Discrete Mathematics 59 (1986) 235–242
North-Holland

CLIQUE NUMBERS OF GRAPHS

Paul ERDÖS
Mathematics Institute, Hungarian Academy of Sciences, 1053 Budapest V, Hungary

Marcel ERNÉ
Mathematics Institute, University of Hannover, D3000 Hannover, Fed. Rep. Germany

Received July 29, 1985
Revised October 11, 1985

For each natural number n, denote by $G(n)$ the set of all numbers c such that there exists a graph with exactly c cliques (i.e., complete subgraphs) and n vertices. We prove the asymptotic estimate

$$|G(n)| = o(2^n \cdot n^{-2/5})$$

and show that all natural numbers between $n+1$ and $2^{n-6n^{5/6}}$ belong to $G(n)$. Thus we obtain

$$\lim_{n\to\infty} \frac{|G(n)|}{2^n} = 0,$$

while

$$\lim_{n\to\infty} \frac{|G(n)|}{a^n} = \infty \quad \text{for all } 0 < a < 2.$$

Many graph-theoretical problems involve the study of *cliques*, i.e., complete subgraphs (not necessarily maximal). In this context the following combinatorial problem arises naturally: For which numbers n and c is there a graph with n vertices and exactly c cliques? For fixed n, let $G(n)$ denote the set of all such 'clique numbers' c. Since each singleton and the empty set are always cliques, we have

$$n < c \leq 2^n \quad \text{for all } c \in G(n).$$

It is easy to check that every integer between $n+1$ and $2^{n/2}$ occurs in $G(n)$ (see the remark at the end of this paper), and a more thorough investigation shows that even all integers between $n+1$ and $2^{2n/3}$ are clique numbers of suitable graphs with n vertices. For small n, the first jumps in $G(n)$ occur between $2^{2n/3}$ and $2^{2n/3} \cdot 2$. Denoting by $c(n)$ the smallest $c > n+1$ with $c \notin G(n)$, we obtain Table 1. (As usual, $\lfloor a \rfloor$ denotes the greatest integer not greater than a, while $\lceil a \rceil$ denotes the least integer not less than a.)

In the higher regions near 2^n, $G(n)$ has large gaps. For example, the only clique numbers above 2^{n-1} are the numbers $2^{n-1} + 2^k$ with $0 \leq k < n$. The number c_1 of ones in the binary expansion of a given number c plays a crucial role for the question whether c is the clique number of a graph with n vertices (see the proof of Theorem 1). As a consequence of the fact that c_1 cannot be too large for

0012-365X/86/$3.50 © 1986, Elsevier Science Publishers B.V. (North-Holland)

Table 1

n	1	2	3	4	5	6	7	8	9
$\lceil 2^{2n/3} \rceil$	3	5	7	11	19	29	47	79	127
$c(n)$	2	3	4	7	11	16	26	41	64
$\lfloor 2^{2n/3} \cdot 2 \rfloor$	3	5	8	12	20	32	50	80	128

$c \in G(n)$, we show that the ratio $|G(n)|/2^n$ tends to zero when $n \to \infty$. But, on the other hand, it will turn out that for all positive reals $a < 2$, the ratio $|G(n)|/a^n$ goes to infinity, and moreover, that all numbers c between $n+1$ and $2^{n-6n^{5/6}}$ belong to $G(n)$. In particular, for each $b < 1$ there is an n_b such that $c(n) > 2^{bn}$ whenever $n > n_b$. Of course, this result disproves the conjecture (suggested by the above table) that $c(n)$ would not exceed $2^{2n/3} \cdot 2$. In order to determine the sets $G(n)$, it suffices to compute, for each natural number c, the smallest n such that there exists a graph with n vertices and c cliques. This is an immediate consequence of the following observation:

$$c \in G(n) \text{ and } n+1 < c \quad \text{implies} \quad c \in G(n+1). \tag{$*$}$$

In fact, if G is a graph with n vertices and $c > n+1$ cliques then G must have at least one edge joining two vertices, say, x and y. Delete this edge, adjoin a new vertex z to G, and join it with all vertices which are already joined with both, x and y. This gives a new graph G' with $n+1$ vertices, but the number of cliques remains the same as for G because each clique of G containing x and y is replaced by a clique of G' containing z. (Cf. Fig. 1.)

Next, we derive an asymptotic upper bound for the cardinality of $G(n)$:

Theorem 1. $|G(n)| = o(2^n \cdot n^{-2/5})$.

Proof. Let G be a graph with n vertices and c cliques. Choose a clique K of maximal size, say, k. Denoting by $\mathscr{C}$ the set of all cliques of the induced subgraph $G - K$, we have

$$c = \sum_{C \in \mathscr{C}} 2^{d_C},$$

where d_C is the number of vertices in K joined with each vertex of C. By

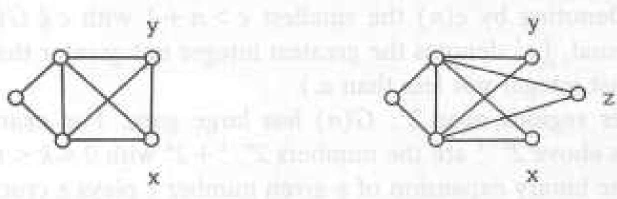

Fig. 1

maximality of K, d_C is not greater than $k-|C|$, whence

$$c \leq \sum_{j=0}^{n-k} \binom{n-k}{j} 2^{k-j} = (\tfrac{3}{4})^{n-k} 2^n.$$

Furthermore, the number c_1 of ones in the binary expansion of c is bounded by the cardinality of $\mathscr{C}$, whence

$$c_1 \leq |\mathscr{C}| \leq 2^{n-k}.$$

Combining both inequalities, we obtain

$$c \cdot c_1^\alpha \leq 2^n, \quad \text{where } \alpha = 2 - \log_2 3 > \tfrac{2}{5}.$$

Now choose an arbitrary real number β with $\frac{2}{5} < \beta < \alpha$, and let

$$m := \lfloor n - \beta \log_2 n + 1 \rfloor.$$

If $c \geq 2^m$, then $c_1 \leq 2^{(n-m)/\alpha} \leq 2^{(\beta/\alpha)\log_2 n} = n^{\beta/\alpha}$. Hence

$$|\{c \in G(n): c \geq 2^m\}| \leq |\{c \in G(n): c_1 \leq n^{\beta/\alpha}\}| \leq \sum_{k=0}^{\lfloor n^{\beta/\alpha} \rfloor} \binom{n}{k}$$

$$\leq n^{1+n^{\beta/\alpha}} = o(2^n \cdot n^{-2/5}) \quad \text{since } \beta/\alpha < 1.$$

On the other hand, we have

$$|\{c \in G(n): c \leq 2^m\}| \leq 2^{n-\beta\log_2 n+1} = o(2^n \cdot n^{-2/5}) \quad \text{since } \beta > 2/5. \quad \square$$

Table 2 suggests that $2^n \cdot n^{-2/5}$ is also a good estimate for small values of $|G(n)|$. Although $|G(n)|$ is of smaller order than 2^n, we shall show in the second part of this paper that $\log_2 |G(n)|$ is asymptotically equal to n.

Table 2

n	1	2	3	4	5	6	7	8	9
$\|G(n)\|$	1	2	4	8	16	30	55	99	178
$\lfloor 2^n \cdot n^{-2/5} \rfloor$	2	3	5	9	16	31	58	111	216

Henceforth let m be a natural number and

$$s := m^{1/6}, \qquad r := \left\lfloor \frac{s-1}{2} \right\rfloor.$$

For any nonempty finite set V of integers, put

$$d(V) := \max V - \min V.$$

We shall use the following version of the 'pigeon-hole principle':

(PP) If W is a set of w integers, then for all natural numbers v with $1 < v < w$ there exists a subset V of W with v elements and $\lfloor (w-1)/(v-1) \rfloor \, d(V) \leq$

$d(W)$; in particular,

$$d(V) \leqslant d(W) \frac{v-1}{w-v}.$$

For the construction of suitable graphs with prescribed clique numbers, we need a somewhat technical definition. Call a set V of nonnegative integers *m-adequate* if the following conditions are satisfied (recall that r and s are functions of m) (cf. Fig. 2):

$$V = V_1 \cup V_2 \quad \text{with} \max V_1 < \min V_2, \quad |V_1| = r^2 + 1 \quad \text{and} \quad |V_2| = 2r,$$
$$m - \max V \geqslant s^5,$$
$$d(V) \leqslant \tfrac{3}{4}s^5,$$
$$d(V_1) \leqslant s^3,$$
$$\min V_2 - \min V \geqslant \tfrac{1}{2}s^4.$$

Our main result is prepared by an auxiliary lemma ensuring that there are enought m-adequate sets.

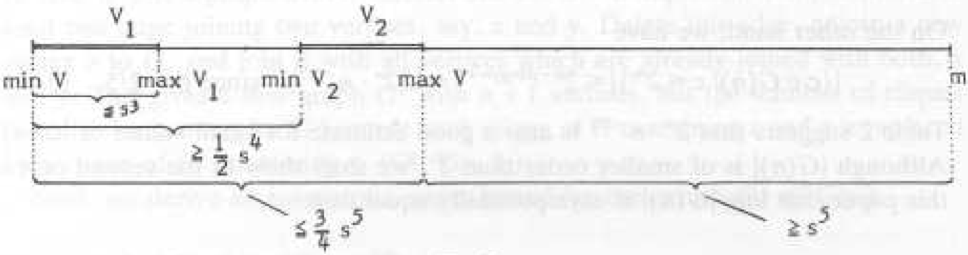

Fig. 2

Lemma. *Every set* $W \subseteq \{0, \ldots, m-1\}$ *with not less than* $2s^5$ *elements contains an m-adequate set.*

Proof. Choosing the $\lfloor s^5 \rfloor$ smallest elements from W, we obtain a subset W_1 with $d(W_1) \leqslant \max W_1 \leqslant m - s^5$. Now (PP) gives a subset W_2 of W_1 with $\lceil \frac{3}{4}s^4 \rceil + 1$ elements such that $d(W_2) \leqslant \frac{3}{4}s^5$. In fact, $2s^5 \leqslant m = s^6$ implies $s \geqslant 2$, whence

$$d(W_1) \frac{\lceil \frac{3}{4}s^4 \rceil}{\lfloor s^5 \rfloor - \lceil \frac{3}{4}s^4 \rceil - 1} \leqslant \frac{(s^6 - s^5)\lceil \frac{3}{4}s^4 \rceil}{s^5 - \frac{3}{4}s^4 - 3} \leqslant \tfrac{3}{4}s^5.$$

The $\lceil \frac{1}{4}s^4 \rceil$ smallest elements of W_2 form a subset W_3. Again by (PP), we can select a subset V_1 of W_3 with $r^2 + 1$ elements and $d(V_1) \leqslant s^3$, because $s \geqslant 2$ and $r = \lfloor \frac{1}{2}(s-1) \rfloor$ implies

$$d(W_3) \frac{r^2}{\lceil \frac{1}{4}s^4 \rceil - r^2 - 1} \leqslant d(W_2) \frac{s^2}{s^4 - s^2} \leqslant \frac{3s^7}{4s^4 - 4s^2} \leqslant s^3.$$

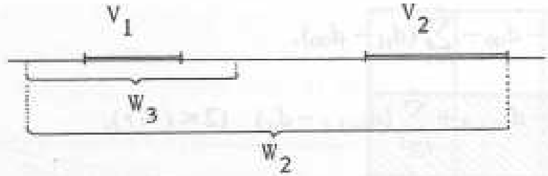

Fig. 3

Finally, let V_2 consist of the $2r$ greatest elements of W_2 (cf. Fig. 3).

Then $W_2 \setminus V_2$ has $\lceil \frac{3}{4}s^4 \rceil + 1 - 2r \geq \lceil \frac{1}{4}s^4 \rceil$ elements (because $s \geq 2$ and $r \leq \frac{1}{2}(s-1)$ yields $\lceil \frac{3}{4}s^4 \rceil - 2r \geq \frac{3}{4}s^4 - s + 1 \geq \frac{1}{4}s^4 + 1$). Thus $V_1 \subseteq W_3 \subseteq W_2 \setminus V_2$ and therefore $\max V_1 \leq \max W_3 < \min V_2$. Moreover,

$$\begin{aligned}\min V_2 - \min V_1 &\geq \min V_2 - \max V_1 + r^2 \geq \min V_2 - \max W_3 + r^2 \\ &\geq |W_2 \setminus (W_3 \cup V_2)| + 1 + r^2 \geq \tfrac{3}{4}s^4 - \tfrac{1}{4}s^4 - 2r + 1 + r^2 \\ &= \tfrac{1}{2}s^4 + (r-1)^2 \geq \tfrac{1}{2}s^4.\end{aligned}$$

Hence $V = V_1 \cup V_2$ has the required properties. □

Now we can prove

Theorem 2. *For all natural numbers n and c with $n < c \leq 2^{n-6n^{5/6}}$ there is a graph with exactly n vertices and c cliques.*

Proof. Let $c = 2^m + \sum_{d \in W} 2^d$, with $W \subseteq \{0, \ldots, m-1\}$. Furthermore, let $\mathcal{V}$ be a maximal collection of pairwise disjoint m-adequate subsets of W. By the lemma, the remainder $\bar{W} = W \setminus \bigcup \mathcal{V}$ contains less than $2s^5$ elements where $s = m^{1/6}$. Now a graph G with exactly c cliques is constructed as follows. First, form an m-element clique M. Second, choose a family $\{G_V : V \in \mathcal{V}\}$ of pairwise disjoint $(2r+1)$-sets outside of M. Consider one such $G_V = \{x_1, \ldots, x_r, y_1, \ldots, y_r, z\}$ and make it a bipartite graph by joining each x_i with each y_j. The m-adequate set $V = V_1 \cup V_2$ is labelled in form of an $(r+1) \times (r+1)$ array such that

$$\begin{aligned}
&V_1 = \{d_{00}\} \cup \{d_{ij} : 1 \leq i, j \leq r\} && (|V_1| = r^2 + 1),\\
&V_2 = \{d_{i0} : 1 \leq i \leq r\} \cup \{d_{0j} : 1 \leq j \leq r\} && (|V_2| = 2r),\\
&d_{00} < d_{ij} < d_{i-1,j} < d_{i-1,0} < d_{i0} && (2 \leq i \leq r, 1 \leq j \leq r),\\
&d_{00} < d_{ij} < d_{i,j-1} < d_{0,j-1} < d_{0j} && (1 \leq i \leq r, 2 \leq j \leq r).
\end{aligned}$$

Now we define an integer-valued $(r+1) \times (r+1)$ matrix (s_{ij}) by setting

$$\begin{aligned}
&s_{00} := d_{00}, &&\\
&s_{ij} := d_{ij} - d_{00} && (1 \leq i, j \leq r),\\
&s_{10} := d_{10} - d_{00} - \sum_{j=1}^{r} (d_{1j} - d_{00}), &&
\end{aligned}$$

$$s_{01} := d_{01} - d_{00} - \sum_{i=1}^{r} (d_{i1} - d_{00}),$$

$$s_{i0} := d_{i0} - d_{i-1,0} + \sum_{j=1}^{r} (d_{i-1,j} - d_{ij}) \quad (2 \leq i \leq r),$$

$$s_{0j} := d_{0j} - d_{0,j-1} + \sum_{i=1}^{r} (d_{i,j-1} - d_{ij}) \quad (2 \leq j \leq r).$$

Then we have

$$s_{ij} \geq 0 \quad (0 \leq i, j \leq r). \tag{1}$$

This is clear for $i = j = 0$ and for $i + j > 1$. By definition of m-adequate sets, we obtain

$$s_{10} \geq \min V_2 - \min V - r \cdot d(V_1) \geq \tfrac{1}{2}s^4 - rs^3 > 0, \quad \text{since } r < \tfrac{1}{2}s.$$

The same inequality holds for s_{01}. Next, one proves by induction

$$d_{i0} = d_{00} + \sum_{k=1}^{i} s_{k0} + \sum_{j=1}^{r} s_{ij} \quad (1 \leq i \leq r),$$
$$d_{j0} = d_{00} + \sum_{k=1}^{j} s_{0k} + \sum_{i=1}^{r} s_{ij} \quad (1 \leq j \leq r). \tag{2}$$

Third, we have the inequality

$$\sum_{i=0}^{r} \sum_{j=0}^{r} s_{ij} \leq m. \tag{3}$$

In fact,

$$\begin{aligned}
s_{00} + \sum_{i=1}^{r}\sum_{j=1}^{r} s_{ij} + \sum_{i=1}^{r} s_{i0} + \sum_{j=1}^{r} s_{0j} &= (2) \\
&= d_{00} + \sum_{i=1}^{r}\sum_{j=1}^{r} s_{ij} + d_{r0} - d_{00} - \sum_{j=1}^{r} s_{rj} + d_{0r} - d_{00} - \sum_{i=1}^{r} s_{ir} \\
&= \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} (d_{ij} - d_{00}) + (d_{r0} - d_{rr}) + d_{0r} \\
&\leq (r-1)^2\, d(V_1) + d(V) + \max V \\
&\leq \frac{s^2}{4} s^3 + \tfrac{3}{4}s^5 + m - s^5 = m.
\end{aligned}$$

On account of (1) and (3), we can choose a family of pairwise disjoint subsets S_{ij} (cf. Fig. 4) of M with s_{ij} elements $(0 \leq i, j \leq r)$. Join x_i with all points of the set

$$X_i = \bigcup_{k=0}^{i} S_{k0} \cup \bigcup_{j=1}^{r} S_{ij} \quad (1 \leq i \leq r),$$

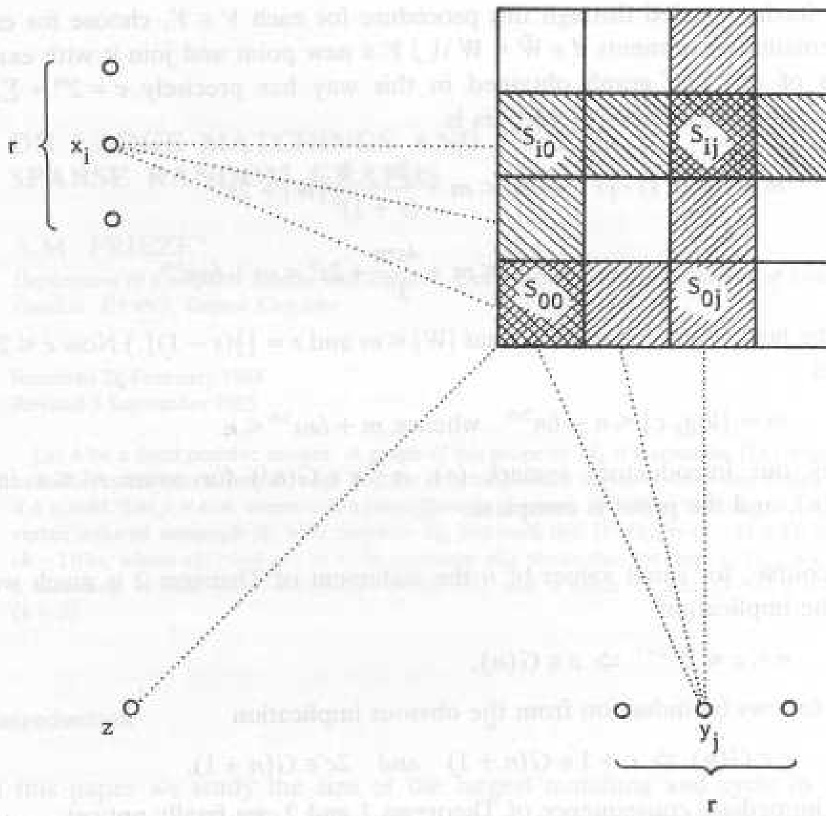

Fig. 4

and join y_j with all points of the set

$$Y_j = \bigcup_{k=0}^{j} S_{0k} \cup \bigcup_{i=1}^{r} S_{ij} \quad (1 \leq j \leq r).$$

By (2), we have

$$|X_i| = d_{i0} \quad (1 \leq i \leq r),$$
$$|Y_j| = d_{0j} \quad (1 \leq j \leq r).$$

Furthermore, the number of points joined with both, x_i and y_j, is

$$|S_{00} \cup S_{ij}| = d_{00} + s_{ij} = d_{ij} \quad (1 \leq i, j \leq r).$$

Finally, join the remaining point z of G_V with the points of S_{00} and recall that $|S_{00}| = d_{00}$. Then the number of cliques containing at least one point from G_V amounts to

$$\sum_{i=0}^{r} \sum_{j=0}^{r} 2^{d_{ij}} = \sum_{d \in V} 2^d.$$

After having carried through this procedure for each $V \in \mathcal{V}$, choose for each of the remaining exponents $d \in \bar{W} = W \setminus \bigcup \mathcal{V}$ a new point and join it with exactly d points of M. The graph obtained in this way has precisely $c = 2^m + \sum_{d \in W} 2^d$ cliques, and the number of vertices is

$$m + (2r+1) \cdot |\mathcal{V}| + |\bar{W}| < m + \frac{2r+1}{(r+1)^2}|W| + 2s^5$$

$$\leq m + \frac{4sm}{s^2} + 2s^5 = m + 6m^{5/6}.$$

(For the last inequality, observe that $|W| \leq m$ and $r = \lfloor \frac{1}{2}(s-1) \rfloor$.) Now $c \leq 2^{n-6n^{5/6}}$ implies

$$m = \lfloor \log_2 c \rfloor \leq n - 6n^{5/6} \quad \text{whence } m + 6m^{5/6} \leq n.$$

But by our introductory remark (*), $n < c \in G(n')$ for some $n' \leq n$ implies $c \in G(n)$, and the proof is complete. □

Of course, for small values of n the statement of Theorem 2 is much weaker than the implication

$$n < c \leq 2^{n/2+1} \Rightarrow c \in G(n),$$

which follows by induction from the obvious implication

$$c \in G(n) \Rightarrow c + 1 \in G(n+1) \quad \text{and} \quad 2c \in G(n+1).$$

As an immediate consequence of Theorems 1 and 2, we finally notice:

Corollary.

$$\lim_{n\to\infty} \frac{|G(n)|}{2^n} = 0, \quad \textit{but} \quad \lim_{n\to\infty} \frac{|G(n)|}{a^n} = \infty \quad \textit{for } 0 < a < 2.$$