the points A, B, C, D to one of the other points.

Let E be a vertex different from A, B, C, D. Suppose that E is adjacent to at least three points from the set $\{A, B, C, D\}$. Without loss of generality we may assume that these points are A, B and C. Then there are 5 edges joining the points A, B, C, E: EA, EB, EC, AB and BC. Thus we get a contradiction.

It follows that a point E different from A, B, C and D cannot be adjacent to more then two points from the set $\{A, B, C, D\}$.

Therefore the maximal number is equal to:

$$\left[\frac{(n-4)^2}{4}\right] + 2(n-4) + 4 = \left[\frac{(n-4)^2 + 2 \cdot 4 \cdot (n-4) + 4^2}{4}\right] = \left[\frac{n^2}{4}\right].$$

This completes the proof. ■

The following problem was proposed at IMO 2003.

Problem 4. Let A be a 101-element subset of the set

$$S = \{1, 2, \ldots, 1000000\}.$$

Prove there exist numbers $t_1, t_2, \ldots, t_{100}$ in S such that the sets:

$$A_j = \{x + t_j \mid x \in A\}, \quad j = 1, 2, \ldots, 100$$

are pairwise disjoint.

Solution Let $a_1 < a_2 < \cdots < a_{101}$ be the elements of A.

Let G be a graph with S the set of its vertices, having an edge between vertices i and j if and only if the sets $B = \{a_1 + i, a_2 + i, \dots, a_{101} + i\}$ and $C = \{a_1 + j, a_2 + j, \dots, a_{101} + j\}$ are disjoint. The graph has 1000000 vertices. For an arbitrary vertex k to be joined with l, $k - l$ should not be equal to any of the numbers $a_i - a_j$ $(i \geq j)$. Obviously, there are $101 \cdot 100$ of the numbers $a_i - a_j$ $(i \geq j)$. Thus any vertex x has degree at least $1000000 - 101 \cdot 100$.

Therefore the graph G has at least $\dfrac{1000000(1000000 - 101 \cdot 100)}{2}$ edges. The problem asks to prove that there is a 100-clique in the graph. By *Turán*'s theorem, there is a k-clique in a graph with n vertices if the number of edges is strictly greater than

$$M(n, k) = \frac{k-2}{k-1} \cdot \frac{n^2 - r^2}{2} + \frac{r(r-1)}{2},$$

where we have taken r to be the remainder of n when divided by $k - 1$. In our case $n = 1000000$, $k = 100$ and $r = 1$. We can check that the number of edges $\dfrac{1000000^2}{2} - \dfrac{100 \cdot 101 \cdot 1000000}{2}$ is greater than $M(1000000, 100) = \dfrac{98}{99} \cdot \dfrac{999999999999}{2}$ and thus we are done. ■

We continue with a problem from BMO 2008 Shortlist that can be solved using Turán's theorem.

Problem 5. In some country there are $n \geq 5$ cities operated by two airline companies. Every two cities are operated in both directions by at most one of the companies. The government introduces a restriction that all round trips that a company can offer should have at least six cities. Prove that there are no more than $\left[\dfrac{n^2}{3}\right]$ flights offered by these companies.

Solution. Consider the graph G with n vertices representing connections between them operated by two companies. The condition of the problem is equivalent to the fact that there do not exist circuit subgraphs C_3, C_4, C_5. Assume to the contrary that there are at least $\left[\dfrac{n^2}{3}\right] + 1$ edges in the graph G. Using *Turán*'s theorem we conclude there exists a complete subgraph $K_4 = \{A_1, A_2, A_3, A_4\}$ of the graph G, with all its edges colored in two colors (blue and red). As there are no circuit subgraphs C_3, C_4, C_5 in G, the only possible coloring is the following: the edges A_1A_2, A_2A_3, A_3A_4 are colored blue and A_1A_3, A_2A_3, A_2A_4 are colored red.

First of all we prove that we get contradiction for $n = 5, 6, 7, 8$. Extract from the graph G four vertices A_1, A_2, A_3, A_4 of the subgraph K_4 and observe that each of the remaining $n - 4$ vertices has at most two connections with these 4 vetices. If there will be three connections than two of them will be of the same colour and they together will form with vertices of K_4 will form the subgraphs C_3, C_4 of C_5. There are at most $\dfrac{(n-4)(n-5)}{2}$ edges between $n - 4$ remaining vertices. Thus there are in total at most:

$$6 + 2(n-4) + \frac{(n-4)(n-5)}{2}$$

edges of the graph G. But we can check that for $5 \leq n \leq 8$:

$$6 + 2(n-4) + \frac{(n-4)(n-5)}{2} \leq \frac{n^2}{3}.$$

So the statement is true for $n = 5, 6, 7, 8$.

Now we prove it by mathematical induction, using the above result as a base case. We apply the same idea. We assume the contrary and find the subgraph K_4 whose existence ensured by *Thuran*'s theorem. By the induction hypothesis there will be no more than $\frac{n^2}{4}$ edges between the remaining $n-4$ vertices. Thus in the graph G there will be at most $6 + 2(n-4) + \frac{(n-4)^2}{3}$ edges.

But for $n > 8$:

$$6 + 2(n-4) + \frac{(n-4)^2}{3} \leq \frac{n^2}{3} \Leftrightarrow 6n - 12 + (n-4)^2 \leq n^2 \Leftrightarrow 4 \leq 2n.$$

We get a contradiction and thus we are done. ■

Further we will prove a lemma which can be used in many problems and at the same time is interesting in itself.

Lemma. *Let G be a graph with v vertices $a_1, a_2, \ldots, a_v$. Suppose it does not contain a complete bipartite subgraph $K_{m,n}$. Denote by d_i the degree of the vertex a_i, and denote by d the arithmetic mean of the degrees of all vertices. Then the following inequalities hold*

$$\sum_{i=1}^{v} \prod_{j=0}^{m-1} (d_i - j) \leq (n-1) \prod_{j=0}^{m-1} (v - j); \qquad \prod_{j=0}^{m-1} (d - j) \leq (n-1) \prod_{j=0}^{m-1} (v - j).$$

Proof. i) Consider $D(G)$ - the doubling of the graph G. We claim a graph G contains a bipartite compete subgraph of the form $K_{m,n}$ if and only if the doubling of the graph also contains a complete bipartite graph of the form $K_{m,n}$.

Denote by $F_1, F_2, \ldots, F_v$ the vertices of the graph G and denote by $A_1, A_2, \ldots, \ldots, A_n$, $B_1, B_2, \ldots, B_n$ the vetices of the graph $D(G)$. Suppose that graph G contains a complete bipartite subgraph with the points F_{i_1},F_{i_2}, …, F_{i_n}, $F_{j_1}, F_{j_2}, \ldots, F_{j_n}$. Hence the points $F_{i_1}, F_{i_2}, \ldots, F_{i_n}$ are adjacent to the points $F_{j_1}, F_{j_2}, \ldots, F_{j_n}$. By the definition of doubling graph the points A_{i_1}, $A_{i_2}, \ldots, A_{i_n}$ are adjacent to the points $B_{i_1}, B_{i_2}, \ldots, B_{i_n}$. Thus we get a complete bipartite graph. Similarly we can prove the inverse.

Let S_i be the set of vertices adjacent to the vertex B_i. Then S_i has d_i elements. It follows that the doubling of the graph G doesn't contain complete bipartite graphs of the form $K_{m,n}$. Because we don't want to get a complete bipartite graph of this form in the $D(G)$ the number of the sets S_i containing all m arbitrary vertices $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$ must be at most $n-1$. We can choose these m arbitrary vertices in $\binom{v}{m}$ ways. Hence the number of possible $(m+1)$-tuples $(S_i, A_{i_1}, A_{i_2}, \ldots, A_{i_m})$, where S_i contains the points $A_1, A_2, \ldots, A_n$, is at most $(n-1)\binom{v}{m}$. On the other hand S_i contains d_i vertices. Hence it was counted $\binom{d_i}{m}$ times. Therefore the number of $(m+1)$-tuples $(S_i, A_{i_1}, A_{i_2}, \ldots, A_{i_n})$ is equal to $\sum \binom{d_i}{m}$. Therefore we have the inequality:

$$\sum \binom{d_i}{m} \le (n-1)\binom{v}{m}.$$

Multiplying both sides by $m!$ we get:

$$\sum d_i(d_i-1)\cdots(d_i-m+1) \le (n-1)v(v-1)(v-2)\cdots(v-m+1).$$

ii) Suppose that $P(x) = x(x-1)\cdots(x-m+1)$ For $d \le m-1$ the inequality is obvious. Let's calculate the minimal value of the expression $ddsumP(d_i)$. We shall prove that the minimum is achieved when all variables are equal. The inequality written in the form $\sum P(d_i) \ge vP(d_i)$, or:

$$\frac{P(d_1)+\cdots+P(d_v)}{v} \ge P\left(\frac{d_1+\cdots+d_v}{v}\right)$$

is similar to *Jensen* inequality. But the function $P(x)$ isn't convex. It is however convex for $x \ge m-1$. Let's try to replace the numbers $d_1, d_2, \ldots, d_v$ with another numbers in order to get rid of the numbers that are smaller than $m-1$. At the same time we would like to keep the sum of the numbers unchanged.

Suppose that d_i and d_j are numbers such that $d_i < m-1 \le d < d_j$. Let's observe that $P(1) = P(2) = \cdots = P(m-1) = 0$. The function $P(x)$ is increasing for $x \ge m-1$. Hence we get:

$$P(d_i) + P(d_j) = P(d_j) \ge P(d_j + d_i - m + 1) = P(d_i + d_j - m + 1) + P(m-1).$$

It follows that we can replace the numbers d_i and d_j with the numbers at least equal to $m-1$: $d_i + d_j - m + 1$ and $m-1$. We observe that the sum of the numbers didn't change. Continuing in the same way we will get a sequence of the numbers bigger than $m-2$. We can apply *Jensen* inequality to them.

The following two problems are applications of the above result.

Problem 6. Suppose that the two parts of a complete bipartite graph have w and $v-w$ vertices respectively. Suppose that this graph doesn't contain complete bipartite subgraphs of the form $K_{2,n}$. Prove that the number of edges is at most $\dfrac{v+\sqrt{v^2+8(n-1)(v-2)w(v-w)}}{4}$.

Solution. Denote by $d_1, d_2, \ldots, d_w$ the degrees of first part's vertices and denote by S their sum. Obviously the sum of degrees of first part's vertices is equal to the sum of degrees of second part's vertices. Using the lemma we get:

$$\sum_{i=1}^{w} d_i(d_i-1) \le (n-1)(v-w)(v-w-1) \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^{w} d_i^2 - S - (n-1)(v-w)(v-w-1) \le 0.$$

It follows from a well-known inequality that: $\dfrac{\sum_{i=1}^{w} d_i^2}{w} \ge \dfrac{S}{w^2}$.

Using this result we get the following inequality:

$$S^2 - wS - (n-1)w(v-w)(v-w-1) \le 0.$$

By considering the second part we get a similar inequality:

$$S^2 - (v-w)S - (n-1)w(v-w)(w-1) \leq 0.$$

Adding these two inequalities we get: $2S^2 - vS - (n-1)vw(v-2) \leq 0$. Solving this inequality we get: $S \leq \dfrac{v + \sqrt{v^2 + 8(n-1)(v-2)w(v-w)}}{4}$.

The number of edges is half the sum of degrees of vertices of the graph, but the sum of degrees of vertices of a part is also half the sum of degrees of vertices of the graph. It means that the number of edges is equal to the sum of degrees of vertices of a part. Thus the number of edges is equal to S. It completes the proof of the problem. ■

Problem 7. Suppose that a graph with v vertices doesn't contain a complete bipartite graphs of the form $K_{2,m}$. Prove the inequality:

$$\frac{v\left(1 + \sqrt{1 + 4(n-1)(v-1)}\right)}{4}.$$

Solution. Using the lemma we get an inequality for the average degree of the vertices:

$$d(d-1) \leq (n-1)(v-1), \qquad d^2 - d - (n-1)(v-1) \leq 0.$$

Using this inequality we get $d \leq \dfrac{1 + \sqrt{1 + 4(n-1)(v-1)}}{2}$.

Because the number of edges is equal to $\dfrac{vd}{2}$, we are done. ■

REFERENCES

[1] M. Bălună, M. Becheanu, ..., *Zece lecţii alese de matematică elementară*, S. S. M. R. (1998).
[2] R. Diestel, *Graph Theory*, Springer-Verlag (1997).
[3] D. Djukić, V. Janković, I. Matić, N. Petrović, *The IMO Compendium.*
[4] I. Tomescu], *Probleme de combinatorică şi teoria grafurilor.*
[5] ∗ ∗ ∗, *Problems of the Sankt-Petersburg Mathematical Olympiad* (2001).