GRAPH SPANNERS §

(Revised version, July 88)

David Peleg

Alejandro A. Schäffer

Computer Science Department

Stanford University, Stanford, CA 94305

Abstract

Given a graph $G = (V, E)$, a subgraph $G' = (V, E')$ is a *t-spanner* of G if for every $u, v \in V$, the distance from u to v in G' is at most t times longer than that distance in G. This paper presents some results concerning the existence and efficient constructability of sparse spanners for various classes of graphs, including general undirected graphs, undirected chordal graphs and general directed graphs.

Keywords: Spanners, Stretch factor, Chordal graphs.

§ The work of the first author is supported in part by a Weizmann Fellowship and by contract ONR N00014-85-C-0731. During the academic year the work of the second author is supported primarily by a Fannie and John Hertz Foundation Fellowship and in part by contract ONR N00014-85-C-0731 and by grant NSF CCR 87-04170. Part of the work was done while the second author was a Research Student Associate at IBM Almaden Research Center.

1. Introduction

In this paper, we study the concept of graph spanners that arises in several recent works on theoretical analysis of computer networks. Given a connected simple graph $G = (V, E)$, a subgraph $G' = (V, E')$ is a *t-spanner* of G if for every $u, v \in V$,

$$\frac{dist(u, v, G')}{dist(u, v, G)} \leq t,$$

where $dist(u, v, H)$ denotes the distance from u to v in H. We refer to t as the *stretch factor* of G'.

Spanners appear to be the underlying graph structure in various constructions in distributed systems and communication networks. Close relationships were established between the quality of spanners for a given undirected graph (in terms of the stretch factor, t, and the number of spanner edges, $|E'|$) and the time and communication complexities of any synchronizer for the network based on this graph [PU1]. Efficient routing schemes use the edges of a sparse spanner for message routing in order to maintain succinct routing tables [PU2].

In the above-mentioned applications, one is mainly interested in finding sparse spanners with small stretch factor. In particular, a family $\mathcal{F}$ of graphs is said to have an *optimal spanner* if for every n-vertex graph $G \in \mathcal{F}$ there is an $O(1)$-spanner with $O(n)$ edges. This paper presents some upper and lower bounds concerning the existence and efficient constructability of sparse or optimal spanners for various families of graphs.

The notion of spanners appears in other forms in the literature. Bhatt, Chung, Leighton, and Rosenberg [BCLR] consider efficient embeddings of graphs in graphs, with the ultimate goal of selecting a *universal network* as the underlying topology of a universal parallel machine. One of the parameters studied is the *dilation* of the embedding, which is similar to our stretch factor. The dilation of embedding a network G in a network H has immediate implication on the time delay incurred by simulating G on a universal machine based on H. There, however, the subgraph of H onto which G is embedded need not be a subgraph of G. Moreover, the major size consideration in [BCLR] is the size of the universal network H, and not the number of edges of H needed to embed G.

In this paper, we assume that the distance along each edge is 1. One may also consider graphs with a non-trivial distance function on the edges. Chew [C] and Dobkin, Friedman, and Supowit [DFS] study the existence of sparse spanners for the special case where G is a complete graph whose vertices are points in the plane and $dist(u, v)$ is the distance from the point u to the point

v in some L_p metric. The Delaunay triangulation (see [PS] for a definition) in the same metric chosen for the distances is a natural candidate to be a spanner, since it has $O(|V|)$ edges and tends to contain edges between points that are close together. In order for the triangulation to be well-defined it is traditional to assume that no four points are cocircular. Chew [C] shows that if distances are measured in the L_1 metric, then the L_1 Delaunay triangulation is a $\sqrt{10}$-spanner. Dobkin, Friedman and Supowit [DFS] show that if the distances are measured in the Euclidean (L_2) metric, then the corresponding Delaunay triangulation is a $\phi\pi$-spanner, where ϕ is the golden ratio.

The rest of the paper is organized as follows. Section 2 concerns general undirected graphs, and gives upper and lower bounds on sparse spanners for graphs in this class. Section 3 handles (undirected) chordal graphs. Section 4 discusses the directed case.

2. Sparse spanners for general undirected graphs

The following lemma gives an equivalent, and sometimes more convenient to use, formulation of the definition of t-spanners.

Lemma 2.1: *The subgraph $G' = (V, E')$ is a t-spanner of the graph $G = (V, E)$ iff for every $(u, v) \in E$, $dist(u, v, G') \leq t$.* ▮

A natural question concerns the existence and efficient constructability of sparse spanners for general undirected graphs. As one might expect, we have

Theorem 2.2: *The problem of determining, for a given graph $G = (V, E)$ and two integers $t, m \geq 1$, whether G has a t-spanner with m or fewer edges, is NP-complete.*

Proof: Membership in NP is immediate. The problem is complete for NP even for fixed $t = 2$. This can be proved by a reduction from the Edge Dominating Set problem for bipartite graphs (see problem (GT2) in [GJ]). Consider an instance of this problem, consisting of a bipartite graph $H = (X, Y, A)$, $X = \{x_i \mid 1 \leq i \leq n_x\}$, $Y = \{y_i \mid 1 \leq i \leq n_y\}$, and an integer $K \geq 1$. An *edge dominating set (EDS)* for H is a subset $A' \subseteq A$ such that for every edge in $A \setminus A'$ there is an adjacent edge in A'. The problem is to decide whether H has an EDS of size at most K. We reduce this instance of the EDS problem to the following instance of our 2-spanner problem.

Construct a graph $G(H) = (V, E)$ as follows. For every pair of vertices $x_i, x_j \in X$ create a new vertex x_{ij}, and similarly for the vertices in Y. Let

$$V = X \cup Y \cup \{x_{ij} \mid 1 \leq i < j \leq n_x\} \cup \{y_{ij} \mid 1 \leq i < j \leq n_y\},$$

$$\begin{aligned} E_{XY} &= \{(x_i, x_j), (x_i, x_{ij}), (x_j, x_{ij}) \mid 1 \leq i < j \leq n_x\} \\ &\quad \cup \{(y_i, y_j), (y_i, y_{ij}), (y_j, y_{ij}) \mid 1 \leq i < j \leq n_y\} \end{aligned}$$

and let

$$E = A \cup E_{XY}.$$

Fix

$$m = K + n_x(n_x - 1) + n_y(n_y - 1).$$

Define the *normal* subset of E_{XY} as

$$N = \{(x_i, x_j), (x_i, x_{ij}) \mid 1 \leq i < j \leq n_x\} \cup \{(y_i, y_j), (y_i, y_{ij}) \mid 1 \leq i < j \leq n_y\},$$

$$|N| = n_x(n_x - 1) + n_y(n_y - 1).$$

See Figures 2.1a and 2.1b.

By Lemma 2.1, a *solution* to the 2-spanner problem on $G(H)$ is any subset of edges $E' \subseteq E$, $|E'| \leq m$ satisfying the property that for every edge $(u, v) \in E \backslash E'$ there is a vertex $w \in V$ such that $(u, w), (w, v) \in E'$. A *normal solution* for $G(H)$ is a solution $E' = N \cup A'$ where $A' \subseteq A$. Observe that if $E'' = B \cup A'$, $B \subseteq E_{XY}$, $A' \subseteq A$ is a solution for $G(H)$ then $E' = N \cup A'$ is a normal solution for $G(H)$ and $|N| \leq |B|$. The reduction is completed upon noting that $E' = N \cup A'$, $A' \subseteq A$ is a normal solution for $G(H)$ iff A' is an EDS for H. See Figure 2.1c. ▮

Despite Theorem 2.2, it is possible to study the existence question in terms of its asymptotic behavior.

A *k-partition* of a graph $G = (V, E)$ is a collection of clusters

$$P = \{(c_i, C_i) \mid C_i \subseteq V,\ c_i \in C_i,\ 1 \leq i \leq \ell\}$$

such that the clusters C_i form a complete disjoint partition of V (i.e., $V = \bigcup_{i=1}^{\ell} C_i$ and $C_i \cap C_j = \emptyset$ for every $1 \leq i < j \leq \ell$), the subgraph induced by each cluster C_i is connected, and for every $u \in C_i$, $dist(u, c_i) \leq k$. The *density* of P, denoted $\rho(P)$, is the number of pairs $C_i, C_j \in P$ such that C_i and C_j are connected by an edge.

Lemma 2.3 [A]: *For every n-vertex graph G and for every $1 < k < n$ there exists a (polynomial-time-constructable) $(\log_k n)$-partition P with density $\rho(P) \leq (k-1)n$.* ∎

Theorem 2.4: *For every n-vertex graph G and for every $1 < k < n$, there exists a (polynomial-time-constructable) $(4\log_k n + 1)$-spanner with at most kn edges.*

Proof: The claim is derived in a straightforward way from a construction of Awerbuch [A] (used there to construct a "γ synchronizer"). Construct a partition P as in Lemma 2.3. For every cluster $(C_i, c_i) \in P$ construct a single-source, shortest-paths spanning tree rooted at c_i. The set of edges E' of the spanner contains all the edges of these trees (there are less than n such edges). In addition, for every pair of neighboring clusters C_i, C_j select a single "intracluster" edge e_{ij} connecting the clusters and include it in E'. By Lemma 2.3 the total number of edges in E' does not exceed kn. Finally we have to verify that the resulting subgraph indeed has stretch factor $4\log_k n + 1$. Consider an edge (u, w) in G and suppose that u belongs to the cluster C_i and w belongs to C_j. The radius of each cluster in the partition is at most $r = \log_k n$. If $i = j$ then the path from u to w through c_i is of length at most $2r$. Otherwise, there is a path going from u to c_i (in at most r steps), from c_i through e_{ij} to c_j (in at most $r + 1 + r$ steps), and from c_j to w (in at most r steps), so the total length is at most $4r + 1$. ∎

Fixing $k = 2$ we get:

Corollary 2.5: *For every n-vertex graph G there exists a (polynomial-time-constructable) $O(\log n)$-spanner with $O(n)$ edges.* ∎

Similarly, by taking $k = n^{1/r}$ for fixed $r \geq 1$ we get

Corollary 2.6: *For every n-vertex graph G and for every fixed $r \geq 1$, there exists a (polynomial-time-constructable) $(4r + 1)$-spanner with $O(n^{1+1/r})$ edges.* ∎

A complementing lower bound can also be derived from the arguments of [A], demonstrating that for some graphs, the best possible improvements are only a constant factor from the construction of Corollary 2.6.

Lemma 2.7 [Bo]: *For every $r \geq 1$, there exist (infinitely many) n-vertex graphs $G = (V, E)$ with girth $g \geq r$ and $|E| \geq \frac{1}{4} n^{1+1/r}$.* ∎

Lemma 2.8: *For every $t \geq 1$ and $r \geq t + 2$ and for every graph $G = (V, E)$ with girth $g \geq r$, the only t-spanner of G is G itself.*

Proof: The existence of another spanner G' with some edge $(u,v) \in E$ omitted would imply the existence of an alternative path of length t connecting u to v, which together with the edge (u,v) gives a cycle of length $t+1 < r$. ∎

Theorem 2.9: *For every $r \geq 3$, there exist (infinitely many) n-vertex graphs $G=(V,E)$ for which every $(r-2)$-spanner requires $\Omega(n^{1+1/r})$ edges.* ∎

However, for various restricted graph families one can do appreciably better than is implied by this lower bound. Clearly every class of graphs with $O(|V|)$ edges has an optimal spanner, as in such a class each graph G can be taken as its own 1-spanner. This covers various graph topologies commonly used for communication networks, including bounded-degree and planar graphs (e.g., rings, meshes, trees, butterflies, and cube-connected cycles). Likewise, for bounded-diameter graphs, any rooted shortest-paths tree will do as an optimal spanner. Thus the problem remains interesting only for graph classes in between. The family of hypercubes is considered in [PU1]. Since the d-dimensional cube has girth 4, by Lemma 2.8 it does not have a sparse 2-spanner. Nevertheless, optimal spanners do exist for the hypercubes.

Lemma 2.10 [PU1]: *For every $d \geq 0$ the d-dimensional cube has a 3-spanner with fewer than $7 \cdot 2^d$ edges.* ∎

3. Spanners for chordal graphs.

A graph is *chordal* if it contains no induced cycles of length greater than 3. A graph is *split* if its vertices can be partitioned into two sets X and Y such that X induces a clique and Y is an independent set. Every split graph is chordal [Go]. Every connected chordal graph is either a tree or has girth 3; furthermore, any connected induced subgraph of a chordal graph is also chordal, so a lower bound similar to Theorem 2.9 does not apply.

We begin this section by showing that for every n-vertex chordal graph there exists a 2-spanner with $O(n^{1.5})$ edges, and that there are infinitely many n-vertex chordal graphs (in fact, split graphs) for which every 2-spanner requires $\Omega(n^{1.5})$ edges. However, the situation changes drastically when allowing a slight increase in the stretch factor. We prove that for every n-vertex chordal graph there exist a 3-spanner with $O(n \log n)$ edges and a 5-spanner with $O(n)$ edges.

Let $\mathcal{K}(G)$ be the collection of maximal cliques in a given graph $G=(V,E)$. For every maximal clique $K \in \mathcal{K}(G)$ select a vertex $v_K \in K$. Set $E_K = \{(v_K, v) \mid v \in K, v \neq v_K\}$ and $E' =$

$\bigcup_{K\in\mathcal{K}(G)} E_K$. The resulting subgraph $G' = (V, E')$ is called a *star spanner* of G. Note that every selection of vertices v_K in the cliques uniquely determines a star spanner for G.

Lemma 3.1: *Every star spanner $G' = (V, E')$ for G is a 2-spanner.*

Proof: Consider any edge $(u, v) \in E$. This edge belongs to some maximal clique K of G. If u or v is v_K then $(u, v) \in E'$. Otherwise, $(v_K, u), (v_K, v) \in E'$. In both cases $dist(u, v, G') \le 2$, and the claim follows by Lemma 2.1. ▮

Theorem 3.2: *For every n-vertex chordal graph $G = (V, E)$ there is a 2-spanner $G' = (V, E')$ with $O(n^{1.5})$ edges.*

Proof: Consider an n-vertex chordal graph $G = (V, E)$. G can have at most n maximal cliques [Go]. We now describe how to construct a star spanner $G' = (V, E')$ for G with $O(n^{1.5})$ edges. Start with the set $\mathcal{KS}$ containing all maximal cliques of G ($|\mathcal{KS}| \le n$) and $E' = \emptyset$. For every $v \in V$ let $f(v, \mathcal{KS})$ denote the number of cliques $K \in \mathcal{KS}$ such that $v \in K$. Select the vertices v_K one by one in two phases. The first phase consists of iteratively executing the following loop.

While there is a vertex v such that $f(v, \mathcal{KS}) \ge \sqrt{n}$ **do**:

> For every $K \in \mathcal{KS}$ s.t. $v \in K$, select v as v_K, add E_K to E' (removing duplicate edges) and remove K from $\mathcal{KS}$.

Each iteration of the first phase reduces $|\mathcal{KS}|$ by at least $\sqrt{n}$, so the number of iterations is at most $|\mathcal{KS}|/\sqrt{n} \le \sqrt{n}$. Also, in each such iteration E' is augmented by exactly the edges adjacent to the vertex v, so by the end of this phase E' contains at most $n^{1.5}$ edges.

When the first phase ends, the remaining collection $\mathcal{KS}$ satisfies $f(v, \mathcal{KS}) < \sqrt{n}$ for every $v \in V$. In the second phase, the selection of the vertices v_K (and the corresponding edge sets E_K) for the remaining cliques $K \in \mathcal{KS}$ is completed arbitrarily. The number of edges added in this phase is bounded by $\sum_{K\in\mathcal{KS}}(|K| - 1) < \sum_{v\in V} f(v, \mathcal{KS}) < n^{1.5}$. Thus at the end of the process $|E'| \le O(n^{1.5})$. ▮

This result is complemented by a matching lower bound.

Lemma 3.3: *There exist (infinitely many) n-vertex chordal graphs $G = (V, E)$ for which every 2-spanner requires $\Omega(n^{1.5})$ edges.*

Proof: Let q be any prime and let $m = q^2 + q + 1$ and $n = 2m$. Construct a set of $2m$ vertices by letting $X = \{x_i \mid 1 \le i \le m\}$, $Y = \{y_i \mid 1 \le i \le m\}$ and $V = X \cup Y$. Let $L = \{L_i \mid 1 \le i \le m\}$ be the collection of lines of some finite projective plane of order q over the set of points X. Construct a split graph $G = (V, E)$ as follows. For every $1 \le i \le m$ connect $y_i \in Y$ to all the vertices $x \in L_i$, and make X into a clique. See Figures 3.1a and 3.1b. Let E_c denote the set of "crossing" edges in G (i.e., edges (x, y) such that $x \in X$ and $y \in Y$). Since $|L_i| = q + 1$ for every $1 \le i \le m$,

$$|E_c| = m(q+1) = q^3 + 2q^2 + 2q + 1.$$

Consider a 2-spanner $G' = (V, E')$ for G. For every crossing edge $(x, y) \in E_c$, G' has to include either (x, y) or $(x, z), (z, y)$ for some $z \in X$. We now claim that G' must contain a distinct edge for each crossing edge in E_c. This claim completes the proof of the lemma, since it implies that overall, $|E'| \ge |E_c| = q^3 + 2q^2 + 2q + 1 \in \Omega(n^{1.5})$.

The claim is proved by analyzing the contents of E' and "charging" a distinct edge in E' to every crossing edge in E_c. For every $y_i \in Y$ let $Z_i \subseteq L_i$ be the set of vertices to which y_i is connected in G'. Consider some crossing edge $e = (y_i, x) \in E_c$, $x \in L_i$. If $x \in Z_i$ we charge e to itself, since $e \in E'$. In case $x \in L_i \setminus Z_i$, by the definition of a 2-spanner, E' must contain an edge (x, z) for some $z \in Z_i$. This edge is "charged" to e. It remains to verify that (x, z) is not charged to any other edge in E_c. Clearly it cannot be charged to any other edge (y_i, x') for $x' \in L_i$. Now suppose it is charged to some edge $(y_j, x') \in E_c$ where $j \ne i$. Then x' has to be either x or z, and both (y_j, x) and (y_j, z) have to appear in E_c. Hence $x, z \in L_j$, which contradicts the fact that L_i and L_j are two different lines in a projective plane.

The construction can be extended to general n by using Bertrand's postulate (cf. [NZ; Sect. 8.3]) to select an appropriate prime $\sqrt{n}/4 < q < \sqrt{n}/2$. ∎

The 3-spanner construction for chordal graphs is based on the following separator theorem of Gilbert, Rose, and Edenbrandt.

Theorem 3.4 [GRE]: *Every n-vertex chordal graph G contains a maximal clique C such that if the vertices in C are deleted from G, every connected component in the graph induced by any remaining vertices is of size at most $n/2$.* ∎

An $O(|E|)$-time algorithm for finding a separating clique C satisfying the conditions of the theorem is also given in [GRE].

We now describe an algorithm to construct a 3-spanner for a chordal graph $G = (V, E)$. Begin by partitioning G into separating cliques recursively, using the following algorithm.

Define $G_{11} := G$.

Let $\mathcal{G}_1 := \{G_{11}\}$.

Set $i := 1$.

While $\mathcal{G}_i \neq \emptyset$ **do**

Set $\mathcal{G}_{i+1} := \emptyset$.

Suppose $\mathcal{G}_i = \{G_{i1}, \ldots, G_{ir}\}$. For each $G_{ij} \in \mathcal{G}_i$ let C_{ij} be a maximal separating clique for G_{ij} as prescribed in Theorem 3.4.

Add each (nonempty) connected component of $G_{ij} \setminus C_{ij}$ to $\mathcal{G}_{i+1}$.

Let $\mathcal{C}_i = \{C_{ij} \mid 1 \leq j \leq r\}$.

Set $i := i + 1$.

Let P be the set of ordered pairs $\langle v, C_{ht} \rangle$ such that v is a vertex in some clique C_{is} with $i > h$ and v has a neighbor in the clique C_{ht}.

The spanner G' contains two disjoint sets of edges E_1 and E_2. The set of edges E_1 includes a star-spanner based on the maximal separating cliques we chose. The set E_2 includes one edge corresponding to each element of P. If $\langle v, C_{ht} \rangle$ is an element of P we choose one neighbor of v in C_{ht}, call it w_{ht}, and put the edge (v, w_{ht}) in E_2.

Lemma 3.5: *The graph $G' = (V, E_1 \cup E_2)$ has $O(n \log n)$ edges.*

Proof: Since each vertex occurs in exactly one of the separating cliques, $|E_1| < n$.

Let ℓ be the number of iterations of the loop in the partition algorithm. By Theorem 3.4, any graph $G_{ij} \in \mathcal{G}_i$ can have at most $n/(2^{i-1})$ vertices, so $\ell \leq \log_2 n + 1$.

Let G_{is} be any member of $\mathcal{G}_i$. For any h such that $1 \leq h < i$ there is exactly one graph in $\mathcal{G}_h$ that contains the graph G_{is} as an induced subgraph, and none of the other graphs in $\mathcal{G}_h$ are even adjacent to G_{is}. Thus there are at most $n \log_2 n$ ordered pairs $\langle v, C_{ht} \rangle$ in P. Since each member of E_2 corresponds to a member of P, $|E_2| \leq n \log_2 n$. ∎

Lemma 3.6: *The graph $G' = (V, E_1 \cup E_2)$ is a 3-spanner for G.*

Proof: Consider an arbitrary edge $(u, v) \in E$. If both endpoints u, v are in the same separating clique then the edge can be replaced by a path consisting of at most 2 edges of E_1 as in the proof of Lemma 3.1.

Any two graphs in the same $\mathcal{G}_i$ cannot have edges between them. Therefore if the endpoints of (u, v) are not in the same separating clique, it must be the case that u belongs to some separating clique C_{ht} and v belongs to another separating clique C_{is} such that $i \neq h$. We can assume without loss of generality that $i > h$. The ordered pair $\langle v, C_{ht} \rangle$ must be in P. The spanner G' contains an edge $(v, w_{ht}) \in E_2$ such that w_{ht} belongs to the clique C_{ht}. Even if $w_{ht} \neq u$, G' contains a path of at most 2 edges in E_1 from w_{ht} to u, which completes the proof. ▮

Theorem 3.7: *For every n-vertex chordal graph $G = (V, E)$ there is a 3-spanner $G' = (V, E')$ with $O(n \log n)$ edges.* ▮

The 5-spanner construction for chordal graphs is based on the following characterization of chordal graphs as intersection graphs due to Buneman and Gavril.

Theorem 3.8 [Bu, Ga]: *$G = (V, E)$ is a chordal graph if and only if there exists an unrooted, undirected tree T over the vertex set $\mathcal{K}(G)$ such that for any vertex $v \in V$ the set of maximal cliques containing v induces a subtree of T (and not just a subforest). The subtrees corresponding to distinct vertices $v, w \in V$ intersect if and only if v and w are adjacent.* ▮

Example: Figure 3.2a depicts a chordal graph $G = (V, E)$ over $V = \{1, \ldots, 8\}$. The set of maximal cliques of G is $\mathcal{K}(G) = \{K_i \mid 1 \leq i \leq 5\}$ where $K_1 = \{1, 2, 3\}$, $K_2 = \{2, 3, 4\}$, $K_3 = \{2, 4, 5\}$, $K_4 = \{1, 6\}$ and $K_5 = \{6, 7, 8\}$. Figure 3.2b depicts a possible tree T as in Theorem 3.8. We root and direct the tree for later use.

Assume that G is connected; if not, we consider each component separately. As noted above, an n-vertex chordal graph has at most n maximal cliques, so a tree satisfying the conditions of Theorem 3.8 has at most n nodes.

The following lemma is implicit in [Bu74, Ga74]:

Lemma 3.9: *Any two maximal cliques that are adjacent in the tree share a vertex.*

Proof: Let K_1 and K_2 be adjacent maximal cliques containing the vertices v_1 and v_2 respectively. Break the tree T into two trees T_1 and T_2 by cutting the edge (K_1, K_2), so that $K_1 \in T_1$ and $K_2 \in T_2$. Since we assumed that G is connected we can find a simple path in G from v_1 to v_2. This path must contain an edge (v_3, v_4) such that v_3 belongs to a maximal clique in T_1, while v_4 belongs to a maximal clique in T_2. The edge (v_3, v_4) must itself belong to some maximal clique

K_3. If K_3 belongs to T_1, then by the subtree property for vertices v_4 belongs to every maximal clique on the unique path from K_3 to K_2 including K_1. Similarly, if K_3 belongs to T_2, then v_3 belongs to every maximal clique on the unique path from K_3 to K_1, including the maximal clique K_2. ∎

Note that the other direction does not necessarily hold; two maximal cliques that share a vertex need not be adjacent.

We start the 5-spanner construction for G by building a tree T on the vertex set $\mathcal{K}(G)$ as prescribed in Theorem 3.8. Next we root the tree at some arbitrary maximal clique R, and orient all edges of the tree away from R. (In Figure 3.2b we choose the maximal clique K_1 as the root.) For each maximal clique $K \in \mathcal{K}(G)$, let $d(K)$ be the *distance* from R to K in the tree or alternatively the *depth* of K. We say that maximal clique K_1 is *shallower* than maximal clique K_2, and K_2 is *deeper* than K_1, if $d(K_1) < d(K_2)$.

In each maximal clique K, we choose one vertex v_K to be its *leader*. For each maximal clique K, we also identify one ancestor of K, which we call its *supplier* and denote $s(K)$. The leaders and suppliers are determined, as explained below, using an auxiliary arc-labeled, directed multigraph $H = (\mathcal{K}(G), A)$. The node set of H is the same as that of the tree T. The multigraph H contains an arc from K_1 to K_2 with label v if and only if

i. K_1 is a proper ancestor of K_2, and

ii. the maximal cliques K_1 and K_2 both contain the vertex v. (Note that *proper* ancestor means that K_1 is an ancestor of K_2, but $K_1 \neq K_2$.)

Figure 3.2c depicts the multigraph H corresponding to the graph G and the tree T of Figures 3.2a and 3.2b.

Assume that the vertices of G are numbered in some arbitrary fashion. We define v_R, the leader of the root R, to be the lowest numbered vertex in R. The leader of a non-root maximal clique K is determined as follows. Let $\mathcal{P}(K)$ be the set of all shortest paths from R to K in the graph H. Let $A_f(K)$ be the set of labeled arcs that occur as final arcs in a shortest path belonging to the set $\mathcal{P}(K)$. Let $A_t(K)$ be the set of arcs in $A_f(K)$ whose tail has minimum depth. Let $a(K)$ be the member of $A_t(K)$ with the lowest vertex label, v. This v is chosen to be the leader of K.

For each non-root maximal clique K, let the length of the paths in $\mathcal{P}(K)$ be $p(K)$. Define $p(R) := 0$. Set the supplier of R to be $s(R) := R$. The supplier of a non-root maximal clique K is the deepest ancestor K' of K that satisfies $p(K') < p(K)$.

The following lemmas state some simple properties of the multigraph H, of leader vertices, and of supplier cliques.

Lemma 3.10: *If H contains an arc $K_1 \to K_2$ with label w, it also contains arcs $K_1 \to K$ and $K \to K_2$ with label w for any node K that is a proper descendant of K_1 and a proper ancestor of K_2.*

Proof: By definition of H, the maximal cliques K_1 and K_2 must both contain the vertex w. Because of the subtree property in Theorem 3.8, any maximal clique K on the path from K_1 to K_2 in the tree T must also contain w. Hence, H must contain the arcs $K_1 \to K$ and $K \to K_2$ with label w. ∎

Lemma 3.11: *If maximal clique K_1 is an ancestor of maximal clique K_2, then $p(K_1) \leq p(K_2)$.*

Proof: Let P_2 be a shortest path from R to K_2 in the multigraph H. Let K_0 be the last node on P_2 such that K_0 is an ancestor of K_1. If $K_0 = K_1$, then $p(K_1) < p(K_2)$. Otherwise, it follows from Lemma 3.10 that we can replace the segment of P_2 following K_0 with an arc $K_0 \to K_1$. This yields a path from R to K_1 that is no longer than P_2. ∎

Lemma 3.12: *If $K_1 = s(K_2)$, then for every maximal clique K, that is a proper descendant of K_1 and an ancestor of K_2, $K_1 = s(K)$.*

Proof: By Lemma 3.11, $p(K) \leq p(K_2)$. Since K is an ancestor of K_2, but a proper descendant of $s(K_2)$, and K is not chosen as $s(K_2)$, it must be the case that $p(K) = p(K_2)$. Since K_1 is the supplier of K_2, $p(K_1) < p(K_2)$. Therefore K_1 is the deepest ancestor of K such that $p(K_1) < p(K)$. Hence $K_1 = s(K)$. ∎

We will need the following corollary of Lemma 3.12.

Corollary 3.13: *Let $K_1, \ldots, K_r$ be any path in the tree T, such that K_1 is a supplier for some $K_i, i > 1$. Then K_1 is a supplier for K_2, and the maximal cliques in $\{K_2, K_3, \ldots, K_r\}$ whose supplier is K_1 must be consecutive.* ∎

For any vertex w in G, let $C(w)$ be the shallowest maximal clique in T that contains w. This maximal clique is well-defined because if two maximal cliques contain w, their least common ancestor must also contain w, by the subtree property of Theorem 3.8.

Lemma 3.14: *For any maximal clique K, its leader v_K belongs to its supplier $s(K)$. If $s(K) \neq R$, then $v_{s(K)} \neq v_K$.*

Proof: Let P be the shortest path in H from R to K whose final arc $a(K)$ determines v_K; that is, the label of $a(K)$ is v_K. Let J be the tail of $a(K)$. By construction of H, the maximal clique J must contain the vertex v_K. The path P without its final arc is a path of length $p(K) - 1$ from R to J; in particular $p(J) < p(K)$. Therefore by definition of $s(K)$, J must be an ancestor of $s(K)$. Since v_K belongs to both J and K, it must also belong to $s(K)$ by the subtree property in Theorem 3.8.

Suppose, seeking to establish a contradiction, that $s(K) \neq R$ and $v_{s(K)} = v_K$. Let P' be the shortest path in H from R to $s(K)$ whose final arc determines $v_{s(K)}$; that is, the label of $a(s(K))$ is $v_{s(K)}$; $v_{s(K)}$ is the same as v_K by assumption. We can replace the final arc of P' with an arc having the same tail and label, but having head K instead of $s(K)$. This replacement would yield a path in H from R to K having the same length as P' which has length $p(s(K))$ by construction. But the defintion of suppliers implies that $p(s(K)) < p(K)$, provided $K \neq R$, a contradiction. ▮

Lemma 3.15: *If maximal clique K_1 is an ancestor of maximal clique K_2, and $s(K_1) \neq s(K_2)$, then $v_{K_2} \notin s(K_1)$.*

Proof: Suppose, seeking to establish a contradiction, that v_{K_2} does occur in $s(K_1)$. Let P be a shortest path in H from R to $s(K_1)$; by definition, P is of length $p(s(K_1))$. Since we assumed $v_{K_2} \in s(K_1)$, the graph H contains an arc $s(K_1) \rightarrow K_2$ with label v_{K_2}. This implies that

$$p(K_2) \leq p(s(K_1)) + 1.$$

By Lemma 3.11, we have that

$$p(K_1) \leq p(K_2).$$

By definition of suppliers,

$$p(s(K_1)) < p(K_1).$$

¿From the last three inequalities and the fact that p takes only integer values, it follows that

$$p(K_1) = p(K_2).$$

It now follows from Lemma 3.11 and the definition of suppliers that $s(K_1)$ must also be the supplier of K_2, a contradiction. ∎

The 5-spanner contains of two sets of edges, E_3 and E_4, that play different roles. However, the sets E_3 and E_4 are not disjoint, so some edges play both roles.

Recall that for any vertex w in G, $C(w)$ is the shallowest maximal clique in T that contains w. For every vertex $w \neq v_R$, E_3 contains an edge $(w, v_{C(w)})$.

For every maximal clique K, such that $v_K \neq v_R$, we include the edge $(v_{s(K)}, v_K)$ in E_4. It follows from Lemma 3.14 that this edge exists in the graph G.

Lemma 3.16: *The graph $G' = (V, E_3 \cup E_4)$, which is a subgraph of G, has at most $2n - 2$ edges.*

Proof: For every $v \neq v_R$, there is at most one edge of the form (v, v_K) in E_3, so $|E_3| \leq n - 1$. For each maximal clique $K \neq R$, the set E_4 contains one edge $(v_{s(K)}, v_K)$. All the edges of E_4 connect the leader of a maximal clique to the leader of the supplier of that maximal clique, so $|E_4| \leq n - 1$. ∎

Lemma 3.17: *The graph $G' = (V, E_3 \cup E_4)$ is a 5-spanner for G.*

Proof: Let $(u, w) \in E$ be given. Recall that $C(u)$ (respectively, $C(w)$) is the *unique* maximal clique of minimum depth (in T) that contains the vertex u (resp., w). Without loss of generality, assume that $d(C(u)) \leq d(C(w))$. Since (u, w) is an edge of G, the vertices u and w must occur together in some maximal clique. Therefore the maximal clique $C(w)$ must be a (not necessarily proper) descendant of the maximal clique $C(u)$, and u must occur in $C(w)$.

We bidirectionally grow the desired path of length at most 5 from both u and w until we reach a common vertex. Observe that either $u = v_{C(u)}$ or the edge $(u, v_{C(u)}) \in E_3$. Similarly, either $w = v_{C(w)}$ or the edge $(w, v_{C(w)}) \in E_3$. If $u = v_{C(w)}$ then we are done, so for the remainder of the proof, suppose that $u \neq v_{C(w)}$. It is sufficient to exhibit a path of length at most 3 from $v_{C(u)}$ to $v_{C(w)}$.

If $v_{C(u)} = v_{C(w)}$, we can immediately complete the desired path, so for the remainder of the proof suppose that $v_{C(u)} \neq v_{C(w)}$. By Lemma 3.14, either $v_{C(u)} = v_{s(C(u))}$, or the edge $(v_{s(C(u))}, v_{C(u)}) \in E_4$. Similarly, either $v_{C(w)} = v_{s(C(w))}$, or the edge $(v_{s(C(w))}, v_{C(w)}) \in E_4$. Therefore it is sufficient to show that $v_{s(C(u))}$ and $v_{s(C(w))}$ are either identical or adjacent in the spanner.

If $s(C(u)) = s(C(w))$, we can immediately complete the desired path. So, for the remainder of the proof suppose that $s(C(u)) \neq s(C(w))$. By Lemma 3.15, $v_{C(w)}$ cannot occur in the supplier clique $s(C(u))$.

Let the path in T from $s(C(u)) = J_1$ to $C(w) = J_r$ be

$$J_1 \to J_2 \to \cdots \to J_{r-1} \to J_r.$$

Let J_q be the first maximal clique in this path such that $s(J_q) \neq J_1$; J_q is well-defined because $s(J_r) \neq J_1$. Let $J_p = C(u)$. By Corollary 3.13, $s(J_q) = J_{q-1}$ and $p < q$. Since u occurs in J_p and u occurs in $J_r = C(w)$, it follows from the subtree property that u must occur in the supplier clique J_{q-1}. Thus H contains a path of length $p(J_{q-1}) + 1$ from R to J_r ending with an arc $J_{q-1} \to J_r$ labeled u, which implies $p(J_r) \leq p(J_{q-1}) + 1$. Since $s(J_q) = J_{q-1}$, we have $p(J_q) \geq p(J_{q-1}) + 1$. By Lemma 3.11,

$$p(J_r) = p(J_{r-1}) = \cdots = p(J_{q+1}) = p(J_q) = p(J_{q-1}) + 1.$$

Hence J_{q-1} is the deepest ancestor of J_r with $p(J_{q-1}) < p(J_r)$, and J_{q-1} must be the supplier of $J_r = C(w)$.

Figure 3.3 summarizes the relative order (along the path from J_1 to J_r) of some of the maximal cliques we consider in this proof, and also shows some of the identities that we assumed or proved.

By choice of q, $s(J_{q-1}) = J_1 = s(C(u))$. Therefore it must be the case that $v_{s(C(u))} = v_{J_{q-1}}$, or that we included the edge $(v_{s(C(u))}, v_{J_{q-1}})$ from the maximal clique $s(C(u))$ in E_4. Thus $v_{s(C(u))}$ and $v_{s(C(w))}$, which is the same as $v_{J_{q-1}}$, are either identical or connected by an edge of E_4.

To summarize, the graph G contains a (not necessarily simple) path from w to $v_{C(w)}$, to $v_{s(C(w))}$, to $v_{s(C(u))}$, to $v_{C(u)}$, to u such that each consecutive pair is either identical or connected by an edge in $E_3 \cup E_4$, which completes our proof. ∎

Combining Lemmas 3.16 and 3.17 yields:

Theorem 3.18: *Every n-vertex chordal graph $G = (V, E)$ has a 5-spanner $G' = (V, E')$ with $O(n)$ edges that can be constructed in polynomial time.* ∎

4. The directed case

The situation for directed graphs is generally harder than for undirected ones, in the sense that sparse spanners for general directed graphs do not always exist. The NP-completeness result is naturally extended to this case. It is also clear that there are n-vertex, $\Omega(n^2)$-edge directed graphs that have at most one directed path between any two nodes (e.g., the complete bipartite graph (X, Y, E) with $|X| = |Y| = n/2$ and with all edges directed from X to Y), and for such graphs, no subgraph is a spanner. Moreover, the following lower bound proof implies that for any fixed $t \geq 1$ there are infinitely many dense graphs with high (strong) edge-connectivity for which one cannot do any better than taking the entire graph as its own t-spanner.

Lemma 4.1: *For every $t \geq 1$ and $n \geq t$ there exists an n-vertex directed graph G with $\Omega(n^2/t^2)$ edges and $O(n/t^2)$ directed edge-connectivity for which the only t-spanner is G itself.*

Proof: Let t and n be given, and construct the following directed graph $G = (V, E)$. Let $V = \{0, \ldots, n-1\}$. Think of the vertices of V as arranged in a ring. Fix $p = \lfloor \frac{n-t}{t^2} \rfloor$. For every $0 \leq \ell \leq p$ let

$$E_\ell = \{i \to (i + \ell t + 1) \bmod n) \mid 0 \leq i \leq n-1\},$$

and take $E = \bigcup_{0 \leq \ell \leq p} E_\ell$. See Figure 4.1. We claim that any t-spanner of G must contain the entire edge-set E, which is of size $\Omega(n^2/t^2)$. To see this, consider omitting any edge $i \to i' \in E_\ell$ $(0 \leq \ell \leq p)$ from E. The omitted edge satisfies $i' \equiv (i + \ell t + 1) \pmod n$. In order for the resulting graph to be a t-spanner there has to be some sequence of edges $e_1, \ldots, e_m$ $(2 \leq m \leq t)$ leading from i to i'. Assuming $e_j \in E_{\ell_j}$ for every $1 \leq j \leq m$, the edge sequence leads from i to

$$i'' = \big(i + \sum_{1 \leq j \leq m} (\ell_j t + 1)\big) \bmod n = \big(i + m + t \sum_{1 \leq j \leq m} \ell_j\big) \bmod n.$$

Note that even if the sequence contains t edges from E_p, the maximum distance it may traverse is $t + t^2 p \leq n$, hence it is not possible to complete a whole lap around the ring. Therefore the edge sequence has to satisfy

$$m + t \sum_{1 \leq j \leq m} \ell_j = 1 + t\ell.$$

This leads to an immediate contradiction, since no matter how we choose the sequence, $2 \leq m \leq t$. ∎

Theorem 4.2: *For every $t \geq 1$ there are infinitely many n-vertex directed graphs for which every t-spanner requires $\Omega(n^2/t^2)$ edges.* ∎

Acknowledgements

We wish to thank Nati Linial and Ernst Mayr for helpful discussions and comments.

References

[A] B. Awerbuch, Complexity of Network Synchronization, *J. of the ACM* **32**, (1985), pp. 804–823.

[BCLR] S. Bhatt, F. Chung, F. Leighton and A. Rosenberg, Optimal Simulations of Tree Machines, *27th IEEE Symp. on Foundations of Computer Science*, pp. 274–282, 1986.

[Bo] B. Bollobás, *Extremal Graph Theory*, Academic Press, 1978.

[Bu] P. Buneman, A Characterization of Rigid Circuit Graphs, *Discr. Math.* **9**, (1974), pp. 205–212.

[C] L.P. Chew, There Is a Planar Graph Almost as Good as the Complete Graph, *Proc. 2nd Symp. on Computational Geometry*, pp. 169–177, 1986.

[DFS] D.P. Dobkin, S.J. Friedman and K.J. Supowit, Delaunay Graphs are Almost as Good as Complete Graphs, *28th IEEE Symp. on Foundations of Computer Science*, pp. 20–26, 1987.

[GJ] M.R. Garey and D.S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*, W.H. Freeman and Co., 1979.

[Ga] F. Gavril, The Intersection Graphs of Subtrees in Trees are Exactly the Chordal Graphs, *J. Comb. Th. B.* **16** (1974), pp. 47–56.

[GRE] J.R. Gilbert, D.J. Rose and A. Edenbrandt, A Separator Theorem for Chordal Graphs, *SIAM J. Alg. Disc. Meth.* **5** (1984), pp. 306–313.

[Go] M.C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, 1980.

[NZ] I. Niven and H.S. Zuckerman, *An Introduction to the Theory of Numbers*, J. Wiley & Sons Inc., 1972.

[PU1] D. Peleg and J.D. Ullman, An Optimal Synchronizer for the Hypercube, *6th ACM Symp. on Principles of Distributed Computing*, pp. 77–85, 1987.

[PU2] D. Peleg and E. Upfal, A Tradeoff Between Space and Efficiency for Routing Tables, *20th ACM Symp. on Theory of Computing*, pp. 43–52, 1988.

[PS] F.P. Preparata and M.I. Shamos, *Computational Geometry*, Springer-Verlag, 1985.