Available online at www.sciencedirect.com
SCIENCE DIRECT®
DISCRETE APPLIED MATHEMATICS
ELSEVIER
Discrete Applied Mathematics 140 (2004) 223–239
www.elsevier.com/locate/dam

The integrity of a cubic graph

A. Vince

Department of Mathematics, University of Florida, Gainesville, FL 32611-8105, USA

Received 26 August 2002; received in revised form 23 June 2003; accepted 16 July 2003

Abstract

Integrity, a measure of network reliability, is defined as

$$I(G) = \min_{S \subset V} \{|S| + m(G - S)\},$$

where G is a graph with vertex set V and $m(G-S)$ denotes the order of the largest component of $G-S$. We prove an upper bound of the following form on the integrity of any cubic graph with n vertices:

$$I(G) < \tfrac{1}{3}n + \mathrm{O}\left(\sqrt{n}\right).$$

Moreover, there exist an infinite family of connected cubic graphs whose integrity satisfies a linear lower bound $I(G) > \beta n$ for some constant β. We provide a value for β, but it is likely not best possible. To prove the upper bound we first solve the following extremal problem. What is the least number of vertices in a cubic graph whose removal results in an acyclic graph? The solution (with a few minor exceptions) is that $n/3$ vertices suffice and this is best possible.

Keywords: Integrity; Network reliability

1. Introduction

There are several measures of the reliability of a communication network. An elegant and simple one is called the integrity of the network, a concept introduced by Barefoot, Entringer and Swart in 1987 [9]. The motivation is as follows. Model the network as a graph. To disrupt the network a terrorist attempts to remove a small set of vertices (or

E-mail address: vince@math.ufl.edu (A. Vince).

doi:10.1016/j.dam.2003.07.002

edges) such that the remaining connected components are small. Formally, the *integrity* of a graph G with vertex set V is defined as

$$I(G) = \min_{S \subset V} \{|S| + m(G - S)\},$$

where $m(G-S)$ denotes the order of the largest component of $G-S$. The *edge integrity* is defined similarly as

$$I'(G) = \min_{S \subset E} \{|S| + m(G - S)\},$$

where E is the edge set of G. The goal in constructing networks resilient to attack is to find infinite families of graphs for which the integrity is large. The complete graphs have this property, but they are "expensive" in the sense that many edges must be incorporated. For this reason, it is natural to consider the integrity of k-regular graphs, for a fixed integer k.

There is a substantial literature on integrity, but most of the papers are concerned with calculating the integrity of particular graphs. For example, an easy result cited in the survey article on integrity by Bagga et al. [7] gives the integrity and edge integrity of a path P_n with n vertices:

$$I(P_n) = \lceil 2\sqrt{n+1} \rceil - 2, \quad I'(P_n) = \lceil 2\sqrt{n} \rceil - 1. \tag{1}$$

There are results on the interrelations between integrity and other graph parameters [18], on the computational complexity of computing integrity [17] and results on 2-regular graphs [5]. Conspicuously absent from the literature are general results on the integrity of k-regular graphs for fixed $k > 2$, and on constructing infinite families of k-regular graphs with large integrity. In this paper we concentrate on cubic graphs. One result (Theorem 7 of Section 3) gives an upper bound, linear in the number of vertices, on the integrity and edge integrity of cubic graphs. Basically we prove that

$$I(G) < \frac{n}{3} + \mathrm{O}(\sqrt{n}) \quad I'(G) < \frac{n}{2} + \mathrm{O}(\sqrt{n})$$

for connected cubic graphs G with n vertices.

The proof of the above result is based on the solution of the following extremal problem. What is the least number of vertices in a cubic graph whose removal results in an acyclic graph? It was brought to our attention by a knowledgeable referee that this problem, for general as well as cubic, graphs has received attention since at least 1974. Background on the problem appears at the beginning of Section 2. For cubic graphs the result in this paper is an improvement over past results. The solution (with a few minor exceptions) is that $n/3$ vertices suffice and this is best possible. The proof also appears in Section 2 (Theorem 2).

Concerning the question of infinite families of cubic graphs with large integrity, it is somewhat surprising that there exists an infinite family $\mathscr{G}$ of connected cubic graphs with integrity linear in the number of vertices, i.e., there is a positive constant β such that $I(G) > \beta n$ for every in $G \in \mathscr{G}$ (Theorems 8 and 9). The proof relies on expander graphs as explained in Section 4.

2. An extremal graph problem

All graphs are assumed to be simple, i.e., without multiple edges or loops. In fact, the main result in this section is false, in general, for graphs with multiple edges. The problem is to determine the minimum number of vertices in a connected cubic graph G such that their removal results in an acyclic graph. When we say that a vertex is *removed* from a graph, that vertex and all incident edges are removed. The notation $G-S$ is used for the graph obtained by removing set S of vertices from G. A minimal cardinality set S of vertices such that $G-S$ is acyclic will be called a G-set and $f(G)$ will denote the number of elements in a G-set.

The problem of determining the number $f(G)$ of vertices in a G-set has appeared in the literature in various guises. In the computer science literature a set of vertices such that their removal results in an acyclic graph is referred to as a *vertex feedback set* [8,20–22,25] and in the graph theory literature as a *decycling set* [11,12]. The problem of finding a minimum such set is known to be NP-hard for general networks, but there are polynomial time approximation algorithms and polynomial time algorithms for particular families of graphs relevant to applications (see references above). Also several papers address the problem of finding the maximum number $a(G)$ of vertices of G that induce a forest; this is an equivalent problem since $a(G)=n-f(G)$, where n is the number of vertices of G. Alon, Kahn, and Seymour [3] proved that for graphs with average degree $d \geqslant 2$ the upper bound $f(G) \leqslant 2n/(d+1)$ holds. Concerning cubic graphs, in 1974 Jaeger [19] proved that $f(G) \geqslant \lceil (n+2)/4 \rceil$ for connected cubic graphs. Concerning an upper bound Bondy et al. [14] proved that $f(G) \leqslant (3n+2)/8$ for any connected graph all of whose vertices have degree at most 3 and $n \geqslant 5$. Alon et al. [4] proved a result whose corollary (Corollary 1.6) is that $f(G) \leqslant 3n/8$ for any triangle-free graph all of whose vertices have degree at most 3. Actually, in the prior paper of Bondy, Hopkins and Staton it is proved that $f(G) \leqslant (n+1)/3$ for any triangle-free graph all of whose vertices have degree at most 3. This was improved slightly by Lu and Zheng [28] to $f(G) \leqslant n/3$ for the same class of graphs with two exceptions, thus confirming a conjecture of Speckenmeyer [27]. The result in this paper (Theorem 2) is an improvement over prior results on cubic graphs in that the $n/3$ upper bound is proved, with a few given exceptions, without assuming that the graph is triangle-free. This is required for subsequent results in the paper. It is interesting to note that it remains open [1] whether $f(G) \leqslant 3n/8$ for any n vertex, planar, bipartite graph.

Note that there are infinitely many connected cubic graphs such that $f(G)=n/3$, where n denotes the number of vertices of G. Figs. 1A and B can be generalized in the obvious way. To leave the graph in Figure 1A acyclic, at least one vertex must be removed from each triangle. Hence, at least $n/3$ vertices must be removed from G to obtain an acyclic graph. In fact, there are infinitely many connected cubic graphs G on n vertices such that $f(G)=3n/8$. To leave the graph in Fig. 1B acyclic at least one vertex must be removed from each triangle lying on the cycle and at least two vertices from each copy of the subgraph in Fig. 1C. Hence at least $3n/8$ vertices must be removed to leave the graph acyclic.

The graph in Fig. 1C will be denoted K_4^+ throughout this paper. Likewise, the graphs in Figs. 1D and E will be denoted K_4^{++} and $2K_4$, respectively. We use the

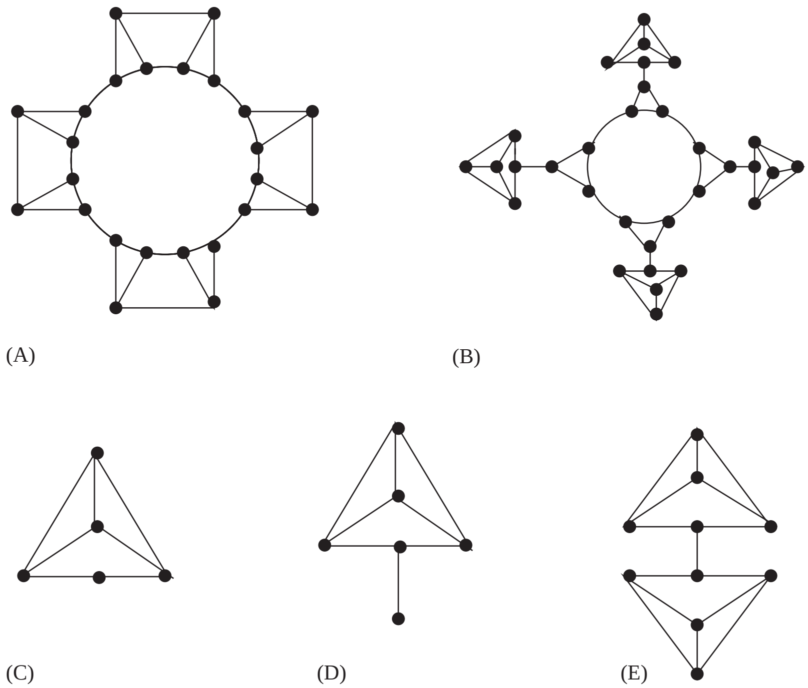

Fig. 1.

term *exceptional* for any of the following graphs: K_4, K_4^+, $2K_4$, any connected cubic graph with 8 vertices, and any connected graph containing K_4^{++} as a subgraph. Note that $f(K_4)=2$, $f(K_4^+)=2$ and $f(2K_4)=4$. Any graph that is not exceptional will be called *ordinary*. The remarks above show that Theorem 2 below is best possible.

Lemma 1. *Any connected cubic graph of order* 8 *has* $f(G)=3$. *Any connect graph of order* 8 *whose vertices have degree at most* 3, *but not all* 3, *has* $f(G)\leqslant 2$.

Proof. We prove the statement about connected, cubic graphs and leave the easy proof of the last statement as an exercise.

We first show that $f(G)>2$. If a set S consisting of any 2 vertices is removed, 6 vertices and at least 6 edges remain. Hence at least one connected component of $G-S$ has at least as many edges as vertices. This implies that $G-S$ contains a cycle.

To show that $f(G)\leqslant 3$, let S be a set of two vertices at a distance at least 3 apart. Then $G-S$ has 6 vertices, each of degree 2. If $G-S$ is a 6-cycle, then removing one addition vertex results in an acyclic graph. There is only one connected graph with 8 vertices, up to isomorphism, where $G-S$ is the disjoint union of two 3-cycles. For this particular graph $f(G)=3$. $\square$

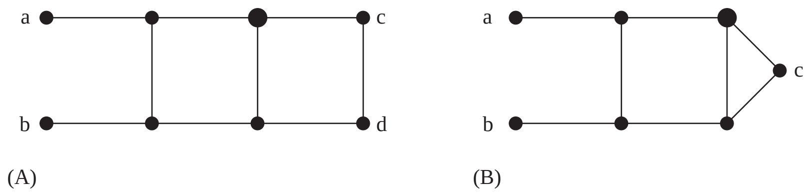

Fig. 2. Case 1.

Theorem 2. *Let G be a connected graph with n vertices, all of degree at most* 3. *If G is ordinary, then $n/3$ vertices can be removed from G such that the remaining graph is acyclic. If G is exceptional, then $3n/8$ vertices suffice unless $G = K_4$, K_4^+ or $2K_4$.*

Proof. All graphs in this proof will have degrees at most 3, even if not explicitly stated. The first statement in the theorem will be proved by induction on n. The second statement will be proved afterwards. Note that $f(G) \leqslant n/3$ for any graph (including the empty graph) with $n \leqslant 1$. Assume that G is any ordinary graph with $n > 1$ vertices and that the statement is true for all graphs with less than n vertices. The induction is somewhat delicate, requiring five cases.

Case 1. The graph G contains the graph in Fig. 2A or B as a subgraph. The following pairs of vertices are assumed non-adjacent: a, c and b, d in Fig. 2A and a, c and b, c in Fig. 2B.

Remove the four non-labeled vertices in Fig. 2 and add edges ac and bd in Fig. 2A (ac and bc in Fig. 2B) to form a connected graph G' with $n-4$ vertices. Let S' be a G'-set, and let S be the union of S' and the "large" vertex in Fig. 2A or 2B. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. By induction $f(G') \leqslant (n-4)/3$ if G' is ordinary. Therefore $f(G) \leqslant f(G') + 1 \leqslant (n-4)/3 + 1 < n/3$. Concerning the exceptional cases, if $G' = K_4$, then G, having 8 vertices, is also exceptional. For the case $G' = K_4^+$, G has 9 vertices and $f(G) = f(K_4^+) + 1 = \frac{9}{3}$. If G' has 8 vertices, then G has 12 vertices and, by Lemma 1, $f(G) = f(G') + 1 = 4 = \frac{12}{3}$. If G' is exceptional containing K_4^{++}, then either G also contains K_4^{++}, hence is exceptional, or G consists of two ordinary graphs G_1 (with n_1 vertices) and G_2 (with n_2 vertices) joined by a bridge. (A *bridge* is an edge whose removal from the graph results in an additional component.) In this case, by induction $f(G) = f(G_1) + f(G_2) \leqslant n_1/3 + n_2/3 = n/3$.

Case 2. The graph G has at least one vertex of degree less than 3.

If G has a vertex v of degree 1, then remove v to obtain a graph G'. If G' is ordinary, then by induction $f(G) = f(G') \leqslant (n-1)/3 \leqslant n/3$. The cases where G' is exceptional are easily checked.

Assuming no vertices of degree 1, let v be a vertex of degree 2 with adjacent vertices a and b. We consider two cases. First assume that a and b are adjacent. Remove vertices v, a, b to obtain a graph G' with $n-3$ vertices, at least one vertex of degree less than 3 (at least one in each component if G' has two components). Let S be the union of a G'-set and vertex a. If G' is ordinary, then by induction

$f(G) \leqslant f(G') + 1 \leqslant (n-3)/3 + 1 = n/3$. If G' is exceptional, then G must contain K_4^{++} and hence also be exceptional.

Now assume that a and b are not adjacent. Form a new graph G' by removing vertex v from G and adding an edge joining a and b. If G' is ordinary, then by induction $f(G) = f(G') \leqslant (n-1)/3 < n/3$. So assume that G' is exceptional. If $G' = K_4$ then $G = K_4^+$ is again exceptional. If $G = K_4^+$, then G has 6 vertices. It is easy to verify that the few graphs G on 6 vertices satisfy $f(G) \leqslant 2 = \frac{6}{3}$. If G' is a cubic graph with 8 vertices then G has 9 vertices and, according to Lemma 1, $f(G) = f(G') = 3 = \frac{9}{3}$. Finally consider the case where G' consists of a graph G_0 with K_4^{++} attached. If G also contains K_4^{++} then G is also exceptional. Otherwise G_0 is ordinary and by induction $f(G) = f(G_0) + 2 \leqslant (n-6)/3 + 2 = n/3$.

In light of Case 2, we can now assume that G is cubic in Cases 3–5.

Case 3. The graph G has a bridge.

Denote the two components after removing the bridge by G_1 and G_2 with number of vertices n_1 and n_2, respectively. If they are both ordinary, then by induction $f(G) = f(G_1) + f(G_2) \leqslant n_1/3 + n_2/3 = n/3$. If G_1 is exceptional, then by Case 2 either $G_1 = K_4^+$ or G_1 contains K_4^{++}; in either case G is exceptional.

Case 4. The graph G has an edge cutset consisting of two edges as in Fig. 3 where G_0 has 4, 6 or 8 vertices.

Consider first the case where G_0 has 6 vertices. It is easy to check that, for any graph G_0 on 6 vertices (each of degree at most 3) there is a set S_0 consisting of 2 vertices such that $G_0 - S_0$ is acyclic and, moreover, one of the vertices in S_0 can be

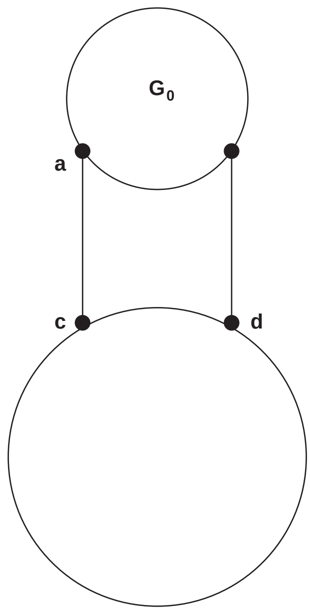

Fig. 3. Case 4.

assigned arbitrarily. In Fig. 3, choose vertex a to be in S_0. Remove G_0 from G to obtain a graph G' on $n-6$ vertices. If G' is disconnected, then the situation reverts to Case 3. Let S' be a G'-set and let $S = S' \cup S_0$. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. Since G' has two vertices of degree less then 3, G' is either ordinary or contains K_4^{++}. In the exceptional case, G is again exceptional. If G' is ordinary, then by induction $f(G) \leqslant f(G') + 2 \leqslant (n-6)/3 + 2 = n/3$.

Next consider the case where G_0 has 8 vertices. By Lemma 1 there is a set S_0 of 2 vertices such that $G_0 - S_0$ is acyclic. If c is not adjacent to d in Fig. 3, then remove G_0 from G and an add edge cd to obtain a connected cubic graph G' with $n-8$ vertices. Let S' be a G'-set and let $S = S' \cup S_0$. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. If G' is ordinary, then by induction $f(G) \leqslant f(G') + 2 \leqslant (n-8)/3 + 2 < n/3$. Concerning the cases where G' is exceptional, if $G' = K_4$ then G has 12 vertices and $f(G) = f(K_4) + 2 = \frac{12}{3}$. If G' has 8 vertices, then G has 16 vertices and $f(G) = f(G') + 2 = 5 < \frac{16}{3}$. Finally, if G' contains K_4^{++}, then G has a bridge which has already been considered in Case 3.

If c is adjacent to d in Fig. 3, then remove G_0 and vertices c, d from G to obtain a graph G' with $n-10$ vertices. If G' is disconnected, then the situation reverts to Case 3. Let S' be a G'-set and let $S = S' \cup S_0 \cup \{c\}$. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. Since G' has two vertices of degree less then 3, G' is either ordinary or contains K_4^{++}. In the exceptional case, G is again exceptional. If G' is ordinary, then by induction $f(G) \leqslant f(G') + 3 \leqslant (n-10)/3 + 3 < n/3$.

Finally consider the case where G_0 has 4 vertices. If c is adjacent to d in Fig. 3, then the graph G' induced by G_0 and vertices $\{c, d\}$ has 6 vertices attached to the rest of G by two edges, a situation already considered. If c is not adjacent to d, then remove G_0 from G and add an edge cd to obtain a connected cubic graph G' with $n-4$ vertices. There is a vertex v_0 in G_0 such that $G_0 - v_0$ is acyclic. Let S' be a G'-set and let $S = S' \cup \{v_0\}$. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. If G' is ordinary then by induction $f(G) = f(G') + 1 \leqslant (n-4)/3 + 1 < n/3$. Concerning the cases where G' is exceptional, if $G' = K_4$, then G is also exceptional, having 8 vertices. If G' has 8 vertices, then G has 12 vertices and, by Lemma 1, $f(G) = f(G') + 1 \leqslant 3 + 1 = \frac{12}{3}$. If G' contains K_4^{++}, then either G also contains K_4^{++}, and hence is exceptional, or G has a bridge, a situation already considered in Case 3.

Case 5. Let v be any vertex of G that is not a cutvertex. Such a vertex must exist in a cubic graph. The vertex v is incident with either 0, 1, 2 or 3 triangles. Once these cases have been considered, the proof of the first statement in Theorem 2 is complete. If 3 triangles are incident with v, then $G = K_4$; if v is incident with exactly 2 triangles, then G contains a graph that has already been considered in Case 3. Thus G must contain one of the graphs in Figs. 4A or B. In each figure, it is possible that some vertices of degree 1 coincide.

Case 5A. Graph G contains the graph in Fig. 4A.

It may be assumed that a_1 is not adjacent to a_2, b_1 not adjacent to b_2, and c_1 not adjacent to c_2. Otherwise, if say a_1 is adjacent to a_2, since v is not a cutvertex of G, neither is a. So choose a instead of v as the non-cutvertex. Then G contains the graph in Fig. 4B, and Case 5B below applies.

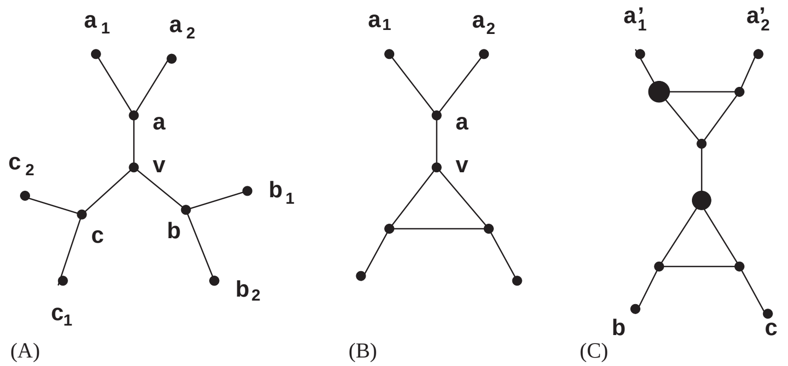

Fig. 4. Case 5.

Remove the vertices v, a, b, c from G and add 3 edges a_1a_2, b_1b_2, c_1c_2 to form a connected cubic graph G' with $n-4$ vertices. If any two of the following three unordered pairs are equal: $\{a_1,a_2\}$, $\{b_1,b_2\}$, $\{c_1,c_2\}$, then G' has a multiple edge, a case we consider later. Let S' be a G'-set and let $S=S'\cup\{v\}$. That $G'-S'$ is acyclic implies that $G-S$ is acyclic. The inductive argument is now exactly as in Case 1 except when G' is an exceptional graph containing K_4^{++}. In this case if G is not itself exceptional, then it is easy to check that G must be of the form already considered in Case 4.

It remains to consider what happens if at least two of the following three unordered pairs are equal: $\{a_1,a_2\}$, $\{b_1,b_2\}$, $\{c_1,c_2\}$. If all three are equal then $G=K_{3,3}$, a graph of order 6, for which $f(K_{3,3})=2=\frac{6}{3}$. If two pairs are equal, say $a_1=b_1$ and $a_2=b_2$, then choose a as the non-cutvertex instead of v. Then a_1 is adjacent to a, to b, and to a third vertex, say c_1; similarly a_2 is adjacent to a to b, and to, say c_2. If no two of the vertices c, c_1, c_2 are identical, then the situation is exactly as in the paragraph above. If all three are equal, then $G=K_{3,3}$ of order 6. If exactly two are equal, then it is easy to check that G is of the form already considered in Case 4.

Case 5B. Graph G contains the graph in Fig. 4B.

The situations $b=c$ or $a_1=b$ or $a_2=c$ are covered by Cases 1,3 or 4. The situation b adjacent to c is also covered by Case 1. So we can assume that the 4 vertices of degree 1 are distinct and b and c are non-adjacent.

If a_1 is not adjacent to a_2, then remove v and the three adjacent vertices from G and add 2 edges a_1a_2 and bc to form a connected cubic graph G' with $n-4$ vertices. The inductive argument is now identical to that of Case 5A.

If a_1 is adjacent to a_2 then, referring to Fig. 4C, we may again assume that $b\neq c$ and that b is not adjacent to c. Similarly we may assume that $a_1'\neq a_2'$ and that a_1' is not adjacent to a_2'. Also it may be assumed that $\{a_1',a_2'\}$ is not equal to $\{b,c\}$; otherwise G is a graph already considered in Case 4.

Remove the 6 vertices on the two triangles in Fig. 4C and add edge bc to form a connected (no longer cubic) graph G' with $n-6$ vertices. Let S' be a G'-set

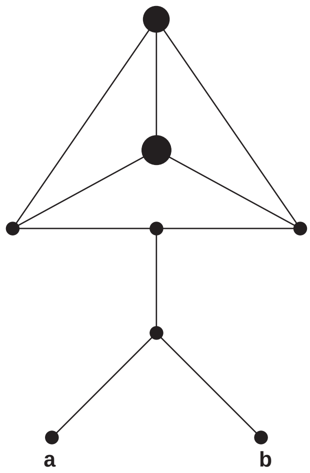

Fig. 5.

and let S be the union of S' and the "large" vertices in Fig. 4C. That $G' - S'$ is acyclic implies that $G - S$ is acyclic. Since G' has two vertex of degree 2, it is either ordinary or an exceptional graph containing K_4^{++}. If G' is ordinary, then by induction $f(G) \leqslant f(G') + 2 \leqslant (n-6)/3 + 2 = n/3$. If G is an exceptional graph containing K_4^{++}, then it is easy to check that G must also contain K_4^{++} or be a graph already considered in Case 4.

It has now been shown that $f(G) \leqslant n/3$ for ordinary graphs. It only remains to show that $f(G) \leqslant 3n/8$ for all other graphs except K_4, K_4^+ and $2K_4$. The proof again is by induction on n. It obviously holds for $n = 0, 1$. So assume $n > 1$ and that the statement holds for all graphs with less than n vertices. We have already proved that the statement is true for ordinary graphs. By Lemma 1 we have $f(G) = 3n/8$ for all graphs with 8 vertices. It only remains to show that $f(G) \leqslant 3n/8$ for graphs other than $2K_4$ containing K_4^{++}. Referring to Fig. 5, there are two cases.

If a and b are adjacent, then remove K_4^{++} and vertices a and b to obtain a graph G' with $n - 8$ vertices. If S' is a G'-set, then let S be the union of S', the two "large" vertices in Fig. 5 and vertex a. Then $G' - S'$ acyclic implies $G - S$ acyclic. It is not possible that $G' = K_4$ or $G' = K_4^+$ or $G' = 2K_2$, so by induction $f(G) \leqslant f(G') + 3 \leqslant 3(n-8)/8 + 3 = 3n/8$.

If a and b are not adjacent (or if either a or b do not exist), then remove K_4^{++} and add an edge ab (if both a and b exist) to obtain a graph G' with $n - 6$ vertices. If S' is a G'-set, then let S be the union of S' and the two "large" vertices in Fig. 5. Then $G' - S'$ acyclic implies $G - S$ acyclic. If $G' \neq K_4$, $G' \neq K_4^+$ and $G' \neq 2K_4$, then by induction $f(G) = f(G') + 2 \leqslant 3(n-6)/8 + 2 < 3n/8$. If $G' = K_4$, then $G = 2K_4$. If $G' = K_4^+$, then G has 11 vertices and $f(G) = 4 < (3 \cdot 11)/8$. If $G' = 2K_4$, then G has 16 vertices and $f(G) = 6 = (3 \cdot 16)/8$. $\square$

Theorem 2 gives a tight upper bound on $f(G)$ for a connected cubic graph G. Obtaining a tight lower bound is easier.

Theorem 3. *Let G be a connected cubic graph with n vertices. If S is a set of vertices such that $G-S$ is acyclic, then $|S| \geqslant n/4+1/2$. Moreover, infinitely many graphs attain this bound.*

Proof. Let $|S|$ be a G-set. The number $3|S|$ counts each edge not in $G-S$ at least once. Also the number of edges in the acyclic graph $G-S$ is at most $n-|S|-1$, the number of edges in a tree on $n-|S|$ vertices. Then for the total number of edges in G we have

$$3n/2 \leqslant 3|S| + (n-|S|-1) = n+2|S|-1,$$

which implies that $|S| \geqslant (n+2)/4$.

To see that infinitely many graphs attain this bound, consider any tree T on $(3n-2)/4$ vertices (assuming $n \equiv 2 \bmod 4$) with no vertex of degree greater than 3. Add to T a set S consisting of $(n+2)/4$ independent vertices, each adjacent to exactly 3 vertices in T, thus forming a graph with $|S|+|V(T)|=(n+2)/4+(3n-2)/4=n$ vertices. This can be done (in many ways) as long as the total number of edges is $3n/2$, as required of a cubic graph. But this is so because: $3|S|+|E(T)|=3(n+2)/4+((3n-2)/4-1)=3n/2$. □

The edge version of the extremal graph problem asks for the minimum number of edges in a connected cubic graph G such that their removal results in an acyclic graph. (When an edge is removed, its endpoints remain.) The result in the edge case is almost trivial, but we state it for use in the next section.

Lemma 4. *Let G be a connected cubic graph with n vertices. The minimum number of edges in a set S such that $G-S$ is acyclic is exactly $n/2+1$.*

Proof. Assume that $G-S$ is acyclic. Add edges to $G-S$ if necessary to form a spanning tree T of G. Since T has $n-1$ edges, $3n/2-(n-1)=n/2+1$ edges must be removed from G to obtain T. Hence at least that number must be in S. □

3. Upper bound on the integrity of cubic graphs

What we seek are constants c and c' such that $I(G) \leqslant cn$ and $I'(G) \leqslant c'n$ for all connected cubic graphs on n vertices. Theorem 7 below states that $\frac{1}{3}$ is such a constant for integrity and $\frac{1}{2}$ for edge integrity, up to $O(\sqrt{n})$.

The proof of Theorem 7 requires the solution to the extremal problem in Section 2 and two lemmas concerning the integrity of trees. The first lemma states that, of all trees on n vertices, the integrity is maximized for a path. This result is a subject of [10], but we include our short proof for completeness. The edge version of Lemma 5 is false, even if restricted to trees where each vertex has degree at most 3. Lemma 6 below, however, suffices for our purposes.

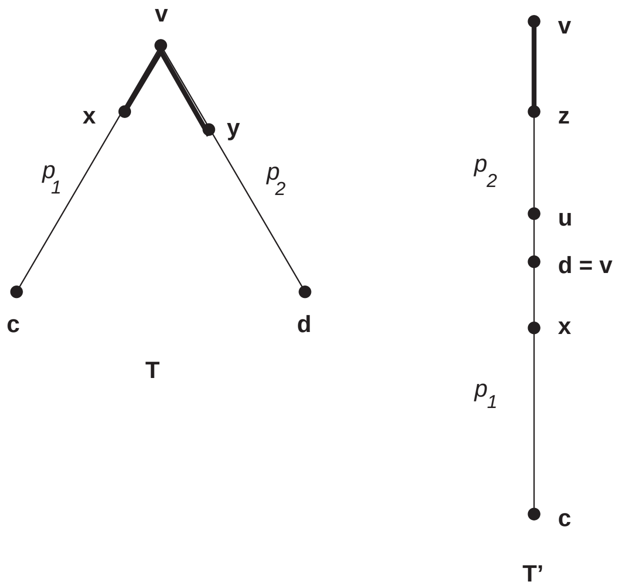

Fig. 6.

Lemma 5. *Of all trees with n vertices, the path has the maximum integrity.*

Proof. The proof is by induction on the number $w = w(T)$ of vertices of degree 1 in T. If $w = 2$, then T is a path. So let T be a tree with $w > 2$ and assume the theorem is true for all trees with smaller w.

Regarding T as rooted at one of its vertices of degree 1, let v be any vertex of degree at least 3 with no descendents of degree greater than 2. The subgraph of T induced by the descendents of v is the disjoint union of paths. Let p_1 and p_2 be any two of these paths. Define a new tree T' obtained from T by removing p_1 and attaching it to the end of p_2 as in Fig. 6. By abuse of language, we allow the notation p_1, p_2 for the subgraphs of both T and T' corresponding to these paths. Note that T and T' have the same number of vertices, but $w(T') = w(T) - 1$. By the induction hypothesis $I(T') \leqslant I(P_n)$, so it only remains to show that $I(T) \leqslant I(T')$.

Let S' be a set of vertices of T' such that $|S'| + m(T' - S') = I(T')$. For simplicity, denote $m := m(T' - S')$. Define a set S of vertices of T as follows. Assume that there are k (possibly 0) vertices in S' that lie on p_2 in T'. Choose a set $S_0 = \{x_1, x_2 \ldots\}$ of vertices on p_2 in T as follows. Starting from the leaf d, let x_i be the $i(m + 1)$st vertex. In other words, successive vertices in S_0 are separated along the path p_2 by m vertices not in S_0. The set S_0 should have as many such vertices as fit on p_2, but not more than k. If k vertices do not fit, then add vertex v to S_0. Let S_1 be the set of vertices of T that correspond to the set of vertices of S' in T' that do not lie on p_2, and let $S = S_0 \cup S_1$. By definition $|S| \leqslant |S'|$. To complete the proof it only remains to show that $m(T - S) \leqslant m(T' - S')$.

By the way S was defined, all components of $T-S$ except possibly the one containing v, have at most $m(T'-S')$ vertices. If $v\in S$, then there are no exceptions and the proof is complete. Otherwise, to verify the one exception, it suffices to show that the number a of vertices in the boldface section of T in Fig. 6 (vertices v, x and y excluded) is at most the number b of vertices in the boldface section of T' (vertices v and z excluded). In Fig. 6, vertices $x, y\in S$ and $z\in S'$, but the boldface sections contain no vertices of S or S'. Also the boldface sections do not contain v. If p denotes the number of vertices on path p_2 and q denotes the number of vertices between v and x in T, then $a=p-k(m+1)+q=p-[(k-1)m+k+(m-q)]\leqslant b$. The first equality is obtained by subtracting the number of vertices between (and including) y and the leaf d from the number of vertices on p_2. The inequality is obtained by subtracting the number of vertices between (and including) b and z from the number of vertices on p_2. Note that the number of vertices between consecutive vertices of S' on p_2 in T' is at most m, and the number of vertices between d (included) and the last vertex u in S' along p_2 (not included) is at most $m-q$. □

Lemma 6. *If T is a tree with no vertex of degree greater than* 3, *then $I'(T)<2\sqrt{2n}$.*

Proof. Consider T rooted at any vertex of degree less than 3. If x is any node in the tree, denote by $g(x)$ the number of descendents of x plus 1 (for x itself). Let m be a positive integer. If $1<m\leqslant n/2$, then we claim that there exists a node x such that $m\leqslant g(x)<2m$ and $g(p(x))\geqslant 2m$ where $p(x)$ is the parent of x. Note that x cannot be the root. Assuming that the claim is true, then deleting the edge from x to its parent divides T into two connected components, one, say S_1, of order less than $2m$ and the other called T_1. Now perform the procedure again on T_1 to obtain S_2 and T_2. Repeat until either $m>|V(T_i)|/2$. At termination at stage k, the components $S_1, S_2, \dots, S_k$ and T_k have order less than $2m$. The number k of edges removed from T is one less than the number c of components. So by the definition of edge integrity $I'(T)\leqslant(2m-1)+(c-1)$. Since $n=\sum_{i=1}^{k-1}|S_i|+|S_k|+|T_k|=\sum_{i=1}^{k-1}|S_i|+|T_{k-1}|\geqslant(k-1)m+2m$, we have $c=k+1\leqslant n/m$. This implies that $I'(T)\leqslant(2m-1)+(n/m-1)$. Let $m=\left\lfloor\sqrt{n/2}\right\rfloor$. Note that $1<\lfloor n/2\rfloor\leqslant n/2$ unless $n\leqslant 3$, in which case the lemma can be easily checked. Then $I'(T)<2\sqrt{2n}$, which takes a little elementary checking.

To prove the claim, mark a node x with B (too big) if $g(x)\geqslant 2m$ and mark it L (too little) if $g(x)<m$. The claim is proved if T has an unmarked node. It can be assumed that such a node x satisfies $g(p(x))\geqslant 2m$; otherwise replace x by $p(x)$, which also must be unmarked. If it is still the case that $g(p^2(x))\geqslant 2m$, then replace $p(x)$ by $p^2(x)$. This process must halt because $g(r)=n\geqslant 2m$ where r is the root. Note that the root is marked B and that each leaf is marked L. Also, if x is marked B, then not both of its (at most 2) children can be marked L. Starting at any leaf x, perform the following procedure. If x is marked L then replace x by its parent; if x is marked B then replace x by a child not marked L; if x is unmarked then stop. If the algorithm terminates, then the claim is proved. Assume that there is a B node in the sequence. A B node is always followed by a child not marked L. Therefore we eventually arrive at an unmarked node or a B leaf, which is impossible. So either there is an unmarked

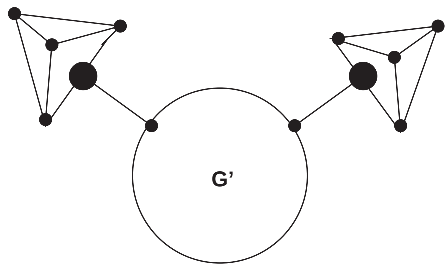

Fig. 7.

node in the sequence or no B nodes are encountered. An L node is always followed by its parent, so we arrive at the root marked L. But the root is a B node. $\square$

Theorem 7. *If G is a connected cubic graph with n vertices, then*

$$I(G) < \frac{n}{3} + \frac{2}{3}\sqrt{6n} + \frac{1}{3} \qquad I'(G) < \frac{n}{2} + 2\sqrt{2n} + 1.$$

Proof. Consider edge integrity first. By Lemma 4 there is a set S consisting of at most $n/2+1$ edges whose removal results in a tree T. Then by Lemma 6 we have $I'(G) \leqslant n/2 + 1 + I'(T) < n/2 + 1 + 2\sqrt{2n}$.

By Theorem 2 the same reasoning shows that $I(G) \leqslant n/3 + I(T)$ when G is ordinary. In this case T is a tree with $2n/3$ vertices and, by Lemma 5 and Eq. (1) in the introduction,

$$I(G) \leqslant \left\lfloor \frac{n}{3} \right\rfloor + \left\lceil 2\sqrt{\left\lceil \frac{2n}{3} \right\rceil + 1} \right\rceil - 2 < \frac{n}{3} + \frac{2}{3}\sqrt{6n}. \tag{2}$$

There remains the case where G is exceptional. It is easy to verify that the formula holds for $G = K_4, K_4^+$ and any cubic graph with 8 vertices, so assume that G contains K_4^{++} as a subgraph. Remove every copy of K_4^{++} appearing as a subgraph of G. Assume that there are k such copies. Fig. 7 shows the case $k=2$. What remains is either an ordinary graph G' or K_4^+. In the later case Theorem 7 can be easily checked, so assume that G' is ordinary.

Call a set X of vertices such that $I(G) = |X| + m(G-X)$ a *G-integrity set*. If S' is a G'-integrity set, then consider the union S of S' and the k additional "large" vertices shown in Fig. 7. Note that $m(G-S) \geqslant 4$. If $m(G'-S') \geqslant 4$, then applying inequality (2) to the ordinary graph G' yields

$$I(G) \leqslant k + I(G') < k + \left(\frac{n-5k}{3}\right) + \frac{2}{3}\sqrt{6(n-5k)} < \frac{n}{3} + \frac{2}{3}\sqrt{6n}$$

and the proof is complete.

If $m = m(G' - S') < 4$, then the first inequality of (2) applied to G' yields

$$I(G) \leqslant k + I(G') < \frac{n}{3} + \frac{2\sqrt{6n}}{3} - \left(1 + \frac{2k}{3}\right) + (4 - m),$$

which is less than or equal to $n/3 + 2\sqrt{6n}/3 + 1/3$ unless $m = 1$ and $k \leqslant 2$. If $m = 1$ and $k \leqslant 2$, then choose three vertices on K_4^{++} (instead of one) to be in S so that also $m(G - S) = 1$ and, from the first inequality of (2) applied to G', we have

$$I(G) \leqslant 3 + I(G') < 3 + \left(\frac{n - 5k}{3}\right) + \frac{2}{3}\sqrt{6n} - 1 \leqslant \frac{n}{3} + \frac{2}{3}\sqrt{6n} + \frac{1}{3}. \quad \square$$

4. Existence of cubic graphs with large integrity

Theorem 7 shows that the integrity of a cubic graph G is bounded above by a linear function in the number of vertices of G. There is no a priori reason to believe, however, that there actually exist infinite families of cubic graphs with anywhere near this large integrity. Our proof that such families do exist depends on expander graphs.

For a set A of vertices of a graph G, denote the set of *neighbors* of A by $N(A) := \{v \in V(G) \mid v \,\mathrm{adj}\, u \in A\}$ and denote the *boundary* of A by $\partial A := N(A) - A$. We say that a graph G on n vertices has the (α, β)-*expanding property* if, for any set A of at most αn vertices, $|\partial A| \geqslant \beta |A|$.

Theorem 8. *For a connected cubic graph G on n vertices with the (α, β)-expanding property, we have*

$$I(G) \geqslant n \min\left(\alpha, \frac{\beta}{3 + \beta}\right), \qquad I'(G) \geqslant n \min(\alpha, \beta/2).$$

Proof. Concerning the second inequality, consider any set S of edges of G and let $A_1, A_2, \ldots, A_m$ be the connected components that result when the edges in S are removed from G. Let $|A_i|$ denote the number of vertices in the component A_i. Assume that $|A_i| \leqslant \alpha n$ for each i. If d_i denotes the number of edges with exactly one endpoint in A_i, then, using the (α, β)-expanding property,

$$2|S| = \sum_i d_i \geqslant \sum_i |\partial A_i| \geqslant \sum_i \beta |A_i| = \beta n.$$

Therefore either $G - S$ has a component of size greater than αn or else $|S| \geqslant \beta n/2$. In either case $I'(G) := \min_{S \subset E} \{|S| + m(G - S)\} \geqslant n \min(\alpha, \beta/2)$.

Concerning the first inequality of Theorem 8, consider any set S of vertices of G and let $A_1, A_2, \ldots, A_m$ be the connected components that result when the vertices in S are removed from G. Assume that $|A_i| \leqslant \alpha n$ for each i. Using the (α, β)-expanding property,

$$3|S| \geqslant \sum_i |\partial A_i| \geqslant \sum_i \beta |A_i| = \beta(n - |S|).$$

Therefore either $G - S$ has a component of size greater than αn or else $|S| \geqslant \beta n/(3 + \beta)$. In either case $I(G) = \min_{S \subset V} \{|S| + m(G - S)\} \geqslant n \min(\alpha, \beta n/(3 + \beta))$. $\square$

Because of their many applications, expanders have received considerable attention in recent years. Various definitions of expander are in common use. One definition is that a k-regular graph on n vertices is called a *β-expander* if, for all subsets A of vertices with $|A| \leqslant n/2$, we have $|\partial A| \geqslant \beta|A|$. In other words, take $\alpha = \frac{1}{2}$ in the definition of (α, β)-expanding property. Every connected k-regular graph is such a β-expander for some β. Of interest are infinite families of expanders where k and β are fixed and n goes to infinity. It is not too difficult, using random graphs, to show that such families exist [23], but finding explicit constructions has proved extremely difficult. (See, for example, the exposition of Lubotzky [23] of a method using the representation theory of semi-simple Lie groups.)

Ramanujan graphs are used in the proof below. A k-regular graph is a *Ramanujan graph* if the second largest (in absolute value) eigenvalue λ_1 of its adjacency matrix is not more than $2\sqrt{k-1}$. The value $2\sqrt{k-1}$ is optimum for an infinite family $\{G_i\}$ of k-regular graphs according to the following result of Alon–Boppama (see [2,24]): $\liminf_{i\to\infty} \lambda_1(G_i) \geqslant 2\sqrt{k-1}$.

Theorem 9. *There exist an explicitly constructed infinite family $\mathscr{G}$ of connected cubic graphs such that $I(G) > 0.038n$ and $I'(G) > 0.058n$ for all $G \in \mathscr{G}$.*

Proof. The proof relies on known constructions of Ramanujan graphs by Lubotzky et al. [24] and independently by Margulis [26]. In particular, an infinite family $\mathscr{G}$ of such graphs have been explicitly constructed for $k = 3$ by Chiu [15]. The expansion of a k-regular graph can be bounded from below in terms of λ_1. Substituting $k = 3$ and $\lambda_1 = 2\sqrt{k-1}$ in such a result of Chung [16, Lemma 6.3] yields

$$\frac{|\partial A|}{|A|} > \frac{1 - |A|/n}{8 + |A|/n}$$

for any subset A of vertices of any graph in the family $\mathscr{G}$. Let $\alpha_0 = (\sqrt{297} - 17)/4 > 0.058$. A simple calculation shows that if $|A|/n \leqslant \alpha_0$, then $|\partial A|/|A| > (1 - |A|/n)/(8 + |A|/n) \geqslant 2\alpha_0$. In the terminology of Theorem 8, this means that each graph in $\mathscr{G}$ is an $(\alpha_0, 2\alpha_0)$-expander. Theorem 8 then implies the bound $I'(G) > 0.058n$. To obtain the bound on $I(G)$, let $\alpha_0 = (\sqrt{684} - 26)/4 > 0.038$. If $|A|/n \leqslant \alpha_0$, then $|\partial A|/|A| > (1 - |A|/n)/(8 + |A|/n) \geqslant 3\alpha_0/(1 - \alpha_0)$. In the terminology of Theorem 8, this means that each graph in $\mathscr{G}$ is an $(\alpha_0, 3\alpha_0/(1 - \alpha_0))$-expander. Theorem 8 then implies $I'(G) > 0.038n$. □

5. Open questions

It is impressive that the upper and lower bounds on the integrity given in Theorems 7 and 9 are both linear in the number of vertices. There is a large gap, though, between the two constants. Atici [6] conjectures that the upper bound on $I(G)$ can be decreased to $n/4 + o(n)$. However, we are unable even to prove the following more modest statement.

Question 10. *For any sufficiently large connected cubic graph G with n vertices, there is a set S of $n/4$ vertices such that each component of $G-S$ contains at most $n/4$ vertices.*

Perhaps better constants can be obtained in Theorem 9 using Bollobás' model [13] for random k-regular graphs, although the proof would not be constructive.

Acknowledgements

The author thanks Prof. M. Atici for valuable discussions on graph integrity.

References

[1] M. Albertson, R. Hass, A problem raised at the DIMACS graph coloring week, 1998.
[2] N. Alon, Eigenvalues and expanders, Combinatorica 6 (1986) 83–96.
[3] N. Alon, J. Kahn, P.D. Seymour, Large induced degenerate subgraphs, Graphs Combin. 3 (1987) 203–211.
[4] N. Alon, D. Mubayi, R. Thomas, Large induced forests in sparse graphs, J. Graph Theory 38 (2001) 113–123.
[5] M. Atici, Integrity of regular graphs and integrity graphs, J. Combin. Math. Combin. Comput. 37 (2001) 27–42.
[6] M. Atici, personal communication.
[7] K.S. Bagga, L.W. Beineke, W.D. Goddard, M.J. Lipman, R.E. Pippert, A survey of integrity, Discrete Appl. Math. 37/38 (1992) 13–28.
[8] R. Bar-Yehuda, D. Geiger, J. Naor, R.M. Roth, Approximation algorithms for the vertex feedback set problem with applications to constraint satisfaction and Bayesian inference, Proceedings of the Fifth Annual ACM–SIAM Symposium on Discrete Algorithms, Arlington, VA, 1994, ACM, New York, 1994, pp. 344–354.
[9] C.A. Barefoot, R. Entringer, H.C. Swart, Vulnerability in graphs—a comparative survey, J. Combin. Math. Combin. Comput. 1 (1987) 12–22.
[10] C.A. Barefoot, R. Entringer, H.C. Swart, Integrity of trees and powers of cycles, Congr. Numer. 58 (1987) 103–114.
[11] S. Bau, L.W. Beineke, G. Du, Z. Liu, R.C. Vandell, Decycling cubes and grids, Utilitas Math. 59 (2001) 129–137.
[12] L.W. Beineke, W. Lowell, R.C. Vandell, Decycling graphs, J. Graph Theory 25 (1997) 59–77.
[13] B. Bollobás, Random Graphs, Academic Press, London, 1985.
[14] J.A. Bondy, G. Hopkins, W. Staton, Lower bounds for induced forests in cubic graphs, Canad. Math. Bull. 30 (1987) 193–199.
[15] P. Chiu, Cubic Ramanujan graphs, Combinatorica 12 (1992) 275–285.
[16] R.K. Chung, Spectral graph theory, CBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence, RI, 1997.
[17] L.H. Clark, R.C. Entringer, M.R. Fellows, Computational complexity of integrity, J. Combin. Math. Combin. Comput. 2 (1987) 179–191.
[18] W. Goddard, H.C. Swart, Integrity in graphs: bounds and basics, J. Combin. Math. Combin. Comput. 7 (1990) 139–151.
[19] F. Jaeger, On vertex-induced forests in cubic graphs. Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing, Florida Atlantic University, Boca Raton, FL, 1974, pp. 501–512;
F. Jaeger, Congr. Numer., Vol. X, Utilitas Math., Winnipeg, Man., 1974.

[20] Y.D. Liang, M. Chang, Minimum feedback vertex sets in cocomparability graphs and convex bipartite graphs, Acta Inform. 34 (1997) 337–346.
[21] E.L. Lloyd, M.L. Soffa, C. Wang, On locating minimum feedback vertex sets, J. Comput. System Sci. 37 (1988) 292–311.
[22] C.L. Lu, C.Y. Tang, A linear-time algorithm for the weighted feedback vertex problem on interval graphs, Inform. Process. Lett. 61 (2) (1997) 107–111.
[23] A. Lubotsky, Discrete Groups, Expanding Graphs and Invariant Measures, Progress in Mathematics, Birkhäuser, Berlin, 1994.
[24] A. Lubotsky, R. Phillips, P. Sarnak, Ramanujan graphs, Combinatorica 8 (1988) 261–278.
[25] F.L. Luccio, Almost exact minimum feedback vertex set in meshes and butterflies, Inform. Process. Lett. 66 (1998) 59–64.
[26] G.A. Margulis, Explicit group-theoretical constructions of combinatorial schemes and their applications to the design of expanders and concentrators, Problemy Peredachi Informatsii (1988) 39–46.
[27] E. Speckenmeyer, On feedback vertex sets and nonseparating independent sets in cubic graphs, J. Graph Theory 12 (1988) 405–412.
[28] M.L. Zheng, X. Lu, On the maximum induced forests of a connected cubic graph without triangles, Discrete Math. 85 (1990) 89–96.