Graphs with three distinct distance eigenvalues

Yuke Zhang and Huiqiu Lin
School of Mathematics, East China University of Science and Technology,
Shanghai 200237, P.R. China

Abstract

In this paper, some special distance spectral properties of graphs are considered. Concretely, we recursively construct an infinite family of trees with distance eigenvalue $-1$, and determine all $\{C_3, C_4\}$-free connected graphs with three distinct distance eigenvalues of which the smallest one is equal to $-3$, which partially answers a problem posed by Koolen, Hayat and Iqbal [Linear Algebra Appl. 505 (2016) 97–108]. Furthermore, we characterize all trees with three distinct distance eigenvalues.

Keywords: Distance eigenvalues, distinct eigenvalues

2000 Mathematics Subject Classification: 05C50

1 Introduction

Let $G = (V(G), E(G))$ be a connected graph with vertex set $V(G) = \{v_1, v_2, \ldots, v_n\}$ and edge set $E(G)$. The distance between $v_i$ and $v_j$, denoted by $d_{i,j}$, is the length of a shortest path from $v_i$ to $v_j$. The distance matrix of $G$, denoted by $D(G)$, is the $n \times n$ real symmetric matrix whose $(i, j)$-entry is $d_{i,j}$. The multiset $\{\lambda_1^{m_1}(D(G)), \ldots, \lambda_k^{m_k}(D(G))\}$ is called the distance spectrum of $G$, and denoted by $\text{Spec}(D(G))$, where $\lambda_1(D(G)) > \cdots > \lambda_k(D(G))$ are all distinct eigenvalues of $D(G)$, and $m_i$ is the multiplicity of $\lambda_i(D(G))$.

The original study of distance eigenvalues was launched around the trees and can be traced back to the paper of Graham and Pollack [9] in which they obtained a very interesting and excellent result that the determinant of the distance matrix of a tree depends only on the number of vertices, and not on the structure of the tree.

Theorem 1.1 (9). Let $T$ be a tree with $n$ vertices and $D(T)$ be its distance matrix. Then

$$\det(D(T)) = (-1)^{n-1}(n-1)2^{n-2}.$$
from Merris [25], which stated a relation between the distance eigenvalues and Laplacain eigenvalues of a tree.

**Theorem 1.2** ([25]). Let \( G \) be a tree of order \( n \). Let \( \lambda_1(D(G)) \geq \cdots \geq \lambda_n(D(G)) \) be the eigenvalues of \( D(G) \) and let \( \mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq 0 \) be the Laplacain eigenvalues of \( G \).

Then

\[
0 > \frac{-2}{\mu_1} > \lambda_2(D(G)) \geq \frac{-2}{\mu_2} \geq \cdots \geq \lambda_{n-1}(D(G)) \geq \frac{-2}{\mu_{n-1}} \geq \lambda_n(D(G)).
\]

Since then, the study of distance eigenvalues of a graph has received much more attention. These two theorems proved by Graham and Pollack [9] and Merris [25] are instructive and also play important roles in our follow-up proofs. For more results on the distance matrix and its spectral properties, we refer the reader to two surveys [1, 21].

It is well-known that the distance spectrum of any graph contains at least two distinct eigenvalues since the diagonal entries of the distance matrix are all 0. So another problem concerning graph eigenvalues is characterizing graphs with some distinct distance eigenvalues. This problem can be considered into two types. One is determining graphs with some prescribed distance eigenvalues. The other is considering graphs with some distinct distance eigenvalues but without given values.

For the former, Lu, Huang and Huang [23] determined all graphs whose distance matrices have exactly two eigenvalues (counting multiplicity) different from \(-1\) and \(-3\) and some other related results see [13, 27]. Besides, the results in [24] and [22] stated that the multiplicities of distance eigenvalues \(-1\) and \(-2\) in a threshold graph (a graph contains no induced \( C_4, P_4 \) or \( 2K_2 \)) and a cograph (a graph contains no induced \( P_4 \)), respectively. In general connected graphs, Li and Meng [18] characterized the graphs with certain multiplicity of the distance eigenvalue \(-2\). We focus on the following problem posed in [16].

**Problem 1** (Problem 6.2 in [16]). Determine the connected graphs with three distinct distance eigenvalues \( \lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G)) \) such that \( \lambda_3(D(G)) = -3 \).

Let \( \mathcal{F} \) be a family of graphs. We say a graph \( G \) is \( \mathcal{F} \)-free if it does not contain any \( F \in \mathcal{F} \) as a subgraph. In this paper, we consider Problem 1 in \( \{C_3, C_4\} \)-free connected graphs.
Theorem 1.3. Let $G$ be a \{\!\{C_3, C_4\}\!\} -free connected graph. Then $G$ contains three distinct distance eigenvalues $\lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G))$ such that $\lambda_3(D(G)) = -3$ if and only if $G$ is the Petersen graph.

For the latter, Indulal [14] showed that the complete graph is the only graph that contains exactly two distinct distance eigenvalues. Lin, Hong, Wang and Shu [20] found two types of graphs with exactly three distinct distance eigenvalues. Besides, Koolen, Hayat and Iqbal [16] showed that the graphs with exactly three distinct distance eigenvalues fall into four classes. Based on these four classes [16], we determine all trees with three distinct distance eigenvalues as an application of Theorem 1.1 and Theorem 1.2 in this paper.

Theorem 1.4. Any tree contains exactly three distinct distance eigenvalues if and only if it is a star.

The distance eigenvalues are closely linked to the structure of a graph. We consider the structure of a tree containing some specific value as its distance eigenvalue. For all $n$-vertex trees, it is known [2, 10] that the number of the Laplacian eigenvalues distributed in $[2, n]$ is at most $\lfloor \frac{n}{2} \rfloor$. We obtain an analog that the number of the distance eigenvalues distributed in $[-1, 0)$ is at most $\lfloor \frac{n}{2} \rfloor$ by Theorem 1.1. So it is significant to characterize the trees with $-1$ as their distance eigenvalues. In [17] and [23], authors showed that distance eigenvalue $-1$ can determine some special structure of a graph, respectively. However, these results cease to be effective in trees.

In this paper, we structure a series of trees with $-1$ as their distance eigenvalues, which also provide a sufficient condition for trees to have $-1$ as their distance eigenvalues. Denote by $T(4n + 2)$ the set of all trees with $4n + 2$ vertices obtained by adding an edge joining an endvertex of $P_4$ to a vertex of a tree in $T(4n - 2)$ (see Fig. 1). Note that the choices of the tree $T \in T(4n - 2)$ and joining vertex $v \in T$ are arbitrary. Clearly, $T(2) = \{P_2\}$ and $T(6) = \{P_6\}$.

![Figure 1: The tree $T \in T(4n + 2)$ where $T' \in T(4n - 2)$ and $v \in T'$.](image)

Theorem 1.5. For any $n \in \mathbb{N}$ and tree $T \in T(4n + 2)$, there always exists $-1$ as an eigenvalue of $D(T)$.
While it should be emphasized that it is not the unique structure to characterize trees with distance eigenvalues $-1$. We find a tree whose distance spectrum contains $-1$ but does not belong to $T(4n + 2)$. Let $S_{a,b}$ be the double star obtained by attaching $a$ and $b$ pendant vertices to the two end vertices of $K_2$, respectively.

**Remark 1.** Let $T$ be a tree with diameter at most three. The value $-1$ is an eigenvalue of $D(T)$ if and only if $T \cong P_2$ or $T \cong S_{2,2}$.

## 2 Preliminaries

We introduce some useful tools in this section.

Hermitian matrices have real eigenvalues. The Cauchy interlace theorem states that the eigenvalues of a Hermitian matrix $A$ of order $n$ are interlaced with those of any principal submatrix.

**Lemma 2.1 (Cauchy Interlace Theorem).** Let $A$ be a Hermitian matrix of order $n$, and let $B$ be a principal submatrix of $A$ of order $m$. If $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ lists the eigenvalues of $A$ and $\mu_1(B) \geq \mu_2(B) \geq \cdots \geq \mu_m(B)$ the eigenvalues of $B$, then

$$\lambda_{n-m+i}(A) \leq \mu_i(B) \leq \lambda_i(A) \quad \text{for } i = 1, \ldots, m.$$ 

Now we give a technical lemma which will be used to calculate distance eigenvalues. Let $M$ be a real $n \times n$ matrix, and let $N = \{1, 2, \ldots, n\}$. Given a partition $\Pi : N = N_1 \cup N_2 \cup \cdots \cup N_k$, the matrix $M$ can be correspondingly partitioned as

$$M = \begin{pmatrix}
M_{1,1} & M_{1,2} & \cdots & M_{1,k} \\
M_{2,1} & M_{2,2} & \cdots & M_{2,k} \\
\vdots & \vdots & \ddots & \vdots \\
M_{k,1} & M_{k,2} & \cdots & M_{k,k}
\end{pmatrix}.$$ 

The quotient matrix of $M$ with respect to $\Pi$ is defined as the $k \times k$ matrix $B_\Pi = (b_{i,j})$ where $b_{i,j}$ is the average value of all row sums of $M_{i,j}$. The partition $\Pi$ is called equitable if each block $M_{i,j}$ of $M$ has constant row sum $b_{i,j}$. Also, we say that the quotient matrix $B_\Pi$ is equitable if $\Pi$ is an equitable partition of $M$.

**Lemma 2.2.** Let $M$ be a real symmetric matrix, and let $B_\Pi$ be an equitable quotient matrix of $M$. Then the eigenvalues of $B_\Pi$ are also eigenvalues of $M$. Furthermore, if $M$ is nonnegative and irreducible, then $\lambda_1(M) = \lambda_1(B_\Pi)$. 

4
3 Graphs with three distinct distance eigenvalues

Koolen, Hayat and Iqbal [16] demonstrated that the graphs with exactly three distinct distance eigenvalues fall into four classes.

**Theorem 3.1** ([16]). Let $G$ be an $n$-vertex connected graph with exactly three distinct distance eigenvalues $\lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G))$, with respective multiplicities $m_1 = 1, m_2, m_3$. Then one of the following holds.

(i) $G$ is complete bipartite;
(ii) $G$ is regular complete multipartite;
(iii) $n$ is odd and $\lambda_1(D(G)) = c\left(\frac{n-1}{2}\right)$ with $3 \leq c \in \mathbb{Z}$, and $m_2 = m_3 \geq 2$;
(iv) $\lambda_1(D(G)), \lambda_2(D(G)), \lambda_3(D(G)) \in \mathbb{Z}$, $\lambda_3(D(G)) \geq 0$ and $\lambda_3(D(G)) \leq -3$.

Based on this useful result, we respectively characterize $\{C_3, C_4\}$-free graphs with three distinct distance eigenvalues and $\lambda_3(D(G)) = -3$ (see Problem 1) and all trees with three distinct distance eigenvalues in this section. We firstly discuss Problem 1 in $\{C_3, C_4\}$-free graphs.

**Lemma 3.1** ([16, 19]). Let $G$ be a connected graph. Then $\lambda_n(D(G)) \leq -d$ where $d$ is the diameter of $G$ and the equality holds if and only if $G$ is a complete multipartite graph.

A Moore graph is a regular graph of diameter $d$ and girth $2d+1$. Consider a $k$-regular Moore graph $G$ of diameter $d = 2$. The unique graphs for $d = 2$ exist for $k = 2, 3, 7$ and possibly for $k = 57$ (which is undecided), but for no other degree.

**Lemma 3.2.** ([10, 15]) Let $G$ be a $k$-regular graph on $n$ vertices with diameter at most 2 and adjacency spectrum $\lambda_1 = k, \lambda_2, \lambda_3, \ldots, \lambda_n$. Then the distance spectrum of $G$ is $2n - 2 - k, -(2 + \lambda_2), -(2 + \lambda_3), \ldots, -(2 + \lambda_n)$.

**Lemma 3.3.** ([11]) A Moore graph with diameter $d = 2$ and degree $k$ contains exactly three distinct adjacency eigenvalues $k, (-1 + \sqrt{4k-3})/2, (-1 - \sqrt{4k-3})/2$.

**Proof of Theorem 1.3** Let $G$ be a $\{C_3, C_4\}$-free connected graph with diameter $d$ and three distinct distance eigenvalues $\lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G))$ such that $\lambda_3(D(G)) = -3$. By Lemma 3.1, we get $-3 = \lambda_3(D(G)) \leq -d$, which implies $d \leq 3$. Note that $d \neq 3$ since the equality condition of Lemma 3.1. So we have the diameter $d \leq 2$. If $d = 1$, then the graph $G$ is complete and $Spec(D(G)) = \{n-1, [-1]^{n-1}\}$, a contradiction.
Thus, the diameter of $G$ is 2. Besides, it is obvious that the girth of $G$ is at most 5, which yields that the girth of $G$ is 5 since $G$ is $\{C_3, C_4\}$-free. So $G$ is a Moore graph with diameter $d = 2$. Referring to Lemmas 3.2 and 3.3 we have $-3 = -(2 + \frac{1 + \sqrt{4k - 3}}{2})$, implying $k = 3$. Thus, $G$ is the Petersen graph.

If $G$ is the Petersen graph, then $\text{Spec}(D(G)) = \{15, [0]^4, [-3]^5\}$. The proof is complete. 

Remark 2. By Theorem 3.1 if $G$ contains three distinct distance eigenvalues $\lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G))$ such that $\lambda_3(D(G)) = -3$, then $\lambda_1(D(G))$ and $\lambda_2(D(G))$ are integral.

Remark 3. Consider $C_3$-free connected graphs in Problem 1. Let $G$ be a $C_3$-free connected graph with three distinct distance eigenvalues $\lambda_1(D(G)) > \lambda_2(D(G)) > \lambda_3(D(G))$ such that $\lambda_3(D(G)) = -3$. Referring to the proof of Theorem 1.3, we only need to discuss one additional case that the girth of $G$ is 4. Since the graph $G$ is connected and $C_3$-free, then $G$ contains $F_1$ or $F_2$ (see Fig. 2) as an induced subgraph. Let $B_{F_1} = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 & 2 \\ 1 & 2 & 2 & 2 & 0 \end{pmatrix}$ and $B_{F_2} = \begin{pmatrix} 0 & 1 & 2 & 1 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 & 1 \\ 2 & 1 & 2 & 1 & 0 \end{pmatrix}$. Note that $B_{F_1}$ or $B_{F_2}$ is a principal submatrix of $D(G)$. Then $\lambda_2(D(G)) \geq \lambda_2(B_{F_1})$ or $\lambda_2(D(G)) \geq \lambda_2(B_{F_2})$ by Lemma 2.1. We obtain $\lambda_2(B_{F_1}) = 0.0841$ and $\lambda_2(B_{F_2}) = 0.3542$ by a simple calculation. Hence, $\lambda_2(D(G)) \geq 1$ as $\lambda_2(D(G))$ is an integer.

![Figure 2: The structure of induced subgraphs $F_1$ and $F_2$.](image)

It is the position to characterize all trees with three distinct distance eigenvalues. Before the proof, we show some useful lemmas. The Laplacian matrix of a graph $G$ is denoted by $L(G) = \Delta(G) - A(G)$, where $\Delta(G) = \text{diag}(d_1, \ldots, d_n)$, and $A(G)$ is the adjacency matrix of $G$. In Theorem 1.2, Merris obtained a relation between the distance eigenvalues and Laplacian eigenvalues of a tree. This relation plays an important role in the proof of Theorem 1.4.
Lemma 3.4 \cite{3}. Let $G$ be a connected graph with diameter $d$. Then $G$ has at least $d + 1$ distinct Laplacian eigenvalues.

An immediate consequence of Theorem \cite{4} is as follows:

Lemma 3.5 \cite{4}. Let $T$ be a tree with $n$ vertices where $n \geq 2$. Then $D(T)$ has 1 positive and $n - 1$ negative eigenvalues.

Proof of Theorem \cite{4}. If $T$ is a star with $n$ vertices, then we have $Spec(D(T)) = \{n - 2 + \sqrt{n^2 - 3n + 3}, n - 2 - \sqrt{n^2 - 3n + 3}, [-2]^{n-2}\}$ by a simple calculation and the condition is obviously sufficient.

Next suppose $T$ is an $n$-vertex tree with exactly three distinct distance eigenvalues $\lambda_1(D(T)) > \lambda_2(D(T)) > \lambda_3(D(T))$, with respective multiplicities $m_1 = 1, m_2, m_3$. By Theorem \cite{5} $T$ is a star if (i) holds. Thus, we assume that the tree $T$ is not a star which implies that $T$ is neither complete bipartite nor regular complete multipartite. Besides, we notice that $\lambda_1(D(T)) > 0 > \lambda_2(D(T)) > \lambda_3(D(T))$ by Lemma \cite{5}. According to the above statements, we rule out the possibility of (i), (ii) and (iv) in Theorem \cite{5} So we only need to consider (iii) in Theorem \cite{5} in the following argument.

Let $n = 2q + 1$ where $q \geq 2$ since $T \cong P_3$ is a star when $q = 1$. Then we have $\lambda_1(D(T)) = cq$ with $3 \leq c \in \mathbb{Z}$, and $m_2 = m_3 = q$. Note that $det(D(T)) = \lambda_1(D(T))\lambda_2(D(T))^{m_2}\lambda_3(D(T))^{m_3} = (-1)^{n-1}(n-1)2^{n-2}$ by Theorem \cite{4} So we obtain

$$(-1)^{2q} \cdot 2q \cdot 2^{q-1} = cq \cdot \lambda_2(D(T))^{q} \cdot \lambda_3(D(T))^{q},$$

which implies

$$4^q = c(\lambda_2(D(T))\lambda_3(D(T)))^q. \tag{1}$$

It is clear that $4 \mid \lambda_2(D(T))\lambda_3(D(T))$ and $\lambda_2(D(T))\lambda_3(D(T))$ equals to 1 or 2, otherwise $\lambda_2(D(T))\lambda_3(D(T)) = 4$ and $c = 1$, contradicting to $3 \leq c \in \mathbb{Z}$. Since the trace of the distance matrix is zero, the sum of distance eigenvalues is zero, and hence we obtain

$$\lambda_1(D(T)) + q\lambda_2(D(T)) + q\lambda_3(D(T)) = 0. \tag{2}$$

If $\lambda_2(D(T))\lambda_3(D(T)) = 1$, then we get $c = 4^q$, $\lambda_1(D(T)) = q \cdot 4^q$. Therefore, we can multiply (1) and (2), obtaining

$$\lambda_2(D(T)) = -4^q + \sqrt{16^q - 4} \quad \text{and} \quad \lambda_3(D(T)) = -4^q - \sqrt{16^q - 4}. \tag{3}$$

7
If \( \lambda_2(D(T))\lambda_3(D(T)) = 2 \), then we have \( c = 2^q \), \( \lambda_1(D(T)) = q \cdot 2^q \). Rearranging equations (1) and (2), we get
\[
\begin{align*}
\lambda_2(D(T)) &= \frac{-2q + \sqrt{4q - 8}}{2} \\
\lambda_3(D(T)) &= \frac{-2q - \sqrt{4q - 8}}{2}
\end{align*}
\] (4)

Armed with the values of \( \lambda_2(D(T)) \) and \( \lambda_3(D(T)) \), it is obvious that \(-2 \notin \text{Spec}(D(T))\).

By Theorem 1.2, \( T \) has at most five distinct Laplacain eigenvalues and the diameter of \( T \) is at most four in view of Lemma 3.4. Hence, the analysis is partitioned in the following three cases according to the diameter \( d \) of the tree \( T \). Firstly suppose \( d \leq 2 \), then \( T \) is a star, a contradiction. Next we consider \( d = 3 \). Note that \(-2 \notin \text{Spec}(D(T))\) and \( T \in T_3 = \{ S_{s+1,t+1} \mid s \in \mathbb{N} \text{ and } t \in \mathbb{N} \} \). Then we conclude that \( T \cong S_{1,1} \cong P_3 \) referring to Lemma 1.2 and \( \text{Spec}(D(P_3)) = \{ 2 + \sqrt{10}, \sqrt{2} - 2, 2 - \sqrt{10}, -2 - \sqrt{2} \} \), contradicting the hypothesis. Finally, we suppose \( d = 4 \). All trees with diameter 4 are shown in Fig. 3. Since \(-2 \notin \text{Spec}(D(T))\), together with Lemma 1.2 implies that \( T \) is either \( T_4^1 \) (\( s = s' = 0, p = k_1 = \cdots = k_t = 1 \)) or \( T_4^2 \) (\( s = s' = p = 0, k_1 = \cdots = k_t = 1 \)). Considering the order of \( T \) is odd, so \( T \cong T_4^2 \) and from now on we shall calculate the distance spectrum of \( T_4^2 \).

Let \( V_i = \{ v \in V(T_4^2) \mid \text{deg}(v) = i \} \). Clearly, \( |V_1| = q, |V_2| = q, |V_4| = 1 \) and \( V(T_4^2) = V_1 \cup V_2 \cup V_4 \) is an equitable partition of \( D(T_4^2) \). So the corresponding quotient matrix is
\[
B(T_1^2) = \begin{bmatrix}
0 & q & 2q \\
1 & 2(q-1) & 3(q-1) + 1 \\
2 & 3(q-1) + 1 & 4(q-1)
\end{bmatrix},
\]

the characteristic polynomial of the matrix \(B(T_1^2)\) is

\[
g(\lambda) = \lambda^3 - (6q - 6)\lambda^2 - (q^2 + 9q - 4)\lambda - 4q.
\]

Since the partition is equitable, then by Lemma 2.2, the characteristic polynomial \(g(\lambda)\) can be expressed as one of the following three forms:

**Form 1.** \(g(\lambda) = (\lambda - \lambda_1(D(T)))(\lambda - \lambda_2(D(T)))(\lambda - \lambda_3(D(T))).\)

We obtain that \(\lambda_1(D(T))\lambda_2(D(T))\lambda_3(D(T)) = 4q\), which implies that \(4q = q \cdot 4^q\) or \(4q = 2q \cdot 2^q\) when \(q \geq 2\), a contradiction.

**Form 2.** \(g(\lambda) = (\lambda - \lambda_1(D(T)))(\lambda - \lambda_2(D(T)))^2.\)

Then we get that \(h_1(q) = \lambda_1(D(T)) + 2\lambda_2(D(T)) = 6q - 6\), which yields that \(\lambda_1(D(T)) + 2\lambda_2(D(T)) - 6q + 6 = 0\). Substituting (3) and \(q \geq 2\), we have

\[
h_1(q) = q \cdot 4^q - 4^q + \sqrt{16q - 4} - 6q + 6
\]

\[
\geq 4^q - 6q + 6
\]

\[
\geq 4^2 - 6 \cdot 2 + 6 = 10 > 0.
\]

This contradicts the equation \(\lambda_1(D(T)) + 2\lambda_2(D(T)) - 6q + 6 = 0\). Next, we consider (4) and get \(h_2(q) = q \cdot 2^q - 2^q + \sqrt{4q - 8} - 6q + 6 = 0\). By a simple calculation, we get \(h_2(2) = 2\sqrt{2} - 2 \neq 0\). So we just need to consider \(q \geq 3\). Then we have

\[
h_2(q) = q \cdot 2^q - 2^q + \sqrt{4q - 8} - 6q + 6
\]

\[
\geq 3 \cdot 2^q - 6q + 6
\]

\[
= 2 \cdot 2^q - 6q + 6
\]

\[
\geq 2 \cdot 2^3 - 6 \cdot 3 + 6 = 4 > 0.
\]

This also contradicts the equation \(\lambda_1(D(T)) + 2\lambda_2(D(T)) - 6q + 6 = 0\).

**Form 3.** \(g(\lambda) = (\lambda - \lambda_1(D(T)))(\lambda - \lambda_3(D(T)))^2.\)

The specific steps are the same as above. We need to prove that there exist integers \(q \geq 2\) to make the equation \(\lambda_1(D(T)) + 2\lambda_3(D(T)) - 6q + 6 = 0\) hold, that is, \(l_1(q) = \)
and induction on \( n \) eigenvalues and the sum of components of each corresponding eigenvector is 0. We apply eigenvalue of \( D \) assertion is trivial. So suppose \( n \neq 1 \).

At the beginning of this section, we characterize a series of trees containing \( -1 \) as their distance eigenvalues. Now we shall prove that any tree in \( T(4n + 2) \) contains \( -1 \) as its distance eigenvalue.

**Proof of Theorem 1.5.** We asset that all trees in \( T(4n + 2) \) have \( -1 \) as their distance eigenvalues and the sum of components of each corresponding eigenvector is 0. We apply induction on \( n \) to prove this assertion. Clearly, \( T(2) = \{P_2\} \) and \( T(6) = \{P_6\} \) and the assertion is trivial. So suppose \( n = k + 1 \) and the assertion holds for \( n \leq k \). For any tree \( T_k \in T(4k + 2) \), we have a vector \( x_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_{4k+1}^{(k)}, x_{4k+2}^{(k)})^T \) to satisfy \( D(T_k)x_k = -x_k \) and \( \sum_{j=1}^{4k+2} x_j^{(k)} = 0 \). Let \( P_4 = w_1w_2w_3w_4 \) be a path on four vertices. We assume that the tree \( T_{k+1} \) is obtained by attaching an edge between the endvertex \( w_1 \) in \( P_4 \) and a vertex in \( T_k \), called \( v_p \), without loss of generality. It is clear that \( T_{k+1} \) belongs to \( T(4k + 6) \). Let \( x_{k+1} = (x_k^T, y_1, y_2, y_3, y_4)^T = (x_1^{(k)}, x_2^{(k)}, \ldots, x_{4k+1}^{(k)}, x_{4k+2}^{(k)}, x_{4k+3}^{(k)}, x_{4k+4}^{(k)}, x_{4k+5}^{(k)}, x_{4k+6}^{(k)})^T \) where \( y_i \) is corresponding to the vertex \( w_i \) of the pendent path \( P_4 \). To prove that \( -1 \) is an eigenvalue of \( D(T_{k+1}) \), it suffices to show that \( D(T_{k+1})x_{k+1} = -x_{k+1} \). By a simple
calculation, we obtain

\[
D(T_{k+1}) = \begin{bmatrix}
D(T_k) & B \\
0 & 1 & 2 & 3 \\
1 & 0 & 1 & 2 \\
2 & 1 & 0 & 1 \\
3 & 2 & 1 & 0
\end{bmatrix}
\]

where

\[
B = \begin{bmatrix}
d_{1,p} + 1 & d_{1,p} + 2 & d_{1,p} + 3 & d_{1,p} + 4 \\
d_{2,p} + 1 & d_{2,p} + 2 & d_{2,p} + 3 & d_{2,p} + 4 \\
\vdots & \vdots & \vdots & \vdots \\
d_{4k+1,p} + 1 & d_{4k+1,p} + 2 & d_{4k+1,p} + 3 & d_{4k+1,p} + 4 \\
d_{4k+2,p} + 1 & d_{4k+2,p} + 2 & d_{4k+2,p} + 3 & d_{4k+2,p} + 4
\end{bmatrix}
\]

This implies that

\[
-x_i^{(k)} = -x_i^{(k)} + 0 \\
= \sum_{j=1}^{4k+2} d_{i,j}x_j^{(k)} + (d_{i,p} + 1)y_1 + (d_{i,p} + 2)y_2 + (d_{i,p} + 3)y_3 + (d_{i,p} + 4)y_4
\]

for \(i = 1, 2, \ldots, 4k + 1, 4k + 2\). Besides, there always holds

\[
-y_1 = -x_p^{(k)} = (-x_p^{(k)} + 0) - x_p^{(k)} - 2x_p^{(k)} + 3x_p^{(k)} \\
= \sum_{j=1}^{4k+2} (d_{j,p} + 1)x_j^{(k)} + y_2 + 2y_3 + 3y_4,
\]

and

\[
-y_2 = x_p^{(k)} = (-x_p^{(k)} + 2 \times 0) + x_p^{(k)} - x_p^{(k)} + 2x_p^{(k)} \\
= \sum_{j=1}^{4k+2} (d_{j,p} + 2)x_j^{(k)} + y_1 + y_3 + 2y_4,
\]

and

\[
-y_3 = x_p^{(k)} = (-x_p^{(k)} + 3 \times 0) + 2x_p^{(k)} - x_p^{(k)} + x_p^{(k)} \\
= \sum_{j=1}^{4k+2} (d_{j,p} + 3)x_j^{(k)} + 2y_1 + y_2 + y_4,
\]

and

\[
-y_4 = x_p^{(k)} = (-x_p^{(k)} + 4 \times 0) + 3x_p^{(k)} - 2x_p^{(k)} - x_p^{(k)} \\
= \sum_{j=1}^{4k+2} (d_{j,p} + 4)x_j^{(k)} + 3y_1 + 2y_2 + y_3.
\]
Thus we conclude that $x_{k+1} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_{4k+1}^{(k)}, x_{4k+2}^{(k)}, -x_p^{(k)}, -x_p^{(k)}, x_p^{(k)})^T$ is the corresponding eigenvector of the eigenvalue $-1$ of $D(T_{k+1})$ and the sum of components of $x_{k+1}$ is 0 as well. This proof is complete because of the arbitrary choices of the tree $T_k \in T(4k + 2)$ and the joining vertex $v_p$ of $T_k$.

Considering the multiplicity of an eigenvalue, we pose the following problem naturally:

For any $n \in \mathbb{N}$ and tree $T \in T(4n + 2)$, is it right that the multiplicity of $-1$ as an eigenvalue of $D(T)$ is one? We start with the more easily addressable issues and give a positive answer to the above problem for $P_{4n+2}$.

**Lemma 4.1.** Let $P_n$ be a path on $n > 2$ vertices. The largest eigenvalue $d_1$ and other eigenvalues $d_i$ of its distance matrix $D(P_n)$ are as follows.

1. $d_1 = 1/(\cosh \theta - 1)$, where $\theta$ is the positive solution of $\tanh(\theta/2) \tanh(n\theta/2) = 1/n$.
2. $d_i = 1/(\cos \theta - 1)$, where (I) $\theta$ is one of the $[(n-1)/2]$ solutions of $\tan(\theta/2) \tan(n\theta/2) = -1/n$ in the interval $(0, \pi)$, or (II) $\theta = (2m - 1)\pi/n$ for $m = 1, \ldots, [n/2]$.

**Theorem 4.1.** Let $P_k$ be a path on $k \geq 2$ vertices. The value $-1$ is an eigenvalue of $D(P_k)$ if and only if $k = 4n + 2$ where $n \in \mathbb{N}$. Moreover, the multiplicity of $-1$ as an eigenvalue of its distance matrix is one.

**Proof.** It follows from Theorem 1.5 that $P_{4n+2} \in T(4n + 2)$ and $-1$ is an eigenvalue of $D(P_{4n+2})$. Conversely, suppose $-1$ is an eigenvalue of $D(P_k)$. Then we have $-1 = 1/(\cos \theta - 1)$ by Lemma 4.1 and it implies that $\theta = \pi/2$ if $\theta$ is in the interval $(0, \pi)$. Considering (I) and (II) in Lemma 4.1 (2), it is obvious that $\tan(\pi/4) \tan(n\pi/4) \neq -1/k$ and $\theta = \pi/2$ when $k = 4m - 2$. We conclude that $-1$ is an eigenvalue of $D(P_k)$ when $k = 4n + 2$ for $n \in \mathbb{N}$.

We show that the multiplicity of $-1$ as an eigenvalue of $D(P_{4n+2})$ is one in the following. The assertion is obviously right for $n = 0$ so we only consider the case for $n \geq 1$. Let $x = (x_1, x_2, \ldots, x_{4n+1}, x_{4n+2})^T \neq 0$ be an arbitrary eigenvector of the eigenvalue $-1$ of $D(P_{4n+2})$. In order to prove this theorem, we need to show that each component of vector $x$ is not zero, that is, $x_i \neq 0$ for any integer $i \in [1, 4n + 2]$. By a simple calculation on $D(P_{4n+2})x = -x$, we have

$$-x_q = \sum_{j=1}^{4n+2} |q - j|x_j$$

$$\quad$$

12
for \( q = 1, 2, \ldots, 4n + 1, 4n + 2 \). Then we have the following equations by (5):

\[
-(x_{2n+1} + x_{2n+2}) = \sum_{q=1}^{2n} (4n + 3 - 2q)(x_q + x_{4n+3-q}) + (x_{2n+1} + x_{2n+2})
\]

and

\[
-(x_{2n} + x_{2n+3}) = \sum_{q=1}^{2n} (4n + 3 - 2q)(x_q + x_{4n+3-q}) + 3(x_{2n+1} + x_{2n+2})
\]

\[
= -2(x_{2n+1} + x_{2n+2}) + 3(x_{2n+1} + x_{2n+2})
\]

\[
= x_{2n+1} + x_{2n+2},
\]

and

\[
-(x_{2n-1} + x_{2n+4}) = \sum_{q=1}^{2n-1} (4n + 3 - 2q)(x_q + x_{4n+3-q}) + 5[(x_{2n} + x_{2n+3}) + (x_{2n+1} + x_{2n+2})]
\]

\[
= -4(x_{2n} + x_{2n+3}) - 3(x_{2n+1} + x_{2n+2}) + 5[(x_{2n} + x_{2n+3}) + (x_{2n+1} + x_{2n+2})]
\]

\[
= (x_{2n} + x_{2n+3}) + 2(x_{2n+1} + x_{2n+2})
\]

\[
= x_{2n+1} + x_{2n+2}.
\]

By parity of reasoning, it is clear that \( -(x_q + x_{4n+3-q}) = x_{2n+1} + x_{2n+2} \), which implies that

\[
-2(x_{2n+1} + x_{2n+2}) = \left(\sum_{q=1}^{2n} (4n + 3 - 2q)(x_q + x_{4n+3-q})\right) = -(4n^2 + 4n)(x_{2n+1} + x_{2n+2}).
\]

Hence

\[
x_q + x_{4n+3-q} = 0, \text{ for } q = 1, 2, \ldots, 2n, 2n + 1.
\]

Furthermore, we also obtain the following equations by (5):

\[
-x_{2n+1} + x_{2n+2} = -\sum_{q=1}^{2n} x_q - x_{2n+1} + x_{2n+2} + \sum_{q=2n+3}^{4n+2} x_q \Rightarrow \sum_{q=1}^{2n} x_q = \sum_{q=2n+3}^{4n+2} x_q,
\]

\[
-x_{2n} + x_{2n+3} = -3\sum_{q=1}^{2n} x_q - x_{2n+1} + x_{2n+2} + 3\sum_{q=2n+3}^{4n+2} x_q = -x_{2n+1} + x_{2n+2},
\]

\[
-x_{2n-1} + x_{2n+4} = -5\sum_{q=1}^{2n-1} x_q - 3x_{2n} - x_{2n+1} + x_{2n+2} + 3x_{2n+3} + 5\sum_{q=2n+4}^{4n+2} x_q
\]

\[
= 5(x_{2n} - x_{2n+3}) - 4(x_{2n+1} - x_{2n+2}) = x_{2n+1} - x_{2n+2}.
\]
and

\[-x_{2n-2} + x_{2n+5} = -7 \sum_{q=1}^{2n-2} x_q - 5x_{2n-1} - 3x_{2n} - x_{2n+1} + x_{2n+2} + 3x_{2n+3} + 5x_{2n+4} + 7 \sum_{q=2n+5}^{4n+2} x_q \]

\[-x_{2n} - x_{2n-1} + x_{2n+3} + x_{2n+4} + x_{2n+1} - x_{2n+2} \]

By parity of reasoning, it is clear that for \( q = 1, 2, \ldots, 2n, 2n + 1 \)

\[
\begin{cases}
-x_q + x_{4n+3-q} = -x_{2n+1} + x_{2n+2} & \text{when } 2n + 1 - q \equiv 0, 1 \pmod{4}, \\
-x_q + x_{4n+3-q} = x_{2n+1} - x_{2n+2} & \text{when } 2n + 1 - q \equiv 2, 3 \pmod{4}.
\end{cases}
\]

(6) and (7) give \( |x_i| = |x_j| \) for any \( i, j = 1, \ldots, 4n + 2 \). Since the vector \( x \) is nonzero, we have \( |x_i| = a \in \mathbb{R} \setminus \{0\} \) for any integer \( i \in [1, 4n + 2] \).

Does the set \( T(4n + 2) \) cover all the trees whose distance spectra contain \(-1\)? The answer is negative. In the following, we characterize all connected trees of diameter at most three with distance eigenvalues that are equal to \(-1\). Besides, we show a special tree whose distance spectrum contains \(-1\) but does not belong to \( T(4n + 2) \).

A pendent vertex of a tree is a vertex of degree one. A pendent neighbor is a vertex adjacent to a pendent vertex.

**Lemma 4.2.** (4) Let \( T \) be a tree with \( n_1 \) pendent vertices and \( n'_1 \) pendent neighbors. Then among the distance eigenvalues of \( T \), \(-2\) as an eigenvalue occurs with multiplicity at least \( n_1 - n'_1 \).

**Proof of Remark** 1 We can easily get by computer that \( \text{Spec}(D(P_2)) = \{1, -1\} \) and \( \text{Spec}(D(S_{2,2})) = \{10, -1, -\frac{5}{2} - \frac{\sqrt{17}}{2}, -\frac{5}{2} + \frac{\sqrt{17}}{2}, [-2]^2 \} \).

Denote by \( T_d \) the set of trees of diameter \( d \). We distinguish the following two cases:

**Case 1.** \( d \leq 2 \).

The result is trivial for the case \( d = 1 \) and \( \text{Spec}(D(K_2)) = \{1, -1\} \). For \( d = 2 \), the graphs in \( T_2 \) are stars and \( T_2 = \{K_{1,t+2} | t \in \mathbb{N}\} \). Let \( J_{t+2} \) be the matrix of order \( t + 2 \) with all entries equal to 1. Then the distance matrix of \( K_{1,t+2} \) is

\[
D = \begin{bmatrix}
2(J - I)_{t+2} & 1 \\
1 & \ddots & 1 \\
1 & 1 & 0
\end{bmatrix}
\]
and the distance characteristic polynomial of $K_{1,t+2}$ is

$$
det(\lambda I - D) = (\lambda + 2)^{t+1} \begin{vmatrix} \lambda - (2t + 2) & -1 \\ -(t + 2) & \lambda \end{vmatrix} = (\lambda + 2)^{t+1} [\lambda^2 - (2t + 2)\lambda - (t + 2)].
$$

If $-1$ is an eigenvalue of $D(K_{1,t+2})$, then $(-1)^2 - (2t + 2)(-1) - (t + 2) = 0$, which implies that $t = -1$, contradicting with $t \geq 0$. Thus, only graph $P_2$ contains $-1$ as its distance eigenvalue among all trees with diameter $d \leq 2$.

**Case 2.** $d = 3$.

Note that $T_3 = \{S_{s+1,t+1} | s \in \mathbb{N} \text{ and } t \in \mathbb{N}\}$. For any tree $T \in T_3$, there is an equitable partition of $D(T)$ and the corresponding quotient matrix is

$$
B(T) = \begin{bmatrix}
2s & 1 & 2 & 3(t + 1) \\
s + 1 & 0 & 1 & 2(t + 1) \\
2(s + 1) & 1 & 0 & t + 1 \\
3(s + 1) & 2 & 1 & 2t
\end{bmatrix},
$$

the characteristic polynomial of the matrix $B(T)$ is

$$
f(\lambda) = \lambda^4 - (2t + 2s)\lambda^3 - (5st + 14s + 14t + 20)\lambda^2 - (4st + 16s + 16t + 32)\lambda - 4s - 4t - 12.
$$

Since the partition is equitable, then by Lemma 2.2, each eigenvalue of the quotient matrix $B(T)$ is also an eigenvalue of the distance matrix $D(T)$. By Lemma 4.2, we obtain that the value $-2$ is an eigenvalue of $D(T)$ for any $T \in T_3$ and its multiplicity is at least $s + t$.

Clearly, $|V(T)| = s + t + 4$ and $-2$ is not a root of the characteristic polynomial $f(\lambda)$ because of $f(-2) = -12st - 12s - 12t - 12 < 0$. So the multiplicity of the eigenvalue $-2$ is $s + t$. Besides, the other four distance eigenvalues of $T$ are the four roots of the equation $f(\lambda) = 0$. If $-1$ is an eigenvalue of $D(T)$, then

$$f(-1) = -st + 1 = 0.$$

It follows that the tree with diameter 3 contains $-1$ as its distance eigenvalue when $t = s = 1$, saying $T \cong S_{2,2}$. □

We find that the trees obtained by successively adding an edge joining an endvertex of $P_4$ to a vertex in $S_{2,2}$ do not contain $-1$ as their distance eigenvalues. Referring to the proof of Theorem 1.5, one reason is that the sum of components of $\mathbf{x}$ is not 0 where $D(S_{2,2})\mathbf{x} = -\mathbf{x}$. 

15
References

[1] M. Aouchiche, P. Hansen, Distance spectra of graphs: A survey, *Linear Algebra Appl.* **458** (2014) 301–386.

[2] R. Braga, V. Rodrigues, V. Trevisan, On the distribution of Laplacian eigenvalues of trees, *Discrete Math.* **313** (**21**) (2013) 2382–2389.

[3] A. Brouwer, W. Haemers, Spectra of Graphs, Springer, Berlin, 2011.

[4] K. Collins, Factoring distance matrix polynomials, *Discrete Math.* **122** (1993) 103–112.

[5] M. Edelberg, M. Garey, R. Graham, On the distance matrix of a tree, *Discrete Math.* **14** (1976) 23–39.

[6] R. Elzinga, D. Gregory, K. Vander Meulen, Addressing the Petersen graph, *Discrete Math.* **286** (2004) 241–244.

[7] C. Godsil, G. Royle, Algebraic Graph Theory, Graduate Texts in Mathematics, 207, Springer-Verlag, New York, 2001.

[8] R. Graham, L. Lovász, Distance matrix polynomials of trees, *Adv. Math.* **29** (1978) 60–88.

[9] R. Graham, H. Pollack, On the addressing problem for loop switching, *Bell Syst. Tech. J.* **50** (1971) 2495–2519.

[10] J. Guo, The $k$th Laplacian eigenvalue of a tree, *J. Graph Theory* **54** (**1**) (2007) 51–57.

[11] A. Hoffman, R. Singleton, On Moore graphs with diameters 2 and 3, *IBM J. Res. Develop.* **4** (1960) 497–504.

[12] H. Hosoya, M. Murakami, M. Gotoh, Distance polynomial and characterization of a graph, *Natur. Sci. Rep. Ochanomizu Univ.* **24** (1973) 27–34.

[13] X. Huang, Q. Huang, L. Lu, Graphs with at most three distance eigenvalues different from $-1$ and $-2$, *Graphs Combin.* **34** (2018) 395–414.
[14] G. Indulal, Sharp bounds on the distance spectral radius and the distance energy of graphs, *Linear Algebra Appl.* 430 (2009) 106–113.

[15] G. Indulal, I. Gutman, A. Vijayakumar, On distance energy of graphs, *MATCH Commun. Math. Comput. Chem.* 60 (2008) 461–472.

[16] J. Koolen, S. Hayat, Q. Iqbal, Hypercubes are determined by their distance spectra, *Linear Algebra Appl.* 505 (2016) 97–108.

[17] P. Křivka, N. Trinajstić, On the distance polynomial of a graph, *Apl. Mat.* 28 (1983) 357–363.

[18] D. Li, J. Meng, The multiplicity of $−2$ as an eigenvalue of the distance matrix of graphs, *Filomat* 32 (2018), no. 10, 3681–3688.

[19] H. Lin, On the least distance eigenvalue and its applications on the distance spread, *Discrete Math.* 338(6) (2015) 868–874.

[20] H. Lin, Y. Hong, J. Wang, J. Shu, On the distance spectrum of graphs, *Linear Algebra Appl.* 439 (2013) 166–169.

[21] H. Lin, J. Shu, J. Xue, Y. Zhang, A survey on distance spectra of graphs, *Adv. Math. (China)* 50(1) (2021) 29–76.

[22] Z. Lou, H. Lin, Distance eigenvalues of a cograph and their multiplicities, *Linear Algebra Appl.* 608 (2021) 1–12.

[23] L. Lu, Q. Huang, X. Huang, The graphs with exactly two distance eigenvalues different from $−1$ and $−3$, *J. Algebraic Comb.* 45 (2017) 629–647.

[24] L. Lu, Q. Huang, Z. Lou, On the distance spectra of threshold graphs, *Linear Algebra Appl.* 553 (2018) 223–237.

[25] R. Merris, The distance spectrum of a tree, *J. Graph Theory* 14 (1990) 365–369.

[26] S. Ruzieh, D. Powers, The distance spectrum of the path $P_n$ and the first distance eigenvector of connected graphs, *Linear Multilinear Algebra* 28 (1990) 75–81.

[27] X. Zhang, Graphs with few distinct $D$-eigenvalues determined by their $D$-spectra, *Linear Algebra Appl.* 628 (2021) 42–55.