ON THE CORE OF A GRAPH†

By FRANK HARARY and MICHAEL D. PLUMMER

[Received 8 October 1965]

1. Introduction

Let G be a graph. A set of points M is said to cover all the lines of G if every line of G has at least one point in M. Let $\alpha(G)$ denote the minimum number of points needed to cover all the lines of G. The core of G is defined as $\bigcup X$, where the union is taken over all sets X of $\alpha(G)$ independent lines of G. The concept of the core of a graph was first introduced by Dulmage and Mendelsohn ((1) (2)), who made it an integral part of their theory of decomposition for bipartite graphs.

In this paper we consider various properties of the core of a graph. Although a bipartite graph will always have a non-empty core, this is not true of graphs in general. We characterize those graphs with empty cores, and at the other extreme, those graphs in which every line is in the core. Next we obtain bounds on the number of lines which may be in the core of a graph, and show that these bounds are best possible. We then show that every integer between these bounds may be realized as the number of lines in the core of some graph. In the last section we obtain several characterizations of trees in which every line is in the core.

2. Basic terminology

In order that this paper be self-contained, it is necessary to define quite a number of concepts. A graph G consists of a finite set V(G) of points together with a collection of lines each of which is an unordered pair of points. If x is the line containing points u and v, then we write x = uv and say that u and v are adjacent and that point u and line x are incident. The degree of a point v, d(v), is the number of lines incident with it. A path joining points u and v is an alternating sequence of distinct points and lines beginning with u and ending with v so that each line is incident with the point before it and the point after it. A cycle is a path containing more than one line together with an additional line joining the first and last points of the path. The length of a path or a cycle is the number of lines in it. A graph G is connected if every two distinct points in G are joined by a path. The distance between two points u and v of a connected

† Work supported in part by the U.S. Air Force Office of Scientific Research under Grant AF-AFOSR-754-65. The authors are indebted to the referee for his helpful suggestions.

graph G is the length of a shortest path joining them, and is denoted by d(u, v). A point v is a *cut-point* of a graph G if the graph obtained from G by deleting v and all lines incident with v is disconnected. A subgraph G of G is a *block* of G if G is a maximal connected subgraph of G with no cut-points.

A tree is a connected graph with no cycles. If v is a point of the tree T, we call v an end-point of T if d(v) = 1, an intermediate point of T if d(v) = 2, and a branch-point of T if d(v) > 2. If T is a tree and v is any point of T, let T_1, \ldots, T_s be the maximal subtrees of T containing v as an end-point. The trees T_1, \ldots, T_s are said to arise by splitting T at v.

A set of points (lines) of G is said to be *independent* if no two of them are joined by a line (point). A set of lines X spans a graph G if every point of G is incident with a line of X. A set of points M is said to cover G if every line in G is incident with at least one point in M. The minimum number of points needed to cover G is called the point-covering number of G and is denoted by $\alpha(G)$. We call M a minimum point cover (m.c.) for G if M covers G and $|M| = \alpha(G)$, where |M| denotes the number of points in M.

A graph is bipartite if V(G) has been partitioned into two non-empty subsets S and T so that every line of G joins a point of S and a point of T. The complete bipartite graph K(m,n) is the bipartite graph obtained by joining every member of a set S of m points with every member of a set T of n points.

The core of a graph G, C(G), is defined as $\bigcup X$, where the union is taken over all sets X containing $\alpha(G)$ independent lines. If every line of G is in the core of G, we say that G is equal to its core, and write G = C(G). Finally we observe that the point-covering number of a graph G is additive on the components of G, and thus the core of G is the union of the cores of the components of G. For this reason, there will be no important loss of generality in studying cores of connected graphs only. In the remainder of this paper, the letter G (with or without subscripts) will denote a connected graph.

3. The core of a graph

Dulmage and Mendelsohn proved that every bipartite graph has a non-empty core. That this is not true for graphs in general may be seen immediately by considering any cycle C_p of length p, where p is odd. Then clearly $\alpha(C_p) = (p+1)/2$. But there are at most (p-1)/2 independent lines in C_p , hence the core of C_p is empty. We now proceed to characterize those graphs with empty cores.

Let M be a minimum point cover for the graph G, and let $U = V(G) \setminus M$. The cover M is said to be *external* if each subset M' of M has the property that $|M'| \leq |U(M')|$, where U(M') is the set of all points in U which are adjacent to some point of M'.

THEOREM 1. The following are equivalent for any graph G:

- (i) C(G) is not empty;
- (ii) G has a minimum cover which is external;
- (iii) every minimum cover for G is external.

Proof. We first show that (i) implies (iii). Suppose that G has an m.c. M which is not external. Then there is a subset $M' = \{v_1, ..., v_r\}, r \leq \alpha(G)$, of M such that |M'| > |U(M')|. Let $U(M') = \{u_1, ..., u_t\}, t < r$. Now let X be a maximum independent set of lines in G. Since no two points of U are adjacent, every line in X must meet at least one point of M. If any line of X meets more than one point of M, then $|X| < \alpha(G)$ and C(G) is empty. So suppose each line of X meets exactly one point of M. Then $|X| \leq t + \alpha(G) - r < \alpha(G)$, and again G has an empty core.

That (iii) implies (ii) is immediate. To show that (ii) implies (i), let $M = \{v_1, ..., v_s\}$ be an m.c. for G, and suppose that M is external.

Let $Y_i = \{u \in V(G) \setminus M : u \text{ is adjacent to } v_i\}$. Then for any k such that $1 \leq k \leq s$, the union of any k distinct sets Y_i contains at least k points since M is external. Hence, by the well-known theorem of P. Hall (3) on systems of distinct representatives, there is a set of s distinct points $\{y_1, \ldots, y_s\}, y_i \in Y_i$. Therefore there exist s independent lines $v_1 y_1, \ldots, v_s y_s$. Hence the core of G is not empty, and the proof of the theorem is complete.

Next we proceed to investigate the structure of the core. Following the terminology introduced by Dulmage and Mendelsohn, let B be a bipartite graph with sets of points S and T, where every line of B joins a point of S and a point of T. Then B is said to be semi-irreducible if B has exactly one m.c. M and either $M \cap S$ or $M \cap T$ is empty. Next, B is said to be irreducible if B has exactly two m.c.'s M_1 and M_2 and either $M_1 \cap S = \emptyset$ and $M_2 \cap T = \emptyset$ or $M_1 \cap T = \emptyset$ and $M_2 \cap S = \emptyset$. Finally, B is irreducible if it is neither irreducible nor semi-irreducible. If B is connected then the preceding definitions simplify as follows: the graph G is semi-irreducible if G has exactly one m.c. M and either M = S or M = T, and G is irreducible if S and S are both m.c.'s of S and S are no other minimum cover.

THEOREM 2. If both points of a line x of G are covered by some minimum cover for G, then x is not in C(G).

Proof. We refer to Theorem 3 of (1) for a proof of the result for bipartite graphs. The same proof carries over to any graph.

Theorem 2 has two immediate corollaries.

COROLLARY 2a. If G has a minimum cover which is not independent, then $G \neq C(G)$.

COROLLARY 2b. If B is a reducible connected bipartite graph, then $B \neq C(B)$.

We may now use these results in the proof of the next theorem.

THEOREM 3. If G has a non-empty core, then C(G) spans G, C(C(G)) = C(G), and the components of the graph C(G) are irreducible or semi-irreducible bipartite subgraphs of G.

Proof. We first show that C(G) spans G. Let X be a set of $\alpha(G)$ independent lines in G. Suppose that X fails to cover the point v_0 . Now since G is connected, there is a line u_0v_0 incident with v_0 in G. Now X covers u_0 , for otherwise $X \cup \{u_0v_0\}$ would be a set of $\alpha(G) + 1$ independent lines in G, which is impossible. Let u_0u_1 be the line of X covering u_0 . Then $[X\setminus\{u_0u_1\}]\cup\{u_0v_0\}$ is a set of $\alpha(G)$ independent lines covering v_0 , and hence C(G) spans G.

Now suppose that C(G) contains an odd cycle S. Let M be any m.c. for G. Then there must be two adjacent points of S which both lie in M, say v_2 and v_3 . Then $v_2v_3 \notin C(G)$ by Theorem 2, and this is a contradiction. Hence C(G) is bipartite.

Now let x be a line in C(G). Then there is a set X of $\alpha(G)$ independent lines of G which contains x. But $X \subseteq C(G)$. Hence $X \subseteq C(C(G))$, and in particular $x \in C(C(G))$. Thus C(C(G)) = C(G).

Now let H_1, \ldots, H_r be the components of C(G). Then

$$\bigcup_{i=1}^{\mathbf{r}} H_i = C(G) = C(C(G)) = \bigcup_{i=1}^{\mathbf{r}} C(H_i),$$

and hence $H_i = C(H_i)$ for each *i*. Thus by Corollary 2b, each H_i is irreducible or semi-irreducible. This completes the proof of the theorem.

We may now use the preceding results to characterize graphs which equal their cores.

THEOREM 4. Let G be any graph. Then G = C(G) if and only if G is irreducible bipartite or semi-irreducible bipartite.

Proof. If G = C(G) then, by Theorem 3, G is an irreducible or a semi-irreducible bipartite graph.

If G is a bipartite graph, then the converse of Theorem 2 also holds (again see Theorem 3 of (1)). Now if G is irreducible or semi-irreducible bipartite, then clearly no m.c. for G can contain both points of any line. Thus every line of G is in the core of G by the converse of Theorem 2, and the proof of Theorem 4 is complete.

We have shown that the core of a graph is an irreducible or semiirreducible bipartite subgraph, if it is non-empty. We now show that the core is maximal with respect to these properties.

THEOREM 5. If G has a non-empty core, and if B is a bipartite subgraph of G which properly contains the core, then B is reducible.

Proof. Assuming the conclusion to be false, suppose that B is irreducible or semi-irreducible. Then each component of B is irreducible or semi-irreducible. Let B_1, \ldots, B_s be these components. Then $C(B_i) = B_i$ by Theorem 4, and hence C(B) = B. Now $C(G) \subseteq B$, hence $\alpha(B) = \alpha(G)$, and thus $C(B) \subseteq C(G)$. But then $B \subseteq C(G)$, and hence B = C(G), contradicting the hypothesis and completing the proof.

4. Interpolation

We have seen that if G has a non-empty core then all the lines of G may be in C(G). In the case of a tree consisting of a single path of odd length p-1, the core has p/2 lines in it. These represent the two extremes of the size of the core. We now turn our attention to the possible number of lines in the core of any given graph. Let c(G) be the number of lines in the core of G. First we establish upper and lower bounds for c(G). We shall make use of the conventional greatest-integer function: if x is any real number, [x] denotes the greatest integer not exceeding x.

THEOREM 6. If a graph G with points p has a non-empty core, then

$$\left\lceil \frac{p+1}{2} \right\rceil \leqslant c(G) \leqslant \left\lceil \frac{p^2}{4} \right\rceil.$$

Proof. The lower bound is clear, since if C(G) is non-empty it must span G. If p is even then there are $p^2/4$ points in K(p/2, p/2), which is irreducible and hence equal to its core. But this is the bipartite graph on p points with the most lines. If p is odd, the same argument applies to K((p+1)/2, (p-1)/2) which is also irreducible, and the theorem is proved.

We now show that if r is any integer between the bounds of Theorem 6, then there is a graph G with p points and r lines in its core.

THEOREM 7. Let r be an integer such that

$$\left[\frac{p+1}{2}\right]\leqslant r\leqslant \left[\frac{p^2}{4}\right].$$

Then there is a graph G with p points and c(G) = r.

Proof. First, suppose that p is odd. If $G_1 \subseteq G_2$ and G_1 and G_2 are both bipartite with $\alpha(G_1) = \alpha(G_2)$, and G_1 is not reducible, then neither is G_2 .

For let M be an m.c. for G_2 . Then M is an m.c. for G_1 . But every m.c. for G_1 is independent, so every m.c. for G_2 is also independent.

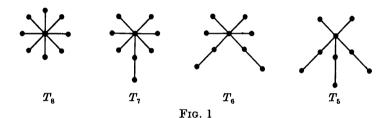
Now let P_p denote a path with p points. Then

$$P_p \subseteq K\!\left(\!\frac{p+1}{2},\!\frac{p-1}{2}\!\right) \quad \text{and} \quad \alpha(P_p) = \frac{p-1}{2} = \alpha\!\left(K\!\left(\!\frac{p+1}{2},\!\frac{p-1}{2}\!\right)\!\right)\!.$$

Also note that P_p is not reducible. Hence every graph G such that

$$P_p \subseteq G \subseteq K\left(\frac{p+1}{2}, \frac{p-1}{2}\right)$$

is not reducible, and $\alpha(G) = (p-1)/2$. Thus every such graph G is equal to its core, as are P_p and K((p+1)/2, (p-1)/2). But one can proceed from P_p to K((p+1)/2, (p-1)/2) by adding one line at a time. Let



 $p-1 \le r \le (p^2-1)/4$, and let G_r be a subgraph of K((p+1)/2, (p-1)/2) obtained from P_p by adding exactly r-p+1 lines to P_p . Then G_r has exactly r lines in its core, since it equals its own core.

Now if $(p+1)/2 \le r \le p-1$, we shall construct a sequence of trees $\{T_r\}$ on p points, where T_r has r lines in its core. If r=p-1, let T_{p-1} be the star with p points, i.e. that tree with one point v of degree p-1 and the remaining points u_1, \ldots, u_{p-1} of degree 1. Now remove line $u_{p-1}v$ and add line $u_{p-1}u_1$. Call the resulting graph T_{p-2} . Then clearly T_{p-2} is a tree (and hence a bipartite graph) with p-2 lines in its core. Now continue to construct the trees T_r inductively as follows: suppose T_{p-j} to have been constructed; to construct T_{p-j-1} one merely deletes line $u_{p-j}v$ from T_{p-j} and adds line $p_{p-j}u_j$. In this way a sequence of trees T_{p-1} , T_{p-2} , ..., $T_{(p+1)/2}$ is obtained. It is easily verified that T_r has exactly r lines in its core. Figure 1 shows the sequence of T_i 's for p=9.

Now suppose that p is even. If $p/2 \le r \le p-1$, the above sequence of trees $\{T_r\}$ suffices. So suppose that $p \le r \le p^2/4$. Let C_p be the cycle on p points. Now

$$C_p \subseteq K\Big(\frac{p}{2},\frac{p}{2}\Big), \quad \alpha(C_p) = \frac{p}{2} = \alpha\Big(K\Big(\frac{p}{2},\frac{p}{2}\Big)\Big),$$

and C_p is not reducible. Hence every graph G such that

$$C_p \subseteq G \subseteq K\left(\frac{p}{2}, \frac{p}{2}\right)$$

is equal to its core. But again one can proceed from C_p to K(p/2, p/2) by adding one line at a time. Hence if $p \le r \le p^2/4$, and if G_r is a subgraph of K(p/2, p/2) obtained by adding r-p lines to C_p , it follows that G_r has exactly r lines in its core. This completes the proof of the theorem.

Finally, we have an upper bound on the number of lines in a graph with a non-empty core. If G_1 and G_2 are disjoint graphs, their join $G_1 + G_2$ consists of their union together with all lines joining a point of G_1 with one of G_2 .

The complete graph K_p with p points has all pairs of points adjacent. Its complement \overline{K}_p is the graph with p points and no lines. Thus the complete bipartite graph K(m,n) is the join $\overline{K}_m + \overline{K}_n$.

THEOREM 8. If G has p points, q lines, and a non-empty core, then (i) $q \leq (p-1)(3p-1)/8$ if p is odd, and (ii) $q \leq p(3p-2)/8$ if p is even.

Proof. First, suppose that p is odd. Let G_0 be a graph on p points with a non-empty core and a maximum number of lines. Let M be a minimum cover for G_0 . Then $1 \leq |M| = r \leq (p-1)/2$. We note that G_0 has at most (r(r-1))/2 + r(p-r) lines, an expression strictly monotonic in r. Hence the maximum total occurs when r = (p-1)/2, and this total is (p-1)(3p-1)/8, which is realized by the graph $G_0 = K_{(p-1)/2} + \overline{K}_{(p+1)/2}$.

If p is even, a similar argument shows that the maximum number of lines in G_0 occurs when r = p/2, and this total is p(3p-2)/8; here $G_0 = K_{p/2} + \overline{K}_{p/2}$.

5. The core of a tree

In §3 we obtained a characterization of graphs which are equal to their cores. In this section we shall obtain several characterizations of trees which are equal to their cores. Whereas the criterion of Theorem 4 is not necessarily easy to apply to a general graph, one can use the results of this section to decide easily whether or not a tree is equal to its core.

Let T be a tree. A path P in T is *bi-ended* if each end-point of P is also an end-point of T. A tree T is called a bc-tree, following (5), if each bi-ended path in T is even in length.

THEOREM 9. If T is a tree and if T = C(T), then T is a bc-tree.

Proof. Suppose that P is any bi-ended path in T of odd length, and that M is an m.c. for T. We may assume that M contains neither endpoint of P. But then since P is of odd length, M must contain two adjacent

312 FRANK HARARY AND MICHAEL D. PLUMMER

points of P, say u and v. Therefore, by Theorem 2, $uv \notin C(T)$. Hence $T \neq C(T)$, contrary to hypothesis.

The next theorem (cf. (5)) tells us that we need check only those bi-ended paths with a common end-point.

THEOREM 10. If v_0 is an end-point of T, and every bi-ended path in T ending in v_0 is of even length, then T is a bc-tree.

We shall also make use of the following lemma. The proof may be found in (4).

LEMMA. Let G be a connected graph with cut point v. Let G_1, G_2, \ldots, G_s be the graphs formed by splitting G at v. Then v is in some minimum cover for G if and only if v is in some minimum cover of at least one graph G_i , $i = 1, 2, \ldots, s$.

Next we point out a property shared by all minimum point covers of bc-trees.

THEOREM 11. Let T be a bc-tree, and let M be a minimum point cover for T. Then M contains no end-point of T, and is an independent set.

Proof. If T has no branch-points, i.e. if T is an even path, the theorem is obvious. Suppose that T has at least one branch-point. We shall prove the first part of the theorem by induction on the number of points in T. If T has 4 points, the theorem is clear. If p > 4, we take as our inductive hypothesis that the theorem is true for bc-trees with fewer than p points. Let v be any end-point of T, and let v be the branch-point of v nearest v.

Case 1. d(v,b) is even. Split T at b, and let T_1, T_2, \ldots, T_s be the trees so obtained, where T_1 is the path $[v,\ldots,b]$, which is even in length. Now by the lemma there is an m.c. M for T containing b if and only if there is an m.c. M_k for some T_k such that $b \in M_k$. But b is an end-point in each T_i , and for each i every bi-ended path in T_i is even in length. Hence since each T_i has fewer points than T we may apply the inductive hypothesis to each of them and conclude that no T_i has an m.c. containing b. Hence, by the lemma, T has no m.c. containing b. Thus the theorem is proved for Case 1.

Case 2. d(v,b) is odd. Let $\{u_0, u_1, ..., u_r\}$, r > 1, be the set of points adjacent to b, where u_0 lies on the path joining v and b. Note that u_0 and v may coincide. Suppose that there is an m.c. M for T which contains v. Then $b \notin M$ since M is minimum, and hence $\{u_0, u_1, ..., u_r\} \subseteq M$. Now u_1

is not an end-point, for if it were, and if $P = [v, ..., u_0, b, u_1]$, then the symmetric difference of M and V(P) covers T and has one point fewer than M, contradicting the minimality of M. Now split T at u_1 . Let the resulting trees be $T_1, ..., T_s$. Now for each i, T_i has only even bi-ended paths, u_1 is an end-point of T_i , and T_i has fewer points than T. Hence, applying the inductive hypothesis to each T_i , we have the result that no m.c. for T_i contains u_1 . Hence, by the lemma, no m.c. for T contains u_1 , and this is a contradiction, completing Case 2 and proving the first part of the theorem.

To prove the second assertion let M be any m.c. for T, and suppose that M is a dependent set, i.e. suppose that two points v, w of M are adjacent in T. Then clearly neither v nor w is an end-point in T. The line vw belongs to some bi-ended path P in T,

$$P = [u_1, u_2, \dots, u_{r-1}, \ u_r = v, \ u_{r+1} = w, \ u_{r+2}, \dots, u_t], \quad 2 \leqslant r \leqslant t-2.$$

Now either r or r+1 is odd; suppose that r is. Split T at u_r , letting T_1, T_2, \ldots, T_s be the resulting trees. Now u_r is an end-point in each T_i . Also, since $u_r \in M$ we know by the lemma that there is some k such that u_r belongs to an m.c. for T_k . But every bi-ended path in T_k is even in length. Hence we may apply the first part of this theorem to T_k and obtain a contradiction.

If r+1 is odd, split T at u_{r+1} and the same type of contradiction may be obtained. This completes the proof of the theorem.

The first of the following two corollaries is immediate.

COROLLARY 11a. If T is a bc-tree then T has a unique minimum point cover.

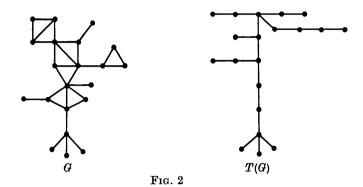
COROLLARY 11b. If T is a bc-tree then T = C(T).

Proof. If T is a bc-tree, Theorem 11 tells us that no m.c. for T covers both ends of any line. Hence, by Theorem 3 of (1), every line of T is in the core, i.e. T = C(T).

We now define as in (5) the block-cutpoint-tree T(G) as the graph whose set of points is the union of the set of blocks and the set of cut-points of G, and in which two points are adjacent if one corresponds to a block B of G, and the other to a cut-point v of G with v belonging to B in G. An example of this construction is shown in Fig. 2.

It is known (5) that every bc-tree is the block-cutpoint-tree of a connected graph, and conversely.

We may now combine the results of this section into the following theorem.



Theorem 12. The following statements are equivalent for any tree T:

- (i) T = C(T);
- (ii) T has a unique minimum cover M; M is independent and contains no end-points of T;
- (iii) if v, u_1 , u_2 , ..., u_p are the end-points of T, then $d(v, u_i)$ is even for each i;
- (iv) T is a bc-tree;
- (v) T is the block-cutpoint-tree of some connected graph G.

REFERENCES

- A. L. Dulmage and N. S. Mendelsohn, 'Coverings of bipartite graphs', Canadian J. Math. 10 (1958) 517-34.
- 'A structure theory of bipartite graphs of finite exterior dimension', Trans. Roy. Soc. Canada, Sect. III, 53 (1959) 1-13.
- 3. P. Hall, 'On representations of subsets', J. London Math. Soc. 10 (1934) 26-30.
- F. HARARY and M. D. PLUMMER. 'On the critical points and lines of a tree', Magyar Tud. Akad. Mat. Kutató Int. Közl., to appear.
- F. Harary and G. Prins, 'The block-cutpoint-tree of a graph', Publ. Math. Debrecen, to appear.

The University of Michigan