Discrete Mathematics 14 (1976) 365–371.

THE MAXIMUM NUMBER OF q-CLIQUES IN A GRAPH WITH NO p-CLIQUE *

Steven ROMAN
Department of Mathematics, University of Washington, Seattle, Wash. 98195, USA

Received 12 May 1975

Let $f(n, p, q)$ be the maximum possible number of q-cliques among all graphs on n nodes with no p-clique. Turán, in 1941, determined $f(n, p, 2)$ for all n and p. For each n and p, he found the unique graph which attains this maximum. In this paper we determine $f(n, p, q)$ for all values of n, p and q. We show that, except for the trivial case $1 \leqslant n < q$, Turán's graph is the unique graph which attains the maximum $f(n, p, q)$ for all q such that $1 < q < p$.

1. Introduction and definition of the problem

Given n nodes, let the complete graph on any p of those nodes be called a p-clique. Let $f(n, p, q)$ be the maximum possible number of q-cliques among all graphs on n nodes with no p-clique. Turán [2], in 1941, in Hungarian, and again in 1954, in English, determined $f(n, p, 2)$, as well as the unique graphs giving these maxima. In 1962, Moon and Moser [1], determined $f(n, 4, 3)$.

In this paper we determine $f(n, p, q)$ for all n, p and q. We show that the graphs of Turán, which we call $T(n, p)$, maximize the number of q-cliques for all values of q, under the restriction that there be no p-clique. Moreover, provided $1 < q < p$ and $q \leqslant n$, $T(n,p)$ is the only graph on n nodes (up to isomorphism) which maximizes the number of q-cliques, under the restriction that there be no p-clique. From now on, we will assume $q < p$. We will dispose of the case $p \leqslant q$ separately in Theorem 1.

* The author is indebted to Professor Micha A. Perles for his helpful suggestions concerning the writing of this paper.

2. A preliminary lemma

To begin we establish the following lemma.

Lemma 1. *If $p-1 \leqslant n$, and if G is any graph on n nodes with $f(n,p,q)$ q-cliques, but no p-clique, then G has a $(p-1)$-clique.*

Proof. Let E be a clique contained in G with a maximal number of nodes; say E is a k-clique. Then $q \leqslant k \leqslant p-1$. If $q \leqslant k < p-1$, take a node of G which is not in E; call it v. Not all edges connecting v with E are in G, for otherwise there would be a $(k+1)$-clique. Let $\{v, u\}$ be an edge, with $u \in E$, which is not in G. Then we may add a q-clique to G as follows. Taken any $q-2$ nodes of E not including u, say $\{u_1, ..., u_{q-2}\}$, together with u and v. Form the q-clique with nodes $\{u_1, ..., u_{q-2}, u, v\}$. Call it F. Now F is not in G since the edge $\{u, v\}$ is not in G. Moreover, we may add F to G without creating a p-clique. To see this, first remark that all the edges added to G contain v. Hence if we do create a p-clique, say on the nodes $\{s_1, ..., s_p\}$, then v is one of the s_i and the other $p-1$ of the s_i must have formed a $(p-1)$-clique before the addition of F. This contradicts the maximality of $k < p-1$ and proves the lemma.

3. The definition of $T(n,p)$

Let $T(n,p)$ be the graph on n nodes defined as follows. If $n \equiv r \pmod{p-1}$, $0 \leqslant r \leqslant p-2$, form r sets with $(n-r)/(p-1)+1$ nodes each, and $p-1-r$ sets with $(n-r)/(p-1)$ nodes each. Call these sets $S_1, ..., S_{p-1}$. Take all possible edges connecting nodes of S_i to nodes of S_j, for $i \neq j$. This is $T(n,p)$. Notice that the total number of q-cliques in $T(n,p)$ is

$$t(n,p,q) = \sum_{k=0}^{q} \binom{r}{k}\left(\frac{n-r}{p-1}+1\right)^k \binom{p-r-1}{q-k}\left(\frac{n-r}{p-1}\right)^{q-k}.$$

We note that if $n = r$, then $t(r,p,q) = \binom{r}{q}$. Also, if $r = 0$, then $t(n,p,q)$ has the particularly simple form,

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

4. The definition of q-cliques of type 1, type 2 and type 3

Let H be any graph on n nodes, $n \geqslant p$, with no p-clique, and with a positive number of q-cliques. Suppose H has a $(p-1)$-clique, say E. Then we can define three types of q-cliques in H with respect to E. Those whose nodes form a subset of the nodes of E, call these cliques of type 1; those whose nodes are disjoint from the nodes of E, call these cliques of type 2, and those cliques which are neither of type 1 nor of type 2, call these cliques of type 3.

It is clear that the three types of cliques form three disjoint classes. It is also clear that H has exactly $\binom{p-1}{q}$ q-cliques of type 1. Furthermore, since H has no p-clique, it has at most $f(n-p+1, p, q)$ q-cliques of type 2. We now compute an upper bound for the number of q-cliques of type 3 in H.

For each choice of a k-clique, $1 \leqslant k \leqslant q-1$, which does not intersect E, consider the number of q-cliques of type 3 which use only those k nodes from outside of E. These q-cliques can involve altogether at most $p-k-1$ of the nodes of E, otherwise there would be a p-clique. Since each one of these q-cliques uses exactly $q-k$ nodes of E, there can be at most $\binom{p-k-1}{q-k}$ such q-cliques. Moreover, there are at most $f(n-p+1, p, k)$ choices of distinct k-cliques which do not intersect E. Therefore, the total number of q-cliques of type 3 is at most

$$\sum_{k=1}^{q-1} \binom{p-k-1}{q-k} f(n-p+1, p, k) .$$

5. The main result

We are now ready to prove the main result.

Theorem 1. *For all positive integers q, $T(n, p)$, as defined in Section 3, has $f(n, p, q)$ q-cliques and no p-clique. Moreover, if $1 < q < p$ and $q \leqslant n$, then $T(n, p)$ is the only graph on n nodes (up to isomorphism) with $f(n, p, q)$ q-cliques and no p-clique.*

Proof. First, it is clear from the definition that $T(n, p)$ contains no p-

clique. Now for the case $p \leqslant q$, $f(n, p, q) = 0$. Also $T(n, p)$ contains no p-clique, hence no q-clique, and thus gives the maximum $f(n, p, q)$ number of q-cliques. In the rest of the proof, we assume $q < p$.

Fix p; we proceed by induction on n, for all q such that $1 \leqslant q < p$.

By way of initial conditions, we consider the case $1 \leqslant n \leqslant p-1$. For this case, $T(n, p)$ is simply the complete graph on n nodes. Clearly, it has the maximum number of q-cliques, for $1 \leqslant q < p$. Moreover, for $1 < q \leqslant n$, the complete graph on n nodes is the only graph with the maximum number of q-cliques.

Now suppose the theorem is true for n. We will show it is also true for $n+p-1$. Choose any $q < p$. First we show that $T(n+p-1, p)$ has $f(n+p-1, p, q)$ q-cliques. Let E be a $(p-1)$-clique of $T(n+p-1, p)$. We compute the number of q-cliques in $T(n+p-1, p)$ of type i, $i = 1, 2, 3$, with respect to E. The number of q-cliques of type 1 is $\binom{p-1}{q}$. The number of q-cliques of type 2 is $f(n, p, q)$. This is because if we remove from $T(n+p-1, p)$ the $(p-1)$-clique E, together with all edges having nodes in E, we are left with $T(n, p)$, and by the induction hypothesis, this has $f(n, p, q)$ q-cliques. Finally, $T(n+p-1, p)$ has

$$\sum_{k=1}^{q-1} \binom{p-k-1}{q-k} f(n, p, k)$$

q-cliques of type 3. This is because, by the induction hypothesis, there are $f(n, p, k)$ choices of k-cliques in $T(n, p)$ (that is, $f(n, p, k)$ choices of k-cliques in $T(n+p-1, p)$ which are disjoint from E), and for each such choice, there are $\binom{p-k-1}{q-k}$ q-cliques of type 3 using only the nodes of that k-clique from outside E. So we see that $T(n+p-1, p)$ has the maximum possible number of q-cliques of each type (see Section 4). Hence $T(n+p-1, p)$ has the maximum total number of q-cliques, namely $f(n+p-1, p, q)$.

We must also show that for $1 < q \leqslant n+p-1$, $T(n+p-1, p)$ is the only graph on $n+p-1$ nodes having $f(n+p-1, p, q)$ q-cliques and no p-clique. So suppose G is a graph on $n+p-1$ nodes with $f(n+p-1, p, q)$ q-cliques and no p-clique. Then we will show G is $T(n+p-1, p)$. Since $p-1 \leqslant n+p-1$, by Lemma 1, G has a $(p-1)$-clique. Call it E. Now since $T(n+p-1, p)$ has the maximum possible number of q-cliques of each type relative to any of its $(p-1)$-cliques, so must G. Hence if we remove from G the $(p-1)$-clique E, together with all edges having nodes in E, we are left with a graph on n nodes with $f(n, p, q)$ q-cliques and no p-clique. We claim this must be $T(n, p)$. If $1 < q \leqslant n$, then the uniqueness part of the induction hypothesis establishes this. If $n = 1$, $T(n, p)$ is the only graph on one node. If $1 < n < q$, then we appeal to the fact that there are

$$\sum_{k=1}^{q-1} \binom{p-k-1}{q-k} f(n, p, k)$$

q-cliques of type 2 in G. Take the term corresponding to $k = 2$. Then there must be $\binom{p-3}{q-2} f(n, p, 2)$ q-cliques of type 3 using exactly one edge from outside E. But $n < q$, together with $q < p$, implies $n < p$, and so $f(n, p, 2) = \binom{n}{2}$. Now in order for there to be $\binom{p-3}{q-2}\binom{n}{2}$ q-cliques of type 3 using exactly one edge from outside E, there must be $\binom{n}{2}$ choices of such edges. That is, the complement of E in G must be the complete graph on n nodes, which in this case, is exactly $T(n, p)$.

So G is $T(n,p)$, together with a $(p-1)$-clique E and some additional edges connecting $T(n, p)$ with E. Now let the nodes of E be $n_1, \ldots, n_{p-1}$, and let the sets described in the definition of $T(n, p)$ be $S_1, \ldots, S_{p-1}$.

Suppose first that no S_i is empty. Then no n_i can be connected to a node in each of the S_i, $i = 1, \ldots, p-1$, since, by the definition of $T(n,p)$, that would produce a p-clique. But n_i must be connected to each node in all but one of the S_i. To see this, suppose n_i was not connected to any node of S_k, nor to some node, say v, of S_l, $l \neq k$. There exists

$$\sum_{k=1}^{q-1} \binom{p-k-1}{q-k} f(n, p, k)$$

q-cliques of type 3, so for $k = 1$, there must exist $\binom{p-2}{q-1} f(n, p, 1) = \binom{p-2}{q-1} n$ q-cliques which use only one node from outside of E. This means that each of the n nodes outside of E must be connected to $p-2$ of the $p-1$ nodes of E. Now choose a node w from S_k. Then w is not in E and is not connected to n_i. Therefore, it must be connected to all n_j, $j \neq i$. Similarly, v must be connected to all n_j, $j \neq i$. But v is connected to w. Therefore, the set of $p-2$ nodes of E, not including n_i, together with v and w, form a p-clique. Hence n_i must be connected to v.

Moreover, if n_i and n_j are two distinct nodes of E, they cannot be connected to all nodes of the same $p-2$ sets from among $S_1, \ldots, S_{p-1}$ because they are connected to each other, and that would produce a p-clique. Therefore we may assume that n_i is connected to all the nodes of each S_j for $j \neq i$. But now we see that sets $S_i \cup \{n_i\}$, $i = 1, \ldots, p-1$, together with the edges of G form exactly $T(n+p-1, p)$. The readers, if there are any, are urged to draw a diagram.

Now suppose that, for $k > 0$, the sets $S_1, \ldots, S_k$ are empty and the sets $S_{k+1}, \ldots, S_{p-1}$ are not empty. Then each S_j, $j \geqslant k+1$, contains exactly one element. Let $S_j = \{s_j\}$ for $j \geqslant k+1$. For each such j, s_j cannot be connected to all of the n_i, $i = 1, \ldots, p-1$. But s_j must be connected to all of the n_i except one, for the same reason as before. By the definition of $T(n, p)$, s_i is connected to s_j for all $i \neq j$. This means that s_i and s_j cannot be connected to the same set of $p-2$ nodes of E. So we may assume that s_i is not connected to n_i for $i = k+1, \ldots, p-1$, and again we see that G is $T(n+p-1, p)$. This completes the proof.

6. The asymptotic behavior of $f(n, p, q)$

We make a final remark about the behavior of $f(n, p, q)$, for fixed p and q, as n approaches infinity. For $n \equiv 0 \pmod{p-1}$,

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

For $n \equiv r \pmod{p-1}$, $r \neq 0$,

$$f(n-r, p, q) \leqslant f(n, p, q) \leqslant f(n-r+p-1, p, q) .$$

So the limit, as n approaches infinity, of the ratio of $f(n, p, q)$ to the total number of possible q-cliques in a graph on n nodes is

$$\lim_{n\to\infty} \binom{n}{q}^{-1} f(n, p, q) = \frac{(p-q) \ldots (p-1)}{(p-1)^q} .$$

This limit is smallest when $p = q+1$, in which case we get $(q-1)!/q^{q-1}$.

References

[1] J.W. Moon and L. Moser, On a problem of Turán, Magyar Tud. Akad. Mat. Kutató Int. Közl. 7 (1962) 283–286.
[2] P. Turán, On the theory of graphs, Colloq. Math. 3 (1954) 19–30.