Discrete Mathematics 118 (1993) 145–156
North-Holland

The total interval number of a graph, I: Fundamental classes

Thomas M. Kratzke
Metron, Inc., Reston, VA 22090, USA

Douglas B. West
University of Illinois, Urbana, IL 61801, USA

Received 30 April 1990
Revised 18 November 1991

Abstract

Kratzke, T.M. and D.B. West, The total interval number of a graph, I: Fundamental classes, Discrete Mathematics 118 (1993) 145–156.

A *multiple-interval representation* of a simple graph G assigns each vertex a union of disjoint real intervals, such that vertices are adjacent if and only if their assigned sets intersect. The *total interval number* $I(G)$ is the minimum of the total number of intervals used in any such representation of G. We obtain the maximum value of $I(G)$ for n-vertex graphs ($\lceil (n^2+1)/4 \rceil$), n-vertex outerplanar graphs ($\lfloor 3n/2-1 \rfloor$), and m-edge connected graphs ($\lfloor (5m+2)/4 \rfloor$).

1. Introduction

Given a collection $\mathscr{F}$ of sets, the *intersection graph* of $\mathscr{F}$ is the undirected simple graph whose vertices correspond to the sets of the collection, with two vertices adjacent if and only if the corresponding sets intersect. The family of sets is then called an *intersection representation* of its intersection graph. The most well-studied class of intersection graphs is the *interval graphs*, which are the intersection graphs obtainable by assigning each vertex a single interval on the real line. More generally, we allow a representation f to assign each vertex a union of intervals on the real line; if G is the intersection graph of this collection of sets, then f is called a *multiple-interval representation* of G. Let $\#f(v)$ be the minimum number of intervals whose union is $f(v)$; note that these intervals are disjoint. If $\#f(v)=k$, we say that $f(v)$ *consists of* k intervals or that v is *assigned* k intervals.

Correspondence to: Douglas B. West, Department of Mathematics, University of Illinois, Urbana, IL 61801, USA.

There are two natural ways to use multiple-interval representations to measure how far a graph is from being an interval graph. Let $\mathscr{I}$ be the collection of all multiple-interval representations of G. Then $i(G)=\min_{f\in\mathscr{I}}\max_{v\in V(G)} \#f(v)$ is called the *interval number* of G; the interval graphs are the graphs with interval number 1. Interval number has been actively studied for a number of years, beginning with [10] and [6]. Another parameter, $I(G)=\min_{f\in\mathscr{I}}\sum_{v\in V(G)}\#f(v)$, is called the *total interval number* of G and can be viewed as minimizing the average number of intervals assigned per vertex instead of the maximum number. In studying $I(G)$, it is helpful in simplifying theorem statements and proofs to allow $\#f(v)=0$, so that isolated vertices contribute nothing to the representation or the count of intervals; we adopt this convention. Also, as in the study of $i(G)$, it is naturalto define the *depth* of a representation to be the maximum number of vertices whose images contain a single point on the line, and then the *depth-r total interval number* $I_r(G)$ is the minimum of $\sum \#f(v)$ over all representations of G with depth at most r.

Although introduced in [6], total interval number was not studied until Aigner and tAndreae [1] obtained extremal results for some fundamental families. In this paper we extend their results, resolving several of their conjectures. In future papers, we will present additional extremal and algorithmic results about this parameter. Most of this work appeared in the dissertation of the first author, accepted in 1987 [8]. Independently, the main results of this first paper were also obtained by Kostochka [7] and announced in the summer of 1988.

Aigner and Andreae obtained the maximum value of $I(G)$ for several classes of graphs on n vertices, including trees ($\lfloor(5n-3)/4\rfloor$), 2-connected outerplanar graphs ($\lfloor 3n/2-1\rfloor$), triangle-free planar graphs ($2n-3$), and triangle-free graphs ($\lceil(n^2+1)/4\rceil$). For the latter three classes, they conjectured that the upper bounds would still hold when the '2-connected' or 'triangle-free' restrictions were removed. We have proved these conjectures, and the proofs for outerplanar and general graphs on n vertices appear in this paper. The proof for planar graphs is quite lengthy and will appear in a later paper in this series. Aigner and Andreae also conjectured that the maximum total interval number for a connected graph with m edges is $\lfloor(5m+2)/4\rfloor$, which we also prove in this paper.

For triangle-free graphs, the total interval number is equivalent to another parameter. Let $t(G)$ be the minimum number of edge-disjoint trails in G such that every edge has an endpoint in at least one trail. Then $I(G)=m+t(G)$ for a triangle-free graph with m edges, which we will show shortly. In this paper, we use this only to provide lower bounds for $I(G)$ in graph classes. In the next paper in this series, it yields the NP-completeness of testing $I(G)=m+1$ even for triangle-free 3-regular planar graphs, a linear-time algorithm to compute $I(G)$ for trees, a characterization of the trees requiring $m+t$ intervals for fixed t, and an alternate proof of the extremal bound for connected graphs.

Because large cliques allow representation of many edges with few intervals, the total interval number tends to be maximized over a class of graphs by a triangle-free graph. Thus, denser classes of n-vertex graphs tend to allow a larger total interval number. In subsequent papers we will obtain the maximum total interval number for cacti ($\lfloor(18n-12)/13\rfloor$) and planar graphs ($2n-3$), which fit naturally in the spectrum of classes between trees, outer-planar graphs, and general graphs. It is interesting to compare these bounds with those known for the interval number. Trivially, $I(G) \leqslant ni(G)$. The maximum values of I/n for trees, cacti, outer-planar, planar, and general graphs are asymptotically 5/4, 18/13, 3/2, 2, and $n/4$. For $i(G)$, the corresponding values are 2, 2, 2, 3, and $\lceil(n+1)/4\rceil$. Only for general graphs is the average number of intervals per vertex not guaranteed to provide savings over the maximum. Also, total interval number provides a finer spectrum of complexity of graph classes than interval number.

In terms of the number of edges m, we have the characterization $I(G)=m+t(G)$ for triangle-free graphs and the extremal bound $I(G) \leqslant \lfloor(5m+2)/4\rfloor$ for connected graphs. These suggest studying classes of graphs with m edges. Here greater density permits greater flexibility in selecting trails, so we expect the maximum of $I(G)$ to decline as we add minimum degree or connectivity requirements. The results listed here will be presented in later papers. The maximum value of $I(G)$ for connected m-edge graphs with minimum degree at least 2 is $\lfloor(9m+1)/8\rfloor$. In general, for minimum degree k there is a best-possible bound $I(G) \leqslant c_d m + O(1)$, where c_d is a sequence decreasing toward 1. For $d \geqslant 3$, we have examples showing that $c_d \geqslant (2d^2+3)/(2d^2+2)$. Connectivity or edge-connectivity k is a stricter requirement than minimum degree k and permits yet smaller upper bounds on $I(G)$. In particular, $I(G) \leqslant \lfloor 10m/9\rfloor$ for 2-connected or 2-edge-connected graphs with m edges. We have sequences of arbitrarily large 3-connected and 3-edge-connected graphs for which $I(G)/m$ is at least $1+1/192$ and $1+1/193.5$, respectively. However, as noted by Aigner and Andreae, $I(G) \leqslant m+1$ for 4-edge-connected graphs.

We close the introduction with the structural characterization of $I(G)$ for triangle-free graphs. Note that, as we traverse a depth-2 representation from left to right, we represent at most one additional edge for each interval encountered after the first, so $I_2(G) \geqslant m+1$, and this bound is achievable if G has a trail containing a vertex of every edge. We call such a graph *trail-coverable*; the trees that are trail-coverable have also been called *caterpillars*. In general, a *vertex cover* of G is a set of vertices that contains an endpoint of every edge of G. Analogously, we refer to an edge-disjoint collection of trails whose vertices together form a vertex cover as a *trail cover*, and we let $t(G)$ be the minimum number of trails in a trail cover of G. A *displayed interval* in a representation contains an open interval that belongs to no other interval of the representation.

Lemma 1. *For any graph G with m edges, $I(G) \leqslant I_2(G)=m+t(G)$. For a triangle-free graph with m edges, $I(G)=m+t(G)$.*

Proof. For triangle-free graphs, $I(G)=I_2(G)$. First we show that $I_2(G) \leqslant m+t$. Let $\{Z_j\}$ be a trail cover. Represent each trail $Z_j=(v_1,\ldots,v_r)$ by choosing r intervals in

$(j-1, j)$ such that the ith interval intersects only the $(i-1)$th and $(i+1)$th intervals (for $2 \leqslant i \leqslant r-1$) and assigning the ith of these intervals to v_i. Note that a given vertex may appear several times in a trail, and that all these intervals are displayed. For each edge not in these trails, assign an interval for one endpoint within the displayed portion of its neighbor in $\cup V(Z_j)$. In the first phase, since the edges of the trails are disjoint, the number of intervals used exceeds the number of edges represented by t. The second phase uses one more interval for each additional edge, so we have represented G with $m+t$ intervals.

Conversely, given an optimal depth-2 representation, we obtain a trail cover consisting of $I_2(G)-m$ edge-disjoint trails. Because no more than two intervals intersect at any point, we can eliminate any intersection of intervals by shortening or deleting one of them without affecting any other intersection. Therefore, we may assume that every edge is represented exactly once. Removal of each nondisplayed interval from an optimal representation leaves a representation of edge-disjoint trails as described above, having deleted one edge for each interval deleted. Furthermore, the vertices of the resulting trails touch all edges of the original graph. If we now shrink each trail to a single vertex by deleting one interval and edge at a time, we have deleted every edge of G and one interval for each of them. What remains is $I_2(G)-m$ intervals, one from each trail in the trail cover. □

In addition to n for the number of vertices of a graph and m for the number of edges, we use $N_G(v)$ for the neighbors of v in G and $x \leftrightarrow y$ to mean 'x is adjacent to y'.

2. Connected graphs with m edges

Lemma 1.1 makes it easy to obtain trees with $I(G) = \lfloor (5n-3)/4 \rfloor$. Take the star $K_{1,\lfloor (n-1)/2 \rfloor}$ and subdivide each edge into a path of two edges. Add one more vertex adjacent to the vertex of high degree if n is even. The resulting tree T_n has $\lceil (n-1)/2 \rceil$ leaves and $\lfloor (n-1)/2 \rfloor$ 2-valent vertices. Since a trail cover must touch each pendant edge, it must have a trail that ends within each path of length 2. Thus, T_n must have at least $\lceil \lfloor (n-1)/2 \rfloor /2 \rceil = \lfloor (n+1)/4 \rfloor$ trails in any trail cover, and this many suffice. Since $m = n-1$, we have $I(T_n) = \lfloor (5n-3)/4 \rfloor$. Note that the representation of T_n minimizing the total number of intervals assigns many intervals to the single high-valent vertex.

Aigner and Andreae [1] showed that no n-vertex tree has a larger interval number than T_n, and they conjectured that T_n also attains the maximum $I(G)$ among connected graphs with $n-1$ edges (including those with fewer than n vertices). Our proof of this contains a proof that T_n maximizes $I(G)$ among n-vertex trees, by a method slightly different from theirs. Note that the lack of any connectivity requirement allows an m-edge graph to have an interval number $2m$, achieved by m disjoint edges.

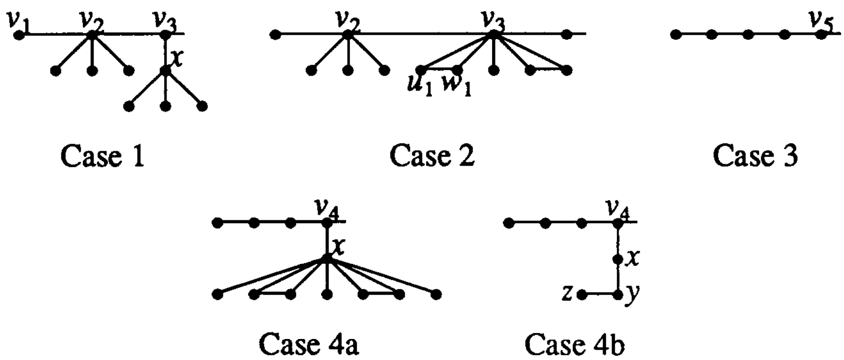

Fig. 1. Cases for establishing the inductive bound.

Theorem 2.1. *A connected graph with m edges has a depth-2 interval representation with at most* $\lfloor (5m+2)/4 \rfloor$ *intervals, if* $m \geqslant 2$*. Furthermore, for* $m \equiv 2 \bmod 4$ *the trees* T_n *defined above are the only connected graphs with* $I_2(G) = \lfloor (5m+2)/4 \rfloor$.

Proof. By induction on m. For connected graphs with up to 3 edges, we check the claim by inspection. In general, we want to prove $4I_2 \leqslant 5m+2$. For the induction step, it suffices to find a set $F \subset E(G)$ such that (a) $|F| \geqslant 4$, (b) F contains the edges of a trail incident to all edges of F, and (c) the subgraph $G' = G - F$ has only one nontrivial component. We then have $4I_2(G) \leqslant 4(I_2(G') + |F| + 1) \leqslant 5(m - |F|) + 2 + 4|F| + 4 = 5m + 6 - |F| \leqslant 5m + 2$.

Among all spanning trees of G, let T be a spanning tree for which the maximum path length (diameter) is largest. Among all longest paths in T, let $P = v_1, v_2, \ldots$ be a longest path that minimizes the distance from the endpoint v_1 to the first vertex v_r with degree at least 3. We will use P to split the edges of T into 2 connected subtrees R_1, R_2 that share a vertex. The subtree R_1 will have at least four edges and be trail-coverable. The desired F will be obtained by enlarging R_1 to include all edges induced by $V_1 = V(R_1)$. By construction, $G - F$ has only one nontrivial component. Since the covering trail in R_1 need not include all of V_1, it is not immediately obvious that F is trail-coverable.

To show this, in each case we will specify a trail in F and describe the additional edges $F' = F - E(R_1)$ that can be induced by V_1. Any other edge induced by V_1 would produce a spanning tree with a longer path than P or a path of the same length with a smaller value of r. The illustrations contain examples of every type of edge that can appear in F' in each case, except for chords of P. The cases depend primarily on the value of r and are illustrated in Fig. 1. In the discussion of cases, 'by the choice of T' means that otherwise we obtain a path longer than P in a spanning tree of G, and 'by the choice of P' means that otherwise we obtain a path of the same length in T with a smaller value of r.

Case 1: $r \in \{2, 3\}$ and T contains a vertex $y \notin P$ with $d_T(v_3, y) = 2$. Let v_3, x, y be the corresponding path in T. Split T at v_3, and let $V_1 = N_T(v_2) \cup N_T(x) \cup \{v_2, x\}$. The choices of T and P forbid other edges in the subgraph of G induced by V_1, except in the following subcases: v_1v_3 can occur if $r=3$, and if $r=3$ and $N_T(x)=\{v_3,y\}$ then v_3y can occur. In either case, the trail v_2, v_3, x covers all of F. (Note: In case 1 we need the existence of y to guarantee $|F|\geq 4$. Also, in this case T may have other paths of length two from v_3; since v_3 remains in the one nontrivial component, this causes no problem.)

Case 2: $r\in\{2,3\}$ and there is $y\notin P$ with $d_T(v_3,y)=2$. Split T at v_4, so $V_1=N_T(v_2)\cup N_T(v_3)$. By the choice of T and P, F contains no edge in $N_T(v_2)$ and no edge between the set $N_T(v_2)\cup\{v_2,v_4\}$ and the set $U=N_T(v_3)-\{v_2,v_4\}$. If U induces a pair of incident edges, such as uw and tu, then t,u,w,v_3 contradicts P. Hence, the only possible edges in F' are $(v_1v_4, v_2v_4\}$ and disjoint matched pairs $\{u_1w_1,\ldots,u_kw_k\}$ with $u_i,w_i\in U$. The covering trail in F is $v_1,v_2,v_3,u_1,w_1,\ldots,v_3,u_k,w_k$, visiting v_3 between each of these pairs. (Note: $v_1\leftrightarrow v_4$ is forbidden if U is nonempty, and $v_2\leftrightarrow v_4$ is forbidden if U induces any edges.)

Case 3: $r\geq 5$. Split T at v_r, and let R_1 be the path $v_1,\ldots,v_r$; the covering trail is R_1 itself. The graph F' may contain arbitrary chords of R_1.

Case 4: The only remaining possibility is $r=4$. Choose $x\in N_T(v_4)$, $x\notin P$, and let X be the set of vertices reachable from x via paths of T not passing through v_4. Split T at v_4, and let $V_1=\{v_1,v_2,v_3,v_4,x\}\cup X$.

Case 4(a): $X-N_T(x)=\emptyset$. If $N_T(x)$ induces a pair of incident edges, such as uw and tu, then t,u,w,x,v_4 contradicts the choice of T. Any edge between $\{v_1,v_2,v_3\}$ and $\{x\}\cup X$ contradicts the choice of T or the choice of P to minimize r. Hence, the only possible edges in F' are chords of v_1,v_2,v_3,v_4, disjoint matched pairs $\{u_1w_1,\ldots,u_kw_k\}$ with $u_i,w_i\in X$, and edges from X to v_4 (if $|X|\leq 2$). The covering trail in F is $v_2,v_3,v_4,x,u_1,w_1,\ldots,x,u_k,w_k$.

Case 4(b): $z\in X-N_T(x)$. Let y be the common neighbor of x and z in T. The choice of P implies $X=\{y,z\}$, and the choice of T forbids edges between $\{v_1v_2v_3\}$ and $\{x,y,z\}$. The only possible edges in F' are chords of v_1,v_2,v_3,v_4 and chords of v_4,x,y,z. The covering trail in F is v_2,v_3,v_4,x,y.

Now suppose $m\equiv 2 \bmod 4$, so that $(5m+2)/4$ is an integer. The procedure above yields an inductive proof that G achieves the extreme value $\lfloor(5n+2)/4\rfloor$ only if $G=T_{m+1}$. The basis step $m=2$ is trivial. Because $(5m+2)/4$ is an integer and the number of intervals added in the iterative step is $|F|+1$, the iteration produces a representation using fewer than $\lfloor(5m+2)/4\rfloor$ intervals, unless exactly four edges are removed from G at every step, until nothing but the two-edge path T_3 remains. Since at each step F contains at least four edges of T in the subtree R_1, this must be all of F at each step, with exactly four edges. The inductive hypothesis says that the single nontrivial component of $G-F$ is T_{m-3}. It remains only to consider the ways in which R_1 can be attached to $G-F=T_{m-3}$.

The possible 5-vertex R_1's, in the cases of $r=3,2,3,5,4$, respectively, are shown in Fig. 2, with the splitting vertex v_s placed rightmost. The cases indicated are those discussed above.

We want to show that only the first case can hold. If there is a single trail incident to all edges of R_1 and to one of the pendant edges of the remaining graph $G-F=T_{m-3}$,

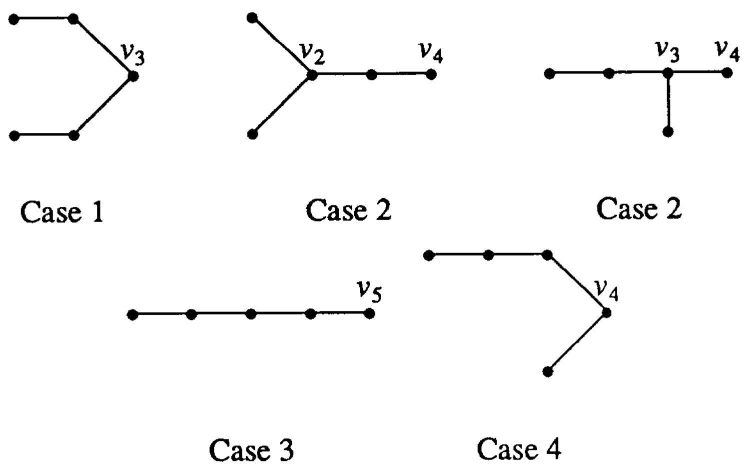

Fig. 2. Cases for achieving the inductive bound.

then that edge need not be covered by the remaining trails, and there is a trail cover of G with fewer disjoint trails than required by T_{m+1}. In this case, the lemma implies $I(G) \leqslant I_2(G) < I(T_{m+1})$. For all cases except case 1, the covering trail of $F = R_1$ ends at the vertex shared by R_1 and R_2, and the trail can be continued (if necessary) to reach a pendant edge of $G - F = T_{m-3}$ without leaving a disconnected subgraph when this larger F' is deleted. This argument applies also in the first case if the vertex of $G - F = T_{m-3}$ identified with v_3 is other than the central vertex of T_{m-3}. Hence, the only way to achieve the maximum is if R_1 is a 5-path, $G - R_1 = T_{m-3}$, and they are incident only by identifying their central vertices. This is precisely the construction $G = T_{m+1}$. (Note: when $m = 2$, the 'central vertex' of T_{m+1} is actually one of the endpoints of the 2-edge path.) □

The characterization of graphs achieving the maximum for other congruence classes mod 4 is messier, but it could be done by a similar analysis. We omit it here because it would be tedious, and because the next paper in this series contains a separate proof that the extremal graph must be a tree, an algorithm for computing the interval number of trees, and a resulting characterization of the trees requiring t trails, which also settles the matter.

3. Graphs on n vertices

Griggs [5] proved that $i(G) \leqslant \lceil (n+1)/4 \rceil$ for an n-vertex graph G, achieved by the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. Andreae [2] showed that this is the unique graph achieving the bound when n is even. Since $I(G) \leqslant ni(G)$, Griggs' bound implies $I(G) \leqslant \lceil (n^2+n)/4 \rceil$. Since $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is triangle-free and trail-coverable, we in fact have

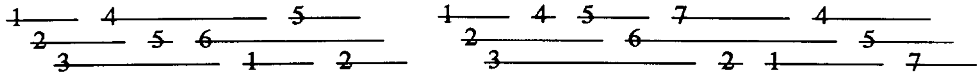

Fig. 3. Depth-3 representations of K_6 and K_7 with $\lfloor n^2/4 \rfloor$ intervals.

$I(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \lceil (n^2+1)/4 \rceil$. The extremal results of Griggs and Andreae on $i(G)$ are difficult, but it is not difficult to show that $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is the unique extremal graph for $I(G)$. In general, we can represent any other G with fewer intervals even when we restrict the depth of the representation to 3 (except for K_4 and K_5).

Theorem 3.1. *If G is an n-vertex graph not in $\{K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}, K_4, K_5\}$, then $I(G) \leqslant I_3(G) < I(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \lceil (n^2+1)/4 \rceil$.*

Proof. As noted above, $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is trail-coverable, triangle-free, and has $\lfloor n^2/4 \rfloor$ edges, so $I(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \lceil (n^2+1)/4 \rceil$. For $n \leqslant 3$, the claims hold by inspection. For K_4, depth 3 prevents four intervals from sharing a common point, so some vertex must get two intervals. More generally, a depth-3 representation with t intervals allows at most $2t-3$ edges to be represented (from left to right, each interval after the first two introduces at most two new edges), so depth-3 representations of K_n require at least $(n^2-n+6)/4$ intervals. Thus, K_4 and K_5 require $\lceil (n^2+1)/4 \rceil$ intervals in depth-3 representations, and this is achievable. For $n=6, 7$, Fig. 3 shows depth-3 representations with $\lfloor n^2/4 \rfloor$ intervals.

With these examples as a basis step, we give an inductive proof. We claim first that any $G \notin \{K_4, K_5, K_6, K_7, K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}\}$ has adjacent vertices u, v such that $G-u-v \notin \{K_4, K_5, K_{\lfloor (n-2)/2 \rfloor, \lceil (n-2)/2 \rceil}\}$. To see this, we may assume first that G is connected, so $\delta(G) \geqslant 1$; let x be a vertex of minimum degree. If $d(x) \leqslant \lfloor (n-1)/2 \rfloor$, we can choose $u \in N_G(x)$ and $v \in N_G(u) - \{x\}$. Otherwise, $\delta(G) \geqslant \lceil n/2 \rceil$, and every pair of adjacent vertices has a common neighbor. This means every edge lies on a triangle. If $d(x) < n-1$, delete the endpoints of any edge not involving x. If $d(x) = n-1$, then $n \geqslant 8$ by our restriction on G.

Now, let $H = G - \{u, v\}$. By the induction hypothesis, $I_3(H) \leqslant \lfloor (n-2)^2/4 \rfloor = \lfloor n^2/4 \rfloor - (n-1)$. We need only prove $I_3(G) \leqslant I_3(H) + (n-1)$. Let f be an optimal depth-3 representation of H. It is easy to represent the edges of G that involve $\{u, v\}$ using at most n additional intervals. We use a pair of intersecting displayed intervals I_u and I_v, and then we place an interval for each other vertex x into $I_u \cap I_v$, $I_u - I_v$, $I_v - I_u$, or nowhere to establish the adjacencies between x and $\{u, v\}$. Note that we are done if any vertex is independent of u and v. Hence, we may assume every other vertex is adjacent to at least one of $\{u, v\}$.

It suffices to show that we can use the rightmost intervals in f to save one interval in the process of representing the edges involving u and v. Let x, y be the vertices involved in the rightmost intersection in f. If either of $\{x, y\}$ has just one neighbor in $\{u, v\}$, then we extend its interval to overlap the first of $\{I_u, I_v\}$. If all four of these edges are present, then we extend $f(x)$ and $f(y)$ rightward to meet I_u, place I_v at the right end of I_u, and insert a small interval for v in the extended part of $f(x) \cap f(y)$. In each case we have added only three intervals for these four vertices. Note also that we never exceed depth 3 in these additions. □

4. Outerplanar graphs on *n* vertices

Aigner and Andreae established the bound of $I(G) \leqslant \lfloor 3n/2-1 \rfloor$ for 2-connected outerplanar graphs on n vertices. Because $I(G) \geqslant m+1$ for any triangle-free graph with m edges, this bound is attained by any triangle-free outerplanar graph with $\lfloor 3(n-2)/2 \rfloor$ edges. Such graphs can be constructed from an n-cycle by adding a matching consisting of $\lfloor n/2-3 \rfloor$ 'parallel' chords. In this section we extend the Aigner–Andreae bound to all outerplanar graphs. We first prove it for 2-connected outerplanar graphs, as they did, but we give a shorter proof for the first lemma. Recall that a *block* is a maximal subgraph with no cut-vertex; it is always an edge or a 2-connected graph.

Lemma 4.1 (Aigner and Andreae [1]). *The edges of a 2-connected outerplanar graph with at least 4 vertices can be 2-colored (using red and blue) so that*

(a) *the edges of the outer face are red,*

(b) *every triangle has at least one blue edge, and*

(c) *every blue edge e belongs to a triangle $T(e)$ consisting of e and two red edges, and these distinguished triangles are pairwise edge-disjoint.*

Proof. By induction on the number of faces. If G has only one bounded face, color all edges red. If G has two bounded faces and either is a triangle, color the chord blue; otherwise, color all edges red. Hence, suppose G has at least three bounded faces. If G has a *peripheral face* F (a face having exactly one non-outer edge) that is not a triangle, color the outer edges of F red, delete them, and apply the induction hypothesis.

Finally, if all peripheral faces are triangles, then for each peripheral face color the outer edges red and the chord blue. For each such blue edge e, the distinguished triangle $T(e)$ is the peripheral triangle containing it. Delete all the edges of these peripheral triangles. For each block H of the outer-planar graph that remains, we have three possibilities: (1) If H is a single edge, color it red. (2) If H is a triangle, its edges cannot all be outer edges of G. Assign blue to an edge e of H that is not an outer edge of G and put $T(e)=H$. (3) If H has at least four vertices, apply the induction hypothesis to it. Since all edges of the 'distinguished' peripheral triangles were deleted before applying the induction hypothesis, the resulting distinguished triangles are edge-disjoint. □

Lemma 4.2 (Aigner and Andreae [1]). *If G is an outerplanar block, then $I(G) \leqslant \lfloor 3n/2-1 \rfloor$.*

Proof. If G has only two vertices, the formula applies; so we may assume $n \geqslant 3$ and G is 2-connected. Let R be the set of red edges in a coloring of the edges of G as guaranteed by Lemma 4.1. We will construct a representation of G using at most $1+|R|$ intervals. This will complete the proof, because a triangle-free outerplanar graph has at most $1+3(n-2)/2=3n/2-2$ edges.

Number the vertices of G as $v_1, \ldots, v_n$ in order around the outside face. If G is more than a simple cycle, we choose v_1 to be a vertex with degree at least 3. Represent the n red edges of the cycle $v_1, \ldots, v_n$ using $n+1$ displayed intervals (the first and last assigned to v_1) such that each interval except the first and last intersects two others. Call these the 'outer intervals'. For red chords that belong to no distinguished triangle, we add a small interval for one endpoint in a displayed portion of the outer interval for the other endpoint.

The remaining edges belong to the edge-disjoint distinguished triangles, each containing one blue edge. Let $T=v_iv_jv_k$ be a distinguished triangle, with v_iv_j its blue edge. Suppose that s of the red edges of T are chords of G and therefore not yet represented; we must represent these and v_iv_j by adding only s new intervals. If $s=2$, place small intersecting intervals for v_i and v_j within the displayed portion of the outside interval for v_k. If $s=1$, let v_jv_k be the outside edge of G in T, and place a small interval for v_i within the intersection of the outside intervals for v_j and v_k.

Finally, if $s=0$, then $T=v_iv_jv_k$ has two outside edges of G, meeting at v_k. Let I be the outside interval for v_k. Note that $d(v_k)=2$, so $k \neq 1$ and no intervals for chords were placed within I. Thus, we can extend the outside intervals for v_i and v_j until they intersect each other within I, providing the desired additional edge (and no others), without adding any intervals. □

Note that distinguished triangles with two outside edges (the troublesome 'peripheral triangles' of Lemma 4.1) cause us to cover up the interval used for a vertex of degree 2. Indeed, if $G=K_4-e$, then $I(G)=4$, but G has no representation with at most five intervals in which every vertex has a displayed interval.

In extending the bound to arbitrary outer-planar graphs, the following lemma is useful and interesting in its own right.

Lemma 4.3. *Every outerplanar graph G with at least two vertices can be extended to an outerplanar block G' on the same vertices such that all edges of $G'-G$ lie on the outer face of G'.*

Proof. First add edges between components until G becomes connected, marking the added edges as new. Note that this graph is still outer-planar, and all the new edges appear twice on the outer face. Next, we iteratively reduce the number of blocks until reaching an outerplanar block G'. Suppose two blocks B_1, B_2 meet at a cut-vertex v. Let uv, vw be consecutive edges along the outer face, with $u \in B_1$, $w \in B_2$. Add the edge uw, marked new, belonging to the outer face. If uv or vw was marked new and is no longer an outer edge, delete it. This reduces the number of blocks and keeps all new edges on the outer face. □

Theorem 4.4. *If G is an outerplanar graph on n vertices, then* $I(G) \leq \lfloor 3n/2 - 1 \rfloor$, *and this is best possible.*

Proof. It remains only to show that the bound holds for outerplanar graphs that are not blocks. To construct a representation for such a graph, we begin by applying Lemma 4.3 to augment G to an outerplanar block G'. We then represent G' within the desired bound using the construction given for blocks. Since all edges of $G' - G$ are outer edges in G', it then suffices to show that we can delete arbitrary outer edges from that representation without increasing the number of intervals.

Suppose we wish to delete the outer edge $v_j v_k$ from G'. In the construction for outerplanar blocks, $v_j v_k$ is represented by the intersection of two successive outer outer intervals. These can be shrunk to delete $v_j v_k$ without disturbing other intersections, unless there is an interval for some v_k properly contained in their intersection. Reviewing the construction, we see that this happens only when $v_j v_k$ belongs to a distinguished triangle with two chords and one outer edge. There is only one such triangle containing this edge, and we can delete $v_j v_k$ by shrinking the outer intervals for v_j and v_k while maintaining their intersection with the interval for v_i. □

It is also possible to prove Theorem 4.4 without Lemma 4.3 by showing that an arbitrary set of edges can be removed from the representation constructed for an outerplanar block; this requires the consideration of additional cases.

5. Additional remarks

In following the development of knowledge about $i(G)$, there are two additional remarks that can be made without difficulty for $I(G)$. These concern the interval number of the random graph and a bound on $I(G)$ involving the maximum vertex degree.

In [4], the possible labeled graphs on n vertices and the possible intersection patterns of t intervals for each of n vertices were counted to show that almost all graphs have an interval number of at least $n/(4 \lg n)$, where lg denotes $\log_2$. Scheinerman [9] obtained the upper bound of $n/(2 \lg n)$ and brought the lower bound up to match it. His upper-bound proof uses a technique of Bollobás [3]. His lower-bound proof uses the fact that almost all graphs have clique number $(2+o(1)) \lg n$ (see [3]) and the fact that using this as a bound on depth greatly reduces the number of intersection patterns. The behavior of the total interval number is analogous, as we now show.

Theorem 5.1. *Almost every graph G satisfies* $I(G) = (\frac{1}{2} + o(1)) n^2/\lg n$.

Proof. Since $I(G) \leqslant ni(G)$, the upper bound follows from Scheinerman's result; alternatively, we could give a direct proof paralleling his argument. For the lower bound, the counting argument is also similar to that for $i(G)$. We need an upper bound on the number of labeled graphs on n vertices that can be generated by representations having depth at most r and having a total of at most t intervals. Because the left-to-right greedy coloring of the intersection graph of a collection of intervals never uses more colors than the depth of the representation, we can view each interval as belonging to one of the levels $1, \ldots, r$, with the intervals on a level being disjoint.

Thus, we can generate a representation with the desired properties as follows. As we move to the right of the current point, the collection of vertices currently being represented changes by making a change in some level. If that level is currently occupied by an interval, the change occurs by encountering the right endpoint of that interval. Otherwise, the change occurs by starting a new interval, in which case it may be an interval for any of the n vertices. The latter type of change occurs t times (allowing fewer than t intervals does not yield any additional graphs). Thus, the number of these representations is $r^{2t}n^t$, which has logarithm $2t \lg r + t \lg n$. If $r = O(\lg n)$, then we can write this as $(t + o(1)) \lg n$. Since the logarithm of the number of labeled n-vertex graphs with clique number at most $(2 + o(1)) \lg n$ is $n^2/2 + o(n^2)$, setting $r = O(\lg n)$ and $t < n^2/(2 \lg n) + o(n^2/\lg n)$ implies that almost all graphs lack such representations. $\square$

Finally, concerning the maximum degree, Griggs and West [6] showed that $i(G) \leqslant \lceil (\Delta(G) + 1)/2 \rceil$, with equality for any triangle-free regular graph. Certainly, the total interval number is unbounded for a fixed maximum degree, but we could try to bound $I(G)/n$ or $I(G)/m$. We conjecture that the obvious $I(G)/n \leqslant \lceil (\Delta(G) + 1)/2 \rceil$ can be improved to $I(G)/n \leqslant (\Delta + 1/\Delta)/2$, which is achieved by disjoint copies of $K_{\Delta,\Delta}$. This graph has $I(G)/m = 1 + 1/\Delta^2$, which we conjecture is also extremal.

References

[1] T. Andreae and M. Aigner, The total interval number of a graph, J. Combin. Theory Ser. B 46 (1989) 7–21.

[2] T. Andreae, On an extremal problem involving interval number, Discrete Appl. Math. 14 (1986) 1–9.

[3] B. Bollobás, The chromatic number of random graphs, Combinatorica 8 (1988) 49–55.

[4] P. Erdős and D.B. West, A note on the interval number of a graph, Discrete Math. 55 (1985) 129–133.

[5] J.R. Griggs, Extremal values of the interval number of a graph, II, Discrete Math. 28 (1979) 37–47.

[6] J.R. Griggs and D.B. West, Extremal values of the interval number of a graph, I, SIAM J. Algebraic Discrete Methods 1 (1980) 1–7.

[7] A.V. Kostochka, in: Proc. 8th All-Union Conf. on Theoretical Cybernetics, 1988.

[8] T.M. Kratzke, The total interval number of a graph, Ph.D. Thesis, Univ. of Illinois, 1987; Coordinated Science Laboratory Research Report UILU-ENG-88-2202.

[9] E.R. Scheinerman, On the interval number of random graphs, Discrete Math. 82 (1990) 105–109.

[10] W.T. Trotter and F. Harary, On double and multiple interval graphs, J. Graph Theory 2 (1978) 137–142.