SPANNING TREES WITH MANY LEAVES[1]

Daniel J. Kleitman[2]
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Douglas B. West[3]
University of Illinois
Urbana, Illinois 61801

Abstract

A connected graph having large minimum vertex degree must have a spanning tree with many leaves. In particular, let $l(n, k)$ be the maximum integer m such that every connected n-vertex graph with minimum degree at least k has a spanning tree with at least m leaves. Then $l(n, 3) \geq n/4 + 2$, $l(n, 4) \geq (2n + 8)/5$, and $l(n, k) \leq n - 3\lfloor n/(k + 1) \rfloor + 2$ for all k. The lower bounds are proved by an algorithm that constructs a spanning tree with at least the desired number of leaves. Finally, $l(n, k) \geq (1 - b\ln k/k)n$ for large k, again proved algorithmically, where b is any constant exceeding 2.5.

1. Introduction

Given a connected simple graph G, suppose we wish to find a spanning tree in G with many leaves. If G is a cycle, we can only guarantee 2 leaves, but we may have better luck if we require that every vertex have degree at least k. To make this precise, let $\mathbf{G}_{n,k}$ denote the collection of connected n-vertex graphs with minimum degree at least k. We wish to determine $l(n,k)$, the the maximum m such that every graph in $\mathbf{G}_{n,k}$ has a tree with at least m leaves. Note that $l(n,2) = 2$.

The question of determining $l(n,k)$ has occured independently to several researchers. For this investigation, the question was raised by Lovasz and Saks [6]. Independently, Payan, Tchuente, and Xuong [7] showed that every 3-regular graph has a tree with at least $n/4$ leaves, and Storer [8] gave the lower bound of $n/4 + 2$ for that case. This was subsequently rediscovered by Linial and Sturtevant [5] and extended to minimum degree 3. Another proof appears in [2]. Storer was motivated by complexity considerations. The problem of finding a spanning tree with maximum number of leaves is NP-complete, even if G is regular of degree 4 [1]. We provide here a simple algorithm to construct a tree with at least $n/4 + 2$ leaves in any $G \in \mathbf{G}_{n,3}$. Extending this approach, we also present an algorithm to construct a tree with at least $(2n+8)/5$ leaves in any $G \in \mathbf{G}_{n,4}$. Finally, we present a simple family of algorithms that provide lower bounds implying $l(n,k) > (1 - \ln k/k)n$. In particular, this means that the fraction of the vertices that can be guaranteed to be leaves in the spanning tree with the most leaves approaches 1 as k grows.

For arbitrary k, a simple construction yields a $G \in \mathbf{G}_{n,k}$ with no tree having more than $n - 3\lfloor n/(k+1) \rfloor + 2$ leaves. When $k \leq 4$ and $k+1$ divides n, this achieves the bound. Griggs and Wu [3] have proved optimality for $k = 5$ (and give an alternate proof for $k = 4$. Linial [5] conjectured that this construction is essentially optimal in general, i.e. that $l(n,k) \geq n - 3n/(k+1) + c_k$ for each k and an appropriate constant c_k. More generally, Linial suspects that a connected graph with degree sequence $d_1 \geq d_2 \geq \cdots \geq d_n \geq 2$ has a spanning tree with at least $\Sigma(d_i - 2)/(d_i + 1)$ leaves.

Hutchinson [4] has investigated spanning forests. If one seeks a forest of c components with many leaves, then the upper and lower bounds presented here still hold, with 2 replaced by $2c$. They were further interested in limiting the diameter of the components, but our methods do not seem relevant to that question.

2. The Upper Bound Construction

THEOREM 1. $l(n,k) \leq n - 3\lfloor n/(k+1) \rfloor + 2$.

Proof. We construct $G_{n,k} \in \mathbf{G}_{n,k}$ having no tree with more than $n - 3\lfloor n/(k+1) \rfloor + 2$ leaves. Let $m = \lfloor n/(k+1) \rfloor$ and $r = n - m(k+1)$. Partition the vertex set $V(G)$ into sets $R_0, \ldots, R_{m-1}$, where $|R_i| = k+1$ for $i \neq 0$ and $|R_0| = k+1+r$. Choose $x_i, y_i \in R_i$. Place edges between all pairs of vertices in R_i except $x_i y_i$. Add the edges $Z = \{x_i y_{(i+1) \bmod m} : 0 \leq i < m\}$, and let $W = \{x_i\} \cup \{y_i\}$.

It suffices to show that any spanning tree T of $G_{n,k}$ has at most $n - 3m + 2$ leaves. Every pair of edges in Z forms an edge cut, so T lacks at most one edge of Z. Suppose first that $x_j y_{j+1} \notin T$; T then contains an x_i, y_i-path in R_i, for each i. This forces a non-leaf in $R_i - W$ for each i, and each vertex of W must be a non-leaf except $\{x_j, y_{j+1}\}$. On the other hand, if T omits no edge of Z, then T lacks an x_i, y_i-path in R_i for one value of i, say j. This forces at least $3(m-1)$ non-leaves in $V - R_j$, and $k \geq 2$ forces an additional non-leaf at x_j or y_j. $\blacksquare$

Note that $G_{n,k}$ contains many copies of the "almost-clique" $K_{k+1} - e$. If this induced subgraph is forbidden, a higher proportion of the vertices must be leaves. In particular, Griggs *et al* [2] have shown that every $G \in \mathbf{G}_{n,3}$ that does not contain $K_4 - e$ has a tree with at least $(n+4)/3$ leaves; this was earlier conjectured in [7]. The proof is more difficult than that of the unrestricted result in the next section.

We also note that when k is even there is another class of graphs where the tree with the most leaves has $n - 3\lfloor n/(k+1) \rfloor + 2$ leaves, as shown by a similar argument. The graph can be described as a cyclic sequence of cliques, in

which each vertex is also joined to every vertex of the clique before and after it. The cliques have sizes $k/2, k/2, 1, k/2, k/2, 1, \cdots$. Note that $G_{n,k}$ can also be described in this way, with the clique sizes being $1, k-1, 1, 1, k-1, 1, \cdots$.

3. The Case $k = 3$.

The lower bound for $k = 3$ appeared in [7] and in [8] for 3-regular graphs. We include a short proof of the general result, different from both of these, to illustrate the method we will use for $k = 4$. Another proof, similar in spirit to this but phrased also in terms of 3-regular graphs, appears in [2].

This and the later proofs grow the desired spanning tree of G via an iterative algorithm. In each case, we let T denote the current tree, with n vertices and l leaves. If x is a leaf of T, then the *out-degree* of x, denoted $d'(x)$, is the number of neighbors it has in $G - T$. The operation of *expansion at x* consists of adding to T the $d'(x)$ edges from x to *all* its neighbors not in T. We grow T by vertex expansion sequences (also called "operations"); this guarantees that all edges from T to $G - T$ are incident to leaves of T.

THEOREM 2. Every $G \in \mathbf{G}_{N,3}$ has a spanning tree with at least $N/4 + 2$ leaves.

Proof. A leaf x of T with $d'(x) = 0$ is *dead*; no expansion is possible at a dead leaf, and it must be a leaf in the final tree. Let m be the number of dead leaves in T. An expansion that makes y a dead leaf *kills y*. We call an expansion sequence *admissible* if its effect on T satisfies the "augmentation inequality" $3\Delta l + \Delta m \geq \Delta n$.

We initialize T to a small subtree and provide a collection of admissible operations to grow T into a spanning tree of G. If G is not 3-regular, we initialize T to be all edges incident to a vertex of maximum degree $\Delta \geq 4$. If G is 3-regular and every edge belongs to a triangle, then $G = K_4$, and the claim holds. Otherwise, G is 3-regular and has an edge in no triangle, and we initialize T to consist of such an edge and the four other edges incident to it.

If T is grown to a spanning tree with L leaves by admissible operations, then all leaves eventually die, and summing the augmentation inequality yields $3(L - \Delta) + L \geq N - \Delta - 1$ if G is not 3-regular, or $3(L - 4) + L \geq N - 6$ if G is 3-regular. These simplify to $4L \geq N + 2\Delta - 1 \geq N + 7$ and $4L \geq N + 6$, respectively. We can improve this to $4L \geq N + 8$ by considering the final admissible operation. For this operation, the augmentation inequality is satisfied with an excess of at least two, because the operation kills at least two final leaves whose death is not usually guaranteed for the operation.

It remains to present a collection of admissible operations of which at least one is always available until T absorbs all vertices, and to verify the statement claimed about the last operation. The three operations we use are illustrated in Fig. 1.
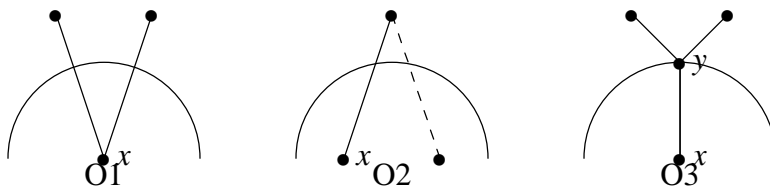

Fig. 1. Operations used when $k = 3$.

O1: If $d'(x) \geq 2$ for some current leaf x, then expanding at x yields $\Delta l = \Delta n - 1 \geq 1$ and $\Delta m \geq 0$.

O2: If $d'(x) \leq 1$ for every current leaf x and some vertex outside T has at least two neighbors in T, then expanding at one of them yields $\Delta l = 0$, $\Delta m \geq 1 = \Delta n$.

O3: If y is the only neighbor of x outside T and y has at least two neighbors not in T, then expanding at x and then y yields $\Delta l = \Delta n - 2 \geq 1$ and $\Delta m \geq 0$.

Because $k = 3$, any neighbor of a vertex in T has at least two neighbors in T or at least two neighbors outside T. This implies that one of O1-3 is available until T becomes a spanning tree. Also, the inequalities they satisfy imply that each is

admissible.

Now consider the final operation. Each of the three operations adds a(t least one) leaf z to T that previously did not belong to T. That leaf has a neighbor w not appearing in the illustration; since this is the last operation, w must have been a non-dead leaf of T. Since z and w both die now, we obtain the needed excess of two dead leaves. ∎

Before leaving this section, we note that the operations used above also yield the following result.

THEOREM 3. If every edge of G belongs to a triangle and $G \neq K_3$, then G has a tree with at least $(|V(G)| + 5)/3$ leaves, and this is best possible.

Proof. We use the same terminology, except that now an operation is *admissible* if it satisfies the augmentation inequality $2\Delta l + \Delta m \geq \Delta n$. Operations O1 and O2 above satisfy this admissibility inequality; we claim they suffice to grow T to a spanning tree. If T does not yet span, then there is an edge xy with $x \in T$, $y \notin T$; xy forms a triangle with some additional vertex z. If $z \notin T$, then O1 applies; if $z \in T$, then O2 applies.

If $G \neq K_3$ and $\Delta(G) < 4$, then $G = K_4$ or $G = K_4 - e$ and the bound holds. Otherwise, G has a vertex of degree at least 4 to use as the center of the initial T. If also $\delta(G) \geq 3$, then again the last operation provides two additional dead leaves, and summing the augmentation inequalities yields $2(L - 4) + L - 2 \geq N - 5$, or $L \geq (N + 5)/3$.

If $\delta(G) = 2$, then the last operation may provide only one additional dead leaf if it is O2 to a 2-valent vertex. However, if G has a 2-valent vertex x, then the edge-in-triangle property leads to a vertex w of degree at least 4 within distance 2 of x. If w is adjacent to x, then beginning at w makes x initially a dead leaf, and we have the same inequality as above. Otherwise, x and w have two common (adjacent) 3-valent neighors u, v. If the initial tree is the star at w plus the edge ux, then we begin with x, v as dead leaves and again get an extra at the end. Now the inequality is $2(L - 4) + L - 3 \geq N - 6$, or again $L \geq (N + 5)/3$. ∎

To show this is best possible, consider the graph $G_{n + \lfloor n/3 \rfloor 3}$ of Section 2, delete one cut-edge, and contract the remaining cut-edges. ∎

4. The Case $k = 4$.

For the case $k = 4$, we will use arbitrarily long expansion sequences as operations. We use the same terminology and notation as above, except that now an expansion sequence (or "operation") is *admissible* if it satisfies the augmentation inequality $4\Delta l + \Delta m \geq 2\Delta n$.

THEOREM 4. Every $G \in \mathbf{G}_{N,4}$ has a spanning tree with at least $(2N + 8)/5$ leaves.

Proof. Again we initialize T to be a small subtree, and we provide a collection of admissible operations to grow T into a spanning tree of G. If we provide an exhaustive set of admissible operations, summing the augmentation inequalities will yield $4(L - c_1) + (L - c_2) \geq 2(N - c_3)$, or $L \geq 2N/5 + c$, where c_1, c_3 are the number of leaves and vertices in the initial tree and c_2 is the number of leaves not counted as dead by summing the general augmentation inequalities. We postpone the discussion of the additive constant.

The first three operations are similar to those used for $k = 3$ and are illustrated in Fig. 2.
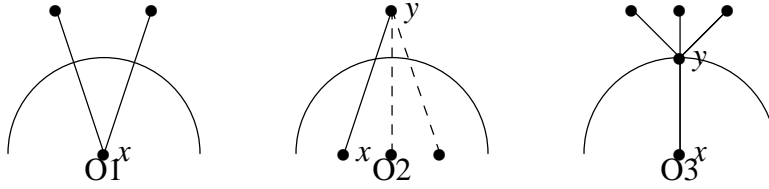
Fig. 2. Elementary operations used when $k = 4$.

O1: If $d'(x) \geq 2$ for some current leaf x, then expanding at x yields $\Delta l = \Delta n - 1 \geq 1$ and $\Delta m \geq 0$.

O2: If $d'(x) \leq 1$ for every current leaf x and some vertex outside T has at least three neighbors in T, then expanding at one of them yields $\Delta l = 0$, $\Delta m \geq 2 = 2\Delta n$.

O3: If y is the only neighbor of x outside T and y has at least three neighbors not in T, then expanding at x and then y yields $\Delta l = \Delta n - 2 \geq 2$ and $\Delta m \geq 0$.

Each of these operations is admissible. If none of O1-3 are available, then every non-dead leaf of T has out-degree one, and its neighbor outside T has two neighbors in T and two neighbors outside T.

The subsequent operations, which involve arbitrarily long expansion sequences, will apply in this case. We consider only *principal* expansion sequences; these expand a single leaf $x = y_0$ of T and then other leaves that do not belong to T before the initial expansion. The *length r* of a principal expansion sequence is the number of expansions outside T. A principal expansion sequence is *live* if each expansion after y_0 introduces two new vertices to the tree.

When O1-3 are not available, a live sequence almost satisfies the augmentation inequality for admissibility. The expansion at y_0 adds one vertex and kills the other neighbor of y_1 in T. Each subsequent expansion in Y increases l and adds two new vertices. Altogether, $4\Delta l + \Delta m = 4r + 1$ and $2\Delta n = 4r + 2$, leaving a deficiency of one in the augmentation inequality.

O4-7 rely on various additional conditions that imply admissibility and are illustrated by example in Fig. 3. For specification of O4-7, let Y be a live sequence of length r, and assume O1-3 are not available. Let W denote the set of leaves introduced by executing Y, and let $U = V(G) - (T \cup Y \cup W)$; U is the set of vertices that would still be outside the tree after executing Y.
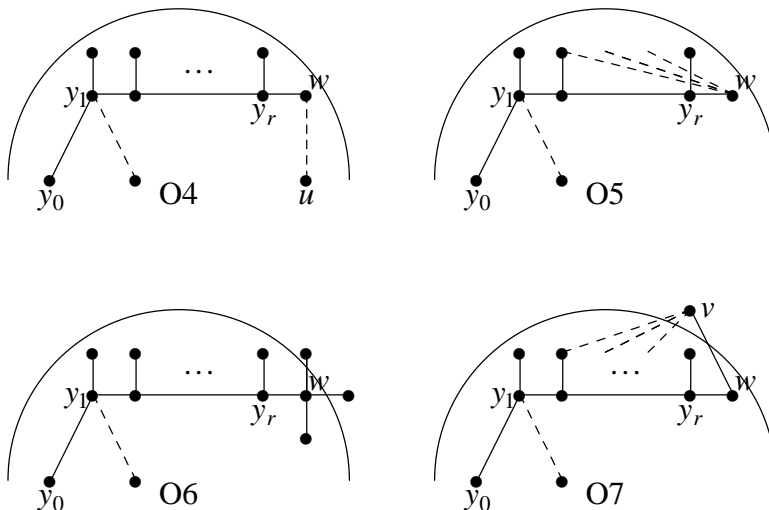

Fig. 3. Complex operations used when $k = 4$.

O4: If some $w \in W$ has a neighbor $u \in T$, then Y is admissible. Executing Y kills u, which increases Δm by one to eliminate the deficiency.

O5: If some $w \in W$ has all its neighbors in $Y \cup W$, then Y is admissible. Executing Y kills w, which increases Δm by one to eliminate the deficiency.

O6: If some $w \in W$ has at least three neighbors in U, then Y followed by (w) is admissible. The final expansion satisfies $4\Delta l - 2\Delta n \geq 2$, which eliminates the deficiency.

O7: If v is the unique neighbor in U for at least four vertices of W, then Y followed by expansion at one of these vertices is admissible. The final expansion kills (at least) three leaves, yielding $\Delta m - 2\Delta n \geq 1$, which eliminates the deficiency.

Next we show that some operation of types O1-7 is always available until T becomes a spanning tree. To prove this, we consider a special class of expansion sequences. A *linear* expansion sequence is a live sequence $Y = (y_0, \ldots, y_r)$ such that, for each $i \geq 1$, y_{i+1} is one of the two leaves introduced by expanding y_i. The illustrations in Fig. 3 suggest linear sequences, although expansion sequences of types O4-7 need not be linear. For a linear sequence, we let z_i denote the other leaf introduced by expanding y_i, and let z_r, w denote the two leaves introduced by expanding y_r. We may refer to w as y_{r+1}. Let $Z = \{z_1, \ldots, z_r\}$ and $W = Y \cup Z \cup \{w\}$. For $1 \leq i \leq r$, let $Y_i = (y_0, \ldots, y_i)$ and $Z_i = \{z_1, \ldots, z_i\}$. We use $R \cdot S$ for the concatenation of two vertex sequences, $N(a)$ for the set of neighbors of vertex a, and $N(S)$ for $\cup_{x \in S} N(x)$.

If O1-3 are unavailable and T does not span G, then any neighbor of T is the end of a linear sequence of length 1; i.e., linear sequences exist. Because G is finite, linear sequences cannot be arbitrarily long. If O1-7 are unavailable, then for a maximal linear sequence it must be true that each leaf introduced by the last expansion has exactly one neighbor in U.

Suppose O1-7 are unavailable, and let $Y = (y_0, \ldots, y_r)$ be a maximal linear sequence. In addition to y_r and one vertex $v \in U$, w has at least two additional neighbors. Because Y is live, these must appear in Z. Suppose $z_t, z_s \in N(w)$, with $t = \min\{i: z_i \in N(w)\}$, so $t < s \leq r$.

We claim that z_t must have exactly one neighbor u not in W. Otherwise, Y is of type O5 (killing z_t) or $Y_t \cdot (z_t)$ is of Type O6. Furthermore, if $u \neq v$, then $Y_t \cdot (z_t, w)$ is of type O6. Hence we may assume $u = v$. If $s < r$, then $Y_{s-1} \cdot (z_t, y_{r+1}, \ldots, y_{s+2})$ is a type O5 sequence killing y_s. Hence we may also assume $s = r$.

Applying the same arguments to $z_r = z_s = w'$, we obtain a neighbor $z_{t'}$ of w' adjacent to $v' \notin W$ (see Fig. 4). If $t = t'$, then $\{w, w', v, v'\} \subset N(z_t)$, and $Y_t \cdot (z_t)$ is of type O6. If $t \neq t'$ and $v \neq v'$, then $Y_{r-1} \cdot (z_t, z_{t'})$ is a type O5 sequence killing y_r. If $t \neq t'$ and $v = v'$, then v is the only neighbor in U for each of $\{z_t, z_{t'}, w, w'\}$, and $Y \cdot (w)$ is of Type O7.
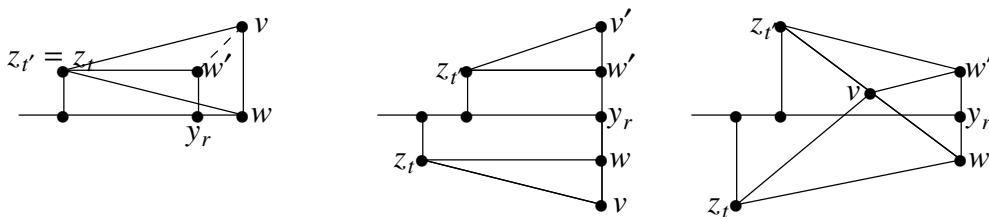

Fig. 4. Resolution of maximal live sequences when $k = 4$.

We have provided an exhaustive set of admissible operations. Now consider the additive constant. Recall that $L = 2N/5 + c$, where $c = (c_2 + 4c_1 - 2c_3)/5$ and c_1, c_2, c_3 are the number of initial leaves, leaves not counted as dead, and initial vertices. As for $k = 3$, each operation illustrated has a leaf incident to another edge not drawn, which again means that the last operation must kill at least two additional leaves (except for O2 and O7, the extra count is always at least four). Since G has minimum degree at least 4, we have $c \geq (2 + 16 - 10)/5 = 8/5$. ∎

It should be noted that there are only two known examples of graphs in $\mathbf{G}_{n,4}$ that have no tree with at least $2N/5 + 2$ leaves. These are the 4-regular graph on 6 vertices and the 4-regular graph on 8 vertices around a circle in which each vertex is joined to the four vertices closest to it. The desired bound asks for 5 and 6 leaves, respectively. On 6 vertices, having 5 leaves

would require a 5-valent vertex, and on 8 vertices, having 6 leaves would require two vertices whose neighborhoods include all the vertices. We conjecture that $2N/5 + 2$ is a lower bound except for these two examples. If G has a vertex of degree at least 5, then starting with the edges incident to it yields $c \geq 2$. If G is 4-regular and has an edge not in a triangle, then starting with its endpoints and their neighbors yields $c_1, c_2, c_3 = 6, 2, 8$ and $c = 2$. Hence any graph that violates this bound is 4-regular and has every edge in a triangle.

5. Larger Values of k.

In general, the conjectured lower bound on $l(n, k)$ is $(k - 2)n/(k + 1) + 2$, except possibly for small exceptions. Whenever k is even, there is a small example that slightly violates this bound. Whenever $k > 2$, we can choose n so that $3k/2 + 2 \leq n < 5(k + 1)/3$, and let G be the graph on n vertices around a circle in which each vertex is adjacent to the k closest vertices, $k/2$ in each direction. Then $(k - 2)n/(k + 1) + 2 > n - 3$, so the bound asks for a tree with $n - 2$ leaves. However, there are no two adjacent vertices whose neighborhoods cover $V(G)$.

The most interesting question, of course, is the coefficient of n in $l(n, k)$. For $k = 5$, Griggs and Wu [3] have proved the conjecture (they also have an alternate proof of the bound for $k = 4$, using a different augmentation inequality for admissibility). For large k, we give a short proof that the coefficient approaches 1. The ease of this argument is attributable to the fact that we are not seeking an optimal algorithm for any individual value of k. By considering more operations, i.e. by making the algorithm more complicated, we could improve the rate of convergence.

THEOREM 5. If k is sufficiently large, then there is an algorithm that constructs a spanning tree with at least $[1 - b\ln k/k]n$ leaves in any graph with minimum degree k, where b is any constant exceeding 2.5.

Proof. We design an algorithm like those above, in which the current tree T is expanded at leaves. We will develop an admissibility inequality that has the form $r\Delta l + \Delta M \geq (r - 1)\Delta n$, where r is a function of k. Here M is a measure of "deadness" for the leaves of the current tree. This is not a physical concept. Rather, the final value of M is a multiple counting of the leaves of the final tree, and the individual changes in M are an amortized distribution of this count over the operations.

The statistic we use to measure "deadness" is $M = \Sigma_{i=0}^{r-1} \alpha_i m_i$, where m_i is the number of leaves of T having i neighbors outside T; the coefficients α_i will be chosen shortly. It is natural to think that a leaf is more dead when it has fewer outside neighbors, so we will require $0 = \alpha_{r-1} \leq \alpha_{r-2} \leq \cdots \leq \alpha_0$. This requirement guarantees that expansion at any leaf with out-degree at least r will satisfy $r\Delta l + \Delta M \geq (r - 1)\Delta n$; the net change in M will be non-negative. Note that it makes sense to assume $r < k$.

If every operation used by the algorithm satisfies $r\Delta l + \Delta M \geq (r - 1)\Delta n$, then beginning with a star at a vertex of degree k and summing the augmentation inequalities yields $r(L - k) + \alpha_0 L \geq (r - 1)(N - k - 1)$, or $L \geq [(r - 1)N + (k + 1 - r)]/(r + \alpha_0) > (1 - (\alpha_0 + 1)/r)N$. We will choose the values of r and $\{\alpha_i\}$ so that the operations are admissible and $(\alpha_0 + 1)/r < b\ln k/k$, as desired.

For each $i < r$, define an *i-operation* be an operation that is performed only when the maximum out-degree of current leaves is i. Each i-operation begins by expansion at a vertex x with $d'(x) = i$. This or additional expansions may add to the tree a vertex y that was an outside neighbor of some z in the current tree with $d'(z) = j \leq i$. The net changes to M for this operation include $-\alpha_i$ for the loss of x as a leaf and $\alpha_{j-1} - \alpha_j$ for the effect of the edge yz on $d'(z)$. It will suffice to consider changes of these types.

Let $c_i = \alpha_{i-1} - \alpha_i$ for all i. If in addition to $\alpha_{r-1} \leq \cdots \leq \alpha_0$ we also have $c_{r-1} \leq \cdots \leq c_1$, then for any i-operation each edge from a new vertex to an old leaf contributes at least c_i to ΔM. Since we lose the contribution from the leaf expanded to begin the operation and ignore the possible gains for the new vertices, it suffices to show $r\Delta l + c_i q - \alpha_i \geq (r - 1)\Delta n$ for each i-operation, where q is the number of non-tree edges from new vertices to old vertices of the tree.

It order to guarantee the desired properties of the operations, we will choose $r = \lfloor k/5 \rfloor$ and $c_i = (r-i)/[i(k-3r)-r]$. Note that this formula for c_i increases as i decreases and that $c_1 < 1$ when $k \geq 5r$.

Let us now specify the i-operations. Let $i < r$ be the maximum out-degree of current leaves, and let x be a current leaf with maximum out-degree. Either we expand at x and stop, which we call Oi, or we expand at x and also at the new neighbor y of x for which the second expansion gives the maximum number of additional leaves; we call the latter Pi. We choose Pi if the number of vertices introduced by the second expansion is more than $3r - i$.

By construction, there is always an operation available to grow T until T spans. For the admissibility of Pi, we have $\Delta l = \Delta n - 2$. Ignoring gains due to possible edges from new vertices to old vertices, it suffices to show that $\Delta n \geq 2r + \alpha_i$. Since $\Delta n > 3r$, this holds when $c_i \leq 1$, since $\alpha_i = \Sigma_{j=i+1}^{r-1} c_j < rc_i$.

For the admissibility of Oi, suppose that y is an outside neighbor of x and a second expansion at y would introduce at most $3r - i$ new vertices. Because y also has at most i neighbors among x and the vertices introduced by expanding at x, it has at least $k - 3r$ neighbors in T besides x. This is true for each outside neighbor of x, so $q \geq i(k-3r)$ for the conditions under which we apply Oi. We have $\Delta l = i - 1$ and $\Delta n = i$, so

$$r\Delta l + c_i q - \alpha_i \geq r(i-1) + c_i(q-r) \geq r(i-1) + (r-i) = (r-1)\Delta n.$$

Finally, we study $\Sigma c_i = \alpha_0$. Since $k \geq 5r$, we have $\Sigma c_i \leq \Sigma_{i=1}^{r-1}(r-i)/r(2i-1)$. Using calculus, we can bound this by $\frac{1}{r}[r-1+\int_1^{r-1}(r-x)dx/(2x-1)]$. With the substitution $u = 2x-1$, we can evaluate the definite integral as $\frac{1}{4}[(2r-1)\ln(2r-3)-(2r-4)]$. Putting this all together yields $\alpha_0 < \frac{1}{4r}[4(r-1)-2(r-2)+(2r-1)\ln(2r-3)] < .5 + .5\ln 2r$. When we replace r by $\lfloor k/5 \rfloor$, we find $1 - (\alpha+1)/r > 1 - b\ln k/k$ for sufficiently large k as long as $b > 2.5$. ∎

This constant b can be reduced by choosing c_i and r to make use of some slack in the argument. In particular, the admissibility of Pi requires only $\Delta n \geq 2r + \alpha_i$, so we can use P$i$ whenever the second expansion introduces more than $2r + \alpha_i - i$ additional vertices. When this fails for all neighbors of x, we have $q \geq i(k-2r-\alpha_i)$. The admissibility of Oi requires only $c_i q - \alpha_i \geq r - i$, so it suffices to define c_i iteratively, with $\alpha_{r-1} = 0$, $c_i = (r-i+\alpha_i)/[i(k-2r-\alpha_i)]$, and $\alpha_{i-1} = \alpha_i + c_i$. We still wish to keep each c_i small to make α_0 of at most logarithmic size, and for this it suffices to have $k - 2r > \beta r$ (i.e., $r = \lfloor k/(2+\beta) \rfloor$) for some constant $\beta > 0$. The aim is then to bound α_0 by some function $f(\beta)\ln r$, which would lead to the constant $\beta f(\beta)$ in place of b. It does not seem worthwhile to pursue the details of this, since better improvements could be generated by considering a larger variety of operations.

Acknowledgement

References

[1] M.L. Garey and D.S. Johnson, *Computers and Intractibility: A Guide to the Theory of NP-Completeness*. (Freeman, 1979), p206.

[2] J.R. Griggs, D.J. Kleitman, and A. Shastri, Spanning trees with many leaves in cubic graphs. Manuscript 1988.

[3] J.R. Griggs and M.-S. Wu, Construction of spanning trees with many leaves, presented at xxth Southeastern Conf. on Graph Theory, Combinatorics, and Computing, Boca Raton (1989).

[4] J.P. Hutchinson, Bounds on the r-domination number, to appear.'

[5] N. Linial and D. Sturtevant, private communication.

[6]   L. Lovasz and M.E. Saks, private communication.

[7]   C. Payan, M. Tchuente, and N.H. Xuong, Arbres avec un nombre maximum de sommets pendants, *Discrete Mathematics* **49** (1984), 267-273.

[8]   J.A. Storer, Constructing full spanning trees for cubic graphs, *Inf. Proc. Letters* **13** (1981), 8-11.