The k-Spectrum of a Graph

P. Erdős & V. T. Sós
Mathematical Institute
Hungarian Academy of Sciences

R.J. Faudree[1]
Department of Mathematical Sciences
Memphis State University

Abstract

The *k-spectrum* $s_k(G)$ of a graph G is the set of integers that occur as the sizes of the induced subgraphs of G of order k. Properties of those sets $S \subseteq \{0, 1, 2, \cdots, \binom{k}{2}\}$ that are the k-spectrum $s_k(G)$ of some graph G will be investigated. Gap theorems, which indicate the distribution of elements in $s_k(G)$, will be proved, and the k-spectra of large order trees will be characterized as the union of two intervals. The number of subsets that are the k-spectrum of a graph will be studied, and extremal problems concerning the k-spectrum will be considered.

1 Introduction

The vertex and edge set of a graph G will be denoted by $V(G)$ and $E(G)$ respectively, and the *order* and *size* of G are the number of elements in $|V(G)|$ and $|E(G)|$. Specialized notation will be introduced as needed. If $S \subseteq V(G)$, then $\langle S \rangle$ will denote the subgraph induced by the vertices in S. For a fixed positive integer k, the *k-spectrum* of a graph G is $s_k(G) = \{|E(\langle S \rangle)| : S \subseteq V(G) \text{ and } |S| = k\}$. If $|V(G)| < k$, then $s_k(G) = \emptyset$, so, $s_k(G) \subseteq \{0, 1, 2, \cdots, \binom{k}{2}\}$. For example, $s_k(K_n) = \{\binom{k}{2}\}$ and $s_k(K_{1,n}) = \{0, k-1\}$ for $n \geq k$. The k-spectrum of a graph was studied in [1]. For small values of k all k-spectra of graphs were determined, and several extremal problems involving the k-spectra of graphs were considered.

There are some obvious properties of the k-spectrum. If H is an induced subgraph of G, then clearly $s_k(H) \subseteq s_k(G)$. Thus, if G is the disjoint union of all nonisomorphic graphs of order k, then clearly $s_k(G) = \{0, 1, 2 \cdots, \binom{k}{2}\}$. If $R \subseteq \{0, 1, 2, \cdots, \binom{k}{2}\}$, and $R^* = \{\binom{k}{2} - r : r \in S\}$, then $s_k(G) = R$ implies $s_k(\overline{G}) = R^*$ for the complement $\overline{G}$ of G. By Ramsey's theorem [4], any large order graph G must have

[1]Research supported by ONR Grant No. N00014-91-J-1085 and NSA Grant No. MDA 904-90-H-4034

a clique of order k or an independent set of order k. Hence, at least one of 0 or $\binom{k}{2} \in s_k(G)$. If both 0 and $\binom{k}{2} \in s_k(G)$, then what other terms must be in $s_k(G)$? A complete answer to this question can be found in [1]. This type of question will be studied in section 2, where gap theorems that give information about the distribution of terms in a k-spectrum will be proved.

In section 3 the number of different subsets of $\{0, 1, 2, \cdots, \binom{k}{2}\}$ that are the k-spectrum of a graph will be studied. Which small collections of sets that are the k-spectrum of a graph will be determined, and which families of graphs are determined by their k-spectrum will be discussed in this section. The k-spectra of large order trees will characterized in section 4. Bounds on the number of possible subsets that are the k-spectrum of a tree will be given, and other extremal problems involving the k-spectrum of a tree will be discussed.

2 The Gap Theorems

Consider the graph $K_n - K_k$ obtained from a K_n by deleting the edges of a K_k. For $n \geq 2k$, it is easy to see that $s_k(K_n - K_k) = \{0, k-1, 2k-3, \cdots, \binom{k}{2} - \binom{j}{2}, \cdots, \binom{k}{2}\}$. If the elements of $s_k(K_n - K_k)$ are ordered using the natural order of the integers, then the maximum "gap" between consecutive terms in the k-spectrum is at most $k-1$, and in fact, the gap becomes smaller as the terms become larger. However, the complement of $K_n - K_k$, which is $K_k \cup \overline{K}_{n-k}$, has the following k-spectrum: $s_k(K_k \cup \overline{K}_{n-k}) = \{0, 1, 3, \cdots, \binom{j}{2}, \cdots, \binom{k}{2}\}$, which has a maximum gap of $k-1$ at the end and smaller gaps at the beginning. More specifically, for consecutive terms $s_1 = \binom{j}{2}$ and $s_2 = \binom{j+1}{2}$ in $s_k(K_k \cup \overline{K}_{n-k})$ the gap is j and j is approximately $\sqrt{2s_1}$ (and $\sqrt{2s_2}$). Likewise, in the complementary graph $K_n - K_k$, the gap between the two terms $s_1 = \binom{k}{2} - \binom{j+1}{2}$ and $s_2 = \binom{k}{2} - \binom{j}{2}$ is j, which is approximately equal to $\sqrt{2\binom{k}{2} - 2s_i}$ for $i = 1$ or 2. These examples indicate what will be proved about the gap structure of the k-spectrum of a graph.

We start by stating and proving an elementary "gap theorem" for the k-spectrum of a graph. If H is a subgraph of G and v is a vertex of H, then $\langle V(H) - \{v\}\rangle$ will be denoted by just $H - v$; in the same way, if u is a vertex of G, then $\langle V(H) \cup \{u\}\rangle$ will be denoted by $H + u$.

THEOREM 1 (Elementary Spectrum Gap) *If $s_k(G) = \{s_1 < s_2 < \cdots < s_r\}$, then $|s_{i+1} - s_i| \leq k-1$ for $1 \leq i < r$. Moreover, $|s_{i+1} - s_i| \leq k-2$ except possibly when $s_i = 0$ or $s_{i+1} = \binom{k}{2}$.*

PROOF: Let $a = s_i$ and $b = s_{i+1}$, and select sets X and Y with k vertices such that $|E(\langle X\rangle)| = a$ and $|E(\langle Y\rangle)| = b$. We can assume that X and Y have been chosen such that $X \cap Y$ is a maximum among all pairs of sets with the above property. Let $X - Y = \{x_1, x_2, \cdots, x_t\}$, $Y - X = \{y_1, y_2, \cdots, y_t\}$, and $X \cap Y = \{z_1, z_2, \cdots, z_{k-t}\}$, and let $H_0 = \langle X\rangle$ and $H_t = \langle Y\rangle$. Define a sequence of graphs starting with H_0 and ending with H_t by letting $H_{i+1} = H_i + y_i - x_i$ for $1 \leq i < t$. By assumption, $|E(H_i)| < a$ or $|E(H_i)| > b$. Observe that $|\,|E(H_i)| - |E(H_{i+1})|\,| \leq k-1$ with equality if and only if one of x_i and y_i has degree $k-1$ and the other has degree 0 in the graphs H_i and H_{i+1} respectively.

First consider the case when $t > 1$. If $|E(H_1)| > b$, then $b - a < |E(H_1)| - |E(H_0)| \leq k-1$. If $|E(H_1)| < a$, then select the first j such that $|E(H_{j+1})| \geq b$ (possibly $j = t-1$). By assumption, $|E(H_j)| < a$, and so again $b - a < |E(H_{j+1})| - |E(H_j)| \leq k-1$. If $t = 1$, then the result follows unless x_1 has degree 0 in H_0 and y_1 has degree $k-1$ in H_1. However, if $H_0 \cong \overline{K}_k$ (or $H_1 \cong K_k$), the result follows. If this does not occur, then there is a vertex, say z_1, such that $0 < d_{H_0}(z_1) < k-1$. Then, for the graph $H' = H_0 - z_1 + y_1$, we have $a < |E(H')| < b$, a contradiction that completes the proof of Theorem 1. □

The examples presented prior to Theorem 1 indicate that the maximum gaps in the k-spectrum decrease as the terms of the spectrum move away from the extremes 0 and $\binom{k}{2}$. We will give an improved gap theorem for s_k. But first we prove a gap theorem for the degree sequence of a graph. If G is a graph of order n with degree sequence $d_1 \leq d_2 \leq \cdots \leq d_n$, then the *gap degree* for G, denoted by $gd(G)$, is the maximum of $d_{i+1} - d_i$ for $1 \leq i < n$. If G is regular, then $gd(G) = 0$, and if G is a star, then $gd(G) = n-2$. The graph $G = K_j \cup \overline{K}_{n-j}$ has $\binom{j}{2}$ edges and $gd(G) = j-1$, and so $gd(G) = \sqrt{2|E(G)| - (j-1)}$. Also, for the complementary graph $\overline{G} = K_n - K_j$, we have $gd(\overline{G}) = \sqrt{\binom{n}{2} - 2|E(G)| + (j-1)}$. These examples indicate the sharpness of the following degree gap result, and parallel the examples given for the k-spectrum of a graph.

THEOREM 2 (Gap Degree) *If G is a graph of order n, then* $gd(G) \leq \max\{\sqrt{2|E(G)|}, \sqrt{\binom{n}{2} - 2|E(G)|}\}$.

PROOF: For $n \leq 5$ it is straightforward to verify Theorem 2, so we will assume that $n \geq 6$. Let G be a graph of order n and size $n^* = |E(G)|$ for which $gd(G)$ is a maximum. We can assume that $|E(G)| \geq n(n-1)/4$, since $gd(G) = gd(\overline{G})$, and the upper bound is symmetric.

Let h and ℓ be the degrees of two consecutive terms in the degree sequence of G such that $h-\ell = gd(G)$. We now partition the vertices of G into two parts, those of degree at least h, which we call *high degree* vertices and denote by H, and those of degree at most ℓ, which we will call *low degree* vertices and denote by L. Replacing an edge in L or an edge between H and L by an edge in H will not lower the gap. If possible, we will do this. The same is true of replacing an edge in L by an edge between L and H, so do this when possible. Thus, we can assume that if there are any edges in L, there is a complete bipartite graph between L and H. We can also assume that if there are any edges between L and H, then H induces a complete graph. Let m be the number of vertices in H, and so there are $n-m$ vertices in L.

We will first consider the case when $\langle L\rangle$ contains some edges, say $q > 0$ edges. In this case, we know that $\langle H\rangle$ is a complete graph, and the edges between H and L form a complete bipartite graph. If $q \geq n-m-1$, then we will change G. Let G' be a graph of order n with a complete subgraph H' of order $m+1$ such that all of the vertices in H' have degree $n-1$, and with an additional $q-(n-m-1)$ edges placed in $L' = G' - H'$ arranged such that this graph is nearly regular (vertices differ in degree by at most 1). Then, $|E(G')| = \binom{m+1}{2} + (m+1)(n-m-1) + q - (n-m-1) = \binom{m}{2} + m(n-m) + q = |E(G)|$. Also, $gd(G') = n-1-(m+1+\lceil \frac{2(q-n+m+1)}{n-m-1}\rceil) = n-1-(m+\lceil\frac{2q}{n-m-1}\rceil - 1) \geq n-1-(m+\lceil 2q/(n-m)\rceil) \geq gd(G)$, since $q \geq n-m-1$. Thus, it is sufficient to consider G' instead of G. A repetition of this change results in a graph G'' with m'' high degree vertices and less than $n-m''-1$ edges between the low degree vertices. Thus, we can assume that G has the property that L has $q < n-m-1$ edges.

We now show that $gd(G) \leq \sqrt{2|E(G)|}$. In fact, $gd(G) \leq n-1-(m+\lceil 2q/(n-m-1))\rceil \leq n-m-1$. To complete the proof of this case, it is sufficient to show that $(n-m-1)^2 \leq 2(\binom{m}{2} + m(n-m) + q)$, and this is equivalent to $4nm - 2m^2 - n^2 + m + 2n + 2q - 1 \geq 0$. However, since $|E(G)| \geq n(n-1)/4$, we have $\binom{m}{2} + m(n-m) + q \geq n(n-1)/4$, and this is equivalent to $4nm - 2m^2 - n^2 - 2m + n + 4q \geq 0$. Thus, it is sufficient to show that $3m + n - 2q - 1 \geq 0$, or (using the fact that $q \leq n-m-2$), it is sufficient to show that $5m - n + 3 \geq 0$. This is true, which completes the proof of the case when L has edges.

From this point on we can assume there are no edges in L. If there are no edges between H and L, then $\ell = 0$, and $gd(G) = h < m$. The number of edges in G is at least $mh/2$, and clearly $h^2 < mh \leq (\sqrt{2|E(G)|})^2$, which completes the proof of this case. We can assume there are edges between H and L and $\langle H\rangle$ is a complete subgraph.

The number of edges between H and L is $cm(n-m)$ for some c with $0 < c \leq 1$. Therefore the "average degree" of a vertex in H relative to L is $c(n-m)$ and the average degree of a vertex in L is cm. There is no loss of generality in assuming that each vertex of L has degree either $\lfloor cm \rfloor$ or $\lceil cm \rceil$, and each vertex in H has degree either $m-1+\lceil c(n-m) \rceil$ or $m-1+\lfloor c(n-m) \rfloor$. Therefore, $gd(G) = m-1+\lfloor c(n-m) \rfloor - \lceil cm \rceil \leq m-1+cn-2cm$. To complete the proof, it is sufficient to show that $(m-1+cn-2cm)^2 \leq m(m-1)+2cm(n-m)$, which is equivalent to $4c^2nm - 4c^2m^2 + 2cm^2 - c^2n^2 + 2cn - 2cm + 2m - 1 \geq 0$. By assumption, $n(n-1)/4 \leq |E(G)| = m(m-1)/2 + cm(n-m)$, which implies $4cnm - 4cm^2 + 2m^2 - 2m - n^2 + n \geq 0$. The fact that $c(1-c)n^2 + 2m + cn - 1 \geq 0$ and the previous inequality immediately gives the required inequality. This completes the proof of Theorem 2. □

We are now prepared, using Theorem 2, to prove an gap theorem for the k-spectrum of a graph.

THEOREM 3 (Spectrum Gap) *If $s_k(G) = \{s_1 < s_2 < \cdots < s_r\}$, then $|s_{i+1} - s_i| \leq \max\{\sqrt{2s_{i+1}}, \sqrt{\binom{k+1}{2} - 2s_i + k}\}$ for $1 \leq i < r$.*

PROOF: The structure of this proof will be identical to that of Theorem 1. Select sets X and Y with k vertices such that $|E(\langle X \rangle)| = s_i$ and $|E(\langle Y \rangle)| = s_{i+1}$. Let $X - Y = \{x_1, x_2, \cdots, x_t\}$ and $Y - X = \{y_1, y_2, \cdots, y_t\}$, and let $H_0 = \langle X \rangle$ and $H_t = \langle Y \rangle$. Define a sequence of graphs by letting $H_{i+1} = H_i + y_i - x_i$ for each $1 \leq i < t$. By assumption, $|E(H_i)| < s_i$ or $|E(H_i)| > s_{i+1}$ for $1 \leq i < t$. Select the first j such that $|E(H_{j+1})| \geq s_{i+1}$, and so $|E(H_j)| \leq s_i$. Consider the graph $H' = H_j \cup H_{j+1}$, which has $k+1$ vertices and between s_{i+1} and $s_i + k$ edges. Let d_j and d_{j+1} be the degrees of x_j and y_j in H'. Then $s_{j+1} - s_j = d_{j+1} - d_j$, and by Theorem 2, $d_{j+1} - d_j \leq \max\{\sqrt{2|E(H')|}, \sqrt{\binom{k+1}{2} - 2|E(H')|}\}$. The required inequality follows from the fact that $s_{j+1} \leq |E(H')| \leq s_j + k$, and the proof of Theorem 3 is complete. □

Previously described examples indicate that the order of magnitude of the bounds given in Theorem 3 cannot be improved.

3 Extremal Problems

Let n_k be the number of subsets of the $2^{\binom{k}{2}+1}$ subsets of $\{0, 1, 2, \cdots, \binom{k}{2}\}$ that are the k-spectrum of a graph. It is clear from Theorem 3 that not all subsets of $\{0, 1, 2, \cdots, \binom{k}{2}\}$ are the k-spectrum of a graph. For any integer $r \leq k-1$ and any selection of nonnegative integers $0 \leq a_1, a_2, \cdots, a_r < k-1$, consider the graph H obtained from a complete $(r+1)$-partite with r parts of order k and one part of order $n - rk$ by adding edges to form a star K_{1,a_i} with a_i edges into the i^{th} part of the complete $(r+1)$-partite graph $(1 \leq i \leq r)$, and making the last part complete. The smallest r terms in $s_k(H)$ are $\{a_1, a_2, \cdots, a_r\}$, and all of the remaining terms are at least $k-1$. This implies $n_k \geq 2^{k-1}$. In fact, this lower bound can be improved. Consider, for example, any selection of t integers $b_1, b_2, \cdots, b_t \subseteq [k, 2k-5] - \{a_1+k-2, a_1+k-1, a_2+k-2, a_2+k-1, \cdots, a_r+k-2, a_r+k-1\}$, where $[k, 2k-5]$ are the integers from k to $2k-5$. In this case let G be the graph obtained from a complete $(r+t+1)$-partite with $r+t$ parts of order k and one complete part of order $n-(r+t)k$ by adding a star K_{1,a_i} into the i^{th} part of the complete $(r+t+1)$-partite graph $(1 \leq i \leq r)$ and adding a nearly regular graph on k vertices and b_j edges in the j^{th} part $(r+1 \leq j \leq r+t)$. It is straightforward to show that the only terms of $s_k(G)$ less than or equal to $2k-5$ are $\{a_1, a_2, \cdots, a_r, b_1, b_2, \cdots, b_t, k-1, a_1+k-2, a_1+k-1, a_2+k-2, a_2+k-1, \cdots, a_r+k-2, a_r+k-1\}$. This implies that there are at least $\sum_{r=1}^{k-1} \binom{k-1}{r} 2^{k-4-2r}$ distinct sets that are the k-spectrum of a graph. Since, $\sum_{r=1}^{k-1} \binom{k-1}{r} 2^{k-4-2r} > (5/2)^{k-1}/16$, we have verified the following rather crude bounds for n_k. This type of construction can be extended to give additional, but not significant, improvements.

THEOREM 4 *For any integer $k \geq 2$, $\frac{1}{16}\left(\frac{5}{2}\right)^{k-1} < n_k < 2^{\binom{k}{2}+1}$.*

For small values of ℓ it is possible to enumerate all the ℓ-sets that the k-spectrum of a graph G of large order.

By [4] any such graph G must have either 0 or $\binom{k}{2}$ in its k-spectrum. Thus, $\{0\} = s_k(\overline{K}_n)$ and $\{\binom{k}{2}\} = s_k(K_n)$ are the only sets with one element that are the k-spectrum of some large order graph G. Note that $s_k(K_2 \cup \overline{K}_{n-2}) = \{0, 1\}$, $s_k(K_{1,n-1}) = \{0, k-1\}$, and $\{\binom{k}{2}, \binom{k}{2}-1\}$, and $\{\binom{k}{2}, \binom{k}{2}-(k-1)\}$ are the k-spectrum of the respective complementary graphs.

Moreover, these are the only sets with precisely 2 elements that are the k-spectrum of some graph G. To see this, let G be a graph for which this is not true. We can assume that $k \geq 3$. With no loss of generality we can assume (by [4]) that $0 \in s_k(G)$, and of course G has at least one edge. If

$\binom{k}{2} \in s_k(G)$, then by Theorem 1 there will be at least 3 terms in the spectrum. If G has a large connected component, then this component of G must contain an induced tree on k vertices, and so $k-1 \in s_k(G)$. If G has no large connected component, then it has many components (at least $k-1$), so $1 \in s_k(G)$.

The graphs $K_{1,2} \cup \overline{K}_{n-3}$, $K_3 \cup \overline{K}_{n-3}$, $K_{1,n-2} \cup K_1$, $K_{2,n-2}$, and $K_2 + \overline{K}_{n-2}$ have as k-spectra $\{0,1,2\}$, $\{0,1,3\}$, $\{0, k-2, k-1\}$, $\{0, k-1, 2k-4\}$, and $\{0, k-1, 2k-3\}$ respectively. Using straightforward techniques these sets, along with the k-spectra that are complementary to them, can be shown to be the only sets with 3 elements that are the k-spectrum of a graph. More generally, there is the following result.

THEOREM 5 *For any fixed positive integer r, the number of k-spectra of a sufficiently large order graph with exactly r terms is bounded by a constant $C = C_r$ (independent of k).*

PROOF: For each large order graph G with $|s_k(G)| = r$ at least one of 0 and $\binom{k}{2}$ is in $s_k(G)$ by [4], and for $k > 2r$ not both, since Theorem 1 implies there are at least $k/2$ terms in $s_k(G)$. Thus, with no loss of generality we can assume that $0 \in s_k(G)$ but $\binom{k}{2} \notin s_k(G)$. Select a graph G of order n such that $s_k(G) = \{a_1, a_2, \cdots, a_r\}$ and a_r is a maximum element in all such k-spectra. If $a_r \le r^2$, then the number of sets with r elements that are the k-spectrum is at most $\binom{r^2}{r}$, which is independent of k. Thus, we assume that $a_r > r^2$. Let A be a set of k vertices of G with a_r edges. If there is a large set, say $n/2$ vertices of G that are not adjacent to any vertex of A, then by [4] there is a large independent set B with the same property. By successively replacing vertices in A of degree at least 1 by vertices in B, more than r terms in $s_k(G)$
will be generated.

We can assume that half the vertices in $\overline{A}$ are adjacent to a vertex in A. This implies that there is a vertex v of A of very large degree (at least $n/2k$). Denote the neighborhood of v by D. If there is a sufficiently large (as a function of r) number of vertices not in D, then by [4] and a bipartite version of Ramsey's theorem (see [2]), there is either a complete bipartite graph $K_{2r,k}$ or the bipartite complement of this graph with the k vertices in D and the $2r$ vertices not in D. Moreover, we can assume that the k vertices are independent and the r vertices form either a clique or an independent set. Denote this graph by H. If H contains the complete bipartite graph, then Theorem 1 implies that $s_k(G)$ has at least $r+1$ terms. If not, then by using v and the vertices of H, at least $r+1$ terms of the k-spectrum can also be generated. This implies v must be adjacent to at least $n - c'_r$ vertices of G, where c'_r is a constant depending only on r.

Let T be all the vertices of G, which like v, are of very large degree. Then each vertex in T has degree at least $n - c'_r$. Also, $|T| < r$, for otherwise, $K_{r,k-r} \subset G$ and Theorem 1 implies $s_k(G)$ has at least $r+1$ terms. Let S be the vertices of G adjacent to each vertex of T. Thus, $|S| \ge n - c^*_r$ for some constant c^*_r depending only on r. If $\langle S \rangle$ has r^2 edges, then there is a subset of S with k vertices and at least r^2 edges. Thus, using the same argument used with the set A, there must be a vertex in S of very large degree, a contradiction. Hence, there are at most r^2 edges in $\langle S \rangle$. If a vertex u not in $S \cup T$ has as many as $2r^2 + r$

adjacencies in S, then there is an independent set of $k+r$ vertices of S such that u is adjacent to precisely r of these vertices. Using u, v, and these $k+r$ vertices, $r+1$ terms of $s_k(G)$ can be generated. Therefore, we can assume that u has less that $2r^2+r$ adjacencies in S. This implies that there is a set R (vertices in $\overline{S}$ and vertices of S of degree at least 1 relative to S) of at most c_r vertices, (c_r depends only on r), such that G differs from the complete bipartite graph between S and T by only the edges in R. Thus, the number of different k-spectra depends only on r. This completes the proof of Theorem 5. □

In some cases the k-spectrum of a graph determines the family of graphs. For $n \geq k \geq 3$, it is true for K_n, which has $s_k(K_n) = \{\binom{k}{2}\}$, and $\overline{K}_n$, which has $s_k(\overline{K}_n) = \{0\}$. The same is true for $K_{1,n-1}$, which has $s_k(K_{1,n-1}) = \{0, k-1\}$. More generally, the following is true.

THEOREM 6 *For $k \geq 3$, $m \geq \lfloor \frac{k}{2} \rfloor$, and $n \geq k$, $s_k(K_{m,n}) = \{0, 1(k-1), 2(k-2), \cdots, \lfloor \frac{k}{2} \rfloor \cdot \lceil \frac{k}{2} \rceil\}$. Also, any graph of sufficiently large order with this k-spectrum must be a member of this family of complete bipartite graphs.*

PROOF: It is straightforward to verify that $s_k(K_{m,n}) = \{0, 1(k-1), 2(k-2), \cdots, \lfloor k/2 \rfloor \cdot \lceil k/2 \rceil\}$. Conversely, assume that G is a graph of large order n with this k-spectrum. First consider the case when G is a bipartite graph. The largest term in the k-spectrum is $\lfloor k/2 \rfloor \cdot \lceil k/2 \rceil$, so clearly G must contain at least $\lfloor k/2 \rfloor$ vertices in each part. If G is not a complete bipartite graph, then clearly $i \in s_k(G)$ for some $0 < i < k-1$.

Assume G is not bipartite. Select the smallest odd cycle of G, say C_r of order r, which is an induced cycle. If $r \geq k+2$, then $k-2 \in s_k(G)$, a contradiction. Thus, we assume that $r \leq k+1$. If $r > 3$, then any vertex not in C_r is adjacent to more than 2 vertices of C_r, and if there are two adjacencies they must be at a distance 2 on the cycle. If $r = 3$, then the number of adjacencies could be 3. Since the order of G is large, there is a large set S of vertices not in C_r that have precisely the same adjacencies in C_r. The graph G does not contain a K_k since $\binom{k}{2} > \lfloor k/2 \rfloor \cdot \lceil k/2 \rceil$. Therefore, by [4], we can assume that S is an independent set with at least k vertices. Using S and C_r, it is easy to show that $1 \in s_k(G)$ (which gives a contradiction) unless $r = 3$ and each vertex of S is adjacent to either 2 or 3 vertices of C_3. Hence, we can assume that $r = 3$. In either case, $k-2$ vertices of S with an appropriate 2 vertices of the C_3 implies that $2k-3 \in s_k(G)$, which give a contradiction except when $k = 6$. A straightforward case analysis shows that graphs containing either $K_2 + K_6$ or $K_3 + K_6$ cannot have a 6-spectrum of $\{0, 5, 8, 9\}$. This completes the proof of Theorem 6. □

The most obvious open question from this section is the determination of the order of magnitude of n_k. It would also be interesting to know of additional families of graphs determined by their k-spectra.

4 The k-Spectrum of a Tree

If T_n is a tree of order $n \geq k$, then $s_k(T_n) \subseteq \{0, 1, 2, \cdots, k-1\}$, since a forest on k vertices can have at most $k-1$ edges. Clearly $k-1 \in s_k(T_n)$, and if $n \geq 2k-1$, then T_n has a independent set of order k. In this case $0 \in s_k(T_n)$ as well. For the star $K_{1,n}$ with $n \geq k$, 0 and $k-1$ are the only elements in the k-spectrum (i.e. $s_k(K_{1,n} = \{0, k-1\}))$.

If $s_k(T_n) = [0, k-1]$, (the integers from 0 to $k-1$) we say the k-spectrum is *tree complete*. There are several conditions on a tree that insure that the k-spectrum is tree complete. Consider, for example, a path. For any $n \geq k$, $s_k(P_n) = [\max\{0, 2k-1-n\}, k-1]$. In particular, if $n \geq 2k-1$, then $s_k(P_n) = [0, k-1]$. To see this, let $P_n = (x_1, x_2, \cdots, x_n)$ be a path with n vertices. Note for any t with $2k - n \leq t \leq k$, that the set $\{x_1, x_2, \cdots, x_t, x_{t+2}, x_{t+4}, \cdots, x_{t+2(k-t)}\}$ induces a graph with $t-1$ edges. Thus $s_k(P_n) \supseteq [\max\{0, 2k-1-n\}, k-1]$. Also, a simple induction proof shows that if $n \leq 2k-1$, then any set of k vertices of P_n will induce a graph with at least $2k-1-n$ edges. This proves the claim. The next result give some elementary conditions on a tree T_n that insure that the k-spectrum is tree complete.

THEOREM 7 *The k-spectrum of a tree T_n is tree complete if (i) T_n has diameter at least $2k-2$, (ii) T_n has at least k independent endedges, or (iii) $\Delta(T_n) \leq k$ and $n \geq \max\{3k-5, 2k-1\}$.*

PROOF: (i) If T_n has diameter at least $2k-2$, then T_n contains a path P_{2k-1}. We have already observed that the k-spectrum of P_n is tree complete if $n \geq 2k-1$, and thus the same is true for T_n.

(ii) For any positive integer $t \leq k$, select a subtree T_t of T_n with t vertices. Then, $k-t$ endvertices from the k independent endedges can be chosen that are not adjacent to any of the vertices in the tree T_t. This set induces a subgraph with $t-1$ edges, which completes the proof of this case.

(iii) Fix some integer r such that $0 \leq r \leq k-2$. We will show that there is a subtree T_{k-r} with $k-r$ vertices and r independent vertices that are not adjacent to any vertex of T_{k-r}. This will insure that $k-r-1 \in s_k(T_n)$. Since, clearly $0 \in s_k(T_n)$ this would imply the k-spectrum is tree complete.

Prune the tree T_n by deleting an endstar (including the center of the star) with the smallest number of vertices. Continue to do this until at least $2r-1$ vertices have been pruned, leaving a tree T'. If in this process, the subtrees being pruned are never stars, then no more than half of the remaining vertices are pruned at any step. Thus, at most $2r - 2 + (n - 2r + 2)/2 = n/2 + r - 1$ vertices are pruned. Hence

T' has at least $n/2 - r + 1 \geq k - r$ vertices, as required. If one of the pruned subtrees is a star, then this tree has at most k edges, and therefore $n \leq 2r - 2 + k \leq 3k - 6$, a contradiction. Hence, we have that the procedure terminates with a tree T' with at least $k - r$ vertices. A set of r independent vertices from the pruned vertices (endvertices of the endstars) can be selected that are not adjacent to any vertex of T'. This completes the proof of this case and of Theorem 7. □

Not every subset of $\{0, 1, 2, \cdots, k-1\}$ is the k-spectrum of a large order tree; in fact, in general very few subsets are. We will give a characterization of the k-spectrum of large order trees, but before doing so we describe a family of examples that indicate the nature of the k-spectrum of a large tree. For positive integers a, b and c, let $T_{a,b,c}$ denote the tree formed from a path P_{a+1}, a star $K_{1,b}$, and a star $K_{1,c}$ by identifying the center of the star $K_{1,b}$, an endvertex of the path P_{a+1}, and an endvertex of the star $K_{1,c}$. Thus if $a, b, c > 0$, then $T_{a,b,c}$ has a total of $a+b+c+1$ vertices, $b+c$ vertices of degree 1, one vertex of degree $b+2$, one vertex of degree c, and the remaining $a-1$ vertices of degree 2. Note that if $a = b = 0$, then $T_{a,b,c}$ is just a star with c edges. Let $n = a+b+c+1$, and denote this tree by T_n. If $a+b < k \leq c$, then it is straightforward to show that $s_k(T_n) = [0, a+b] \cup [k-1-\lceil a/2 \rceil - b, k-1]$. If $r = a + b$ and $s = \lceil a/2 \rceil + b$, then by an appropriate choice of a and b, we have $s_k(T_n) = [0, r] \cup [k-1-s, k-1]$, where $0 \leq \lceil r/2 \rceil \leq s \leq r$.

Before proving the characterization result for the k-spectrum of a large tree we will prove a useful result that will be needed in the proof. We have already shown that if a large order tree T_n does not have a vertex of degree at least k, then the k-spectrum is tree complete. The next result deals with the case of a vertex of large degree.

THEOREM 8 *If $\Delta(T_n) \geq k$, then there exist integers r and s with $0 \leq \lceil r/2 \rceil \leq s \leq r \leq k-1$ such that $s_k(T_n) = [0, r] \cup [s, k-1]$.*

PROOF: Let v be a vertex of T_n of degree $\Delta \geq k$, let $v_1, v_2, \cdots, v_\Delta$ be the vertices of T_n adjacent to v, and let T_{v_i} be the subtree of $T_n - v$ that contains v_i. If m_i is the number of vertices in T_{v_i}, then we can assume that $m_1 \geq m_2 \geq \cdots m_t > 1$ for some $t \geq 0$, and $m_i = 1$ for $i > t$. Clearly $[k-1-t, k-1] \subseteq s_k(T_n)$, since for any $j \leq t$ one can select j vertices (one from each of the first T_{v_i} $(1 \leq i \leq j)$) that are not adjacent to v and a star with $k-1-j$ edges centered at v that contains no vertices in any of the T_{v_i}. If $m = m_1 + m_2 + \cdots + m_t \geq k$, then by selecting appropriate subtrees of some of the T_{v_i}, $1 \leq i \leq t$, and by selecting an appropriate number of vertices of the v_j from the remaining T_{v_j} that do not contain the chosen subtrees, one can insure that $[0, k-t-1] \subseteq s_k(T_n)$. This, along with the fact that $[k-1-t, k-1] \subseteq s_k(T_n)$, implies that $[0, k-1] = s_k(T_n)$, so we assume $m < k$.

The previous observation also implies that $[0, m-t] \subseteq s_k(T_n)$. Let α be the independence number of the forest $F = (T_{v_1} - v_1) \cup (T_{v_2} - v_2) \cup \cdots \cup (T_{v_t} - v_t)$. If $\alpha \geq k$, then just as above, we have $[0, k-1] \subseteq s_k(T_n)$, since we can select an independent set of $j < k$ vertices and add an appropriate star centered at v to get a set of k vertices that induces a graph with $k - 1 - j$ edges. Thus, we assume $\alpha < k$. To complete the proof of Theorem 8, we will verify that $s_k(T_n) = [0, m-t] \cup [k-1-\alpha, k-1]$. It has already been verified that $s_k(T_n) \subseteq [0, m-t] \cup [k-1-\alpha, k-1]$. Let S be a set of k vertices of T_n. If $v \notin S$, then, by definition of T_n, all of the edges of S are in $T_{v_1} \cup T_{v_2} \cup \cdots T_{v_t}$, and so there can be at most $m - t$ edges in the graph induced by S. If $v \in S$, then let β be the number of vertices of S in F. Thus, the graph induced by S contains v, which is the center of a star with at least $k - \beta - 1$ edges. If $\beta \leq \alpha$, then $k - \beta - 1 \geq k - \alpha - 1$, as required. If $\beta \geq \alpha$, then there are at least $\beta - \alpha$ edges induced by the vertices of S in F, since α is the independence number of F. This implies that the number of edges in the graph induced by S is at least $k - \beta - 1 + \beta - \alpha = k - 1 - \alpha$. Note that $\alpha \leq m$ and also $\alpha \geq \lceil m/2 \rceil$. Thus, this completes the proof of the claim and of the Theorem 8. □

An immediate consequence of Theorems 7 and 8 is the following.

THEOREM 9 *If $n \geq \max\{2k-1, 3k-5\}$, then for any tree T_n, there exist integers r and s with $0 \leq \lceil r/2 \rceil \leq s \leq r \leq k-1$ such that $s_k(T_n) = [0, r] \cup [s, k-1]$.*

The characterization of $s_k(T_n)$ from Theorem 9 is not valid for all values of $n \geq k$. For example if $n = k + 1$, then it is easy to determine $s_k(T_n)$. If $(d_1, d_2, \cdots, d_{k+1})$ is the degree sequence of T_{k+1}, then $s_k(T_{k+1}) = \{k - 1 - d_1, k - 1 - d_2, \cdots, k - 1 - d_{k+1}\}$, since any subgraph with k vertices is determined by deleting single vertices. Thus, this family of examples does not satisfy the conclusion of Theorem 9. The condition that $n \geq \max\{2k-1, 3k-5\}$ is not sharp, but it is not clear what the sharp condition should be. The tree T_{2k-4} that contains two adjacent vertices of degree $k - 2$ has $s_k(T_{2k-4}) = \{0, 2, 3, \cdots, k-3, k-1\}$, so $n \geq 2k-3$ is certainly needed to insure the spectrum is the union of at most 2 intervals. It could be $n \geq 2k-1$ is sufficient. Certainly, $n \geq 2k-1$ is needed to insure that $0 \in s_k(T_n)$.

Let t_k denote the number of different k-spectra of trees. There are 2^k subsets of $\{0, 1, 2, \cdots, k-1\}$, but by Theorem 9 less than $\binom{k}{2}$ of these subsets are the k-spectrum of a large order tree T_n. However, this is not true when n is small. For example if $n = k + 1$, it has already been noted that the k-spectrum of the tree T_n is determined by the degree sequence (actually the degree set) of the tree. For any set $S \subseteq \{3, 4, \cdots, \lfloor \sqrt{2k+2} \rfloor - 2\}$, there is a tree with degree set precisely $S \cup \{1, 2\}$, and so we have $2^{\lfloor \sqrt{2k+2} \rfloor - 2} < t_k \leq 2^k$. Such a tree can easily be constructed by attaching stars of appropriate degrees to a path of appropriate length. Since the number of integer partitions of k is bounded above by $2^{\sqrt{ck}}$ (for an appropriate constant c) (see [3]), the lower bound of the previous inequality has the correct order of magnitude for the number of sets that are the k-spectrum of a tree of order $k+1$. This may not be true in general, as the upper and lower bounds for t_k given above are of different orders of magnitude.

For a positive integer r, let $k = 2^r$ and let $n = 2^r + r = k + \log_2 k$. Let T_n be the tree formed from a $K_{1,r}$ by attaching a star with an additional 2^{j-1} edges at the j^{th} endvertex of the $K_{1,r}$. Thus, T_n has order $n = 2^r + r$ with vertices of degree $2, 4, \cdots, 2^{r-1}, r$ and k vertices of degree 1. The k-spectrum of T_n is tree complete. To see this, determine the binary representation of any integer m $(0 \leq m \leq k-1)$. If the j^{th} term of this binary representation is 0, then delete the center of the j^{th} star; otherwise, delete an endvertex of the j^{th} star. This will leave a graph with k vertices and precisely m edges. Moreover, trees with very few less vertices than this cannot have a k-spectrum that is tree complete, as the following result indicates.

THEOREM 10 *For $\epsilon < 1/3$ and k sufficiently large, any tree T_n of order $n < k + \epsilon \log_2 k$ cannot have a k-spectrum that is tree complete.*

PROOF: Assume the k-spectrum of T_n is tree complete. Since $0 \in s_k(T_n)$, there is an independent set A of k vertices in T_n. Let B be the remaining vertices in T_n. The sum of the degrees of the vertices in B is at least k, since $\delta(T_n) \geq 1$, and so the sum of the degrees of the vertices in A is less than $k + 2\epsilon \log_2 k$. This implies that T_n has at least $k - 2\epsilon \log_2 k$ vertices of degree 1, which we will denote by C. Let D denote the remaining vertices, and so $|D| \leq 3\epsilon \log_2 k$.

Any subset of $V(T_n)$ with k vertices is determined by deleting at most $\epsilon \log_2 k$ vertices, say a set C' from C and a set D' from D. The set C' can be selected in at most $2^{3\epsilon \log_2 k} = k^{3\epsilon}$ ways. However, after the set C has been deleted, only the number of vertices of D' that are now of degree 0 in this resulting graph is important as far as the k-spectrum is concerned. Thus, D' can be chosen in at most $\epsilon \log_2 k$ ways (i.e. maximum number of vertices of degree 0 in D'). Therefore, the maximum number of terms in the k-spectrum of T_n is at most $k^{3\epsilon} \epsilon \log_2 k < k$, if $\epsilon < 1/3$ and k is sufficiently large. This completes the proof of Theorem 10. □

Several questions are left unanswered. It would be interesting to determine the order of magnitude of t_k; in particular, is $t_k = 2^{o(k)}$? Determination of the smallest integer $n^* = n(k)$ such that if $n \geq n^*$, then $s_k(T_n)$ is the union of two intervals is also of interest.

References

[1] R. Faudree, R. Gould, M. Jacobson, J. Lehel, and L. Lesniak, *Graph Spectra*, manuscript.

[2] R. L. Graham, B. L. Rothschild, and J. H. Spencer, **Ramsey Theory**, John Wiley Inc., New York, (1980).

[3] M. Hall, **Combinatorial Theory**, John Wiley Inc., New York, (1967).

[4] R. P. Ramsey, *On a Problem of Formal Logic*, **Proc. Lond. Math. Soc.**, 48, (1930), 264-286.