Graphs and Combinatorics 7, 31–37 (1991)

Graphs and Combinatorics

On the Maximal Number of Certain Subgraphs in K_r-Free Graphs

Ervin Györi, János Pach, and Miklós Simonovits

Mathematical Institute of the Hungarian Academy of Sciences, P.O.B. 127, 1364 Budapest, Hungary

Dedicated to Paul Turán on his 80th Birthday

Abstract. Given two graphs H and G, let $H(G)$ denote the number of subgraphs of G isomorphic to H. We prove that if H is a bipartite graph with a one-factor, then for every triangle-free graph G with n vertices $H(G) \le H(T_2(n))$, where $T_2(n)$ denotes the complete bipartite graph of n vertices whose colour classes are as equal as possible. We also prove that if K is a complete t-partite graph of m vertices, $r > t$, $n \ge max(m, r-1)$, then there exists a complete $(r-1)$-partite graph G^* with n vertices such that $K(G) \le K(G^*)$ holds for every K_r-free graph G with n vertices. In particular, in the class of all K_r-free graphs with n vertices the complete balanced $(r-1)$-partite graph $T_{r-1}(n)$ has the largest number of subgraphs isomorphic to K_t $(t < r)$, C_4, $K_{2,3}$. These generalize some theorems of Turán, Erdös and Sauer.

1. Introduction

Let $T_{r-1}(n)$ denote the complete $(r-1)$-partite graph with n vertices whose colour classes are as equal as possible, i.e., each class contains either $\left\lfloor \frac{n}{r-1} \right\rfloor$ or $\left\lceil \frac{n}{r-1} \right\rceil$ vertices. Turán's well-known theorem [8, 9] states that every K_r-free graph G with n vertices contains at most as many edges as $T_{r-1}(n)$ does. Furthermore, if G is different from $T_{r-1}(n)$, then its number of edges $e(G)$ is strictly smaller than $e(T_{r-1}(n))$.

In this paper we consider the following extension of this problem. Given a graph H, and two natural numbers r and n, what is the maximum number of subgraphs isomorphic to H a K_r free graph with n vertices can have? (Notice that if $K_r \not\subseteq H$ then the order of magnitude of this maximum is obviously $cn^{|V(H)|}$. So we are interested either in a sharper asymptotic formula or in an exact result.) Turán's theorem settles the special case when H is a single edge.

To formulate our results, we shall need some notation. For any two graphs H and G, let $\overline{H(G)}$ denote the number of different embeddings $\varphi\colon V(H) \to V(G)$ such that

$$\text{(i) } v_1 \ne v_2 \Rightarrow \varphi(v_1) \ne \varphi(v_2),$$

$$\text{(ii) } v_1 v_2 \in E(H) \Rightarrow \varphi(v_1)\varphi(v_2) \in E(G)$$

for every pair $v_1, v_2 \in V(H)$.

Let $\overline{H(G)}$ denote the number of subgraphs of G isomorphic to H. Evidently, $\overline{H(G)}/H(G)$ is equal to the number of automorphisms of H, provided that $H(G) \neq 0$. Hence, in any class of graphs $\mathcal{G}$, $\overline{H(G)}$ and $H(G)$ attain their maxima for the same $G \in \mathcal{G}$.

Theorem 1. *Let H be a bipartite graph with $m \geq 3$ vertices, containing $\lfloor m/2 \rfloor$ independent edges.*

Then, for every triangle free graph G with $n > m$ vertices, $H(G) \leq H(T_2(n))$, and equality holds if and only if $G \simeq T_2(n)$.

In particular, it follows that in the class of all triangle-free graphs of n vertices $T_2(n)$ contains the largest number of subgraphs isomorphic to P_k (the path of length k), C_{2k} (the cycle of length $2k$), $T_2(k)$ etc. The problem of maximizing the number of odd cycles is radically different (cf. [3, 6]).

Let $K_{r-1}^{(1)}$ and $K_{r-1}^{(2)}$ be two complete subgraphs of a graph H, $|V(K_{r-1}^{(1)})| = |V(K_{r-1}^{(2)})| = r-1$. We call them *adjacent*, if $|V(K_{r-1}^{(1)}) \cap V(K_{r-1}^{(2)})| = r-2$. We say that the $(r-1)$*-skeleton of H is connected*, if for any two vertices $v_1, v_2 \in V(H)$ there is a sequence $K_{r-1}^{(1)}, \ldots, K_{r-1}^{(s)}$ of complete subgraphs of H such that $v_1 \in K_{r-1}^{(1)}$, $v_2 \in K_{r-1}^{(s)}$, and $K_{r-1}^{(i)}$ and $K_{r-1}^{(i+1)}$ are adjacent for every $1 \leq i < s$. The following assertion is a straightforward generalization of Theorem 1.

Theorem 2. *Let $r \geq 3$, and let H be an $(r-1)$-partite graph with $m \geq r-1$ vertices, containing $\lfloor m/(r-1) \rfloor$ vertex disjoint complete subgraphs of $r-1$ vertices. Suppose further that the $(r-1)$-skeleton of each component of H is connected.*

Then, for every K_r-free graph G with n vertices, $H(G) \leq H(T_{r-1}(n))$, and equality holds if and only if $G \simeq T_{r-1}(n)$.

In particular, we obtain that, for every $r-1 \leq k \leq n$, in the class of all K_r-free graphs with n vertices $T_{r-1}(n)$ contains the largest number of subgraphs isomorphic to $T_{r-1}(k)$.

For the more general problem, when we wish to maximize the number of subgraphs isomorphic to a given complete t-partite graph whose classes may have different sizes, we can prove the following.

Theorem 3. *Let K be a complete t-partite graph of m vertices, and let $r > t$, $n \geq max(m, r-1)$ be arbitrary integers.*

Then there exists a complete $(r-1)$-partite graph G^* with n vertices such that, for every K_r-free graph G with n vertices, $K(G) \leq K(G^*)$. Furthermore, if $n \geq m+1$, then $\max_G K(G)$ is attained for complete $(r-1)$-partite graphs only.

Remark 1. The graph G^* in Theorem 3 is not necessarily balanced. In fact, the ratio of the sizes of its smallest and largest classes is not even bounded. Indeed, let K be the complete bipartite graph $K_{a,b}$ whose colour classes are of size a and b, respectively, and let $r = 3$. Then $K_{m,n-m}$ will contain $c_{a,b}(m^a(n-m)^b + m^b(n-m)^a) + O(n^{a+b-1})$ copies of $K_{a,b}$. Hence, if $(a-b)^2 > a+b$, then $T_2(n)$ is clearly not optimal. Furthermore if $a, b \to \infty$, $a \gg b$, then for the optimal $K_{m,n-m}$ we have $m \sim \dfrac{a}{a+b} n$.

Remark 2. Theorem 3 cannot be generalized to every t-partite graph, because one can easily construct a bipartite graph K for which no (complete) bipartite graph G^* can be optimal. An easy example can be obtained by taking two disjoint stars $K_{1,a-2}$ and joining their centres by a path of length 3. Any bipartite graph G contains at most

$$K(G) \le \binom{\frac{n}{2}}{a-2}^2 \frac{n^4}{16} + O(n^{2a-1}) = \frac{1}{2^{2a}(a-2)!^2} n^{2a} + O(n^{2a-1})$$

copies of K. Let us divide now a set of n points into 5 classes $C_0, C_1, \ldots, C_4$, and join every vertex in C_i to every vertex in $C_{i+1} (\text{mod } 5)$. If $|C_0| = \left(1 - \frac{2}{a}\right)n$, $|C_1| = \cdots = |C_4| = \frac{n}{2a}$ and a is sufficiently large, then we obtain a graph G_1 for which $K(G_1)$ is much larger than the above upper bound for the maximum of $K(G)$ over all bipartite graphs G with n vertices.

Remark 3. If $n = m$, then we may have other extremal graphs that are not $r-1$-partite, as well. For example, let $r = 4$, $n = m = 6$, $K = K_{3,3}$. Then, it is easy to show that in the class of all K_4-free graphs G with 6 vertices $\max K_{3,3}(G)$ is attained for $K_{3,3}$, $K_{3,3}$ plus an edge, and $K_{3,2,1}$.

Theorem 3 implies that to determine $\max_{|V(G)|=n} K(G)$ and the extremal graphs is equivalent to maximizing certain polynomials. We mention three particular cases.

Corollary 4. [4, 5, 7] *For $t < r$ and for every K_r-free graph G with $n \ge r-1$ vertices, $K_t(G) \le K_t(T_{r-1}(n))$ and equality holds if and only if $G \simeq T_{r-1}(n)$.*

Proof. The statement is trivial for $n = r-1$. By Theorem 3, for $n > r-1 \ge t$, the extremal graphs are of the form $K_{n_1,n_2,\ldots,n_{r-1}}$. The number of K_t's in $K_{n_1,n_2,\ldots,n_{r-1}}$ is $\sum_{1 \le i_1 < i_2 < \cdots < i_t \le r-1} n_{i_1} n_{i_2} \ldots n_{i_t}$, which is maximal if and only if the n_i's are as equal as possible, i.e. $K_{n_1,n_2,\ldots,n_{r-1}} \simeq T_{r-1}(n)$. □

Corollary 5. *For every K_r-free graph G with $n \ge \max(r-1, 5)$ vertices, $C_4(G) = K_{2,2}(G) \le C_4(T_{r-1}(n))$, and equality holds if and only if $G \simeq T_{r-1}(n)$.*

Proof. The extremal graphs are of the form $K_{n_1,\ldots,n_{r-1}}$. The number of $K_{2,2}$'s in $K_{n_1,\ldots,n_{r-1}}$ is

$$\binom{n}{4} + 2 \sum_{1 \le i_1 < i_2 < i_3 < i_4 \le r-1} n_{i_1} n_{i_2} n_{i_3} n_{i_4} - \sum_{i=1}^{r-1} \left(\binom{n_i}{4} + \binom{n_i}{3}(n - n_i) \right) \qquad (1)$$

The first sum is maximal if and only if the n_i's are as equal as possible. We show that the second sum S is minimal in the same case.

Assume that the n_i's are chosen so that S is minimal, but there are some indices i, j such that $n_i < n_j - 1$. Then increasing n_i and decreasing n_j by 1, S will change by

$$\binom{n_i+1}{4}-\binom{n_i}{4}+\binom{n_j-1}{4}-\binom{n_j}{4}+\binom{n_i+1}{3}(n-n_i-1)-\binom{n_i}{3}(n-n_i)$$
$$+\binom{n_j-1}{3}(n-n_j+1)-\binom{n_j}{3}(n-n_j)$$
$$=\binom{n_i+1}{3}(n-n_i-1)-\binom{n_i}{3}(n-n_i-1)+\binom{n_j-1}{3}(n-n_j)-\binom{n_j}{3}(n-n_j)$$
$$=(n-n_i-1)\binom{n_i}{2}-(n-n_j)\binom{n_j-1}{2}$$
$$=(n-n_i-n_j)\left[\binom{n_i}{2}-\binom{n_j-1}{2}\right]+(n_j-1)\binom{n_i}{2}-n_i\binom{n_j-1}{2}$$
$$=(n-n_i-n_j)\left[\binom{n_i}{2}-\binom{n_j-1}{2}\right]+\frac{1}{2}n_i(n_j-1)(n_i-n_j+1)<0. \qquad \square$$

The following corollary can be proved quite similarly.

Corollary 6. *For every K_r-free graph G with $n \geq \max(r-1, 6)$ vertices, $K_{2,3}(G) \leq K_{2,3}(T_{r-1}(n))$, and equality holds if and only if $G \simeq T_{r-1}(n)$.*

Many related questions are discussed in [1, 2].

2. Proof of Theorem 1

A bipartite graph H is said to have the *T-property* (the *strong T-property*) if, for any natural number $n \geq |V(H)|$, and for any triangle-free graph G with n vertices,

$$H(G) \leq H(T_2(n)),$$

(and equality holds if and only if $G \simeq T_2(n)$).

Using this terminology, our Theorem 1 states that any bipartite graph H having a perfect matching (or $(|V(H)|-1)/2$ independent edges if $|V(H)|$ is odd) has the strong T-property.

Lemma 2.1. *Let H be a bipartite graph, all of whose connected components H_1, $H_2, \ldots, H_k$ have the T-property. Assume that each H_i, except possibly the last one, consists of two equal colour classes.*

Then H has the T-property. Furthermore, if H_1 has the strong T-property, then H has the strong T-property, too.

Proof. It is more convenient to estimate $\overline{H(G)}$, the number of different embeddings of H into G. Set $|V(H_i)| = m_i$. Embedding the connected components of H successively, by our assumptions we obtain

$$\overline{H(G)} \leq \prod_{i=1}^{k} \overline{H_i\left(T_2\left(n - \sum_{j<i} m_i\right)\right)} = \overline{H(T_2(n))}.$$

If H_1 has the strong T-property, then equality can hold only for $G \simeq T_2(n)$. □

Let I_k denote the graph consisting of k independent edges, and let I_k^+ denote the graph obtained from I_k by adding an isolated vertex.

Corollary 2.2. *For every natural number k, the graphs I_k and I_k^+ have the strong T-property.*

Proof. Turán's theorem states that I_1 has the strong T-property. The graph consisting of a single vertex obviously has the T-property. Hence we can apply the previous lemma. □

In view of Lemma 2.1, it is sufficient to prove Theorem 1 in the special case when H is connected. Let G be any triangle-free graph with n vertices.

Assume first that $m = |V(H)| = 2k$, and let $a_1 b_1, \ldots, a_k b_k \in E(H)$ be a perfect matching of H. According to Corollary 2.2, there are $\overline{I_k(G)} \leq \overline{I_k(T_2(n))}$ injections $\varphi\colon \{a_1, \ldots, a_k, b_1, \ldots, b_k\} \to V(G)$ such that $\varphi(a_i)\varphi(b_i) \in E(G)$ for every i. Two such injections φ_1 and φ_2 are called *equivalent*, if

(i) $\varphi_1(a_1) = \varphi_2(a_1)$, and

(ii) $\{\varphi_1(a_i), \varphi_1(b_i)\} = \{\varphi_2(a_i), \varphi_2(b_i)\}$ for every $1 \leq i \leq k$.

In every equivalence class there are exactly 2^{k-1} elements. However, due to the fact that H is connected and G has no triangles, each class contains at most one embedding of H into G, i.e., one injection φ satisfying

$$\varphi(x)\varphi(y) \in E(G) \quad \text{for every} \quad xy \in E(H).$$

Thus,

$$\overline{H(G)} \leq 2^{1-k}\overline{I_k(G)} \leq 2^{1-k}\overline{I_k(T_2(n))} = \overline{H(T_2(n))}.$$

as required. Since I_k has the *strong* T-property, $\overline{H(G)} = \overline{H(T_2(n))}$ if and only if $G \simeq T_2(n)$.

Suppose next that $m = |V(H)| = 2k + 1$. Let $\{a_0, a_1, \ldots, a_k\}$ and $\{b_1, \ldots, k_k\}$ be the colour classes of H, and assume without loss of generality that $a_i b_i \in E(H)$ for every $1 \leq i \leq k$. There are $\overline{I_k^+(G)} \leq \overline{I_k^+(T_2(n))}$ injections $\varphi\colon \{a_0, \ldots, a_k, b_1, \ldots, b_k\} \to V(G)$ such that $\varphi(a_i)\varphi(b_i) \in E(G)$ for every $1 \leq i \leq k$. Two such injections φ_1 and φ_2 are now called *equivalent*, if

(i) $\varphi_1(a_0) = \varphi_2(a_0)$, and

(ii) $\{\varphi_1(a_i), \varphi_1(b_i)\} = \{\varphi_2(a_i), \varphi_2(b_i)\}$ for every $1 \leq i \leq k$.

Each equivalence class has 2^k elements, and it follows just like in the previous case that at most one of them can be an embedding of H into G, as a subgraph. Hence,

$$\overline{H(G)} \le 2^{-k}\overline{I_k^+(G)} \le 2^{-k}\overline{I_k^+(T_2(n))} = \overline{H(T_2(n))}$$

with equality if and only if $G \simeq T_2(n)$. □

3. Proof of Theorem 3

The proof is based on the symmetrization method of Zykov [10]. We split the proof into a series of steps. A graph G will be called *extremal* if $K(G) = \max K(G')$, where the maximum is taken over all K_r-free graphs G' with n vertices.

Lemma 3.1. *There is a complete s-partite extremal graph for some $s \le r - 1$.*

Proof. Suppose that G is an extremal graph containing the maximum number of pairs $\{u, v\}$ of nonadjacent vertices such that $N(u) = N(v)$, where $N(w)$ denotes the set of neighbours of w. We prove that G is a complete s-partite graph for some s, i.e., $N(u) = N(v)$ for any nonadjacent vertices u, v.

Assume that G contains some nonadjacent vertices x and y such that $N(x) \ne N(y)$. Let a, b, c denote the number of K's in G containing x and y, containing x but not containing y, containing y but not containing x, respectively.

Suppose first that $b \ne c$, say, $b > c$. It is clear that deleting the edges incident to y and joining y to the neighbours of x, we obtain another K_r-free graph, a does not decrease and c increases by $b - c > 0$. Hence $K(G)$ increases, contradicting the choice of G.

Suppose next $b = c$. Now, let p and q denote the number of vertices v such that $N(v) = N(x)$ and $N(v) = N(y)$, respectively. Assume, say, $p \ge q$. It is clear again that deleting the edges incident to y and joining y to the neighbours of x, we obtain another K_r-free graph, $b = c$ does not change, a does not decrease (and cannot increase either by, the choice of G). However, the number of pairs $\{u, v\}$ with $N(u) = N(v)$ increases by $p - q + 1 > 0$, a contradiction. □

Lemma 3.2. *There is no complete s-partite extremal graph with $s < r - 1$, provided $n \ge \max(m + 1, r - 1)$.*

Proof. We prove the statement by contradiction. Suppose that G is a complete s-partite extremal graph with classes $V_1, V_2, \ldots, V_s$. Let H be a subgraph of G isomorphic to K.

Suppose that there is a class V_i such that $V(H) \cap V_i \ne \emptyset$, or V_i. Let $u \in V(H) \cap V_i$, $v \in V_i - V(H)$ and let w be a neighbour of u in H. Then, joining v to all the remaining $n - 1$ vertices, we obtain an $s + 1$-partite graph G_0 such that $V(H) - \{w\} \cup \{v\}$ induces a copy of K containing the edge uv. Thus, $K(G_0) > K(G)$, a contradiction.

If $V(H) \cap V_i = \emptyset$ or V_i for $i = 1, \ldots, s$, then one can choose i and j so that $V(H) \cap V_i = V_i$, $V(H) \cap V_j = \emptyset$ and either $|V_i| \ge 2$ or $|V_j| \ge 2$. Pick any $v_i \in V_i$, $v_j \in V_j$. Then the vertex set $(V(H) - \{v_i\}) \cup \{v_j\}$ induces a subgraph H with $V(H) \cap V_k \ne \emptyset$, V_k for $k = i$ or j. □

Lemma 3.3. *All extremal graphs are complete $r - 1$-partite graphs, provided $n \ge \max(m + 1, r - 1)$.*

Proof. Suppose that there is an extremal graph G^* that is not a complete $r-1$-partite graph. The proof of Lemma 3.1 provides an algorithm to turn G^* into a complete s-partite extremal graph, where $s = r - 1$ by Lemma 3.2. Before the last step of this algorithm, we have an extremal graph G which is not complete $r-1$-partite, however, appropriately changing the neighbourhood $N(x)$ of some vertex x, we obtain a complete $r-1$-partite graph. We claim that G is a proper subgraph of a complete $r-1$-partite graph.

If $G - \{x\}$ is complete $r-2$-partite with classes $V_1, V_2, \ldots, V_{r-2}$, and x is joined to all the remaining $n-1$ vertices, then G is complete $r-1$-partite, a contradiction. Thus, x is not joined to all the remaining vertices, and G is a proper subgraph of the complete $r-1$-partite graph whose classes are $V_1, V_2, \ldots, V_{r-2}, V_{r-1} = \{x\}$.

Suppose next that $G - \{x\}$ is a complete $r-1$-partite graph with classes $V_1, V_2, \ldots, V_{r-1}$. If $N(x) \cap V_i \neq \emptyset$ for $i = 1, \ldots, r-1$, then G contains K_r as a subgraph, a contradiction. So, we may assume that, say, $N(x) \cap V_1 = \emptyset$. Then G is a proper subgraph of the complete $r-1$-partite graph whose classes are $V_1 \cup \{x\}, V_2, \ldots, V_{r-1}$.

Adding the missing edges (incident to x) to the graph G, we obtain a complete $r-1$-partite graph G_1, and $K(G_1) = K(G)$ by the extremality of G. Thus, if $xy \in E(G_1) - E(G)$, say, then xy is not contained in any copy of K. Then, by symmetry, no edge joining the classes of x and y is contained in any copy of K. Therefore, deleting these edges, we obtain a complete $r-2$-partite extremal graph, contradicting Lemma 3.2. □

References

1. Bollobás, B.: Extremal Graph Theory. London: Academic Press, 1978
2. Bollobás, B., Nara, C., Tachibana, S.: The maximal number of induced complete bipartite graphs. Discrete Math. **62**, 271–278 (1986)
3. Erdös, P.: Two problems in extremal graph theory. Graphs and Combinatorics **2**, 189–190 (1986)
4. Erdös, P.: On the number of complete subgraphs contained in certain graphs. Publ. Math. Inst. Hungar. Acad. Sci. **7**, 459–464 (1962)
5. Erdös, P.: On the number of complete subgraphs and circuits contained in graphs. Cas. Pestovani Mat. **94**, 290–296 (1969)
6. Györi, E.: Note on the maximum number of C_5's in triangle-free graphs. Combinatorica **9** (1989)
7. Sauer, N.: A generalization of a theorem of Turán. J. Comb. Theory (B) 109–112 (1971)
8. Turán, P.: On an extremal problem in graph theory (in Hungarian). Mat. Fiz. Lapok **48**, 436–452 (1941)
9. Turán, P.: On the theory of graphs. Colloq. Math. **3**, 19–30 (1954)
10. Zykov, A.A.: On some properties of linear complexes (in Russian). Mat. Sbornik N.S. **24**, 163–188 (1949); Amer. Math. Soc. Transl. **79** (1952)

Received: June 22 1989
Revised: December 22, 1989