COMBINATORICA 6 (3) (1986) 201—206

COVERING GRAPHS BY THE MINIMUM NUMBER OF EQUIVALENCE RELATIONS

N. ALON

Received 15 December 1984

An equivalence graph is a vertex disjoint union of complete graphs. For a graph G, let $\mathrm{eq}(G)$ be the minimum number of equivalence subgraphs of G needed to cover all edges of G. Similarly, let $\mathrm{cc}(G)$ be the minimum number of complete subgraphs of G needed to cover all its edges. Let H be a graph on n vertices with maximal degree $\leqq d$ (and minimal degree $\geqq 1$), and let $G=\bar{H}$ be its complement. We show that

$$\log_2 n - \log_2 d \leqq \mathrm{eq}(G) \leqq \mathrm{cc}(G) \leqq 2e^2(d+1)^2 \log_e n.$$

The lower bound is proved by multilinear techniques (exterior algebra), and its assertion for the complement of an n-cycle settles a problem of Frankl. The upper bound is proved by probabilistic arguments, and it generalizes results of de Caen, Gregory and Pullman.

1. Introduction

All graphs considered here are finite, simple and undirected. Let V be a finite set. For an equivalence relation R on V, let $G(R)$ denote its graph, i.e., the graph on V in which $x, y \in V$ are adjacent iff x is in relation with y. We call $G(R)$ an *equivalence graph*. Clearly a graph is an equivalence graph iff it is a vertex disjoint union of complete graphs. An *equivalence covering* of a graph G is a family of equivalence subgraphs of G such that every edge of G is an edge of at least one member of the family. The minimum cardinality of all equivalence coverings of G is the *equivalence covering number* of G, denoted by $\mathrm{eq}(G)$. Similarly, a *clique covering* of G is a family of complete subgraphs of G such that every edge of G is an edge of at least one member of the family. The minimum cardinality of such a family is the *clique covering number* of G, denoted by $\mathrm{cc}(G)$.

Clique covering numbers, which are the subject of extensive literature, were first studied in [4], and equivalence covering numbers were first studied in [3]. Obviously $\mathrm{eq}(G) \leqq \mathrm{cc}(G)$ holds for every graph G. Here we first prove the following:

Theorem 1.1. *Let $G=(V, E)$ be a graph and suppose $U=(u_1, u_2, \ldots, u_s)$, $W=(w_1, w_2, \ldots, w_s)$ are two (not necessarily disjoint) sequences of vertices. If $u_i w_i \notin E$ for all $1 \leqq i \leqq s$ and for all $1 \leqq i < j \leqq s$ either $u_i = w_j$ or $u_i w_j \in E$ then* $\mathrm{eq}(G) \geqq \log_2 s$.

Research supported in part by the Weizmann Fellowship for Scientific Research.
AMS subject classification (1980): 68 E 10, 68 E 05, 05 B 25, 05 C 55

The proof of Theorem 1.1 uses exterior algebra and is similar to the proof of the main result of [1]. Two corollaries of this theorem are the following.

Corollary 1.2. *Let T_n denote the complement of a matching of $n/2$ edges. Then* $\mathrm{eq}(T_n)=\lceil\log_2 n\rceil$ *for all even* $n\geqq 2$.

Corollary 1.3. *Let $\bar{C}_n$ denote the complement of a cycle of length n. Then* $\log_2 n+3\geqq \geqq \mathrm{eq}(\bar{C}_n)\geqq\log_2 n-1$ *for all* $n\geqq 3$.

The analogue of Corollary 1.2, for clique covering number was found by Gregory and Pullman [6] who showed that

$$\mathrm{cc}(T_n) = \min\left\{k\colon\, n \leqq 2\binom{k-1}{\lceil k/2\rceil}\right\} \approx \log_2 n+\frac{1}{2}\log_2\log_2 n.$$

Corollary 1.3 settles a problem of Frankl [5]. Solving a conjecture of Duchet [3], Frankl showed that $3\log_2 n\geqq\mathrm{eq}(\bar{C}_n)\geqq\log_2 n/\log_2\log_2 n$ and asked which of these bounds describes the real asymptotic behavior of $\mathrm{eq}(\bar{C}_n)$.

Combining Theorem 1.1 with some probabilistic arguments we prove the following theorem that describes the asymptotic behavior of $\mathrm{eq}(G)$ and $\mathrm{cc}(G)$ for the complement of any sparse graph.

Theorem 1.4. *Let H be a graph on n vertices with maximal degree $\leqq d$ and minimal degree $\geqq 1$. Let $G=\bar{H}$ be its complement. Then* $\log_2 n-\log_2 d\leqq\mathrm{eq}(G)\leqq\mathrm{cc}(G)\leqq \leqq c(d)\log_2 n$ *where* $c(d)=2e^2(d+1)^2/\log_2 e$.

The upper bound generalizes a result of de Caen, Gregory and Pullman [2], who showed that for the case $d=2$, $\mathrm{cc}(G)=O(\log n)$.

Our paper is organized as follows: in Section 2 we prove Theorem 1.1 and its corollaries. In Section 3 we consider complements of sparse graphs. Section 4 contains some concluding remarks.

2. The proof of Theorem 1.1 and its corollaries

We begin with a brief revision of the algebraic background needed. More details about exterior algebra can be found e.g., in [8].

Let $X=\mathbf{R}^m$ be the m-dimensional real space with the standard basis $e_1, e_2, \dots, e_m$. Put $M=\{1, 2, \dots, m\}$. The exterior algebra $\wedge X$ is a 2^m-dimensional real space, in which X is embedded, equipped with a multilinear associative multiplication $\wedge$. Our proof uses the following basic property of the $\wedge$ product. Suppose $r+s=m$ and $v_1, v_2, \dots, v_r$, $u_1, u_2, \dots, u_s\in X$. Define $v=v_1\wedge v_2\wedge\dots\wedge v_r$ and $u= =u_1\wedge u_2\wedge\dots\wedge u_s$. Then $u\wedge v\neq 0$ if and only if $v_1, \dots, v_r$, $u_1, \dots, u_s$ are independent in X. In particular, if $\{v_1, \dots, v_r\}\cap\{u_1, \dots, u_s\}\neq\emptyset$ then $u\wedge v=0$.

Proof of Theorem 1.1. Let $\{G_1, G_2, \dots, G_k\}$ be an equivalence covering of $G=(V, E)$. We must show that $k\geqq\log_2 s$. For $1\leqq i\leqq k$, G_i is a union of vertex disjoint cliques $\{K_{ij}\}_{j=1}^{r_i}$. Note that for each fixed i, $1\leqq i\leqq k$ the vertex sets of the K_{ij}-s form a partition of V.

For each $1 \leqq i \leqq k$ let $X_i = \mathbf{R}^2$ be a copy of the real plane, and let $\{x_{ij}: 1 \leqq \leqq j \leqq r_i\}$ be vectors in general position in X_i (i.e., every two of them are independent in X_i).

Let $Y = X_1 \wedge X_2 \wedge \ldots \wedge X_k$ be the 2^k-dimensional subspace of the exterior algebra $\wedge (X_1 \oplus \ldots \oplus X_k)$, in which each X_i is naturally imbedded. We now associate with each vertex v of $U \cup W$ a vector $\wedge v \in Y$ as follows: $\wedge v = x_{1j_1} \wedge x_{2j_2} \wedge \ldots \ldots \wedge x_{kj_k}$, where for $1 \leqq i \leqq k$, j_i is the unique index j such that $v \in K_{ij}$.

We then claim that for $1 \leqq i \leqq s$

$$(\wedge u_i) \wedge (\wedge w_i) \neq 0. \tag{2.1}$$

Indeed, since u_i and w_i are not adjacent in G they do not belong to a common clique in the covering. Hence $\wedge u_i$ and $\wedge w_i$ are products of disjoint sets of x-s and (2.1) follows by the general position of the x-s and the properties of the $\wedge$ product.

Similarly, if $1 \leqq i < j \leqq s$ then

$$(\wedge u_i) \wedge (\wedge w_j) = 0. \tag{2.2}$$

Indeed, here $\wedge u_i$ and $\wedge w_j$ are products of non disjoint sets of x-s, implying (2.2).

To complete the proof we show that the set $\{\wedge u_i: 1 \leqq i \leqq s\}$ is linearly independent in Y and thus $s \leqq \dim Y = 2^k$ and $k \geqq \log_2 s$, as needed. Indeed, suppose this is false and let

$$\sum_{i \in I} c_i (\wedge u_i) = 0 \tag{2.3}$$

be a linear dependence, with $c_i \neq 0$ for $i \in I$. Put $l = \max\{i: i \in I\}$. Combining (2.2) and (2.3) we get

$$0 = \sum_{i \in I} c_i (\wedge u_i) \wedge (\wedge w_l) = c_l (\wedge u_l) \wedge (\wedge w_l)$$

contradicting (2.1). This completes the proof. ∎

Proof of Corollary 1.2. Let $v_1, v_2, \ldots, v_n$ be the vertices of T_n, where $v_1v_2, v_3v_4, \ldots \ldots, v_{n-1}v_n$ are the edges of the missing matching. By Theorem 1.1 with $s = n$, $U = (v_1, v_2, v_3, v_4, \ldots, v_{n-1}, v_n)$ and $W = (v_2, v_1, v_4, v_3, \ldots, v_n, v_{n-1})$, we get $\mathrm{eq}(T_n) \geqq \lceil \log_2 n \rceil$. To prove the reverse inequality we construct an equivalence covering of cardinality $k = \lceil \log_2 n \rceil$ of T_n. For $1 \leqq i \leqq n$ let b_i be the binary representation of $i-1$. For a partition $W_1, W_2, \ldots, W_r$ of $\{v_1, \ldots, v_n\}$ let $K(W_1, \ldots, W_r)$ denote the equivalence graph consisting of r vertex disjoint cliques on the sets of vertices $W_1, \ldots, W_r$, respectively. Define $G_1 = K(\{v_1, v_3, v_5, \ldots, v_{n-1}\}, \{v_2, v_4, v_6, \ldots \ldots, v_n\})$. For $2 \leqq j \leqq k$ and $\varepsilon = 0, 1$ define

$W_j^\varepsilon = \{v_i$: the sum mod 2 of the least significant bit and the j-th significant bit of b_i is $\varepsilon\}$ and put $G_j = K(W_j^0, W_j^1)$.

One can check easily that $\{G_1, \ldots, G_k\}$ is an equivalence covering of T_n. This completes the proof. ∎

Proof of Corollary 1.3. Let $v_1, v_2, \ldots, v_n$ be the vertices of $\bar{C}_n$, where $v_1v_2\ldots v_nv_1$ is the missing cycle. By Theorem 1.1 with $s = 2[n/3]$, $U = (v_1, v_2, v_4, v_5, v_7, v_8, \ldots$

$\ldots, v_{3[n/3]-2}, v_{3[n/3]-1})$ and $W=(v_2, v_1, v_5, v_4, v_8, v_7, \ldots, v_{3[n/3]-1}, v_{3[n/3]-2})$

$$\mathrm{eq}(\bar{C}_n) \geqq \log_2(2[n/3]) \geqq \log_2 n - 1$$

for all $n \neq 5$. (For $n=5$ one can check easily that $\mathrm{eq}(\bar{C}_5)=3 \geqq \log_2 5-1$.)

It is worth noting that by applying the algebraic proof of Theorem 1.1 directly to the case of $\mathrm{eq}(\bar{C}_n)$ we can prove a lower bound of $\log_2(n-2)$ if n is even and $\log_2(n-1)$ if n is odd. This is done by associating vectors to all the vertices of $\bar{C}_n$ and showing that the space of linear dependences between them is of dimension $\leqq 2$ for even n and $\leqq 1$ for odd n. We omit the details.

The upper bound for $\mathrm{eq}(\bar{C}_n)$ is proved by a recursive construction analogous to the one used by de Caen, Gregory and Pullman [2] to show that $\mathrm{cc}(\bar{C}_n) \leqq \leqq 2\log_2(n-1)+2$.

Let $\bar{P}_n$ denote the complement of a path on n vertices. Observe that since $\bar{P}_{n-1}$ is an induced subgraph of $\bar{C}_n$, $\mathrm{eq}(\bar{P}_{n-1}) \leqq \mathrm{eq}(\bar{C}_n)$. Similarly, $\mathrm{eq}(\bar{P}_m) \leqq \mathrm{eq}(\bar{P}_n)$ for all $m \leqq n$. One can check easily that $\mathrm{eq}(\bar{C}_n) \leqq \mathrm{eq}(\bar{P}_{n-1})+2$. (Indeed, if $G_1, \ldots, G_r$ form an equivalence cover of a $\bar{P}_{n-1}$ on the vertices $v_1, v_2, \ldots, v_{n-1}$, add another vertex v_n and two equivalence graphs: $K(\{v_n, v_2, v_4, v_6, \ldots\})$ and $K(\{v_n, v_3, v_5, \ldots\})$ to get an equivalence cover of a $\bar{C}_n$.) Similarly, we observe that $\mathrm{eq}(\bar{C}_{2n-2}) \leqq \mathrm{eq}(\bar{P}_n)+1$. Indeed, let $R_1, \ldots, R_r$ be equivalence relations on $\{v_1, \ldots, v_n\}$ and suppose that the equivalence graphs $G(R_1), \ldots, G(R_r)$ form an equivalence cover of the complement of the path $v_1 v_2 \ldots v_n$. Put $V=\{v_1, \ldots, v_n, \bar{v}_2, \ldots, \bar{v}_{n-1}\}$. For $1 \leqq i \leqq r$ let $\bar{R}_i$ be the minimal equivalence relation satisfying $\bar{R}_i \supseteq R_i \cup \{v_j \sim \bar{v}_j \text{ for } 2 \leqq j \leqq n-1\}$. Define also an equivalence graph $G_{r+1}=K(\{\bar{v}_2, \bar{v}_4, \ldots, \bar{v}_{2[(n-1)/2]}, v_3, v_5, \ldots \ldots, v_{2[n/2]-1}\}, \{\bar{v}_3, \bar{v}_5, \ldots, \bar{v}_{2\lceil (n-1)/2\rceil-1}, v_2, v_4, \ldots, v_{2\lceil n/2\rceil-2}\})$. One can easily check that $\{G(\bar{R}_i)\}_{i=1}^r \cup G_{r+1}$ form an equivalence cover of the complement of the cycle $v_1 v_2 v_3 \ldots v_n \bar{v}_{n-1} \bar{v}_{n-2}, \bar{v}_2 v_1$.

The above observations, together with the easy fact that $\mathrm{eq}(\bar{C}_6)=2$, imply that $\mathrm{eq}(\bar{P}_n) \leqq \log_2 n+1$ and $\mathrm{eq}(\bar{C}_n) \leqq \log_2 n+3$ for all $n \geqq 3$. ∎

As noted by Frankl [5], $\mathrm{eq}(\bar{C}_n)$ is not monotone, as $\mathrm{eq}(\bar{C}_5)=3$ and $\mathrm{eq}(\bar{C}_6)=2$. However, the last proof shows that if $m \leqq n$ then $\mathrm{eq}(\bar{C}_m) \leqq \mathrm{eq}(\bar{C}_n)+2$.

3. Complements of sparse graphs

In this Section we prove Theorem 1.4 stated in Section 1. For convenience, we split the proof into two lemmas.

Lemma 3.1. *Let n, d, H and $G=\bar{H}$ be as in Theorem 1.4. Then* $\mathrm{eq}(G) \geqq \log_2 n - -\log_2 d$.

Proof. We prove the lemma by constructing two sequences $U=(u_1, \ldots, u_s)$ and $W=(w_1, \ldots, w_s)$ of vertices of G, where $s=\lceil n/d\rceil$ and U and W satisfy the hypotheses of Theorem 1.1. The lemma will then follow from the conclusion of Theorem 1.1. Suppose $G=(V, E)$. Choose arbitrarily some $w_1 \in V$ and let $u_1 \in V$ satisfy $u_1 w_1 \notin E$ (since the degree of any vertex of H is $\geqq 1$ such a u_1 exists). Suppose $l<s$ and assume that $w_1, w_2, \ldots, w_l$ and $u_1, u_2, \ldots, u_l$ have already been chosen so

that for $1 \leqq i \leqq l$ $u_i w_i \notin E$ and for $1 \leqq i < j \leqq l$ either $u_i = w_j$ or $u_i w_j \in E$. Put

$$\bar{V} = V - \bigcup_{i=1}^{l} \{v \in V\colon\ u_i v \notin E\}.$$

Since $|\{v \in V\colon u_i v \notin E\}| \leqq d$ for all $1 \leqq i \leqq l$, $\bar{V} \neq \emptyset$. Choose $w_{l+1} \in \bar{V}$ and let $u_{l+1} \in V$ satisfy $u_{l+1} w_{l+1} \notin E$. Clearly, for $1 \leqq i < j \leqq l+1$ either $u_i = w_j$ or $u_i w_j \in E$. Thus the two required sequences U and W exist and by Theorem 1.1, $\mathrm{eq}(G) \geqq \log_2 [n/d] \geqq \log_2 n - \log_2 d$. ▮

Note that Lemma 3.1 is best possible. Indeed, Corollary 1.2 shows that it gives the exact result for $d=1$. More generally, it is not difficult to show that if G is the complement of the union of $n/(d+1)$ disjoint stars with d edges each, then $\mathrm{eq}(G) \leqq 1 + \log_2 (n/(d+1))$, less than 1 more than the lower bound supplied by Lemma 3.1.

Lemma 3.2. *Let n, d, H and $G = \bar{H}$ be as in Theorem* 1.4. *Then* $\mathrm{eq}(G) \leqq \mathrm{cc}(G) \leqq c(d) \log_2 n$, *where* $c(d) = 2e^2 (d+1)^2 / \log_2 e$.

Proof. We use probabilistic arguments. Consider the following procedure of choosing a complete subgraph of $G = (V, E)$. In the first phase, pick every vertex $v \in V$ independently, with probability $1/(d+1)$ to get a set W. In the second phase define

$$\bar{W} = W - \{w \in W\colon ww' \notin E \quad \text{for some } w' \in W,\ w' \neq w\}.$$

Clearly $\bar{W}$ is the set of vertices of a complete subgraph of G.

Apply now the above procedure, independently, $k = \lfloor c(d) \cdot \log_2 n \rfloor$ times to get k complete subgraphs $K_1, K_2, \ldots, K_k$ of G. Let us estimate the expected value of the number of edges of G that are not covered by the union of the K_i-s. Let uw be an edge of G and fix i, $1 \leqq i \leqq k$. If u and w were chosen in the first phase of the procedure for generating K_i, and all the vertices in $\{v \in V\colon uv \notin E\} \cup \{v \in V\colon wv \notin E\}$ were not chosen then K_i covers the edge uw. Hence

$$\mathrm{Prob}\,(K_i \text{ covers } uw) \geqq \frac{1}{(d+1)^2} \left(1 - \frac{1}{d+1}\right)^{2d} \geqq \frac{1}{e^2 (d+1)^2}.$$

Hence

$$\mathrm{Prob}\,(\cup K_i \text{ does not cover } uw) \leqq \left(1 - \frac{1}{e^2 (d+1)^2}\right)^k \leqq \exp\left(-k/e^2 (d+1)^2\right).$$

Thus, the expected number of noncovered edges is at most $(n^2/2) \times \exp\left(-k/e^2 (d+1)^2\right) < 1$. Hence, there is at least one choice of k complete subgraphs of G that form a clique covering of G and $\mathrm{cc}(G) \leqq k \leqq c(d) \cdot \log_2 n$, as needed. ▮

The assertion of Lemma 3.2 for $d=2$ (with a somewhat better estimate of the constant), was proved, constructively, in [2]. It seems, however, that the probabilistic method is essential in the proof of the general result.

4. Concluding remarks

1. The algebraic proof of Theorem 1.1 can be applied to prove more general results. Thus, for example, we can prove the following.

Suppose $G=(V,E)$ satisfies the hypotheses of Theorem 1.1. Let $G_1, \ldots, G_r$ be subgraphs of G such that:

(a) Each G_i is a union of cliques $(K_{ij})_{j=1}^{s_i}$ and no vertex of G belongs to more than k of these s_i cliques.

(b) Every edge of G is an edge of at least one G_i.

Then

$$r \geqq \log_2 s/\log_2 \binom{2k}{k}.$$

Theorem 1.1 is the case $k=1$ of this result.

2. Using the method of Katona in [7], we can give pure combinatorial proofs of Corollaries 1.2 and 1.3. We do not know, however, how to prove Theorem 1.1 and its generalization mentioned above without the algebraic method.

References

[1] N. Alon, An extremal problem for sets with applications to graph theory, *J. Combinatorial Theory (A)*, **40** (1985), 82—89.

[2] D. de Caen, D. A. Gregory and N. J. Pullman, Clique coverings of complements of paths and cycles, *Annals of Discrete Math, to appear.*

[3] P. Duchet, *Représentations, noyaux en théorie des graphes et hypergraphes*, Thése de doctoral d'Etat, Université Paris VI (1979).

[4] P. Erdős, A. W. Goodman and L. Pósa, The representation of a graph by set intersections, *Can. J. Math.* **18** (1966), 106—112.

[5] P. Frankl, Covering graphs by equivalence relations, *Annals of Discrete Math.* **12** (1982), 125—127.

[6] D. A. Gregory and N. J. Pullman, On a clique covering problem of Orlin, *Discrete Math.* **41** (1982), 97—99.

[7] G. O. H. Katona, Solution of a problem of Ehrenfeucht and Mycielski, *J. Combinatorial Theory (A)* **17** (1974), 265—266.

[8] M. Marcus, *Finite Dimensional Multilinear Algebra*, Part II, Ch. 4, M. Dekker Inc., New York, 1975.

Noga Alon

Department of Mathematics
Tel Aviv University
Ramat Aviv, Tel Aviv 69978
Israel