5-2021

The Minimum Number of Multiplicity 1 Eigenvalues among Real Symmetric Matrices whose Graph is a Tree

Wenxuan Ding

The Minimum Number of Multiplicity 1 Eigenvalues among Real Symmetric Matrices whose Graph is a Tree

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Mathematics from College of William and Mary

by

Wenxuan Ding

Accepted for Honors
(Honors, no Honors)

Charles R. Johnson
Charles Johnson, Advisor

Eric Swartz

Carl Moody

Williamsburg, VA
May 6, 2021

The Minimum Number of Multiplicity 1 Eigenvalues among Real Symmetric Matrices whose Graph is a Tree

Wenxuan "Olivia" Ding
Advisor: Dr. Charles R. Johnson

Department of Mathematics,
College of William & Mary,
Williamsburg, VA 23187-8795
Email: wding@email.wm.edu

April 2021

Abstract

For a tree T, $U(T)$ denotes the minimum number of eigenvalues of multiplicity 1 among all real symmetric matrices, whose graph is T. It is known that $U(T) \geqslant 2$. A tree is linear if all its vertices of degree at least 3 lie on a single induced path, and k-linear if there are k of these high degree vertices. If T' is a linear tree resulting from the addition of 1 vertex to T, we show that $|U(T') - U(T)| \leqslant 1$. We also determine the exact set of possible values of $U(T) - U(T')$, depending upon the manner in which the vertex is added to T to get T'. These results are then used to give a new bound for $U(T)$, the diameter bound, and to improve an existing bound, $2 + D_2(T)$. Moreover, a new classification of nonlinear trees based on cores is introduced and used to study $U(T)$ for nonlinear trees. Lastly, some results about trees with $U(T) = 2$, the path cover number, and Parter vertices are presented.

Acknowledgement

I must express my deepest gratitude to my advisor Professor Charles R. Johnson, who has instilled in me a love for research and mentored me with tremendous patience. I would also like to thank Professor Eric Swartz and Professor Carl Moody for agreeing to serve on my committee and providing me with valuable feedback.

Contents

Chapter 1

Introduction

We consider real symmetric n-by-n matrices and the partitions of n that are the lists of multiplicities for their eigenvalues. The multiplicities in a multiplicity list can be summarized in two ways: *ordered* and *unordered*. An ordered multiplicity list is based on the numerical values of the underlying eigenvalues; an unordered multiplicity list is the same list, listed in nonincreasing order of the multiplicities. For example, if $n = 15$, and the 15 eigenvalues of the matrix are

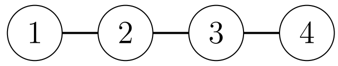
$$-3, -1, -1, 2, 4, 4, 4, 5, 5, 6, 8, 8, 10, 11, 25,$$

the ordered multiplicity list is $(1, 2, 1, 3, 2, 1, 2, 1, 1, 1)$, while the unordered multiplicity list is $(3, 2, 2, 2, 1, 1, 1, 1, 1, 1)$.

Graphs have been used to describe the arrangement of zero and nonzero off-diagonal entries in a matrix. In particular, for a given undirected graph G, let $\mathcal{S}(G)$ denote the collection of real symmetric matrices whose graph is G; $A = (a_{ij}) \in \mathcal{S}(G)$ if and only if $a_{ij}, a_{ji} \neq 0$ if and only if $\{i, j\}$ is an edge of G. (Otherwise, both are 0.) No restriction is placed by G on the diagonal entries of $A \in \mathcal{S}(G)$, except that they must be real. The eigenvalues of A are real, due to the symmetry requirement.

For instance, if G is a path on 4 vertices, then elements in $\mathcal{S}(G)$ are in the form of A, with $a_{ij} \in \mathbb{R}$ and $a_{12}, a_{23}, a_{34} \neq 0$.

$G =$



$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{12} & a_{22} & a_{23} & 0 \\ 0 & a_{23} & a_{33} & a_{34} \\ 0 & 0 & a_{34} & a_{44} \end{bmatrix}$$

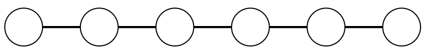
We are primarily interested in the case in which $G = T$ is a tree. A *tree* is a minimally connected undirected graph, i.e., a connected acyclic graph on n vertices with $n - 1$ edges.

In this case, each Hermitian matrix with graph T is unitarily similar to a matrix in $\mathcal{S}(T)$ [JS], so that there is no difference in possible multiplicities if the matrix is Hermitian. Let $\mathcal{L}(T)$ be the collection of all ordered multiplicity lists occurring among matrices in $\mathcal{S}(T)$, the *catalog* for T. ($\mathcal{L}_u(T)$ is the collection of all unordered multiplicity lists.) $U(T)$ denotes the minimum number of 1's among the lists in $\mathcal{L}(T)$. It is known that $U(T)$ is at least 2, corresponding to the smallest and largest eigenvalues [JS]. One of the various proofs of this fact is based on the Perron-Frobenius theorem for nonnegative matrices. However, $U(T)$ can be much greater than 2. For example, $U(P_n) = n$ for the path P_n on n vertices because eigenvalues of irreducible real symmetric tridiagonal matrices are distinct.

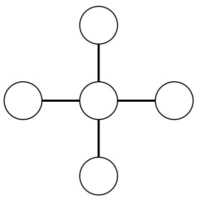
There has been much interest in and progress on determining $\mathcal{L}(T)$ for each tree T. The maximum multiplicity, $M(T)$, is the *path cover number* $P(T)$ (the minimum number of disjoint paths needed to cover every vertex in T), and the minimum number of distinct eigenvalues is at least the *diameter* $d(T)$ (the length of the longest induced path of T, measured by the number of vertices)[JL-D]. Similarly precise information about $U(T)$ would further narrow the possibilities for the catalog $\mathcal{L}(T)$. But little is known about $U(T)$. So, in this thesis, we take up the study of $U(T)$.

Prior work and literature have determined $U(T)$ for special classes of trees such as paths, (simple) stars (i.e., a tree on n vertices having a vertex of degree $n-1$), and generalized stars (i.e., a tree with at most one vertex of degree at least 3) [JS].

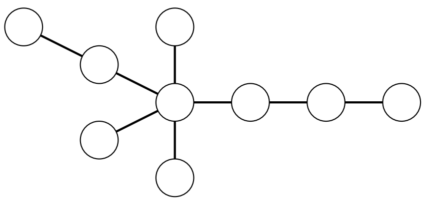
1. A path on n vertices: $U(P_n) = n$.



2. A star on n vertices: $U(S_n) = 2$.



3. A generalized star (g-star): The definition is to be introduced later and the formula presented in Theorem 2.5.



Nevertheless, for general trees with possibly more complicated structures, determining $U(T)$ is not straightforward. Indeed, $U(T)$ seems to be a "residual" quantity after we maximize the sum of multiplicities of multiple eigenvalues (i.e., eigenvalues with multiplicity greater than 1). That is, supposing $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_k$ are multiple eigenvalues of $A \in \mathcal{S}(T)$ with multiplicities $m_1, m_2, m_3, \cdots, m_k$,

$$U(T) = \min_{A \in \mathcal{S}(T)} \{n - \sum_{i=1}^{k} m_i\}.$$

So, $U(T)$ is attained when $\sum_{i=1}^{k} m_i$ is maximized. This somewhat explains why calculating $U(T)$ is difficult as max $\sum_{i=1}^{k} m_i$ is also not determined in a single way.

Remarkably, the authors of *Eigenvalues, Multiplicities and Graphs* [JS] have assembled and made available a database of $\mathcal{L}(T)$ and $U(T)$ for 987 trees on fewer than 13 vertices. We used this database extensively throughout this work to verify conjectures and justify examples.

The organization of this thesis is as follows. Chapter 1 has described the problem and explained why it is of our interest. Chapter 2 cites the necessary background bridging multiplicity theory and graph theory. Chapter 3 delves deeply into linear trees; topological characterizations, incremental changes in $U(T)$ due to vertex addition or deletion, new bounds for $U(T)$, and formulas for certain classes of linear trees are discussed. Chapter 4 considers nonlinear trees from a new perspective. Classification of nonlinear trees into cores based on the diameter facilitates the study of $U(T)$. Chapter 5 concludes with a few separate but worthwhile results about $U(T)$, in which other ideas are presented and some outstanding conjectures are resolved.

Chapter 2

Background

2.1 General background

Throughout, we employ standard submatrix notation. Let A be an n-by-n matrix. If $\alpha \subset \{1, \ldots, n\}$ is an index set, then $A[\alpha]$ is the principal submatrix of A in the rows and columns indexed by α, and $A(\alpha) := A[\{1, \ldots, n\} \backslash \alpha]$. In the case when $\alpha = \{v\}$, we abbreviate $A(\{v\})$ to $A(v)$. Observe that the graph $T[\alpha]$ is the subgraph of T induced by the vertices corresponding to α. Then $A[\alpha] \in \mathcal{S}(T[\alpha])$, and we often think of the matrix and graph interchangeably. We write $m_A(\lambda)$ for the multiplicity of the eigenvalue λ in the matrix A (subscript A sometimes omitted). A classical and fundamental theorem for this study is the interlacing inequalities for Hermitian (real symmetric) matrices.

Theorem 2.1. *(The interlacing inequalities) Let A be an n-by-n Hermitian matrix and let $i \in \{1, \ldots, n\}$. Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of A and let $\mu_1, \ldots, \mu_{n-1}$ be the eigenvalues of $A(i)$. Then*

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \cdots \leq \mu_{n-1} \leq \lambda_n.$$

An immediate consequence of Theorem 2.1 is that for any λ, $\left| m_A(\lambda) - m_{A(i)}(\lambda) \right| \leq 1$. That is, the multiplicity of an eigenvalue of a real symmetric matrix changes by at most 1 if a principal submatrix of size one smaller is extracted.

Let A be a real symmetric matrix whose graph is a tree T. The removal of a vertex v from a tree T corresponds to the removal of a row and column with the same index v from A. When v is deleted from T, a forest of several components $T_1, \ldots, T_{deg_v}$ remains. The corresponding matrix is $A(v) = A\left[T_1\right] \oplus \cdots \oplus A\left[T_{\mathrm{deg}v}\right]$. The Parter-Wiener, etc. theorem is one of the most important tools in the study of eigenvalues, multiplicities, and graphs. The most general form of the theorem is given in [JL-DS03].

Theorem 2.2. *(Parter-Weiner, etc.) Let T be a tree and A a matrix in $\mathcal{S}(T)$. Let $\sigma(A)$ denote the spectrum of A. Suppose that there is a vertex v of T and a real number λ such that $\lambda \in \sigma(A) \cap \sigma(A(v))$. Then*

(1) there is a vertex u of T such that $m_{A(u)}(\lambda) = m_A(\lambda) + 1$;

(2) if $m_A(\lambda) \geq 2$, then $\lambda \in \sigma(A) \cap \sigma(A(v))$ is automatically satisfied and u may be chosen so that $\deg_T(u) \geq 3$ and so that there are at least three components T_1, T_2, and T_3 of $T\backslash u$ such that $m_{A[T_i]}(\lambda) \geq 1, i = 1, 2, 3$; and

(3) if $m_A(\lambda) \geq 1$, then u may be chosen so that $\deg_T(u) \geq 2$ and so that there are two components T_1 and T_2 of $T\backslash u$ such that $m_{A[T_i]}(\lambda) \geq 1, i = 1, 2$

We call a vertex v meeting the requirement in the above theorem a *Parter vertex* of T for λ (or a Parter for λ, for short). In other words, v is a Parter for λ if $m_{A(v)}(\lambda) = m_A(\lambda) + 1$ and λ is an eigenvalue of the submatrices corresponding to least two of the connected components of $T - v$.

A *high degree vertex* (HDV) is a vertex of degree at least 3. A *generalized star (g-star)* is a tree with at most one HDV; moreover, the HDV (or a degree 2 vertex if there is no HDV) is called the *central vertex* of the g-star. A g-star consists of a number of paths (*arms*) hanging from the central vertex.

We often need the concept of an *upward multiplicity list* of a g-star T. Let T be a g-star and call its central vertex v. Let $A \in \mathcal{S}(T)$. We say that an eigenvalue λ of A is an *upward eigenvalue* of A at v if $m_{A(v)}(\lambda) = m_A(\lambda) + 1$. We call the multiplicity of λ in A an *upward multiplicity* of A at v, denoted with a hat. Notice that it is possible that $m_A(\lambda) = 0$ (i.e., λ is not an eigenvalue of A), and the upward multiplicity of λ of A at v is denoted $\hat{0}$. If $q = (q_1, \ldots, q_r)$ is the ordered multiplicity list of A, then we define the *upward multiplicity list* of A at v, which we denote by $\hat{q}$, the list with the same entries as q but in which any upward multiplicity of A at v, q_i is marked as $\hat{q}_i$ in $\hat{q}$. The *complete upward multiplicity list* of A at v is the upward multiplicity list of A augmented with $\hat{0}$'s representing the upward eigenvalues of multiplicity 0. The set of upward multiplicity lists at v that occur among the matrices of $\mathcal{S}(T)$ is called the *upward catalog* for T at v, denoted $\widehat{\mathcal{L}}_v(T)$. Furthermore, we can distinguish *ordered* and *unordered* upward catalogs, which are essentially equivalent except for being comprised of ordered or unordered multiplicity lists. For convenience, we assume $\widehat{\mathcal{L}}_v(T)$ consists of ordered multiplicity lists.

For a g-star, we have the following lemma about its upward eigenvalues.

Lemma 2.3. [JL-DS] *Let T be a g-star with central vertex v. If $A \in \mathcal{S}(T)$ and λ is an eigenvalue of $A(v)$, then $m_{A(v)}(\lambda) = m_A(\lambda) + 1$, i.e., λ is an upward eigenvalue of A.*

From this lemma, we can say that the upward eigenvalues of A (including those with multiplicity 0) are exactly the eigenvalues of $A(v)$ on the pendent arms. A complete upward multiplicity list for a g-star has the form $(1, \hat{q_1}, 1, \hat{q_2}, 1, \cdots, \hat{q_r}, 1)$, in which r upward multiplicities are "interlaced" by $r + 1$ nonupward 1's. Hence every upward multiplicity list of T begins and ends with a nonupward 1. Consecutive nonupward 1's appear when the upward multiplicity in-between is $\hat{0}$ and omitted. In fact, we can generate all the upward multiplicity lists of a g-star.

Theorem 2.4. [JL-DS] *Let T be a g-star on n vertices with central vertex v of degree k and arm lengths $l_1 \geq \cdots \geq l_k$. Then $\hat{q} = (q_1, \cdots, q_r) \in \hat{\mathcal{L}}_v(T)$ if and only if $\hat{q}$ satisfies the following conditions:*

1. *q_i is a nonnegative integer, $1 \leq i \leq r$, and $\sum_{i=1}^{r} q_i = n$;*
2. *if q_i is an upward multiplicity in $\hat{q}$, then $1 < i < r$ and neither q_{i-1} nor q_{i+1} is an upward multiplicity in $\hat{q}$; and*
3. *$(q_{i_1} + 1, \ldots, q_{i_h} + 1)_e \preceq (l_1, \ldots, l_k)^*$, in which $q_{i_1} \geq \cdots \geq q_{i_h}$ are the upward multiplicities of $\hat{q}$ greater than 1, and $(q_{i_1} + 1, \ldots, q_{i_h} + 1)_e$ means that the vector is augmented with e ones so that $e + \sum_{m=1}^{h}(q_{i_m} + 1) = \sum_{m=1}^{k} l_m$. Note that $u \preceq v$ means that v majorizes u, which is defined to be the majorization of partitions of integers.*

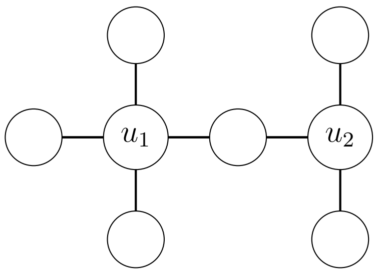
Moreover, an explicit formula for $U(T)$ for g-stars is known.

Theorem 2.5. [JL-DS] *Let T be a g-star with arm lengths $l_1 \geq \cdots \geq l_a$. Then*

$$U(T) = \max\{1 + l_1, 2d(T) - n\}.$$

Lastly, we introduce a construction technique for multiplicity lists: the method of eigenvalue assignments (to subtrees for which possible spectra are known). This construction technique of eigenvalue assignments is an informal visualization and involves Parter vertices and several coincidences of eigenvalues among various subtrees. A *realization* of an assignment verifies the existence of a desired multiplicity list. We use the example in [JS] to illustrate how to use this technique.

Example 2.6. Consider the following tree.



The multiplicity list $(3, 2, 1, 1, 1)$ occurs for T because of the realizable assignment in which u_1 is Parter for α (with $m(\alpha) = 3$) and u_2 is Parter for β (with $m(\beta) = 2$). To be specific, α appears four times in $T - u_1$: on the three neighbors of u_1 and once on the subtree to the right; β appears three times in $T - u_2$: on the two neighbors of u_2 and once on the subtree to the left. Nevertheless, there is no eigenvalue assignment for the multiplicity list $(2, 2, 2, 1, 1)$, hence $(2, 2, 2, 1, 1) \notin \mathcal{L}(T)$. In particular, one of the two HDV's (u_1 and u_2) would have to be Parter for two of the multiplicity 2 eigenvalues. However, neither has enough branches of sufficient size to assign the two eigenvalues a total of six times.

Definition 2.7 states the formalization of eigenvalue assignment. More details can be found in [JS].

Definition 2.7. Let T be a tree on n vertices and let

$$\left(p_1, p_2, \ldots, p_k, 1^{n-\sum_{i=1}^{k} p_i}\right)$$

be a non-increasing list of positive integers, with $\sum_{i=1}^{k} p_i \leqslant n$. The notation 1^l denotes that the last l entries of the list are 1 . These will be the desired eigenvalue multiplicities. Note that some of the p_i's may be 1 . Then, an *assignment* $\mathcal{A}$ is a collection $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_k\}$ of k collections $\mathcal{A}_i$ of subtrees of T, corresponding to eigenvalues with multiplicities p_i, with the following properties.

1. (Specification of Parter vertices) For each i, there exists a set V_i of vertices of T such that
   (1a) Each subtree in $\mathcal{A}_i$ is a connected component of $T - V_i$.
   (1b) $|\mathcal{A}_i| = p_i + |V_i|$.
   (1c) For each vertex $v \in V_i$, there exists a vertex x adjacent to v such that x is in one of the subtrees in $\mathcal{A}_i$.

2. (No overloading) We require that no subtree S of T is assigned more than $|S|$ eigenvalues; define $c_i(S) = |\mathcal{A}_i \cap \mathcal{P}(S)| - |V_i \cap S|$, the difference between the number of subtrees contained in S and the number of Parter vertices in S for the i^{th} multiplicity. Then we require that $\sum_{i=1}^{k} \max(0, c_i(S)) \leq |S|$ for each $S \in \mathcal{P}(T)$. If this condition is violated at any subtree, then that subtree is said to be *overloaded*.

2.2 Linear trees and the Linear Superposition Principle

A tree is called *linear* if all its HDV's lie on a single induced path of the tree. A linear tree with k HDV's is called *k-linear*, and a 1-linear tree is simply a g-star. Let $L(T_1, s_1, T_2, \ldots, s_k, T_{k+1})$ denote the linear tree resulting from connecting the central vertices of the g-stars $\{T_i\}_{i=1}^{k+1}$ by paths of length $\{s_i\}_{i=1}^{k}$, each measured by its number of vertices (≥ 0). Orienting the paths and the tree horizontally, we define the leftmost and rightmost g-stars, namely T_1 and T_{k+1}, *peripheral g-stars* of T. The path consisting of $\{s_i\}_{i=1}^{k}$ and the longest arm of each peripheral g-star is the *central path* of T. Also, if all the vertices of a linear tree are either vertices on the central path or pendent vertices hanging on the central path, the tree is *depth 1*.

The *Linear Superposition Principle (LSP)* is both necessary and sufficient for generating all the multiplicity lists for a linear tree; in other words, for a linear tree, the set of multiplicity lists generated by the LSP is the set of all its multiplicity lists.

Theorem 2.8. *(LSP)* [JLW, JW] *Let $T_1, \ldots, T_{k+1}$ be g-stars and $s_1, \ldots, s_k$ nonnegative integers. Given $\hat{b}_i$ an upward multiplicity list for T_i (with respect to the central vertex), $i = 1, \ldots, k+1$, and $\hat{c}_j$ a list of s_j nonupward ones, $j = 1, \ldots, k-1$, construct augmented lists $b_i^+, i = 1, \ldots, k$, and $c_j^+, j = 1, \ldots, k$, subject to the following conditions:*

1. *all b_i^+ 's and c_j^+ 's have the same length;*

2. *each b_i^+ and c_j^+ are obtained from its corresponding $\hat{b}_i$ and $\hat{c}_j$ by inserting nonupward 0's;*

3. *for each l, the l^{th} element of the augmented lists, denoted $b_{i,l}^+$ and $c_{j,l}^+$, are not all nonupward 0's;*

4. *for each l, arranging the $b_{i,l}^+$'s and $c_{j,l}^+$'s in the order $b_{1,l}^+, c_{1,l}^+, b_{2,l}^+ c_{2,l}^+, \ldots, b_{k,l}^+$, there is at least one upward multiplicity between any two nonupward ones.*

Then $\sum_{i=1}^{k+1} b_i^+ + \sum_{j=1}^{k} c_j^+$, where the addition is termwise, is a multiplicity list for $L(T_1, s_1, \ldots, s_k, T_{k+1})$ generated by the LSP. For a linear tree $T = L(T_1, s_1, T_2, s_2, \ldots, s_k, T_{k+1})$, $\mathcal{L}(T)$ is equal to the set of all candidate multiplicity lists generated by the LSP for T, as above.

Graphically, we can represent this in tabular form as Table 2.1. The LSP is then equivalent to completing the given table so that

1. b_i^+ is the multiplicity list $\hat{b}_i$ along with some added nonupward zeros;

2. c_i^+ contains s_i nonupward ones and the remaining entries are nonupward zeros;

3. no column has all nonupward zeros;

4. if a column contains two nonupward ones, they are separated by an element with upward multiplicity.

	λ_1	λ_2	$\cdots$	$\cdots$	λ_j
b_1^+					
c_1^+					
b_2^+					
c_2^+					
$\vdots$					
c_k^+					
b_{k+1}^+					
sum	a_1	a_2	$\cdots$	$\cdots$	a_j

Table 2.1: The tabular form of the LSP

Example 2.9. Let $T = L(T_1, 2, T_2)$ and let $\hat{b} = (1, \hat{2}, 1)$ and $\hat{c} = (1, \hat{1}, 1, \hat{1}, 1, \hat{1}, 1)$ be upward multiplicity lists of T_1 and T_2, respectively. The following are two ways superimposing to get ordered multiplicity lists for T:

$$\begin{array}{cccccccc} 0 & 1 & \hat{2} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & \hat{1} & 1 & \hat{1} & 1 & \hat{1} & 1 & 0 \\ \hline 1 & 2 & 3 & 2 & 1 & 2 & 1 & 1 \end{array} \quad \text{and} \quad \begin{array}{cccccccccc} 0 & 0 & 1 & 0 & \hat{2} & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \hat{1} & 0 & 1 & \hat{1} & 1 & \hat{1} & 1 \\ \hline 1 & 1 & 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 \end{array}.$$

However, this next superposition is not valid, since the highlighted column violates the condition **4** in Theorem 2.8 (two consecutive nonupward 1's in a column).

$$\begin{array}{ccccccc} 0 & 1 & \hat{2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & \hat{1} & 1 & \hat{1} & 1 & \hat{1} & 1 \\ \hline 1 & 2 & 3 & 3 & 1 & 2 & 1 \end{array}.$$

Nevertheless, the membership of the multiplicity list $(1, 2, 3, 3, 1, 2, 1)$ in $\mathcal{L}(T)$ is not completely ruled out by this particular invalid superposition. As long as there is some valid superposition resulting in $(1, 2, 3, 3, 1, 2, 1)$, then it is in $\mathcal{L}(T)$, perhaps via superimposing other multiplicity lists of g-stars T_1 and T_2.

2.3 Existing results about $U(T)$

There have been prior efforts in providing bounds for $U(T)$. Here we cite the diameter lower bound (Theorem 2.10) and the $2 + D_2$ upper bound (Theorem 2.11). Recall that the *diameter*, denoted $d(T)$, is the length of the longest induced path of T. And, $D_2(T)$ denotes the number of vertices of degree 2 in T.

Theorem 2.10. [JS] *If T is a tree on n vertices, then $U(T) \geqslant 2d - n$.*

Theorem 2.11. [JW] *For a linear tree T, $U(T) \leqslant 2 + D_2$.*

Chapter 3

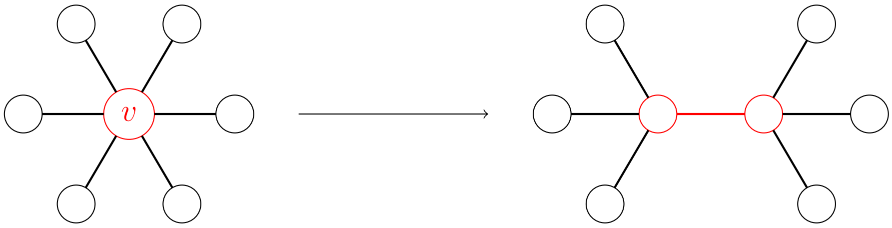
Linear trees

Recall from Section 2.2 that a tree is called linear if all its HDV's lie on a single path. Linear trees have a useful structure that allows decomposition into g-stars and connecting paths. Moreover, it is now known how to generate all multiplicity lists via the *Linear Superposition Principle (LSP)* for linear trees (Theorem 2.8). All trees but one, on no more than 10 vertices, are linear and still by 25, half are linear [JWW]. Nonetheless, direct information about $U(T)$ for linear trees is modest.

This chapter focuses on linear trees. We discuss the topological characterization of linear trees in Section 3.1, which facilitates the understanding of their structure. Then, we use a long Section 3.2 to present results about incremental changes in $U(T)$ due to vertex addition or deletion. Using results in Section 3.2, we give a new upper bound for $U(T)$ for linear trees based on diameter, and improve the previous $2 + D_2$ upper bound in Theorem 2.11 in Section 3.3. Then, in Section 3.4, we look at a specific class of linear trees, 2-linear trees, and determine an explicit formula for its $U(T)$. Finally, we conclude with partial results and a complication in calculating $U(T)$ for k-linear trees in Section 3.5.

3.1 Topological characterization of linear trees

Given a tree, we can add a vertex and obtain a larger tree via either adding a pendent vertex, in which a new edge and a vertex pendent at an existing vertex are added, or *edge subdivision*, in which a new vertex of degree 2 is positioned along an existing edge. Furthermore, another operation that can be performed on trees so that the number of vertices increases by 1 is called *vertex partition*, which splits an existing vertex into two resultant vertices, adjacent to each other, while the vertices that were adjacent to the original vertex are split between the two resultant vertices. An example of vertex partition of vertex v is given as follows.



Notice that these three operations – pendent vertex addition, edge subdivision, and vertex partition – are not mutually exclusive. For example, when we add a pendent vertex at a pendent vertex, it is equivalent to an edge subdivision on the edge adjacent to that pendent vertex. Also, an edge subdivision can be considered as a special case of vertex partition, in which one of the resultant vertices ends up being a degree 2 vertex. The following results characterize linear trees in terms of these operations.

Theorem 3.1. *A tree is linear if and only if the resulting tree from any edge subdivision is linear.*

Proof. Necessity: since edge subdivisions only result in degree 2 vertices, the number and structure of HDV's in a tree are unaffected. Hence, if a tree is linear, i.e., all HDV's lie on a path, the resulting tree from edge subdivisions still has these HDV's lying on a (possibly longer) path. So, this resulting tree is linear. Sufficiency: obviously, given T is a tree, if the resulting tree from any edge subdivision is linear, then T is linear because if it is nonlinear, any resulting tree would be nonlinear. □

Theorem 3.2. *A tree is linear if and only if for each vertex, there is some vertex partition that results in a linear tree.*

Proof. Necessity: suppose a tree T is linear, for any vertex, there is a vertex partition that is equivalent to an edge subdivision. Every edge subdivision results in a linear tree by Theorem 3.1, so necessity is proven. Sufficiency: if T is nonlinear, then for any vertex, no vertex partition leads to a linear tree. It suffices to consider vertex partitions on HDV's because a vertex partition on degree 2 vertices is the same as edge subdivision and that on pendent vertices creates a disconnected graph. Call the resulting tree from vertex partition by T'. Vertex partitions on HDV's could only possibly increase the number of HDV's in T'. In particular, when the degree of the HDV is at least 4, assigning both resultant vertices at least two of the neighbors results in two HDV's in T'; when the degree of the HDV is 3, we have to assign two of the neighbors to one vertex and one neighbor to the other, then the vertex assigned with two neighbors remains an HDV and the only HDV resulting from this vertex partition. Thus, T' still has HDV's that do not lie on a single path, which makes it nonlinear. □

Theorem 3.3. *For a linear tree, of the neighbors of an interior HDV, two are on the central path. A vertex partition of an HDV in a linear tree results in a linear tree if the two successor vertices each get one of the neighbors on the central path. Any vertex partition of a peripheral HDV in a linear tree results in a linear tree.*

Proof. The central path of a linear contains all the HDV's, and it goes through two neighbors of each HDV. Upon a vertex partition of an HDV, if the two successor vertices each get one of the neighbors on the central path, then these two successor vertices are on the central path. Since nothing else in the tree changes, no matter whether the two successor vertices are HDV or degree 2 vertices, the resulting tree is still linear. Moreover, for a peripheral HDV, if the vertex partition results in an additional HDV, it still lies on the central path, so the tree remains linear. □

3.2 Incremental change in $U(T)$ due to vertex addition or deletion

All trees on $n+1$ vertices are simply generated from those on n vertices by appending pendent vertices in all possibly ways. In particular, all linear trees of given diameter may be produced from the path of that diameter. The motivation of this section comes from Theorem 2.1, as the essence of the interlacing inequalities is that the multiplicity of any eigenvalue does not change by more than one upon vertex addition or deletion. We wonder whether the change in $U(T)$ is also bounded. From statistics based upon the database for $\mathcal{L}(T)$ for all trees on fewer than 13 vertices (Table 3.1 and Table 3.2) [In], there are natural conjectures about $U(T')$ vs $U(T)$ when a vertex is added to T to get T' in a certain way.

	-1	0	$+1$
Isolated	0	0	1
Pendent	221	430	554
Degree 2	936	85	0
HDV	226	594	0
Total	1383	1109	555

Table 3.1: Changes in $U(T)$ after adding a pendent vertex, depending on where it is added

-1	0	$+1$
225	490	909

Table 3.2: Changes in $U(T)$ after edge subdivision

Here in this section, our purpose is to prove exactly how $U(T)$ can change when a vertex is added to T, depending upon how the vertex is added – via pendent vertex addition at an HDV, at a degree 2 vertex, and at a pendent vertex, or edge subdivision – and depending upon whether the diameter is increased. The change is proven never to be by more than one, but not all such changes can occur. We determine the exact set of possibilities. Examples that show that all cases not ruled out can occur are given in the Appendix. This investigation is for linear trees, but we suspect the results may apply to all trees. Lastly, we will talk about how $U(T)$ changes upon vertex partition, which is different from pendent vertex addition or edge subdivision.

3.2.1 Any change in $U(T)$ is bounded by 1

Given a tree, we can add a vertex and obtain a larger tree via either adding a pendent vertex or edge subdivision. When a tree is linear and the resulting tree upon vertex addition is also linear, we can use the LSP to prove that the change of $U(T)$ is bounded by 1.

Theorem 3.4. *Let T be a linear tree, and let T' be a linear tree obtained via either adding a pendent vertex, or via edge subdivision in T. Then $|U(T') - U(T)| \leqslant 1$.*

We prove Theorem 3.4 by Lemma 3.5 that $U(T)$ cannot increase by more than 1 and Lemma 3.6 that $U(T)$ cannot decrease by more than 1.

Lemma 3.5. *Let T be a linear tree, and let T' be a linear tree obtained via either adding a pendent vertex, or via edge subdivision in T. Then $U(T') - U(T) \leqslant 1$.*

Proof. Given an arbitrary linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, we may add a vertex in the following six ways while remaining the linearity, as shown in Figure 3.1.

1. Add a pendent vertex to an arm of a g-star T_i for some $1 \leqslant i \leqslant k + 1$;
2. Add a pendent vertex to a central vertex of T_i for some $1 \leqslant i \leqslant k + 1$;
3. Subdivide an edge on an arm of a g-star (in fact, equivalent to the first operation);
4. Subdivide an edge on a connecting path s_i for some $1 \leqslant i \leqslant k$;
5. Add a pendent vertex to a vertex on a connecting path s_i for some $1 \leqslant i \leqslant k$;
6. Add a pendent vertex to a non-pendent vertex on an arm of a peripheral g-star (i.e T_1 or T_{k+1})

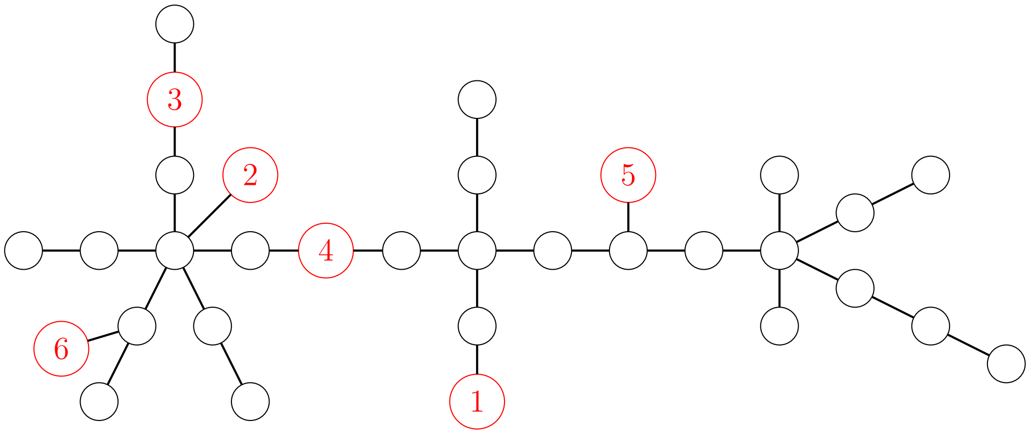


Figure 3.1: Possible ways to add a vertex to a linear tree

Consider a LSP table of T that attains $U(T)$, with upward multiplicity lists b_1^+, c_1^+, $\cdots$, b_k^+, c_k^+, b_{k+1}^+. Adding a vertex to T changes and only changes one of these upward multiplicity lists. As long as we can construct a valid LSP table for T' whose resulting multiplicity list contains no more than $U(T)+1$ 1's, by the sufficiency of the LSP, there is a matrix A such that $U(A) \leqslant U(T)+1$; since $U(T') = \min\{U(A) : A \in \mathcal{S}(T')\}$, then $U(T') \leqslant U(T)+1$ as desired. Now we will examine each of the above operations.

1. Add a pendent vertex to an arm of a g-star T_i for some $1 \leqslant i \leqslant k+1$:
   We may assign the new vertex a distinct eigenvalue from the eigenvalues assigned to arms in T_i, which creates a new nonupward multiplicity 1 eigenvalue. For example, if $b_i = (1, \hat{q}_1, 1, \cdots, 1, \hat{q}_r, 1)$, one possible b_i' could be $(1, \hat{q}_1, 1, \cdots, 1, \hat{q}_r, 1, \hat{0}, 1)$ when the eigenvalue assigned to the new vertex exceeds every other eigenvalue. Keeping the LSP table of T (and hence $U(T)$) the same, we augment it with two more columns at the end for T' as follows.

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_j		λ_{j+1}
b_1^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	0
c_1^+	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\hat{0}$	1
c_k^+ $\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
b_{k+1}^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	0
sum	$\star$	$\star$	$\star$	$\star$	$\cdots$	$\star$	0	1

This is a valid LSP table with $U(T)+1$ multiplicity 1 eigenvalues, so $U(T') \leqslant U(T)+1$.

2. Add a pendent vertex to a central vertex of T_i for some $1 \leqslant i \leqslant k+1$:
   Similarly, we can assign a distinct eigenvalue to this new vertex and construct a LSP table as above, concluding $U(T') \leqslant U(T)+1$.

3. Subdivide an edge on an arm of a g-star:
   Since it is equivalent to adding a pendent vertex to an arm, we skip the discussion.

4. Subdivide an edge on a connecting path s_i for some $1 \leqslant i \leqslant k$:
   When we subdivide an edge on a connecting path s_i, the upward multiplicity list c_i consisting of s_i nonupward 1's becomes c_i' consisting of s_i+1 nonupward 1's. And, keeping the construction of T and hence $U(T)$ the same, we augment it with one more column for T'.

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_j	λ_{j+1}
b_1^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0
c_1^+	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	1
c_k^+ $\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_{k+1}^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0
sum	$\star$	$\star$	$\star$	$\star$	$\cdots$	$\star$	1

   This is a valid LSP table with $U(T)+1$ multiplicity 1 eigenvalues, so $U(T') \leqslant U(T)+1$.

5. Add a pendent vertex to a vertex on a connecting path s_i for some $1 \leqslant i \leqslant k$:
   Now, a connecting path becomes a connecting path, a g-star on 2 vertices, and another connecting path. That is, the upward multiplicity list $c_i = (1, 1, \cdots, 1)$ with s_i 1's and augmenting 0's now becomes three upward multiplicity lists, namely, $(1, 1, \cdots, 1), (1, \hat{0}, 1)$, and $(1, 1, \cdots, 1)$ where the number of 1's totals $s_i + 1$. For example,

	λ_1	λ_2	$\cdots$	$\cdots$	λ_j
b_1^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\vdots$	$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
c_i^+	1	1	1	1	1
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
b_{k+1}^+	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
sum	$\star$	$\star$	$\cdots$	$\star$	$\star$

becomes

	λ_1	λ_2	$\cdots$		λ_h	$\cdots$	λ_{j+1}
b_1^+	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$	$\cdots$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$	$\cdots$
c_i^+	1	1	0	0	0	0	0
$b_i^{\prime +}$	0	0	1	$\hat{0}$	1	0	0
c_{i+1}^+	0	0	0	0	0	1	1
$\vdots$	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$	$\cdots$
b_{k+2}^+	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$	$\cdots$
sum	$\star$	$\star$	$\cdots$	0	1	$\star$	$\star$

As shown above, we may keep the table the same and insert two columns in the middle, which results in an additional 1 in the multiplicity list. This is a valid LSP table with $U(T)+1$ multiplicity 1 eigenvalues, so $U(T') \leqslant U(T)+1$.

6. Add a pendent vertex to a non-pendent vertex on an arm of a peripheral g-star (i.e T_1 or T_{k+1}):
   Notice we can only add a pendent vertex to a non-pendent vertex on an arm of a peripheral g-star (i.e., T_1 or T_{k+1}), otherwise the tree is no longer linear.
   Without loss of generality, we consider b_1^+ for T_1. Since T_1 is a g-star, b_1 is in the form of $(1, \hat{q}_1, 1, \cdots, 1, \hat{q}_r, 1)$. When we add a pendent vertex to a non-pendent vertex on an arm, T_1 becomes T_1' with one fewer arm, T_0, and s_0. Suppose this arm has length l, then to compensate for the loss of this arm, either $\hat{q}_i$ becomes $\hat{q}_i - 1$ for some $\hat{q}_i \geqslant 1$ or $(\hat{0}, 1)$ is removed, and the decrease totals l. Notice that $\hat{q}_i$ cannot decrease by more than 1 since the arm is essentially a path with all eigenvalues with multiplicity 1.
   Now, after the shrinking of T_1 to T_1', we have a g-star T_0' and a connecting path s_0' on $l+1$ vertices in total. We turn one upward multiplicity list b_1 into three upward multiplicity lists b_0', s_0', and b_1' in the table. Since every entry of b_1 can change by at most 1, we can make b_0' and s_0' only consist of 1, $\hat{1}$ and $\hat{0}$ and distribute the 1's (no matter upward or nonupward) to recover b_1. Since the upward entries in b_0' and s_0' need not play a role of separating nonupward 1's in the column, this is manageable. For the extra nonupward 1 resulting from the addition of a vertex, we may make it a separate column, which could produce at most one more multiplicity 1 eigenvalue in the multiplicity list; since everything else is just like the LSP table that attains $U(T)$, $U(T') \leqslant U(T)+1$.

After examining all possible ways of adding a vertex to a linear tree to obtain a new linear tree, we conclude $U(T') \leqslant U(T)+1$, i.e., $U(T)$ does not increase by more than one. □

Lemma 3.6. *Let T be a linear tree, and let T' be a linear tree obtained via either adding a pendent vertex, or via an edge subdivision in T. Then $U(T) - U(T') \leqslant 1$.*

Proof. One might be concerned whether $U(T)$ can decrease by more than 1 when we add a vertex and allow some drastically different forms of LSP tables. In fact, to avoid this trouble, it suffices to show that deleting a vertex from a linear tree does not increase $U(T)$ by more than 1. This is because if deleting a vertex increases $U(T)$ by no more than 1, then the decrease of $U(T)$ by more than 1 via adding a vertex would raise a contradiction as when we delete the added vertex, we cannot restore $U(T)$.

Now, given an arbitrary linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, we may obtain a tree T' with one fewer vertex by either deleting a vertex in some connecting path s_i or deleting a pendent vertex of an arm of some g-star T_i. Suppose we have a LSP table that attains $U(T)$ for T. On the one hand, when a vertex is deleted from some connecting path s_i, then $c_i = (1, 1, \cdots, 1)$ with s_i 1's now becomes a multiplicity list $c_i' = (1, 1, \cdots, 1)$ with $s_i - 1$ 1's. Thus, when we keep everything else in the table the same, we obtain a valid LSP table for T' with only one column affected (i.e., a nonupward 1 is removed). The influence of this column on $U(T)$ is ambiguous: if the resulting multiplicity of the column was 1, then removing the nonupward 1 decreases $U(T)$; if the resulting multiplicity of the column was 2, then removing the nonupward 1 could possibly increase $U(T)$ by 1 by leaving out a new simple eigenvalue; if the resulting multiplicity of the column was greater than 2, then removing the nonupward 1 does not increase $U(T)$ at least. Although the exact effect is not clear, $U(T') \leqslant U(T) + 1$.

On the other hand, when a vertex is deleted from an arm of some g-star T_i of T, consider the upward multiplicity list b_i in a LSP table that attains $U(T)$. The eigenvalue assigned to the vertex was either a nonzero upward eigenvalue (i.e., $\hat{q}_i, \geqslant \hat{1}$) or a upward eigenvalue with multiplicity $\hat{q}_i = \hat{0}$.
In the first case, we can superimpose as T with $b_i' = (1, \hat{q}_1, 1, \cdots, 1, \widehat{q_i - 1}, 1, \cdots, 1)$. This does not violate the LSP. Since only one column is involved, there is a valid LSP table for T' whose resulting multiplicity list has at most one entry different from that of T, which allows us to conclude that $U(T') \leqslant U(T) + 1$.

In the second case, by deleting the vertex with a distinct eigenvalue assigned to it, we essentially change a pair $(\hat{0}, 1)$ in b_i to $(0, 0)$ as demonstrated in the figure .

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_j
b_1^+	$\cdots$	$\cdots$	$\cdots$	$\star$	$\star$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_i^+	$\cdots$	$\cdots$	$\cdots$	$\hat{0}$	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_{k+1}^+	$\cdots$	$\cdots$	$\cdots$	$\star$	$\star$	$\cdots$
sum	a_1	a_2	$\cdots$	$\star$	$\star$	a_j

becomes

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_j
b_1^+	$\cdots$	$\cdots$	$\cdots$	$\star$	$\star$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_i^+	$\cdots$	$\cdots$	$\cdots$	0	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_{k+1}^+	$\cdots$	$\cdots$	$\cdots$	$\star$	$\star$	$\cdots$
sum	a_1	a_2	$\cdots$	$\star$	$\star - 1$	a_j

Now, we can discuss in cases how we can construct a valid LSP table whose resulting multiplicity list has at most $U(T) + 1$ 1's.

1. when the column with the $\hat{0}$ in it has either a trivial upper half above $\hat{0}$ or a trivial lower half below $\hat{0}$ (consisting of only 0's), the above operation does not violate the LSP, so only the column with the nonupward 1's is changed, then it is a valid LSP table for T' with at most one resulting multiplicity different from that of T, which allows us to conclude that $U(T') \leqslant U(T) + 1$.

2. when the column with the $\hat{0}$ in it has a nontrivial upper half above $\hat{0}$ and a nontrivial lower half below $\hat{0}$, and the $\hat{0}$ is the only upward multiplicity separating two nonupward 1's, as follows:

$$\begin{matrix} 1 \\ \hat{0} \\ 1 \end{matrix}$$

we cannot simply change $\hat{0}$ to 0 as indicated above because it violates the LSP's fourth condition.
Notice that if there is some pair like

$$\begin{matrix} 1 & 0 \\ \hat{0} & 1 \\ 1 & 0 \end{matrix} , \quad \begin{matrix} 1 & \hat{1} \\ \hat{0} & 1 \\ 1 & 0 \end{matrix} , \quad \begin{matrix} 1 & 0 \\ \hat{0} & 1 \\ 1 & \hat{1} \end{matrix} , \quad \begin{matrix} 0 & 1 \\ 1 & \hat{0} \\ 0 & 1 \end{matrix} , \quad \begin{matrix} 0 & 1 \\ 1 & \hat{0} \\ \hat{1} & 1 \end{matrix} , \quad \begin{matrix} \hat{1} & 1 \\ 1 & \hat{0} \\ 0 & 1 \end{matrix}$$

we may move one of the nonupward 1's in the column with $\hat{0}$ to the column with the nonupward 1. This does not violate the LSP or change the resulting multiplicity of the column with the nonupward 1. Thus, only the column with $\hat{0}$ might be changed; therefore, $U(T') \leqslant U(T) + 1$.
If none of the above structure appears around b_i, we must have a structure like the following:

$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$
1	$\hat{0}$	1	$\hat{0}$	1	$\hat{0}$	1
$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$

However, given such a structure, neither the first nor the third multiplicity list is complete since the leftmost and rightmost upward $\hat{1}$ must be "interlaced" by nonupward 1's. So, without loss of generality, we have the following structure:

0	1	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	1	0
0	0	1	$\hat{0}$	1	$\hat{0}$	1	$\hat{0}$	1	0	0
1	0	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	1	$\hat{1}$	0	1

We may change the highlighted three columns by shifting the last row without violating the LSP, for example:

1	$\hat{1}$	1
$\hat{0}$	1	0
1	$\hat{1}$	0

becomes

1	$\hat{1}$	1
0	0	0
0	1	$\hat{1}$

In this case, only the leftmost column among the three could change $U(T)$; thus, $U(T') \leqslant U(T) + 1$, completing the discussion.

Again, since we can construct multiplicity lists with at most one more multiplicity 1 eigenvalue, by the sufficiency of the LSP, as the minimum, $U(T') \leqslant U(T) + 1$. □

Like the proof of Lemma 3.5, adding a vertex and deleting a vertex go hand by hand. We can deduce a similar result for vertex deletion.

Corollary 3.7. *Let T be a linear tree, and let T' be a tree obtained by deleting a pendent vertex or by reverse edge subdivision. Then* $|U(T') - U(T)| \leqslant 1$.

Proof. The result follows from Theorem 3.4. If deleting a vertex from a tree changed $U(T)$ by more than 1, when we add this vertex back and recover the tree, $U(T)$ would change by more than 1, which is a contradiction to Theorem 3.4. □

3.2.2 The exact set of possible changes of $U(T)$ when a vertex is added or deleted in a particular way

In fact, when a vertex is added in a particular way, we may give a more accurate description of how $U(T)$ changes. For our purposes, the vertices of a tree are classified into 3 categories based on their degree: a pendent vertex, a degree 2 (*Deg* 2) vertex, or an HDV.

Also, we include the discussion of edge subdivision on an interior connecting path. Before we give further proofs, we display the results in Table 3.3 and Table 3.4. Examples of linear trees such that changes claimed in Table 3.3 actually occur are included in Appendix A.

Type	How the vertex is added		Increases	Stays the same	Decreases
1a	Pendent at an HDV	No Deg 2 vertices before	No	Yes	No
1b		Deg 2 vertices	No	Yes	Yes
2a	Pendent at a Deg 2 vertex		No	Yes	Yes
3a	Pendent at a pendent vertex	The diameter increases	Yes	Yes	Yes
3b		The same diameter	No	Yes	Yes
4	Edge subdivision	Different from 3	Yes	Yes	Yes

Table 3.3: The possible changes of $U(T)$ upon addition of a vertex

Type	How the vertex is deleted		Increases	Stays the same	Decreases
1a	Pendent at an HDV	No Deg 2 vertices after	No	Yes	No
1b		Deg 2 vertices	Yes	Yes	No
2a	Pendent at a Deg 2 vertex		Yes	Yes	No
3a	Pendent at a pendent vertex	The diameter decreases	Yes	Yes	Yes
3b		The same diameter	Yes	Yes	No
4	Reverse edge subdivision	Different from 3	Yes	Yes	Yes

Table 3.4: The possible changes of $U(T)$ upon deletion of a vertex

Proposition 3.8. *Let T be a linear tree, and let T' be the linear tree resulting from adding a vertex to an HDV in T. $U(T)$ stays the same if $D_2 = 0$.*

Proof. This follows from the $2 + D_2$ upper bound in Theorem 2.11. Since adding a pendent vertex to an HDV does not create any degree 2 vertex, $D_2 = D_2' = 0$. Therefore, $2 \leqslant U(T) \leqslant 2 + D_2 = 2$, and $2 \leqslant U(T') \leqslant 2 + D_2' = 2 = U(T)$ □

Theorem 3.9. *Let T be a linear tree, and let T' be the linear tree resulting from adding a pendent vertex at an HDV of T. Then $-1 \leqslant U(T') - U(T) \leqslant 0$.*

Proof. This is a special case (case 2) of Theorem 3.4. It suffices to show that we are able to construct a valid LSP table whose resulting multiplicity list contains no more than $U(T)$ 1's for T'. Given an arbitrary linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, suppose we add a pendent vertex to the central vertex of T_i, $1 \leqslant i \leqslant k+1$.

Consider all LSP tables of T that attain $U(T)$, with upward multiplicity lists b_1^+, c_1^+, $\cdots$, b_k^+, c_k^+, b_{k+1}^+.

If b_i^+ in some LSP table has a nonzero upward multiplicity, then we can simply increase this upward multiplicity by 1 by assigning the associated eigenvalue to the newly added pendent vertex. The LSP is not violated, and the resulting multiplicity list has no more than $U(T)$ 1's for T' because the multiplicity resulting from the only column subject to change cannot change from 0 to 1 as we increase a nonzero upward multiplicity in the column.

If b_i in all such LSP tables has no nonzero upward multiplicity, which means that $b_i = (1, \hat{0}, 1, \hat{0}, 1, \hat{0}, 1, \hat{0}, 1, \cdots, 1, \hat{0}, 1)$, then we may choose to increase the upward multiplicity $\hat{0}$ in a column whose resulting multiplicity is not 0. Therefore, without creating a a new multiplicity 1 eigenvalue, $U(T') - U(T) \leqslant 0$.

If no such upward multiplicity $\hat{0}$ exists, then it must have the form

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	$\cdots$	λ_j
b_1^+	$\cdots$	0	$\cdots$	0	$\cdots$	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
c_i^+ or b_{i-1}^+	$\cdots$	0	$\cdots$	0	$\cdots$	0	
b_i^+	1	$\hat{0}$	1	$\hat{0}$	1	$\hat{0}$	1
c_{i+1}^+ or b_{i+1}^+	$\cdots$	0	$\cdots$	0	$\cdots$	0	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
b_{k+1}^+	$\cdots$	0	$\cdots$	0	$\cdots$	0	$\cdots$
sum	a_1	a_2	$\cdots$	$\cdots$	$\cdots$	$\cdots$	a_j

This cannot be the only LSP table that attains $U(T)$. The nontrivial upward multiplicity list above this multiplicity list has to contain some nonupward 1, and the nonupward 1 cannot appear in the column with nonupward 1's in this multiplicity list, so the nonupward 1 appears in some column between the $\hat{0}$ column and the nonupward 1 column. Then another LSP table that attains $U(T)$ can be created by "merging" this column into the $\hat{0}$ column. And after merging, we can increase the multiplicity of $\hat{0}$ to conclude $U(T') - U(T) \leqslant 0$. □

Second, when we add a pendent vertex to a degree 2 vertex in a linear tree, $U(T)$ does not increase either.

Theorem 3.10. *Let T be a linear tree, and let T' be the linear tree resulting from adding a pendent vertex to a degree 2 vertex in T. Then $-1 \leqslant U(T') - U(T) \leqslant 0$.*

Proof. This is again a special case of Theorem 3.4 (case 5 and 6). It suffices to show that we are able to construct a valid LSP table whose resulting multiplicity list has no more than $U(T)$ 1's for T' in both cases. We prove Theorem 3.10 by Lemma 3.11 that addresses case 5 and Lemma 3.12 that addresses case 6. □

Lemma 3.11. *Let T be a linear tree, and let T' be a linear tree obtained by adding a pendent vertex to a degree 2 vertex on a connecting path s_i for some $1 \leqslant i \leqslant k$ in T. Then $-1 \leqslant U(T') - U(T) \leqslant 0$.*

Proof. This lemma addresses case 5. Given an arbitrary linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, suppose we add a pendent vertex to a degree 2 vertex v on a connecting path s_i in T, there are two possibilities.

1. When v is an interior vertex in s_i:
   The upward multiplicity list c_i consists of s_i nonupward 1's, and we look at the part of c_i^+ around the multiplicity associated with the eigenvalue assigned to v. It is in the form of

$$\cdots \quad 1 \quad 0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0 \quad 1 \quad \cdots$$

After adding a pendent vertex to v, c_i^+ is broken into three multiplicity lists. Without loss of generality, we may skip the the augmented $\hat{0}$'s, part of the LSP table that attains $U(T)$ looks like either of the following with new columns highlighted. Notice that the new columns consist of nonupward 0's except the highlighted two entries.

col_1	col_2	col_3	col_4	col_5
1	0	0	0	0
0	1	$\hat{0}$	1	0
0	0	0	0	1

or

col_1	col_2	col_3	col_4	col_5
1	0	0	0	0
0	1	$\hat{0}$	1	0
0	0	0	0	1

In fact, these two cases are analogous. Without loss of generality, we discuss the situation on the left in more detail. We want to "merge" the original middle column (i.e., col_2) into the newly added column that has $\hat{0}$ (i.e.,col_3). Since col_3 only consists of 0's and one $\hat{0}$, we only need to deal with the nonupward 1 next to $\hat{0}$. To achieve this, we swap the nonupward 1's in col_1 and col_2 like this:

col_1	col_2
1	0
0	1
0	0

becomes

col_1	col_2
0	1
1	0
0	0

Notice that this is valid and the resulting multiplicity of col_1 does not change. Now, we "merge" col_2 into col_3. Then, we move the nonupward 1 in col_5 to the entry under $\hat{0}$ in col_3, and recover col_5 by moving the nonupward 1 in col_4 to col_5. Thus,

col_1	col_2	col_3	col_4	col_5
1	0	0	0	0
0	1	$\hat{0}$	1	0
0	0	0	0	1

becomes

col_1	col_2	col_3	col_4	col_5
0	0	1	0	0
1	0	$\hat{0}$	0	1
0	0	1	0	0

Notice that col_2 and col_4 result in 0 in the LSP table and can be omitted. Since this is a valid LSP table, $U(T') \leqslant U(T)$.

2. When v is not an interior vertex on s_i:
The vertex v (the multiplicity associated with its eigenvalue is highlighted) could be on the left end of s_i, on the right end of s_i, or be s_i itself.

$$0 \quad \cdots \quad 0 \quad \boxed{1} \quad 0 \quad \cdots \quad 0 \quad 1 \quad \cdots$$

or

$$\cdots \quad 1 \quad 0 \quad \cdots \quad 0 \quad \boxed{1} \quad 0 \quad \cdots \quad 0$$

or

$$0 \quad \cdots \quad 0 \quad \boxed{1} \quad 0 \quad \cdots \quad 0$$

We discuss the possibilities of the column with the highlighted nonupward 1.

(1) If the entries above and below the nonupward 1 are both 0,

$$\begin{matrix} 0 \\ 1 \\ 0 \end{matrix}$$

Then since all the upward multiplicities are enveloped by two nonupward 1's, then both in the first and third row, either to the left or to the right of the nonupward 0, there exists a nonupward 1 with only possibly some augmenting 0's in between.

If the nonupward 1's in the first and third row are both to the left or to the right (i.e., on the same side of the middle column), then we can insert the two columns on that side as well. For example, if the nonupward 1's in the first and third row are both to the left; notice that the nonupward 1's in the first and third row cannot be in the same column because of the LSP. We "merge" everything in col_2 into col_4, move the nonupward 1 in col_1 to col_4, and recover col_1 by the the nonupward 1 in col_3.

col_1	col_2	col_3	col_4	col_5	col_6
1	0	0	0	0	0
0	0	1	$\hat{0}$	1	0
0	1	0	0	0	0

becomes

col_1	col_2	col_3	col_4	col_5	col_6
0	0	0	1	0	0
1	0	0	$\hat{0}$	1	0
0	0	0	1	0	0

Notice that col_2 and col_3 result in 0 in the LSP table and can be omitted. Since this is a valid LSP table, $U(T') \leqslant U(T)$.

If the nonupward 1's in the first and third row are not to the same side, without loss of generality, we may assume the nonupward 1 in the first row is to the left and the nonupward 1 in the third row is to the right. Then, we can simply apply the procedure when v is an interior vertex so that $U(T') \leqslant U(T)$.

(2) If the entries above and below the nonupward 1 are not both 0, then we have the following three possibilities:

$$\begin{matrix} \hat{1} \\ 1 \\ 0 \end{matrix}\ , \quad \begin{matrix} 0 \\ 1 \\ \hat{1} \end{matrix}\ , \text{ or } \begin{matrix} \hat{1} \\ 1 \\ \hat{1} \end{matrix}$$

In fact, since all the upward multiplicities, namely, the $\hat{1}$s here, are enveloped by two nonupward 1's, in either the first row or the third row, we can find nonupward 1's on both the left and the right of $\hat{1}$, which guarantees the existence of two nonupward 1's on the same side of the middle column. Therefore, we can use the technique in case 2.(1) to construct a valid LSP table, hence $U(T') \leqslant U(T)$.

Based on the discussion, the proof is complete. □

Lemma 3.12. *Let T be a linear tree, and let T' be the linear tree resulting from adding a pendent vertex to a degree 2 vertex on an arm of a peripheral g-star (i.e T_1 or T_{k+1}), then $-1 \leqslant U(T') - U(T) \leqslant 0$.*

Proof. This lemma addresses case 6. Given an arbitrary linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, and we add a pendent vertex to a degree 2 vertex v on an arm of a peripheral g-star (i.e T_1 or T_{k+1}). Continuing the discussion of case 6 in Lemma 3.5, we need to deal with the extra 1 and we cannot make it a separate column this time. As mentioned before, it does not matter whether we use upward or nonupward 1's to recover b_1. So, we may assume that this extra 1 is an upward $\hat{1}$. Without violating the LSP, we can put this $\hat{1}$ above some nonupward 1 in b_1'. Thus, we not only do not create a new multiplicity 1 eigenvalue that increases $U(T)$, but could possibly reduce $U(T)$ by picking a column resulting in 1 in the multiplicity list if there is any. □

When we add a pendent vertex to a pendent vertex, $U(T)$ must be able to increase; otherwise, $U(T)$ would never increase. However, if we add the pendent in a way such that the diameter stays the same, $U(T)$ does not increase.

Theorem 3.13. *Let T be a linear tree, and let T' be the linear tree resulting from adding a vertex to a pendent vertex in T such that the diameter stays the same. Then $-1 \leqslant U(T') - U(T) \leqslant 0$.*

Proof. We prove this by induction on k-linear trees.

We may first check the base cases. Consider 1-linear trees (i.e., g-stars): suppose we add a vertex to a pendent vertex in T such that the diameter stays the same, then the g-star has at least 3 arms with length $l_1 \geqslant l_2 \geqslant l_3 \geqslant \cdots \geqslant l_k$ and we elongate some arm other than the two longest ones. Since $U(T) = \max\{l_1 + 1, 2d - n\}$ for g-stars by Theorem 2.5, l_1 and d are the same, and n increases by 1, then $U(T') \leqslant U(T)$, as desired.

Now suppose that for some $k \in \mathbb{N}$, given any k-linear tree T, let T' be a linear tree obtained by adding a vertex to a pendent vertex in T such that the diameter stays the same, $U(T') \leqslant U(T)$.

Consider a $(k+1)$-linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$.

If either T_i is a simple star for some $1 < i < k+1$ i or if T_1 or T_{k+1} has at most one arm of length greater than 2, then we can remove all the length 1 arms of T_i to get a k-linear tree such that $U(T') \leqslant U(T)$ upon addition of a vertex to a pendent vertex. Then we can add the length 1 arms back. Since we are adding pendent vertices to a degree 2 vertex or an HDV, $U(T') \leqslant U(T)$ by Theorem 3.9 and Theorem 3.10. If neither happens, then T_i for all $1 < i < k+1$ has at least one arm of length $\geqslant 2$, and T_1 and T_{k+1} both have at least two arms of length greater than 2.

Now, we consider peripheral g-stars, namely T_1 and T_{k+1}.

We may and will choose the peripheral g-star that is no part of any diameter. Notice that it is not only typical for the central path to be the diameter so that each peripheral g-star contributes one arm to the diameter, but it is guaranteed that not both peripheral g-stars attain the diameter. This can be shown by contradiction. Suppose the T_{k+1} has arms with length $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_k$ and T_1 has arms with length $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_j$. The longest path in T_{k+1} has length $l_1 + l_2 + 1$, and that in T_1 has length $m_1 + m_2 + 1$. The path connecting the longest arm in T_1 and the longest arm in T_{k+1} has length greater than or equal to $(l_1 + 1) + (m_1 + 1) = l_1 + m_1 + 2$. Suppose both of peripheral g-stars attain the diameter; that is,

(1) $l_1 + l_2 + 1 = m_1 + m_2 + 1 = d$; and

(2) $d \geqslant l_1 + m_1 + 2$.

From (1), we attain $l_1 = d - l_2 - 1$ and replace l_1 in (2), getting the following:

$$\begin{aligned} d &\geqslant (d - l_2 - 1) + m_1 + 2 \\ 0 &\geqslant m_1 - l_2 + 1 \\ l_2 &\geqslant m_1 + 1 \end{aligned}$$

Therefore, $l_1 + l_2 + 1 \geqslant l_2 + l_2 + 1 \geqslant (m_1 + 1) + (m_1 + 1) + 1 \geqslant m_1 + m_2 + 3 > m_1 + m_2 + 1$, which is a contradiction to equation (1) such that $l_1 + l_2 + 1 = m_1 + m_2 + 1 = d$.

Without loss of generality, we consider T_{k+1} with arm length sequence $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_m$ as the desired peripheral g-star that does not attain the diameter within itself. Keeping only the longest arm of T_{k+1}, we "trim" a $(k+1)$-linear tree T back to a k-linear tree so that

$U(T') \leqslant U(T)$ upon addition of a vertex to a pendent vertex by the induction hypothesis. It suffices to show that when we add all other arms of T_{k+1} back, $U(T') \leqslant U(T)$.

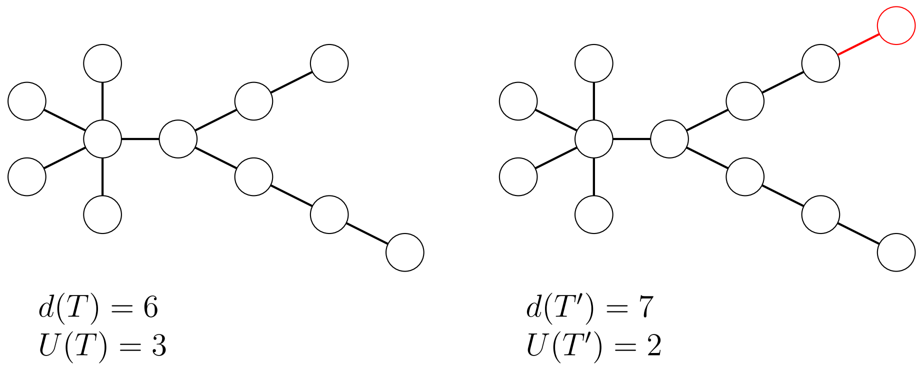
Notice that we primarily care about the second longest arm (its existence is guaranteed) that either creates a new $(\hat{0}, 1)$ pair or increases the upward multiplicity $\hat{0}$ to upward $\hat{1}$. For the rest of the arms (not necessarily existent), we can simply increase the upward multiplicity $\hat{1}$ to upward $\hat{2}$, which does not produce new multiplicity 1 eigenvalues.

Therefore, without loss of generality, suppose the arm of length l_2 is the only other arm of T_{k+1}. Together with the hypothesis that the diameter does not increase, $l_2 < d - l_1$. Consider the LSP table of "trimmed" T, the last upward multiplicity list b_{k+1} is in the form of $(1, \hat{0}, 1, \hat{0}, \cdots, 1, \hat{0}, 1)$ with l_1 upward $\hat{0}$'s and $l_1 + 1$ nonupward 1's. Since $l_2 \leqslant l_1$, there are enough upward $\hat{0}$'s for which we can increase multiplicity, namely, from $\hat{0}$ to $\hat{1}$. Also, in Chapter 6.2 of [JS], it is shown that the minimum number of distinct eigenvalues for a tree T, denoted $c(T)$, is greater than or equal to the diameter, i.e., $c(T) \geqslant d(T)$.

Therefore, since $l_2 < d - l_1$, besides the column containing $l_1 + 1$ nonupward 1's in the LSP construction, there are at least $d - l_1 - 1$ columns such that the resulting multiplicity is greater than or equal to 1. Therefore, there are enough columns in which we can put upward $\hat{0}$'s at the bottom and increase from upward $\hat{0}$'s to upward $\hat{1}$'s without worrying about creating new simple eigenvalues. Thus, when we recover the second longest arm of T_{k+1}, $U(T') \leqslant U(T)$, completing the proof for $(k+1)$-linear trees.

Therefore, by the principle of mathematical induction, the proof is complete. □

Remark 3.14. If we add a vertex to a pendent vertex such that the diameter is increased, it is possible that $U(T')$ is less than $U(T)$ although it is a rare occurrence. From the database of all trees on fewer than 13 vertices, such trees appear only three times. Here we give one example.



In fact, we have an analogous conjecture for edge subdivision when the diameter is not increased.

Conjecture 3.15. Let T be a linear tree, and let T' be the linear tree resulting from an edge subdivision in T such that the diameter stays the same. Then $-1 \leqslant U(T') - U(T) \leqslant 0$.

Note that edge subdivision and pendent vertex addition are equivalent in many cases, so a substantial part of Conjecture 3.15 has been accounted for by Theorem 3.13. However, they differ when we subdivide an interior edge. In particular, to prove Conjecture 3.15, it suffices to consider linear trees whose central path is not the diameter and subdivide edges on the portion of the central path that is not part of any diameter. This is because if we subdivide an edge not on the central path, it is on some arm of a g-star, hence the edge subdivision is the same as pendent vertex addition. On the other hand, the central path cannot be the diameter because by assumption, the edge subdivision does not increase the diameter.

We have not had a complete proof of Conjecture 3.15, but if the following statement, Conjecture 3.16, holds, then Conjecture 3.15 will follow.

Conjecture 3.16. For a linear tree T whose central path is not a diameter, there exists a linear tree T_0 resulting from the removal of some pendent vertex v such that $d(T_0) = d(T)$, and $U(T_0) = U(T)$.

Proof of Conjecture 3.15 (assuming Conjecture 3.16). We use the minimal counterexample argument. Without loss of generality, we only consider linear trees that satisfy the hypothesis in Conjecture 3.16. Suppose the smallest linear tree T such that $U(T') > U(T)$ where T' results from an edge subdivision of T that does not increase the diameter is on n vertices.

Then, by Conjecture 3.16, there is a linear tree $T_0 = T - v$ for some pendent vertex v such that $d(T_0) = d(T)$ and $U(T_0) = U(T)$. Then, since T_0 is on $n - 1$ vertices, by the minimal counterexample hypothesis, for the linear tree T_1 resulting from the corresponding edge subdivision on T_0, which does not increase the diameter, $U(T_1) \leqslant U(T_0)$. Then for T' obtained from adding the pendent vertex v to T_1, $U(T') \leqslant U(T_1)$ by Theorem 3.13. A contradiction arises because $U(T') \leqslant U(T_1) \leqslant U(T_0) = U(T)$ whereas we assume $U(T') > U(T)$.

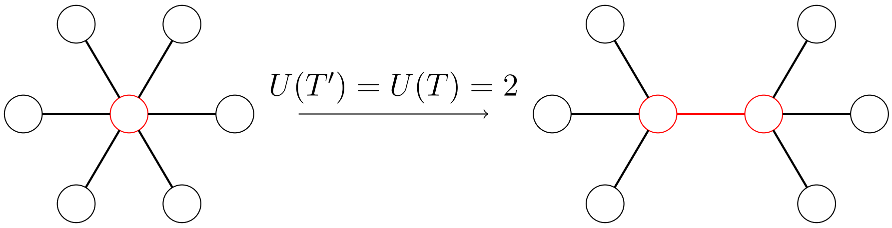
Hence, $-1 \leqslant U(T') - U(T) \leqslant 0$. □

3.2.3 Comparison of changes in $U(T)$ upon pendent vertex addition, edge subdivision, and vertex partition

We have talked about three topological operations, namely, pendent vertex addition, edge subdivision, and vertex partition. Notice that the latter categories are broader and might include part of the former categories in some sense, but we will identify an operation in an as accurate as possible way, i.e., if an operation can be viewed as pendent vertex addition, we will not describe it as edge subdivision or vertex partition even if we could.

Recall from Theorem 3.4 that the change in $U(T)$ upon pendent vertex addition and edge subdivision on linear trees is bounded by 1. However, vertex partition might change $U(T)$ by an arbitrarily large amount in both directions. We demonstrate by the following three constructive examples.

1. An example of vertex partition where $U(T') = U(T)$:

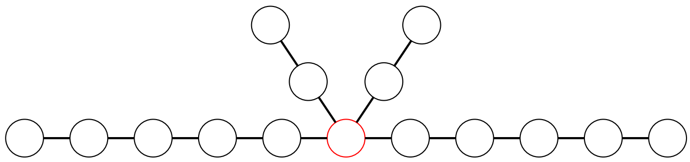


2. A generalized example of vertex partition where $U(T') = U(T) - l$:

   Let T be a g-star with four arms of length (k, k, l, l) where $k > l$, and let T' be a double g-star $DL(k, k; l, l)$ with two arms of length k adjacent to one HDV and two other arms of length l adjacent to the other HDV. In this case, $U(T') - U(T) = -l$.

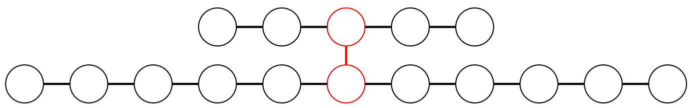
   We display a concrete example such that $U(T') = U(T) - 2$. Notice that T and T' are not in the database as they have more 12 vertices; however, we can still calculate $U(T)$ and $U(T')$ by using the formula for $U(T)$ for g-stars (Theorem 2.5) and 2-linear trees (Theorem 3.24 to be introduced in Section 3.4 [DIJ]).

$T =$



$U(T) = 6$

$T' =$



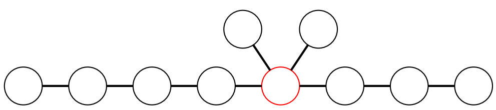
$U(T') = 4$

3. A generalized example of vertex partition where $U(T') = U(T) + l$:

   Let T be a g-star with four arms of length $(k, l+1, 1, 1)$ where $k > l+1$, and let T' be a double g-star $DL(k, 1; l+1, 1)$ with two arms of length k and length 1 adjacent to one HDV and two other arms of length $l+1$ and length 1 adjacent to the other HDV. In this case, $U(T') - U(T) = l$.

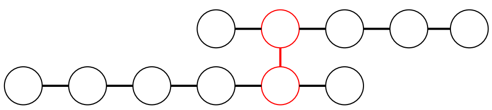
   We display a concrete example $U(T') = U(T) + 2$.

$T =$



$U(T) = 5$

$T' =$



$U(T') = 7$

3.3 New bounds for $U(T)$

In Section 3.3, using in part the ideas of Section 3.2, new bounds and refined bounds are given for $U(T)$ for linear trees. In particular, $U(T) \leqslant d(T)$, with equality only when T is a path. Also, bounds in terms of $D_2(T)$, the number of degree 2 vertices, are improved.

3.3.1 The diameter upper bound for $U(T)$

Lemma 3.17. *For a linear tree T, $U(T) = d(T)$ if and only if T is a path.*

Proof. Sufficiency is immediate: $U(P_n) = n = d(P_n)$. We prove the necessity by contrapositive using induction on the number of vertices beyond the diameter. Suppose T is not a path. We start from a path P of length equal to the diameter $d(T)$. Since T is not a path, we need to add some vertex to a degree 2 interior vertex on the path. Notice that when we add this pendent vertex, the path P becomes a g-star, denoted T_0, and $U(T_0) = \max\{2d - n, l_1 + 1\}$ by Theorem 2.5. Since $2d - n = 2d - (d + 1) = d - 1$ and $l_1 + 1 \leqslant (d - 2) + 1 = d - 1$, $U(T_0) = d - 1 = U(P) - 1$. Therefore, the disparity between $U(T_0)$ and $d(T)$ is 1.

Now, we can recover T from T_0 through adding a pendent vertex to degree 2 vertices, adding a pendent vertex to an HDV, and adding a pendent vertex to a pendent vertex without increasing the diameter. By Theorem 3.9, 3.10, and 3.13, $U(T) \leqslant U(T_0) = d - 1$ and the disparity is never zero again. Therefore, $U(T) < d(T)$, completing the proof for necessity. ☐

Remark 3.18. The proof for the necessity in Lemma 3.17 starts from a path of length equal to the diameter and recovers the linear tree by adding vertices. In fact, we can also trim the linear tree down to a path of length equal to the diameter in a way such that $U(T)$ does not decrease by the results in Table 3.4. This establishes the desired upper bound as well.

Theorem 3.19. *For any linear tree T, $U(T) \leqslant d(T)$; moreover, $U(T) \leqslant d(T) - 1$ unless T is a path.*

Proof. Given a linear tree T with diameter $d(T)$, we can "trim" the tree down to a path of length $d(T)$ by deleting pendent vertices at HDV's, at degree 2 vertices, and at pendent vertices such that $d(T)$ is not decreased. By applications of Theorem 3.9, Theorem 3.10, and Theorem 3.13 (results a,b, and c listed in Table 3.4), $U(T)$ does not decrease during the "trimming" process. Since $U(P_d) = d$, then $U(T) \leqslant U(P_d) = d$.

Moreover, since $U(T) = d(T)$ if and only if T is a path by Lemma 3.17, $U(T) < d(T)$ if T is not a path. Hence, $U(T) \leqslant d(T) - 1$ unless T is a path. ☐

3.3.2 The $1 + D_2$ upper bound for $U(T)$

Besides the diameter upper bound, we are able to improve the $2 + D_2$ upper bound in Theorem 2.11 by giving a characterization of linear trees with $U = 2 + D_2$.

Theorem 3.20. *For a linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, $U(T) = 2 + D_2$ if and only if T is depth 1 and either of the following is true:*

(a) $s_i = 0$ for all $1 \leqslant i \leqslant k$; or

(b) the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$.

Notice that if a linear tree is depth 1, we put all the HDVs in the central path and hang the pendent vertices onto it; also, $T_2, \cdots, T_k$ are necessarily simple stars, but T_1 and T_{k+1} could have at most one arm of length greater than 1 as part of the central path. Since the proof is involved, we prove the characterization by Lemma 3.21 proving necessity and Lemma 3.22 proving sufficiency.

Lemma 3.21. *For a linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, $U(T) = 2 + D_2$ only if T is depth 1 and either of the following is true:*

(a) $s_i = 0$ for all $1 \leqslant i \leqslant k$; or

(b) the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$.

Proof. It suffices to show the following two statements: for a linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$,

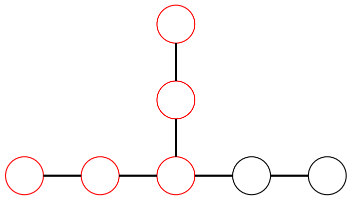
1. if T is not depth 1, then $U(T) < 2 + D_2$; and
2. suppose T is depth 1, if $s_i \neq 0$ for some $1 \leqslant i \leqslant k$ and the degree of the central vertex of T_j is greater than or equal to 4 for some $1 < j < k+1$, then $U(T) < 2 + D_2$.

First, we can prove statement 1. Suppose T is not depth 1, then there are two possibilities:

1.1 Either that T_1 or T_{k+1} has more than one arm of length greater than 1; or

1.2 that some T_i of $T_2, \cdots, T_k$ has one arm of length greater than 1.

In case 1.1 , regardless of the number or the length of arms, it suffices to show that $U(T) < 2 + D_2$ in the "minimum" part where T_1 has two arms of length equal to 2. This is because the "minimum" part is always a subgraph; more arms or longer arms result in correspondingly larger D_2. The "minimum" part is shown as highlighted vertices in the following tree.



Recall the LSP table related to the degree conjecture that attains $U(T) = 2 + D_2$ where d_i denotes the degree of HDVs in the tree. Since $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, we have $d_i \geqslant 3$ for $1 \leqslant i \leqslant k+1$.

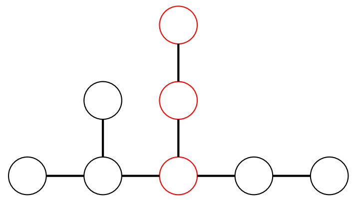
	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_{j-2}	λ_{j-1}	λ_j	$\cdots$
b_1^+	1	$\widehat{d_1 - 2}$	1	0	0	0	0	0	$\cdots$
b_2^+	0	1	$\widehat{d_2 - 3}$	1	0	0	0	0	$\cdots$
b_3^+	0	0	1	$\widehat{d_3 - 3}$	1	0	0	0	$\cdots$
$\cdots$	0	0	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$
b_k^+	0	0	0	0	1	$\widehat{d_k - 3}$	1	0	$\cdots$
b_{k+1}^+	0	0	0	0	0	1	$\widehat{d_{k+1} - 2}$	1	$\cdots$
sum	1	$d_1 - 1$	$d_2 - 1$	$d_3 - 1$	$\cdots$	$d_k - 1$	$d_{k+1} - 1$	1	$\cdots$

With this structure of multiplicity lists of T_1 through T_{k+1}, to accommodate the connecting paths and longer arms of g-stars (with exact one-to-one correspondence to degree 2 vertices), we can simply append either nonupward 1 (for the vertices on the connecting paths) or $(\hat{0}, 1)$ (for longer arms of g-stars) at the end. So, this construction provides an upper bound of $U(T) \leqslant 2 + D_2$.

Therefore, in a tree with the "minimum" part, the two degree 2 non-pendent vertices are viewed as two $(\hat{0}, 1)$'s in the LSP table, contributing two 1's in the resulting multiplicity list. To show that $U(T) < 2 + D_2$, it suffices to find another valid LSP table that combines some 1's.

Now, since T_1 has two arms of length 2, by the eigenvalue assignment principle, it is valid for T_1 to have multiplicity list $(1, \widehat{d_1 - 2}, 1, \hat{1}, 1)$ instead of $(1, \widehat{d_1 - 2}, 1, \hat{0}, 1, \hat{0}, 1)$ with two $(\hat{0}, 1)$ corresponding to the two degree 2 vertices. We can move the upward $\hat{1}$ into the column that produces the highlighted 1 on the right and append the nonupward 1 as usual. This leads to two fewer 1's in the resulting multiplicity list, which justifies $U(T) < 2 + D_2$.

In case 1.2, suppose that some T_i of $T_2, \cdots, T_k$ has one arm of length greater than 1. Again for the argument of the "minimum" part, we may assume the length of the only arm of T_i is 2 so that T_i looks like the following highlighted g-star:



Since T_i is not a peripheral g-star, b_i is not the top or bottom upward multiplicity list in the LSP table. Therefore, instead of appending $(\hat{0}, 1)$ at the end producing a 1 and contributing to $U(T)$, we can put the upward $\hat{0}$ in the column that produces the right highlighted 1 in the last line, move the nonupward 1 on the right in b_1 also to this column, and then move the nonupward 1 on the left and upward $\widehat{d_1 - 2}$ to the right by two entries. It will look like the following:

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_{j-2}	λ_{j-1}	λ_j			λ_{j+h}
b_1^+	1	$\widehat{d_1 - 2}$	1	0	0	0	0	0	$\cdots$	0	0
b_2^+	0	1	$\widehat{d_2 - 3}$	1	0	0	0	0	$\cdots$	0	0
b_3^+	0	0	1	$\widehat{d_3 - 3}$	1	0	0	0	$\cdots$	0	0
$\cdots$	0	0	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$	$\hat{0}$	1
b_k^+	0	0	0	0	1	$\widehat{d_k - 3}$	1	0	$\cdots$	0	0
b_{k+1}^+	0	0	0	0	0	1	$\widehat{d_{k+1} - 2}$	1	$\cdots$	0	0
sum	1	$d_1 - 1$	$d_2 - 1$	$d_3 - 1$	$\cdots$	$d_k - 1$	$d_{k+1} - 1$	1	$\cdots$	0	1

becomes

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_{j-2}	λ_{j-1}	λ_j			λ_{j+h}
b_1^+	0	0	1	$\widehat{d_1 - 2}$	0	0	0	1	$\cdots$	0	0
b_2^+	0	1	$\widehat{d_2 - 3}$	1	0	0	0	0	$\cdots$	0	0
b_3^+	0	0	1	$\widehat{d_3 - 3}$	1	0	0	0	$\cdots$	0	0
$\cdots$	0	0	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	$\hat{0}$	$\cdots$	0	1
b_k^+	0	0	0	0	1	$\widehat{d_k - 3}$	1	0	$\cdots$	0	0
b_{k+1}^+	0	0	0	0	0	1	$\widehat{d_{k+1} - 2}$	1	$\cdots$	0	0
sum	0	1	$d_2 - 1$	$d_1 + d_3 - 3$	$\cdots$	$d_k - 1$	$d_{k+1} - 1$	2	$\cdots$	0	1

Therefore, in case 1.2, when some T_i of $T_2, \cdots, T_k$ has one arm of length greater than 1, $U(T) < 2 + D_2$.

Now, we will prove statement 2. Suppose T is depth 1, $s_i \neq 0$ for some $1 \leqslant i \leqslant k$, and the degree of the central vertex of T_j is greater than or equal to 4 for some $1 < j < k+1$. Then for some non-peripheral g-star multiplicity list b_j, instead of $(1, \widehat{d_j - 3}, 1)$, we may have $(1, \widehat{d_j - 4, 1, \hat{0}}, 1)$.

For s_i, if $i \leqslant j$, then we can make the column that produces the right highlighted 1 in the last line result in a multiplicity 2 eigenvalue by putting the nonupward 1 in c_i and the upward $\hat{0}$ into it. Thus,

$$\begin{array}{c} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \\ \hline 1 \end{array} \quad \text{becomes} \quad \begin{array}{c} 0 \\ \vdots \\ 1 \\ \vdots \\ \hat{0} \\ 1 \\ \hline 2 \end{array}$$

Therefore, absorbing at least two 1's in the multiplicity list, $U(T) \leqslant 2 + D_2 - 2 = D_2 < 2 + D_2$.

For s_i, if $i > j$, then we can put the nonupward 1 on the right of b_1, the upward $\hat{0}$ from b_j, and the nonupward 1 from c_i in one column to produce a 2 in the multiplicity list. Again, we need to shift the two left entries in b_1 two entries right like what we did in case 1.2. Then, $U(T) < 2 + D_2$ as desired. □

Lemma 3.22. *For a linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$, $U(T) = 2 + D_2$ if T is depth 1 and either of the following is true:*

(a) $s_i = 0$ for all $1 \leqslant i \leqslant k$; or

(b) the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$.

Proof. It suffice to show the following two statements: for a linear tree $T = L(T_1, s_1, \cdots, s_k, T_{k+1})$,

1. if T is depth 1 and $s_i = 0$ for all $1 \leqslant i \leqslant k$, then $U(T) = 2 + D_2$; and
2. if T is depth 1 and the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$, then $U(T) = 2 + D_2$;

First, suppose T is depth 1 and $s_i = 0$ for all $1 \leqslant i \leqslant k$.

Repeating that for the sake of the proof, T_i or T_k has at most one arm of length greater than one, we may view a g-star with two arms of length more than one as a g-star and a connecting path, hence in which $s_1 \neq 0$.

Suppose the single long arm of T_1 has length l_1 and that of T_{k+1} has length l'_1. Therefore, $2 + D_2 = 1 + (l_1 - 1) + (l'_1 - 1) = l_1 + l'_1$. Notice that T_1 has at least l_1 upward eigenvalues hence $l_1 + 1$ nonupward 1's, and T_{k+1} has at least l'_1 upward eigenvalues hence $l'_1 + 1$ nonupward 1's. Since $s_i = 0$ for all $1 \leqslant i \leqslant k$, there is no chance to combine 1. Therefore, there are $(l_1 + 1) + (l'_1 + 1) - 2 = l_1 + l'_1$ 1's in the multiplicity list since two 1's are used for the construction with $\widehat{d_2 - 2}$ and $\widehat{d_k - 2}$. Hence, since there is no "more efficient" LSP table, $U(T) = l_1 + l_1' = 2 + D_2$.

Second, suppose T is depth 1 and the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$.

Again, in a similar spirit, $2 + D_2 = \sum_{i=1}^{k} s_i + (l_1 - 1) + (l_1' - 1)$. Since the degree of the central vertex of T_j is 3 for all $1 < j < k + 1$. , then $d_j - 3 = 0$ for all $1 < j < k + 1$, and the LSP construction will look like the following:

	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_{j-2}	λ_{j-1}	λ_j	$\cdots$
b_1^+	1	$\widehat{d_1 - 2}$	1	0	0	0	0	0	$\cdots$
b_2^+	0	1	$\hat{0}$	1	0	0	0	0	$\cdots$
b_3^+	0	0	1	$\hat{0}$	1	0	0	0	$\cdots$
$\cdots$	0	0	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0	0	$\cdots$
b_k^+	0	0	0	0	1	$\hat{0}$	1	0	$\cdots$
b_{k+1}^+	0	0	0	0	0	1	$\widehat{d_{k+1} - 2}$	1	$\cdots$
sum	1	$d_1 - 1$	2	2	$\cdots$	2	$d_{k+1} - 1$	1	$\cdots$

Since every upward multiplicity is used except possibly the "useless" $\hat{0}$ in the top/bottom multiplicity lists b_1 or b_{k+1}, there is no chance we can combine the nonupward 1's from connecting path or arms of peripheral g-stars but appending them at the end. Therefore, $U(T) = \sum_{i=1}^{k} s_i + (l_1 - 1) + (l_1' - 1) = 2 + D_2$, completing the proof. □

Theorem 3.23. *If a linear tree T is not depth 1, then $U(T) \leqslant D_2 + 1$.*

Proof. This result follows from Theorem 3.20. □

3.4 A formula for $U(T)$ for 2-linear trees

$U(T)$ has been characterized for 1-linear trees [JL-DS], and our purpose here is to determine $U(T)$ for 2-linear trees. Note that any tree with only two HDV's is 2-linear. This section involves joint work with Matthew Ingwersen, who was a REU student at William & Mary during the summer of 2019. So, we only state the formula here. In fact, the four quantities in Theorem 3.24 are each a lower bound for $U(T)$ for 2-linear trees: the upward $\hat{0}$ bound, length difference bound for T_1, length difference bound for T_2, and the diameter bound. Moreover, the maximum among these 4 lower bounds is exactly $U(T)$. Proofs can be found in [DIJ]. This formula actually gets used in the discussion in Section 3.2.3.

Theorem 3.24. *Let $T = L(T_1, s, T_2)$ be a 2-linear tree. Let n_i be the number of vertices in T_i, $i = 1, 2$. Let l_1 (m_1) be the length of the longest arm of T_1 (T_2). And let $z(T_1) =$*

$\max\{0, l_1 - l_2 - \cdots - l_a\}$ *and* $z(T_2) = \max\{0, m_1 - m_2 - \cdots - m_b\}$. *Then*

$$U(T) = \max \left\{ \begin{array}{c} 2 + z(T_1) + z(T_2) + s, \\ l_1 + 1 - \left\lfloor \dfrac{n_2 - z(T_2) - 1}{2} \right\rfloor, \\ m_1 + 1 - \left\lfloor \dfrac{n_1 - z(T_1) - 1}{2} \right\rfloor, \\ 2d(T) - n_1 - n_2 - s \end{array} \right\}.$$

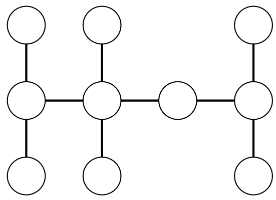
3.5 Calculating $U(T)$ for k-linear trees

Moving from $U(T)$ for 2-linear trees, we discuss $U(T)$ for k-linear trees with $k > 2$, to note that a similar approach for $U(T)$ presents clear difficulties. Also, special cases of k-linear case are discussed in this section.

3.5.1 Complications in extending lower bounds for $U(T)$ for 2-linear trees

One might wonder if the formula in Theorem 3.24 can be generalized for any k-linear trees. We use Example 3.25 to explain why $U(T)$, even just for 3-linear trees, is much more complicated.

Example 3.25. Let $T = L\,(T_1, 0, T_2, 1, T_3)$ be the 3-linear tree shown as follows.



Observe that the following is a valid linear superposition to construct a multiplicity list of T; thus $U(T) = 2$.

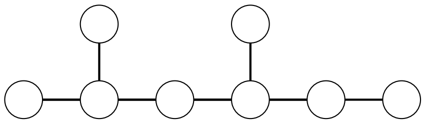
$$\begin{array}{cccccc} & 1 & \hat{1} & 1 & & \\ 1 & \hat{0} & 1 & \hat{0} & 1 & \\ & 1 & & & & \\ & & & 1 & 1 & 1 \\ \hline 1 & 2 & 2 & 2 & 2 & 1 \end{array}$$

However, the 3-linear version of the upward $\hat{0}$ lower bound is $2 + z(T_1) + s_1 + z(T_2) + s_2 + z(T_3) = 3$, which is no longer a lower bound for $U(T)$. The reason is that k-linear trees with $k > 2$ have more flexibility for the LSP, which makes a smaller $U(T)$ possible. Details can be found in [DIJ].

3.5.2 Special cases of k-linear trees

We have attempted to determine $U(T)$ for special k-linear trees: vines.

Definition 3.26. *Vines are segregated binary trees in which every degree 3 vertex is adjacent to at least one degree 1 vertex. For example,*



Proposition 3.27. *Suppose T is a k-linear vine. $U(T) = 2d - n = 2 + D_2 = n - 2k$.*

Proof. From Theorem 2.11, $U(T) \leqslant 2 + D_2$, and from Theorem 2.10, $U(T) \geqslant 2d - n$. Since they are an upper bound and a lower bound, we show that $U(T) = 2 + D_2 = 2d - n$. On the diameter, there are 2 degree 1 vertices on the ends, k degree 3 vertices , and $D_2 = d - 2 - k$. Also, $n = d + k$. Since $D_2 = d - 2 - k$, $k = d - 2 - D_2$, then $n = d + k = d + d - 2 - D_2$. That is, after rearranging terms, $2 + D_2 = 2d - n$. Now, $U(T) = n - 2k$ because $U(T) = 2 + D_2 = 2 + d - 2 - k = d - k = (d + k) - 2k = n - 2k$. □

Alternatively, we can prove that $U(T) = n - 2k$ using Lemma 7.7.2 in [JS].

Lemma 3.28. *Let T be a vine on n vertices. If $p = (m_1, \cdots, m_r, 1, \cdots, 1)$ is a partition of n, with $m_1 \geqslant \cdots \geqslant m_r \geqslant 2$ and such that $\sum_{i=1}^{r}(m_i - 1)$ is no more than the number of degree 3 vertices in T, then $p \in \mathcal{L}(T)$.*

Proof of Proposition 3.27. Because of the structure of vines, $\sum_{i=1}^{r}(m_i - 1) \leqslant k$; we may choose $m_i = 2$ for $1 \leqslant i \leqslant r = k$. Then the multiplicity list is $p = (2, 2, \cdots, 2, 1, 1, \cdots, 1)$. The number of 1 in p is $U(T) = n - 2k$. □

Chapter 4

Nonlinear trees

The next in a natural line of thought about $U(T)$ is to consider nonlinear trees. When the tree is small, linear trees dominate; for example, all trees on 10 or fewer vertices but one are linear. However, as the number of vertices in the tree increases, nonlinear trees grow far more rapidly than linear ones. For example, 12.9% of the trees on 15 vertices are nonlinear, 38.3% of the trees on 20 vertices are nonlinear, and 62.8% of the trees on 25 vertices are nonlinear. In fact, linear trees form an asymptotically vanishing subset of all trees [JWW].

As we move into the vast territory of nonlinear trees, little is known. In particular, the primary tool we used for $U(T)$ for linear trees, the LSP, is not available for nonlinear trees. To get started, we classify nonlinear trees by diameter d ($\geqslant 5$) and look at the minimal elements (*cores*) of nonlinear trees with that diameter. In Section 4.1, we introduce the concept of cores and give a combinatorial result for counting non-isomorphic cores with a certain diameter. Then, in Section 4.2, we determine $U(T)$ for the diameter 5 and 6 nonlinear trees by careful examination, give an explicit formula for $U(T)$ for cores of higher diameter, and then propose a conjecture for $U(T)$ for any nonlinear tree.

4.1 Cores of nonlinear trees

We propose a new way to classify nonlinear trees by the diameter and core.

4.1.1 Definition, properties, and observations

Definition 4.1. *Cores* of diameter d nonlinear trees ($d \geqslant 5$) are the minimal nonlinear trees (with respect to the number of vertices) with that diameter, up to isomorphism. Such a core is called a *diameter d core.*

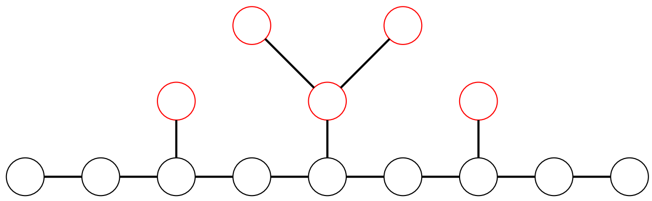
Lemma 4.2. *Any diameter d core contains only 4 HDV's.*

Proof. A nonlinear tree needs its HDV's to not lie on one path. Thus, there are at least 4 HDV's in any nonlinear tree. Now, it suffices to show that a diameter d core contains at most 4 HDV's. Suppose a diameter d core, say T, has more than 4 HDV's, then there is some HDV, say v, which is not "crucial" for nonlinearity of T. That is, the rest of the HDV's still don't lie on a path. Then while the tree remaining nonlinear, we make v degenerate to a lower degree vertex without decreasing the diameter d. The diameter, as the longest path in the tree, is not necessarily unique, but we may choose one of them and stick with it. It goes through at most 2 vertices adjacent to v. Since $deg(v) \geqslant 3$, there is some vertex u adjacent to v, which is not on the diameter. Therefore, when we remove u, the connected component containing v, with fewer vertices than T, is still a nonlinear tree with diameter d, which contradicts the assumption that T is diameter d core. □

Now, we give a characterization of diameter d cores. For a tree T, define $n(T)$ to be the number of vertices in T. In standard graph theory, $n(T) = |V(T)|$, the size of the vertex set of T.

Proposition 4.3. *A nonlinear tree T with diameter d is a diameter d core if and only if the number of vertices in T is $d + 5$.*

Proof. To facilitate understanding, we display a constructive example here, as a prototype of cores.



First, if we start with a path of length d (the path with black vertices), then adding the 5 red vertices results in a nonlinear tree with diameter d. Since a core, say T, has the fewest vertices, then $n(T) \leqslant d + 5$. Second, $n(T) \geqslant d + 5$. By Lemma 4.2, among the 4 HDV's in T, the diameter goes through 3 of them at best. Then, 5 vertices – the last HDV that is not on the diameter, at least 2 of its neighbors, and one neighbor for each of at least two HDV's that are on the path – are not on the diameter. So, $n(T) \geqslant d + 5$. In conclusion, $n(T) = d + 5$. □

Remark 4.4. We classify nonlinear trees according to diameter because cores group nonlinear trees in a systematic way. We remark the following.

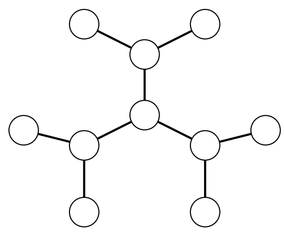
1. For a given diameter, there are finitely many cores, up to isomorphism.
2. Each diameter d core generates an infinite family of diameter d nonlinear trees via a sequence of pendent vertex additions of vertices with the diameter remaining the same.

3. The union of the families generated by all the diameter d cores is the set of all nonlinear trees with diameter d.

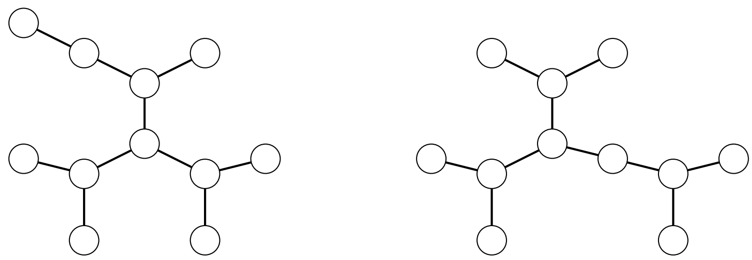
4. The families of two nonisomorphic diameter d cores can overlap with one another.

Example 4.5. To build some intuition, illustrate the concept, and transition to Algorithm 4.6, we enumerate all the diameter 5, 6, and 7 cores. Note that diameter 7 cores are listed in a different fashion from diameter 5 and 6 cores, for which we will get into more details in Section 4.1.2 and 4.1.3.

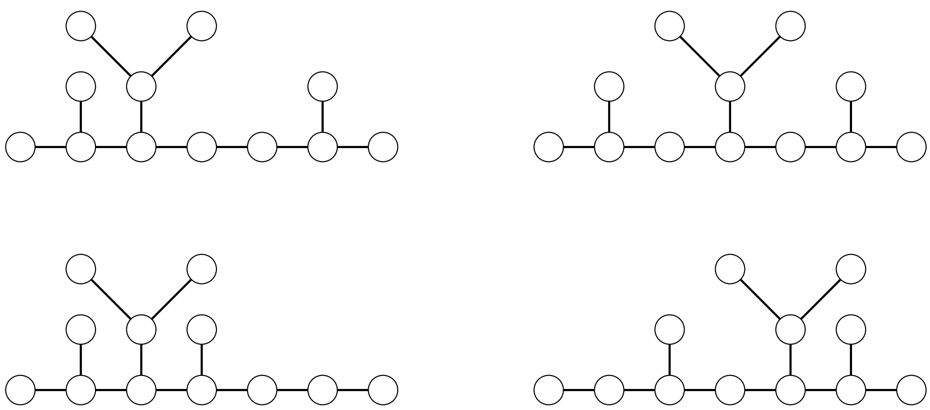
1. Diameter 5 core (on 10 vertices): there is 1 such core.

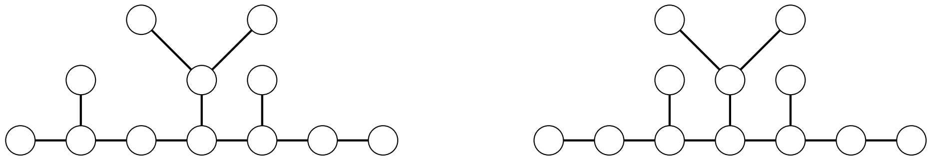


2. Diameter 6 core (on 11 vertices): there are 2 such cores, up to isomorphism.



3. Diameter 7 core (on 12 vertices): there are 6 such cores, up to isomorphism.

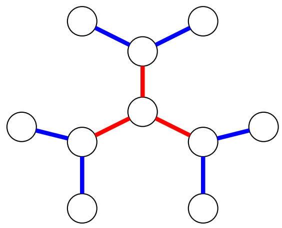




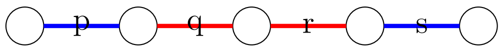
4.1.2 An algorithm for generating diameter d cores

Since there are a finite number of diameter d cores, up to isomorphism, we introduce an algorithm for generating all of them from the smallest nonlinear tree.

Algorithm 4.6. Let T be the smallest nonlinear tree on 10 vertices, shown as follows. There are two types of edges in T: R(ed) edges that are adjacent to the central vertex and B(lue) edges that are adjacent to a certain pendent. We call them R_1, R_2, R_3 and $B_1, B_1', B_2, B_2', B_3, B_3'$.



By the minimality of cores and Proposition 4.3, diameter d cores are obtained by adding $d-5$ vertices to T in a way such that the addition of each vertex increases the diameter by 1. In fact, up to isomorphism, it equates to doing edge subdivision $d-5$ times on one diameter of T. So, generating cores boils down to what we can do on one diameter of T. Without loss of generality, pick a diameter of T (a 5-path), shown as follows, where p (resp. q, r, s) denotes the number of edge subdivisions operated on that edge B_1 (resp. R_1, R_2, B_2).

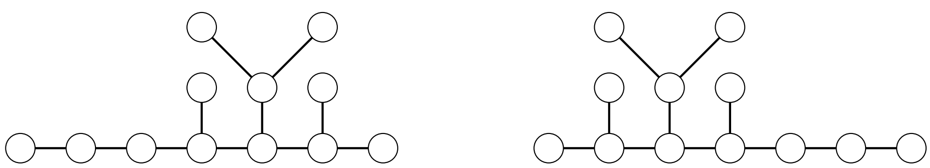


Define a string of length 4, say (p, q, r, s), whose entries are nonnegative integers. The edge subdivision of such a string corresponds to is $B_1^p R_1^q R_2^r B_2^s$, i.e., we subdivide B_1 p times, R_1 q times, R_2 r times, and B_2 s times. The set of all the cores of diameter d equals all possible nonnegative string partitions of $d-5$, i.e., $p+q+r+s = d-5$.

Proposition 4.7. *The Algorithm is well-defined.*

Proof. We explain why only considering edge subdivisions on one diameter is enough. Adding a pendent vertex to a degree 2 vertex or HDV does not increase the diameter, and adding a pendent vertex to a pendent vertex is equivalent to subdividing the edge adjacent to the pendent vertex. After an edge B_i is subdivided for some i, the diameter does not increase when subdividing B_i'; moreover, after edge subdivision is applied to edges with subscripts $i \neq j$, we cannot do edge subdivision on edges with subscript r because it does not increase the diameter. □

We revisit the six diameter 7 cores in Example 4.5. By Algorithm 4.6, we consider string partitions of $d - 5 = 7 - 5 = 2$. Top to bottom, and left to right, the cores correspond to $(0, 0, 2, 0)$, $(0, 1, 1, 0)$, $(0, 0, 0, 2)$, $(1, 1, 0, 0)$, $(0, 1, 0, 1)$, and $(1, 0, 1, 0)$. One might wonder if some partitions such as $(2, 0, 0, 0)$ are missing. The answer is no because $(2, 0, 0, 0) \cong (0, 0, 0, 2)$. In fact, the "forward" and "backward" equivalent strings are isomorphic, i.e., $(p, q, r, s) \cong (s, r, q, p)$. Furthermore, this "forward" and "backward" equivalence is the only isomorphism for cores generated by Algorithm 4.6. The isomorphism can be explained by the symmetry between edges B_1 and B_2 as well as that between edges R_1 and R_2. Instead of going through much details, we demonstrate two diameter 7 cores corresponding to $(2, 0, 0, 0)$ and $(0, 0, 0, 2)$ as an example, for which the isomorphic relationship becomes obvious.



This observation of isomorphism motivates the next section: how many cores are there for a given diameter, up to isomorphism?

4.1.3 Counting non-isomorphic cores

We define $\mathcal{C}(d)$ to be the collection of cores of distinct diameter d, up to isomorphism. And $|\mathcal{C}(d)|$, the cardinality of $\mathcal{C}(d)$, denotes the number of non-isomorphic diameter d cores.

Proposition 4.8.

$$|\mathcal{C}(d)| = p_1 + 2p_2 + 4p_3 + 6p_4 + 12p_5$$

where $p_i, 1 \leqslant i \leqslant 5$, denotes the number of partitions of one of the following patterns: $(a, a, a, a), (a, b, b, b), (a, a, b, b), (a, b, b, c)$, and (a, b, c, d) where a, b, c, d are distinct nonnegative integers.

Moreover, define $k = d - 5$. Then,

$$p_1(k) = \begin{cases} 1 & \text{if } k \pmod 4 = 0 \\ 0 & \text{otherwise} \end{cases},$$

$$p_2(k) = \frac{k - k \pmod 3}{3} + 1 - p_1,$$

$$p_3(k) = \begin{cases} \lceil \frac{k}{4} \rceil & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}.$$

The following is a generating function of p_4:

$$\frac{x^7(1+2x+3x^2)}{(1-x^2)(1-x^3)(1-x^4)}.$$

The following is a generating function of p_5:

$$\frac{x^4}{(1-x)(1-x^2)(1-x^3)(1-x^4)}.$$

Proof. First, for the formula of $|\mathcal{C}(d)|$, since the length of strings is 4, there are 5 patterns of partitions, namely, $(a,a,a,a), (a,b,b,b), (a,a,b,b), (a,b,b,c)$, and (a,b,c,d) where a,b,c,d are distinct nonnegative integers. To avoid double counting of forward and backward equivalent strings, we determine the form and the number of non-isomorphic cores (up to different order) for a partition of each pattern in Table 4.1. Moreover, since p_i for $1 \leqslant i \leqslant 5$ denotes the number of partitions for each pattern, the number of diameter d cores is a linear combination, i.e., $|\mathcal{C}(d)| = p_1 + 2p_2 + 4p_3 + 6p_4 + 12p_5$.

Pattern	Nonisomorphic partitions	Number of partitions
(a,a,a,a)	(a,a,a,a)	1
(a,b,b,b)	(a,b,b,b) (b,a,b,b)	2
(a,a,b,b)	(a,a,b,b) (b,a,a,b) (a,b,a,b) (a,b,b,a)	4
(a,b,b,c)	(a,c,b,b) (b,a,c,b) (a,b,c,b) (a,b,b,c) (c,a,b,b) (b,a,b,c)	6
(a,b,c,d)	(a,b,c,d) (a,b,d,c) (a,c,b,d) (a,c,d,b) (a,d,b,c) (a,d,c,b) (b,a,d,c) (b,a,c,d) (b,c,a,d) (b,d,a,c) (c,a,b,d) (c,b,a,d)	12

Table 4.1: Nonisomorphic cores for each pattern

Second, we give closed form formulas for p_1, p_2, and p_3 and provide generating functions for p_4 and p_5. For p_1, a partition of $k = d-5$ in the form of (a,a,a,a) is possible only when k is divisible by 4. And, if it exists, $a = \frac{k}{4}$ is uniquely determined. For p_2, we can have more than 1 partition of k in the form of (a,b,b,b) (depending on the magnitude of k) by choosing $b = 0$ and $a = k$, $b = 1$ and $a = k-3$, $b = 2$ and $a = k-6$, and so on, subject to $a, b \geqslant 0$. Because we require a and b to be distinct integers, we subtract p_1 to account for the double counting when $a = k - 3b = b$. For p_3, again, such a partition in the form of

(a, a, b, b) only exists when k is even. And if it exists, we vary b from 0 to $\lceil \frac{k}{4} \rceil$ so that a, b are uniquely determined. For p_4 and p_5, a closed form formula is not straightforward, so we calculate them manually from $k = 0$ through $k = 10$. The integer sequences are found in the online database [OEIS], and generating functions are provided. □

Using Proposition 4.8, we may calculate $|\mathcal{C}(d)|$ as the linear combination of p_1 through p_5 for diameter d cores such that $5 \leqslant d \leqslant 21$. The results are displayed in Table 4.2. Then, we put the integer sequence of $|\mathcal{C}(d)|$ into the online database [OEIS] and found a match. The resulting generating function for $|\mathcal{C}(d)|$ is given in Proposition 4.9.

d	k	$\lvert\mathcal{C}(d)\rvert$	p_1	p_2	p_3	p_4	p_5
5	0	1	1	0	0	0	0
6	1	2	0	1	0	0	0
7	2	6	0	1	1	0	0
8	3	10	0	2	0	1	0
9	4	19	1	1	1	2	0
10	5	28	0	2	0	4	0
11	6	44	0	3	2	3	1
12	7	60	0	3	0	7	1
13	8	85	1	2	2	8	2
14	9	110	0	4	0	11	3
15	10	146	0	4	3	11	5
16	11	182	0	4	0	17	6
17	12	231	1	4	3	17	9
18	13	280	0	5	0	23	11
19	14	344	0	5	4	23	15
20	15	408	0	6	0	30	18
21	16	489	1	5	4	31	23

Table 4.2: $|\mathcal{C}(d)|$ for diameter d cores with $5 \leqslant d \leqslant 21$

Proposition 4.9. *The generating function for* $|\mathcal{C}(d)|$ *is*

$$\frac{(1+k^2)}{(1-k)^2(1-k^2)^2} \qquad \text{where } k = d - 5.$$

4.2 $U(T)$

After laying the groundwork for the cores and exploring some nice properties, we return to the discussion of $U(T)$ for nonlinear trees.

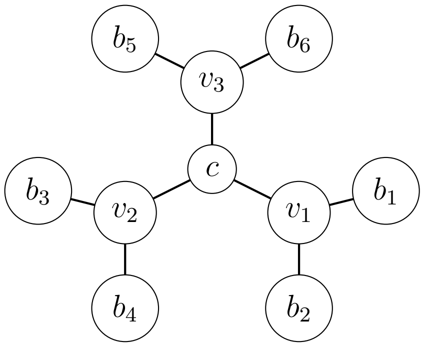
4.2.1 Diameter 5 and 6 nonlinear trees

In Section 4.2.1, through an extensive examination, we find that for any diameter 5 or 6 nonlinear tree, $U(T) = 2$, in Theorem 4.10 and 4.11. The proof is based on expansion of cores and eigenvalue assignment.

Theorem 4.10. *If T is a diameter 5 nonlinear tree, then $U(T) = 2$.*

Proof. A diameter 5 nonlinear tree consists of a central vertex, say c, and k stars, say $S_1, S_2, \cdots, S_k$, attached to it ($k \geqslant 3$) [JM]. Notice that c can also be adjacent to single pendent vertices, 2-paths, or the middle vertex of 3-paths. We reuse the term and call the centers of stars or 3-paths $S_1, S_2, \cdots, S_k$ *peripheral HDV's*, as opposed to the central vertex, i.e., the HDV in the center.

The smallest 10-vertex nonlinear tree, T_0, is embedded in any diameter 5 nonlinear tree. In other words, any diameter 5 nonlinear tree T can be obtained from T_0 by adding pendent vertices to peripheral HDV's and/or adding pendent vertices, 2-paths, 3-paths, or stars to c. It is known from the database that $U(T_0) = 2$ since the multiplicity list (4, 2, 2, 1, 1) is realizable. To realize (4, 2, 2, 1, 1) where multiple eigenvalues are λ, α, and β, we can have the following eigenvalue assignment:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{b_5\}, \{b_6\}, \{c\}\}$ with $V_1 = \{v_1, v_2, v_3\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{v_3, b_5, b_6\}\}$ with $V_2 = V_3 = \{c\}$ for α and β.

$T_0 =$ 

We will show that this multiplicity list (4, 2, 2, 1, 1) can be updated as we obtain T from T_0 so that $U(T) = 2$. We display the eigenvalue assignment for four possible ways to expand T_0; then the assignment for any diameter 5 nonlinear trees is just a combination of these four ways.

1. Adding a pendent vertex u to some v_i, without loss of generality, say v_2. We can increase m_λ so as to realize (5, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{b_5\}, \{b_6\}, \{c\}, \{u\}\}$ with $V_1 = \{v_1, v_2, v_3\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4, u\}, \{v_3, b_5, b_6\}\}$ with $V_2 = V_3 = \{c\}$ for α and β.

2. Adding a pendent vertex u to c: we can increase m_α (or equivalently, m_β) so as to realize (4, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{b_5\}, \{b_6\}, \{c, u\}\}$ with $V_1 = \{v_1, v_2, v_3\}$ for λ.
$\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{v_3, b_5, b_6\}, \{u\}\}$ with $V_2 = \{c\}$ for α.
$\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{v_3, b_5, b_6\}\}$ with $V_3 = \{c\}$ for β.

3. Adding a 2-path $\{u_1, u_2\}$ to c: we can increase m_α and m_β so as to realize (4, 3, 3, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{b_5\}, \{b_6\}, \{c, u_1, u_2\}\}$ with $V_1 = \{v_1, v_2, v_3\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{v_3, b_5, b_6\}, \{u_1, u_2\}\}$ with $V_2 = V_3 = \{c\}$ for α and β.

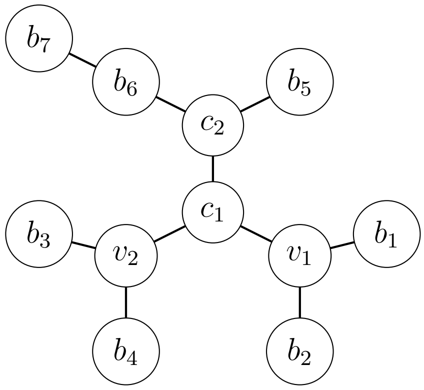
4. Adding a star on k vertices to c: since a star can be expanded by adding pendent vertices as case 1, we consider the smallest case when the middle vertex u_1 of a 3-path $\{u_2, u_1, u_3\}$ is adjacent to c; we can increase m_λ, m_α and m_β so as to realize (5, 3, 3, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{b_5\}, \{b_6\}, \{c\}, \{u_2\}, \{u_3\}\}$ with $V_1 = \{v_1, v_2, v_3, u_1\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{v_3, b_5, b_6\}, \{u_1, u_2, u_3\}\}$ with $V_2 = V_3 = \{c\}$ for α and β.

Therefore, the valid eigenvalue assignment implies a realizable multiplicity list, hence $U(T) \leqslant 2$; since $U(T) \geqslant 2$ for all trees, $U(T) = 2$. □

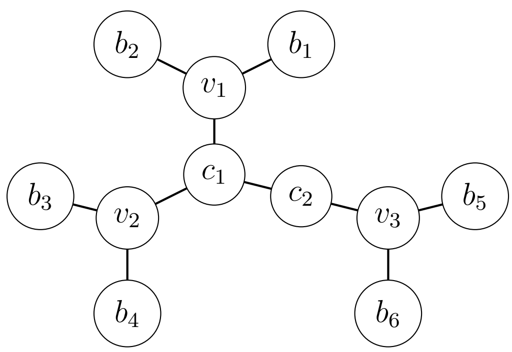
Theorem 4.11. *If T is a diameter 6 nonlinear tree, then $U(T) = 2$.*

Proof. A diameter 6 nonlinear tree consists of two central vertices, c_1 and c_2 connected by an edge to some stars, $S_1, S_2, \cdots, S_k$ [JM]. Like diameter 5 nonlinear trees, c_i can also be adjacent to single vertices, 2-paths, or the middle vertex of 3-paths. Recall in Example 4.5, there are two diameter 6 cores, say T_1 and T_2. $U(T_1) = U(T_2) = 2$ because the multiplicity list (3, 2, 2, 2, 1, 1) is realizable for both of them. Define the multiple eigenvalues as λ, α, β, and γ. We display the eigenvalue assignment for T_1 and T_2.

$T_1 =$



$T_2 =$



For T_1: $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, c_2, b_5, b_6, b_7\}\}$ with $V_1 = \{v_1, v_2\}$ for λ. $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β. $\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

For T_2: $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ. $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β. $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{v_3\}$ for γ.

Any diameter 6 nonlinear tree T can be obtained from T_1 or T_2 by a sequence of vertex addition. We will again show that this multiplicity list (3, 2, 2, 2, 1, 1) can be updated as we obtain T from T_1 or T_2 so that $U(T) = 2$.

We display the eigenvalue assignment strategy for all the possible ways to expand T_1 and T_2. First, for T_1, there are 5 ways to add one additional vertex, 2 ways to add two additional vertices, and 2 ways to add three or more vertices. We examine them one by one.

There are 5 ways to add one additional vertex.

1. Adding a pendent vertex u to some v_i, say v_1: we can increase m_λ so as to realize (4, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{u\}, \{c_1, c_2, b_5, b_6, b_7\}\}$ with $V_1 = \{v_1, v_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2, u\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4, u\}\}$ with $V_4 = \{c_2\}$ for γ.

2. Adding a pendent vertex u to b_5: we can increase m_λ so as to realize (4, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{b_5, u\}, \{b_6, b_7\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, u, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5, u\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

3. Adding a pendent vertex u to c_1: we can increase m_α (or equivalently, m_β) so as to realize (3, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, u, c_2, b_5, b_6, b_7\}\}$ with $V_1 = \{v_1, v_2\}$ for λ.
   $\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7\}, \{u\}\}$ with $V_2 = \{c_1\}$ for α.
   $\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7\}\}$ with $V_3 = \{c_1\}$ for β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{c_1, u, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

4. Adding a pendent vertex u to c_2: we can increase m_γ so as to realize (3, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, c_2, u, b_5, b_6, b_7\}\}$ with $V_1 = \{v_1, v_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, u, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{u\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

5. Adding a pendent vertex u to b_6: we can increase m_λ so as to realize (4, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{b_5\}, \{b_6, b_7, u\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, u, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
$\mathcal{A}_4 = \{\{u\}, \{b_7\}, \{c_1, c_2, b_5, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{b_6\}$ for γ.

There are 2 ways to add two additional vertices.

1. Adding a 2-path $\{u_1, u_2\}$ to c_1: we can increase m_α and m_β so as to realize (3, 3, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{u_1, u_2, c_1, c_2, b_5, b_6, b_7\}\}$ with $V_1 = \{v_1, v_2\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{u_1, u_2\}, \{c_2, b_5, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
$\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4, u_1, u_2\}\}$ with $V_4 = \{c_2\}$ for γ.

2. Adding a 2-path $\{u_1, u_2\}$ to c_2: we can increase m_λ so as to realize (5, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{b_5\}, \{b_6, b_7\}, \{u_1, u_2\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, u_1, u_2, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
$\mathcal{A}_4 = \{\{u_1, u_2\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

There are two ways to add three additional vertices.

1. Adding a 3-path $\{u_3, u_1, u_2\}$ to c_1 with the middle vertex u_1 adjacent to c_1: we can increase m_λ, m_α, and m_β so as to realize (4, 3, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, c_2, b_5, b_6, b_7\}, \{u_2\}, \{u_3\}\}$ with $V_1 = \{v_1, v_2, u_1\}$ for λ.
$\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{u_1, u_2, u_3\}, \{c_2, b_5, b_6, b_7\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
$\mathcal{A}_4 = \{\{b_5\}, \{b_6, b_7\}, \{c_1, v_1, u_1, u_2, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

2. Adding a 3-path $\{u_3, u_1, u_2\}$ to c_2 with the middle vertex u_1 adjacent to c_2: we can increase m_λ and either m_α or m_β so as to realize (5, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{b_5\}, \{b_6, b_7\}, \{u_3, u_1, u_2\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
$\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7\}, \{u_2\}, \{u_3\}\}$ with $V_2 = \{c_1, u_1\}$ for α.
$\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, b_7, u_1, u_2, u_3\}\}$ with $V_3 = \{c_1\}$ for β.
$\mathcal{A}_4 = \{\{u_3, u_1, u_2\}, \{b_6, b_7\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

Second, for T_2, there are four ways to add one more vertex, two ways to add two vertices, and two ways to add three or more vertices. Again, we examine them one by one.

There are four ways adding one additional vertex.

1. Adding a pendent vertex u to some v_i, say v_1: we can increase m_λ so as to realize (4, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6\}, \{u\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2, u\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4, u\}\}$ with $V_4 = \{v_3\}$ for γ.

2. Adding a pendent vertex u to c_1: we can increase m_α (or equivalently, m_β) so as to realize (3, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, u\}, \{v_3, b_5, b_6\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{u\}, \{c_2, b_5, b_6, v_3\}\}$ with $V_2 = \{c_1\}$ for α .
   $\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3\}\}$ with $V_3 = \{c_1\}$ for β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4, u\}\}$ with $V_4 = \{v_3\}$ for γ.

3. Adding a pendent vertex u to c_2: we can increase m_λ so as to realize (4, 2, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6\}, \{u\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, u, b_5, b_6, v_3\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4, u\}\}$ with $V_4 = \{v_3\}$ for γ.

4. Adding a pendent vertex u to v_3: we can increase m_γ so as to realize (3, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6, u\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3, u\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{u\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{v_3\}$ for γ.

There are two ways to add two additional vertices.

1. Adding a 2-path $\{u_1, u_2\}$ to c_1: we can increase m_α and m_β so as to realize (3, 3, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1, u_1, u_2\}, \{v_3, b_5, b_6\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{u_1, u_2\}, \{c_2, b_5, b_6, v_3\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4, u_1, u_2\}\}$ with $V_4 = \{v_3\}$ for γ.

2. Adding a 2-path $\{u_1, u_2\}$ to c_2: we can increase m_λ and either m_α or m_β so as to realize (4, 3, 2, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:

$\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6\}, \{u_1, u_2\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
$\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, u_1, u_2\}, \{b_5\}, \{b_6\}\}$ with $V_2 = \{c_1, v_3\}$ for α .
$\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3, u_1, u_2\}\}$ with $V_3 = \{c_1\}$ for β.
$\mathcal{A}_4 = \{\{u_1, u_2\}, \{v_3, b_5, b_6\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

There are two ways to add three additional vertices.

1. Adding a 3-path $\{u_3, u_1, u_2\}$ to c_1 with the middle vertex u_1 adjacent to c_1: we can increase m_λ, m_α, and m_β so as to realize (4, 3, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{u_2\}, \{u_3\}, \{c_1, c_2, v_3, b_5, b_6\}\}$ with $V_1 = \{v_1, v_2, u_1\}$ for λ.
   $\mathcal{A}_2 = \mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{u_3, u_1, u_2\}\{c_2, b_5, b_6, v_3\}\}$ with $V_2 = V_3 = \{c_1\}$ for α and β.
   $\mathcal{A}_4 = \{\{b_5\}, \{b_6\}, \{c_1, c_2, v_1, b_1, b_2, v_2, b_3, b_4, u_1, u_2, u_3\}\}$ with $V_4 = \{v_3\}$ for γ.

2. Adding a 3-path $\{u_3, u_1, u_2\}$ to c_2 with the middle vertex u_1 adjacent to c_2: we can increase m_λ and either m_α or m_β so as to realize (4, 3, 3, 2, 1, 1). Specifically, an eigenvalue assignment can be as follows:
   $\mathcal{A}_1 = \{\{b_1\}, \{b_2\}, \{b_3\}, \{b_4\}, \{c_1\}, \{v_3, b_5, b_6\}, \{u_3, u_1, u_2\}\}$ with $V_1 = \{v_1, v_2, c_2\}$ for λ.
   $\mathcal{A}_2 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, u_1, u_2, u_3\}, \{b_5\}, \{b_6\}\}$ with $V_2 = \{c_1, v_3\}$ for α .
   $\mathcal{A}_3 = \{\{v_1, b_1, b_2\}, \{v_2, b_3, b_4\}, \{c_2, b_5, b_6, v_3\}, \{u_2\}, \{u_3\}\}$ with $V_3 = \{c_1, u_1\}$ for β.
   $\mathcal{A}_4 = \{\{u_3, u_1, u_2\}, \{v_3, b_5, b_6\}, \{c_1, v_1, b_1, b_2, v_2, b_3, b_4\}\}$ with $V_4 = \{c_2\}$ for γ.

Therefore, using the similar strategy of eigenvalue assignment, we prove that a multiplicity list with two 1's is realizable. As two is the lower bound for $U(T)$ for trees, $U(T) = 2$. □

As for a concluding remark, when the diameter of a nonlinear tree is as small as 5 or 6, the structure of the tree is greatly limited, despite the fact that there are infinitely many such diameter 5 and 6 nonlinear trees. We suspect that $U(T) = 2$ holds true for all diameter 7 nonlinear trees; no counterexample is found yet, but there seems no easy path to a proof. This conjecture is formally stated in Section 4.2.3, in a more generalized form for any diameter d nonlinear tree. However, it is certain that not every diameter 8 nonlinear tree has $U(T) = 2$.

4.2.2 $U(T)$ for cores

Cores are the minimal nonlinear trees with a certain diameter. In some sense, they are "linear-like" because they are back to linear trees by simply removing one vertex. Since we know linear trees well, we take up this advantage and determine the $U(T)$ for cores.

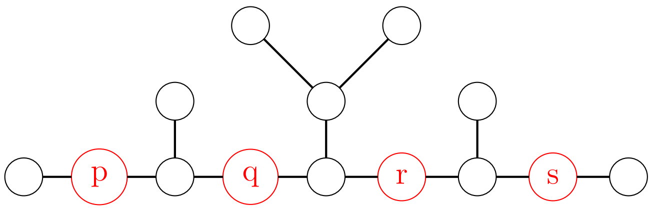
Theorem 4.12. *If T is a diameter d core, then*

$$U(T) = \begin{cases} 2 & \text{if } d \leqslant 7 \\ d-5 & \text{if } d \geqslant 8 \end{cases}.$$

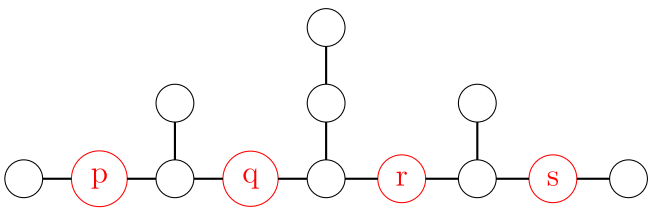
Proof. If T is a diameter d core such that $d < 7$, then either $d = 5$ or $d = 6$. By Theorem 4.10 and 4.11, all diameter 5 and 6 nonlinear trees have $U = 2$; hence, such cores have $U(T) = 2$.

Now, suppose T is a diameter d core such that $d \geqslant 7$, then edge subdivision in Algorithm 4.6 are applied at least twice to one diameter. That is, using the previous notation, $p + q + r + s \geqslant 2$ where p, q, r, and s denote the number of edge subdivisions on four edges of a diameter, i.e., the number of additional vertices in-between (highlighted).

A core looks like the following.



Since a core is a minimal nonlinear tree, there are multiple ways for it to go back to a linear tree upon removal of one vertex. We consider the following 3-linear tree $T_l = L(T_1, q, T_2, r, T_3)$ where T_1 is a g-star with two arms of length $l_1 = p + 1$ and $l_2 = 1$, T_2 is a (degenerate) g-star with one arm of length 2, and T_3 is a g-star with two arms of length $m_1 = s + 1$ and $m_2 = 1$.



We will use the LSP to show that the multiplicity list $(2, 2, 2, 2, 1, 1, \cdots, 1) \in \mathcal{L}(T_l)$. We choose the multiplicity lists $b_1 = (1, \hat{1}, 1, \hat{0}, 1, \hat{0}, 1, \cdots, \hat{0}, 1)$ for T_1, $b_2 = (1, \hat{0}, 1, \hat{0}, 1)$ for T_2, and $b_3 = (1, \hat{1}, 1, \hat{0}, 1, \hat{0}, 1, \cdots, \hat{0}, 1)$ for T_3. Moreover, s_1 and s_2, as our convention, are multiplicity lists with q and r nonupward 1's. Since 1's in the multiplicity lists are no surprise and can be easily appended at the end, we focus on how to superimpose to get four multiplicities of 2, say τ, λ, β, and γ. Since $p + q + r + s \geqslant 2$, at least one of p, q, r and s is a positive integer. Because of the symmetry of this 3-linear tree, $p \neq 0$ is analogous to $s \neq 0$, and $q \neq 0$ is analogous to $r \neq 0$. The key is pairing the upward $\hat{1}$ in b_1 and b_3 with another upward 1 for λ and β, and using the two upward $\hat{0}$'s in b_1 to connect two nonupward 1's for τ and γ. For illustration, we demonstrate two superpositions in

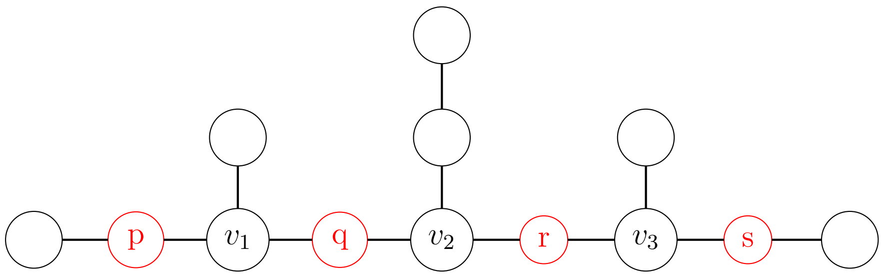
Table 4.2.2: when $q = 2$ and when $p = 2$. In fact, the superposition works whenever $q \geqslant 2$ or $p \geqslant 2$. Of course, it is possible that two of p, q, r and s are equal to 1, but those cases are not much different.

		τ	λ	β	α	γ	
b_1^+		1	$\hat{1}$			1	
s_1				1	1		
b_2^+	1	$\hat{0}$	1			$\hat{0}$	1
s_2							
b_3^+		1		$\hat{1}$		1	
	1	2	2	2	1	2	1

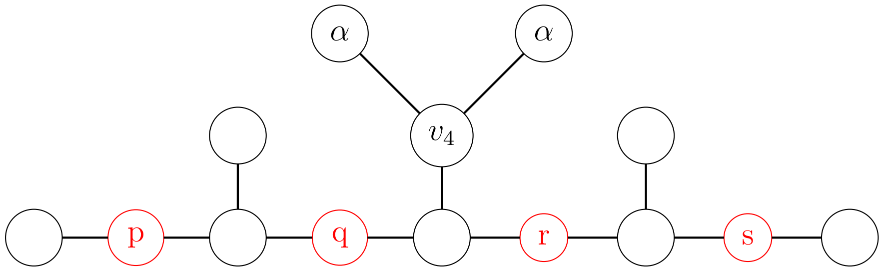
		τ	λ	β		α		γ	
b_1^+		1	$\hat{1}$	1	$\hat{0}$	1	$\hat{0}$	1	
s_1									
b_2^+	1	$\hat{0}$	1					$\hat{0}$	1
s_2									
b_3^+		1		$\hat{1}$				1	
	1	2	2	2		1		2	1

Table 4.3: The left case is when $q = 2$, and the right case is when $p = 2$

From the LSP, we know that $(2, 2, 2, 2, 1, 1, \cdots, 1) \in \mathcal{L}(T_l)$. In fact, since the LSP is necessary and sufficient, there exists an eigenvalue assignment for multiple eigenvalues corresponding to the LSP construction. In particular, we use v_1 as the Parter vertex for λ and assign λ to two arms of T_1, i.e., the pendent vertex and the path adjacent to v_1; we use v_3 as the Parter vertex for β and assign β to two arms of T_3, i.e., the pendent vertex and the path adjacent to v_3; and we use v_2 as the Parter vertex for τ and γ and assign τ and γ to the 2-path arm of T_2, the subtree to the right of v_2, and the subtree to the left of v_2.



Now, we consider the nonlinear tree T, by adding the removed vertex back to T_l. Keeping the rest of the eigenvalue assignment unchanged, we use v_4 as the Parter vertex for a multiplicity 1 eigenvalue for T_l, say α, such that $\tau < \alpha < \gamma$. That is, α's existence is guaranteed because $p + q + r + s \geqslant 2$ (see Table 4.2.2); moreover, α appears once (nonupwardly) in T_l.



With v_4 as the Parter vertex, we assign α to the two pendent vertices adjacent to v_4 and the rest of tree once so that $m(\alpha) = 2$. Also, α is chosen in a way that the three eigenvalues occurring on the 3-path arm of T_2 are α, τ, and γ. Thus, we attain a new multiplicity 2 eigenvalue while keeping the original four unchanged. Hence, the multiplicity list $(2, 2, 2, 2, 2, 1, 1, \cdots, 1) \in \mathcal{L}(T)$. So, $U(T) \leqslant n - 2 \times 5 = d + 5 - 10 = d - 5$. Since the diameter lower bound in Theorem 2.10 requires $U(T) \geqslant 2d - n = 2d - (d + 5) = d - 5$, we have equality $U(T) = d - 5$ for cores with $d \geqslant 7$, completing the proof. □

4.2.3 An upper bound for $U(T)$ for diameter d nonlinear trees

Lastly, we make a note about a potential upper bound for $U(T)$ for any diameter d nonlinear tree. Since we know $U(T)$ for cores by Theorem 4.12, it is conjectured that $U(T)$ for the core is an upper bound for all the trees in the family generated by this core. If we can attain the results for linear trees in Section 3.2.2 about the change of $U(T)$ upon vertex addition for nonlinear trees – namely, $U(T)$ does not increase when a vertex is added to an HDV (Theorem 3.9), a degree 2 vertex (Theorem 3.10), and a pendent vertex without increasing the diameter (Theorem 3.13) – then Conjecture 4.13 will follow. In fact, although it is quite plausible that the results of incremental changes in $U(T)$ hold for nonlinear trees, we do not have an analogous version of the LSP for nonlinear trees, so the proof is still unclear. Some efforts have been put into using the Implicit Function Theorem, but more work needs to be done.

Conjecture 4.13. If T is a diameter d core, then for any T' in the family generated by T, $U(T') \leqslant U(T) = d - 5$ when $d \geqslant 7$.

Chapter 5

Further results and future questions

This chapter records a few separate but worthwhile results that came up in our study. Section 5.1 gives a partial answer to the question of which trees have $U(T) = 2$. Section 5.2 relates $U(T)$ with the path cover number. Section 5.3 provides a counterexample for a conjecture.

5.1 Trees with $U(T) = 2$

As we know, $U(T)$ for any tree is bounded below by 2. An interesting question to ask is which trees have $U(T) = 2$. Inspired by the classification for trees based on the diameter, we determine the classes of trees with $U(T) = 2$ when the diameter is 3 and 4. Diameter 5 and 6 trees are discussed as well. Notice that diameter 3 and 4 trees are necessarily linear, and nonlinear trees start to appear among trees with at least diameter 5.

Theorem 5.1. *Let T be a tree with diameter 3. Then, $U(T) = 2$ if and only if T contains an HDV.*

Proof. Suppose T is a tree with diameter 3. Sufficiency: suppose T contains an HDV, then T is a simple star for which $(1, n-2, 1)$ is a multiplicity list. Therefore, $U(T) = 2$. Necessity: suppose $U(T) = 2$ and there is no HDV for contradiction. Then T is a 3-path for which $U(T) = 3$ as only $(1, 1, 1)$ is a possible multiplicity list. □

Theorem 5.2. *Let T be a tree with diameter 4. Then, $U(T) = 2$ if and only if T contains two HDV's.*

Proof. Suppose T is a tree with diameter 4. Sufficiency: suppose T contains 2 HDV's, then T is a double star for which $(1, n-2, 1)$ is a possible multiplicity list by assigning every pendent the same eigenvalue. Therefore, $U(T) = 2$. Necessity: suppose $U(T) = 2$ and there is no or 1 HDV for contradiction. Then T is either a 4-path for which $U(T) = 4$ as the only possible multiplicity list for T is $(1, 1, 1, 1)$ or a g-star for which the formula for $U(T)$ determines $U(T) \geqslant l_1 + 1 = 2 + 1 = 3$. □

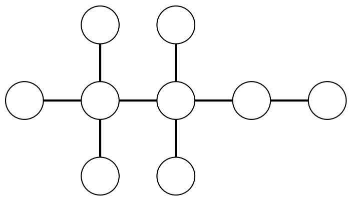
Theorem 5.3. *Let T be a tree with diameter 5, if T contains three or more HDV's, then $U(T) = 2$.*

Proof. We consider linear trees and nonlinear trees with diameter 5.

First, suppose T is a diameter 5 linear tree with three or more HDV's, then T has a subgraph T_0 such that every interior vertex of a 5-path has a pendent vertex hanging on it besides the vertices on the path. There are multiple ways to determine $U(T_0) = 2$, one of which is via $2 + D_2$ upper bound. Then, by Theorem 3.9, 3.10, and 3.13 we proved before, when we add pendent vertices to HDV's, degree 2 vertices, and pendent vertices without increasing the diameter to recover T from T_0, U does not increase. Since the starting point is $U(T_0) = 2$, then $U(T) = 2$. This completes the discussion for diameter 5 linear trees.

Second, suppose T is a diameter 5 nonlinear tree, then the HDV condition is automatically satisfied, and $U(T) = 2$ by Theorem 4.10. □

Remark 5.4. Let T be a tree with diameter 5. Notice that containing 3 or more HDV's is a sufficient but not necessary condition for $U(T) = 2$. For example, the following diameter 5 tree has 2 HDV's but $U(T) = 2$. In fact, among 83 diameter 5 trees on fewer than 13 vertices with $U(T) = 2$, 47 of them contain three or more HDV's whereas the rest do not.



Theorem 5.5. *Let T be a tree with diameter 6, if T contains four or more HDV's, then $U(T) = 2$.*

Proof. The proof for Theorem 5.5 is similar to that of Theorem 5.3, so it is omitted here. □

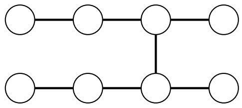
5.2 $U(T)$ and the path cover number

The path cover number P is the minimum number of disjoint paths needed to cover every vertex. In fact, in a tree T, it is shown that the maximum multiplicity, $M(T)$, is the path cover number $P(T)$ [JL-D]. We notice that the path cover number limits $U(T)$, and $U(T)$ for trees with small path cover number can be determined.

This section extends the discussion in Section 3.4 and applies the formula for $U(T)$ for 2-linear trees. The path cover number of 2-linear trees can be arbitrarily large, but when it is small, the formula of calculating $U(T)$ can be simpler. That is, among the four quantities in Theorem 3.24, fewer are needed.

Proposition 5.6. *For a tree T such that $P(T) \leqslant 2$, $U(T) = 2d - n$.*

Proof. When $P(T)=1$, T is a path; therefore, $U(T)=n$ and $d(T)=n$. So, $U(T)=2d-n$. When $P(T)=2$, $T=L(T_1,0,T_2)$ is either a double path or a g-star with three arms. When T is a g-star with 3 arms, $d(T)=l_1+l_2+1$, and by Theorem 2.5, $U(T)=\max\{l_1+1,2d-n\}$. Then $U(T)=2d-n$ because $2d-n=2(l_1+l_2+1)-(l_1+l_2+l_3+1)=l_1+(l_2-l_3)+1\geqslant l_1+1$. When T is a double path, T consists of two paths that are connected by an edge, such as the following.



Since a double path is a 2-linear tree, by Theorem 3.24, it suffices to show two inequalities, namely, $2d(T)-n_1-n_2-s\geqslant 2+z(T_1)+z(T_2)+s$ and $2d(T)-n_1-n_2-s\geqslant l_1+1-\left\lfloor\frac{n_2-z(T_2)-1}{2}\right\rfloor$. The inequality for the other difference bound $m_1+1-\left\lfloor\frac{n_1-z(T_1)-1}{2}\right\rfloor$ is similar.

First, when T is a double path, $z(T_1)=l_1-l_2$, $z(T_2)=m_1-m_2$, $n=l_1+l_2+m_1+m_2+2$, and $s=0$. So, $2d(T)-n_1-n_2-s\geqslant 2+z(T_1)+z(T_2)+s$ boils down to $2d-(l_1+l_2+m_1+m_2+2)\geqslant 2+(l_1-l_2)+(m_1-m_2)$. There are three possibilities for d: $d=l_1+m_1+2$, $d=l_1+l_2+1$, and $d=m_1+m_2+1$.

When $d=l_1+m_1+2$, then $2d-(l_1+l_2+m_1+m_2+2)=2(l_1+m_1+2)-(l_1+l_2+m_1+m_2+2)=2+(l_1-l_2)+(m_1-m_2)$. The inequality holds. When $d=l_1+l_2+1\geqslant l_1+m_1+2$, it implies that $l_1\geqslant l_2\geqslant m_1+1\geqslant m_2+1$. Then, $[2d-(l_1+l_2+m_1+m_2+2)]-[2+(l_1-l_2)+(m_1-m_2)]=2(l_1+l_2+1)-(l_1+l_2+m_1+m_2+2)-[2+(l_1-l_2)+(m_1-m_2)]=2l_2-2m_1-2=2(l_2-m_1-1)\geqslant 0$. Thus, the inequality holds. The case when $d=m_1+m_2+1$ is similar as the case when $d=l_1+l_2+1$.

Second, we will show that $2d(T)-n_1-n_2-s\geqslant l_1+1-\left\lfloor\frac{n_2-z(T_2)-1}{2}\right\rfloor$. when T is a double path, $\left\lfloor\frac{n_2-z(T_2)-1}{2}\right\rfloor=\left\lfloor\frac{m_1+m_2+1-(m_1-m_2)-1}{2}\right\rfloor=m_2$. So, this inequality boils down to $2d-(l_1+l_2+m_1+m_2+2)\geqslant l_1+1-m_2$. Again, we discuss three possibilities for d: $d=l_1+m_1+2$, $d=l_1+l_2+1$, for $d=m_1+m_2+1$.

When $d=l_1+m_1+2\geqslant l_1+l_2+1$, then it implies that $m_1\geqslant l_2-1$. Thus, $[2d-(l_1+l_2+m_1+m_2+2)]-(l_1+1-m_2)=2(l_1+m_1+2)-(l_1+l_2+m_1+m_2+2)-(l_1+1-m_2)=m_1+1-l_2=m_1-(l_2-1)\geqslant 0$. So, the inequality holds. When $d=l_1+l_2+1\geqslant l_1+m_1+2$, it implies that $l_1\geqslant l_2\geqslant m_1+1\geqslant m_2+1$. Then, $[2d-(l_1+l_2+m_1+m_2+2)]-(l_1+1-m_2)=2(l_1+l_2+1)-(l_1+l_2+m_1+m_2+2)-(l_1+1-m_2)=l_2-m_1-1=l_2-(m_1+1)\geqslant 0$. So, the inequality holds. When $d=m_1+m_2+1$, it implies that $m_1\geqslant m_2\geqslant l_1+1\geqslant l_2+1$. Then, $[2d-(l_1+l_2+m_1+m_2+2)]-(l_1+1-m_2)=2(m_1+m_2+1)-(l_1+l_2+m_1+m_2+2)-(l_1+1-m_2)=m_1+2m_2-1-2l_1-l_2=2(m_2-l_1)+(m_1-l_2)-1\geqslant 2+1-1=2$. So, the inequality holds.

In conclusion, $2d-n$ is greater than or equal to the other three lower bounds for this double path; therefore, $U(T)=2d-n$. □

Corollary 5.7. *For a 2-linear tree $T = L(T_1, s, T_2)$ such that $s > 0$ and $P(T) = 3$, $U(T) = 2d - n$.*

Proof. Suppose a 2-linear tree $T = L(T_1, s, T_2)$. $P(T) = 3$ and $s > 0$ imply that T_1 has arm lengths $l_1 \geqslant l_2 > 0 = l_3$ and T_2 has arm lengths $m_1 \geqslant m_2 > 0 = m_3$. It suffices to show that $2d - n$ is the maximal among the four quantities in Theorem 3.24 under all three possibilities of d.

Before we start, we can simplify the four quantities. Since $z(T_1) = \max\{l_1 - l_2, 0\} = l_1 - l_2$ and $z(T_2) = \max\{m_1 - m_2, 0\} = m_1 - m_2$,

$$2 + z(T_1) + z(T_2) + s = 2 + (l_1 - l_2) + (m_1 - m_2) + s$$

Also, by counting, $n = 2 + l_1 + l_2 + s + m_1 + m_2$, $n_1 = 1 + l_1 + l_2$, and $n_2 = 1 + m_1 + m_2$; hence, we have

$$l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor = l_1 + 1 - \left\lfloor \frac{(1 + m_1 + m_2) - (m_1 - m_2) - 1}{2} \right\rfloor = l_1 + 1 - m_2$$

Since the other difference bound is symmetric, we omit it here by assuming without loss of generality that $l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor \geqslant m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$. And finally, the diameter bound is

$$2d - n = 2d - (2 + l_1 + l_2 + s + m_1 + m_2).$$

Moreover, for T, $d = \max\{l_1 + m_1 + 2 + s, l_1 + l_2 + 1, m_1 + m_2 + 1\}$. Among these three possibilities, when $d = l_1 + m_1 + 2 + s$, it means that (1) $l_1 + m_1 + 2 + s \geqslant l_1 + l_2 + 1$ and (2) $l_1 + m_1 + 2 + s \geqslant m_1 + m_2 + 1$. By rearranging the terms, we obtain two inequalities:

$$m_1 - l_2 + s + 1 \geqslant 0 \tag{5.1}$$

$$l_1 - m_2 + s + 1 \geqslant 0 \tag{5.2}$$

Similarly, when $d = l_1 + l_2 + 1$, we have inequalities:

$$l_2 - m_1 - s - 1 \geqslant 0 \tag{5.3}$$

$$l_1 - m_1 + l_2 - m_2 \geqslant 0 \tag{5.4}$$

when $d = m_1 + m_2 + 1$, we have inequalities:

$$m_2 - l_1 - s - 1 \geqslant 0 \tag{5.5}$$

$$m_1 - l_1 + m_2 - l_2 \geqslant 0 \tag{5.6}$$

The proof consists of two parts: $2d - n \geqslant 2 + z(T_1) + z(T_2) + s$ and $2d - n \geqslant l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor$.

First, we compare $2d - n$ and $2 + z(T_1) + z(T_2) + s = 2 + (l_1 - l_2) + (m_1 - m_2) + s$.

When $d = l_1 + m_1 + 2 + s$, we plug it in and obtain $2d - n = 2(l_1 + m_1 + 2 + s) - (2 + l_1 + l_2 + s + m_1 + m_2) = l_1 + m_1 + 2 - l_2 - m_2 + s$. Then, for comparison, we take the difference of $2d - n$ and $2 + z(T_1) + z(T_2) + s$. That is, $(2d - n) - (2 + z(T_1) + z(T_2) + s) = (l_1 + m_1 + 2 - l_2 - m_2 + s) - (2 + (l_1 - l_2) + (m_1 - m_2) + s) = 0$.

When $d = l_1 + l_2 + 1$, $2d - n = 2(l_1 + l_2 + 1) - (2 + l_1 + l_2 + s + m_1 + m_2) = l_1 + l_2 - m_1 - m_2 - s$. The difference is $(2d - n) - (2 + z(T_1) + z(T_2) + s) = (l_1 + l_2 - m_1 - m_2 - s) - (2 + (l_1 - l_2) + (m_1 - m_2) + s) = 2l_2 - 2m_1 - 2s - 2 = 2(l_2 - m_1 - s - 1) \geqslant 0$ by Inequality 5.3.

When $d = m_1 + m_2 + 1$, $2d - n = 2(m_1 + m_2 + 1) - (2 + l_1 + l_2 + s + m_1 + m_2) = m_1 + m_2 - l_1 - l_2 - s$. The difference is $(2d - n) - (2 + z(T_1) + z(T_2) + s) = (m_1 + m_2 - l_1 - l_2 - s) - (2 + (l_1 - l_2) + (m_1 - m_2) + s) = 2m_2 - 2l_1 - 2s - 2 = 2(m_2 - l_1 - s - 1) \geqslant 0$ by Inequality 5.5. Thus, $2d - n \geqslant 2 + z(T_1) + z(T_2) + s$.

Second, we compare $2d - n$ and $l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor = l_1 + 1 - m_2$.

When $d = l_1 + m_1 + 2 + s$, we have shown that $2d - n = l_1 + m_1 + 2 - l_2 - m_2 + s$. Then, the difference of $2d - n$ and $l_1 + 1 - m_2$ is $(2d - n) - (l_1 + 1 - m_2) = (l_1 + m_1 + 2 - l_2 - m_2 + s) - (l_1 + 1 - m_2) = m_1 - l_2 + s + 1 \geqslant 0$ by Inequality 5.1.

When $d = l_1 + l_2 + 1$, $2d - n = l_1 + l_2 - m_1 - m_2 - s$. Then, the difference of $2d - n$ and $l_1 + 1 - m_2$ is $(2d - n) - (l_1 + 1 - m_2) = (l_1 + l_2 - m_1 - m_2 - s) - (l_1 + 1 - m_2) = l_2 - m_1 - s - 1 \geqslant 0$ by Inequality 5.3.

When $d = m_1 + m_2 + 1$, $2d - n = m_1 + m_2 - l_1 - l_2 - s$. The difference is $(2d - n) - (l_1 + 1 - m_2) = (m_1 + m_2 - l_1 - l_2 - s) - (l_1 + 1 - m_2) = m_1 + 2m_2 - 2l_1 - l_2 - s - 1 = (m_2 - l_1 - s - 1) + (m_1 + m_2 - l_1 - l_2) \geqslant 0$ by Inequality 5.5 and Inequality 5.6.

Therefore, $2d - n \geqslant l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor$. Finally, we can conclude that $2d - n$ is the maximal among the four quantities; hence, by Theorem 3.24, $U(T) = 2d - n$. $\square$

Nevertheless, when $s = 0$ for a 2-linear tree $T = L(T_1, s, T_2)$ with $P(T) = 3$, $U(T)$ can be larger than $2d - n$. But we may still reduce the complexity of the formula in Theorem 3.24.

Corollary 5.8. *For a 2-linear tree $T = L(T_1, 0, T_2)$ such that $P(T) = 3$, denote the arm lengths of T_1 by $l_1 \geqslant l_2 \geqslant l_3 > 0$, and the arm lengths of T_2 by $m_1 \geqslant m_2 > 0$, then*

$$U(T) = \max \begin{Bmatrix} 2 + z(T_1) + m_1 - m_2, \\ l_1 + 1 - m_2, \\ 2d(T) - n \end{Bmatrix}.$$

Proof. Notice that $U(T)$ is not as simple as that in Lemma 5.7. In fact, among the four quantities in Theorem 3.24, we need three of them. Recall, if $T = L(T_1, s, T_2)$ is any

2-linear tree, then

$$U(T) = \max \left\{ \begin{array}{c} 2 + z(T_1) + z(T_2) + s, \\ l_1 + 1 - \left\lfloor \dfrac{n_2 - z(T_2) - 1}{2} \right\rfloor, \\ m_1 + 1 - \left\lfloor \dfrac{n_1 - z(T_1) - 1}{2} \right\rfloor, \\ 2d(T) - n_1 - n_2 - s \end{array} \right\}.$$

Indeed, for the special 2-linear tree specified in Lemma 5.8, this formula can be reduced to a simpler form.

First, we quickly show this simpler formula is derived from the general formula. $2 + z(T_1) + z(T_2) + s = 2 + z(T_1) + m_1 - m_2$ because $z(T_2) = m_1 - m_2$ and $s = 0$. In addition, $l_1 + 1 - \left\lfloor \frac{n_2 - z(T_2) - 1}{2} \right\rfloor = l_1 + 1 - \left\lfloor \frac{(m_1+m_2+1)-(m_1-m_2)-1}{2} \right\rfloor = l_1 + 1 - m_2$. Lastly, $2d(T) - n_1 - n_2 - s = 2d(T) - n$ is the same.

Second, we show that each of three quantities, $2 + z(T_1) + m_1 - m_2$, $l_1 + 1 - m_2$, and $2d(T) - n$ could be a unique maximum, which implies that they are necessary components of the formula. Examples suffice.

For $2 + z(T_1) + m_1 - m_2 > \max\{l_1 + 1 - m_2, 2d(T) - n\}$, consider a 2-linear tree $T = L(T_1, 0, T_2)$ such that $l_1 = 2, l_2 = l_3 = 1, m_1 = 4$, and $m_2 = 2$. Then $2 + z(T_1) + m_1 - m_2 = 2 + \max\{2 - 1 - 1, 0\} + 4 - 2 = 4$, whereas $l_1 + 1 - m_2 = 2 + 1 - 2 = 1$, and $2d(T) - n = 3$.

For $l_1 + 1 - m_2 > \max\{2 + z(T_1) + m_1 - m_2, 2d(T) - n\}$, consider a 2-linear tree $T = L(T_1, 0, T_2)$ such that $l_1 = 3, l_2 = 2, l_3 = 1$, and $m_1 = m_2 = 1$. Then $l_1 + 1 - m_2 = 3 + 1 - 1 = 3$, whereas $2 + z(T_1) + m_1 - m_2 = 2 + \max\{3 - 2 - 1, 0\} + 1 - 1 = 2$, and $2d(T) - n = 2$.

For $2d(T) - n > \max\{l_1 + 1 - m_2, 2 + z(T_1) + m_1 - m_2\}$, consider a 2-linear tree $T = L(T_1, 0, T_2)$ such that $l_1 = l_2 = l_3 = 1$, and $m_1 = m_2 = 4$. Then $2d(T) - n = 5$, whereas $l_1 + 1 - m_2 = 1 + 1 - 4 = -2$, and $2 + z(T_1) + m_1 - m_2 = 2 + \max\{1 - 1 - 1, 0\} + 4 - 4 = 2$.

Third, we show that $m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$ is never a unique maximum; hence, it can be removed from the general formula for the 2-linear trees. Suppose $m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$ is a unique maximum among the four quantities in Theorem 3.24 for contradiction. Then,

$$m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor > 2 + z(T_1) + m_1 - m_2$$

which implies that $m_2 > z(T_1) + m_1 + 1 + \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$. Also,

$$m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor > 2d(T) - n \geqslant 2(m_1 + m_2 + 1) - n = m_1 + m_2 - l_1 - l_2 - l_3$$

which implies that $m_2 < 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor + l_1 + l_2 + l_3$.

Putting them together, we obtain

$$1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor + l_1 + l_2 + l_3 > z(T_1) + m_1 + 1 + \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$$

Then, by rearranging the terms,

$$1+l_1+l_2+l_3 > z(T_1)+m_1+1+2\times \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor \geqslant z(T_1)+m_1+(n_1-z(T_1)-1) = m_1+n_1-1$$

That is,

$$1 + l_1 + l_2 + l_3 > m_1 + n_1 - 1 = m_1 + (1 + l_1 + l_2 + l_3) - 1 = m_1 + l_1 + l_2 + l_3$$

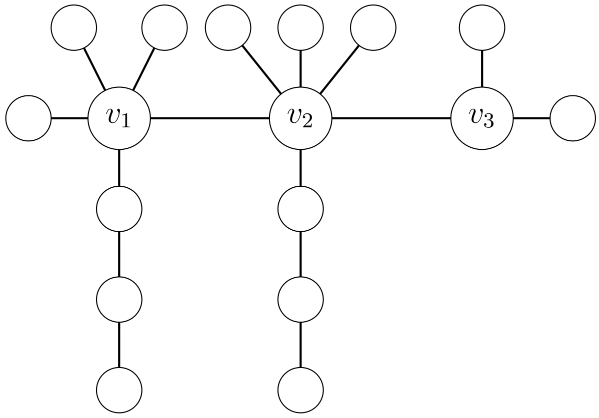
A contradiction arises, because $m_1 \geqslant 1$. Hence, $m_1 + 1 - \left\lfloor \frac{n_1 - z(T_1) - 1}{2} \right\rfloor$ cannot be a unique maximum, so it may be removed from the formula. In conclusion, the formula proposed in Corollary 5.8 is justified for 2-linear trees such that $s = 0$, and $P(T) = 3$. □

5.3 $U(T)$, HDV, and Parter vertices

It is known that when $M(T)$ is attained, every HDV in T is a Parter vertex [JS]. Thus, one may wonder whether this also happens when $U(T)$ is attained. It is certainly common that every HDV is Parter for some multiple eigenvalue to drive down the number of multiplicity 1 eigenvalues, but it is possible that some HDV does not have to be a Parter. Example 5.10 illustrates this.

Proposition 5.9. *For a tree T, when $U(T)$ is attained, it is possible that some HDV is not a Parter vertex for any multiple eigenvalue.*

Example 5.10. Let T be a 3-linear tree as follows. We denote the 3 HDV's in T by v_1, v_2, and v_3.



By the following LSP table and Theorem 2.8, the multiplicity list $(3, 3, 3, 2, 2, 2, 1, 1)$ is realizable, so $U(T) = 2$. Suppose the multiple eigenvalues are λ_1, λ_2, λ_3, β_1, β_2, and β_3 with $m(\lambda_i) = 2$ and $m(\beta_i) = 3$. Notice that the two $\hat{0}$'s from row b_3^+ end up in columns of multiplicity 0 in Table 5.10. The underlying eigenvalue assignment for for this LSP table has v_3 as no Parter for any multiple eigenvalue. In particular, according to Theorem 2.2, v_2 is a Parter for β_1, β_2, and β_3, and each β_i appears four times after the removal of v_2: on one of v_2's pendent neighbors, the 3-vertex arm, the subtree on the left, and the 3-path containing v_3 on the right. On the other hand, v_1 is a Parter for λ_1, λ_2, and λ_3, and each λ_i appears three times after the removal of v_1: on one of v_1's pendent neighbors, the 3-vertex arm, and the subtree on the right. Thus, v_3 is not used as a Parter for any multiple eigenvalue here.

		λ_1	β_1		λ_2	β_2		λ_3	β_3	
b_1^+	1	$\hat{1}$	1	0	$\hat{1}$	1	0	$\hat{1}$	1	0
s_1										
b_2^+	0	1	$\hat{1}$	0	1	$\hat{1}$	0	1	$\hat{1}$	1
s_2										
b_3^+	0	0	1	$\hat{0}$	0	1	$\hat{0}$	0	1	0
	1	2	3	0	2	3	0	2	3	1

Table 5.1: The LSP table for multiplicity list $(3, 3, 3, 2, 2, 2, 1, 1)$

Nevertheless, a weaker statement is still plausible.

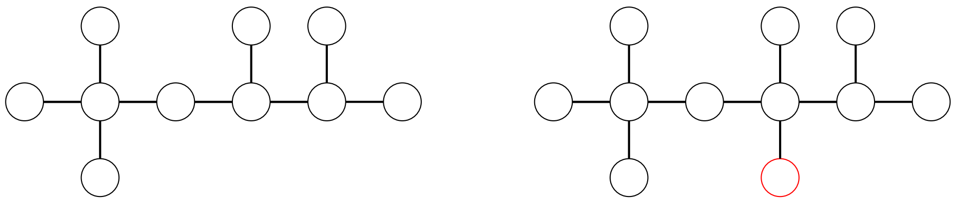
Conjecture 5.11. For a tree T, when $U(T)$ is attained, there exists some eigenvalue assignment such that every HDV is Parter for some multiple eigenvalue.

Appendix A

Examples of the possible changes in $U(T)$ in Table 3.3

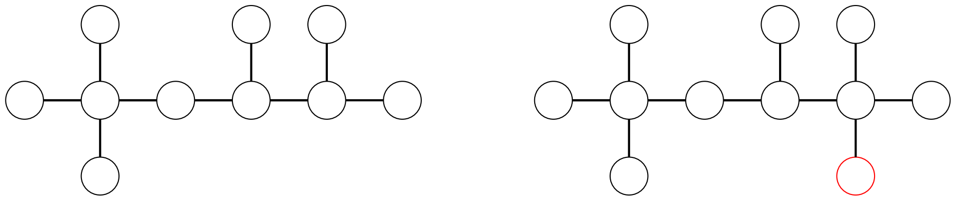
(1) At an HDV: $U(T)$ could stay the same or decrease.

Example A.1. Examples where $U(T')$ is less than $U(T)$:



$U(T) = 3$ $\qquad$ $U(T) = 2$

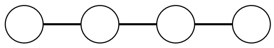
Example A.2. Examples where $U(T') = U(T)$



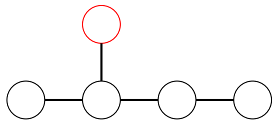
$U(T) = 3$ $\qquad$ $U(T) = 3$

(2) At a degree 2 vertex: $U(T)$ could stay the same or decrease.

Example A.3. Examples where $U(T')$ is less than $U(T)$:

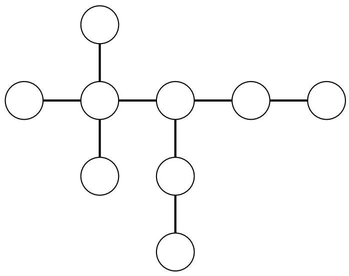


$U(T) = 4$

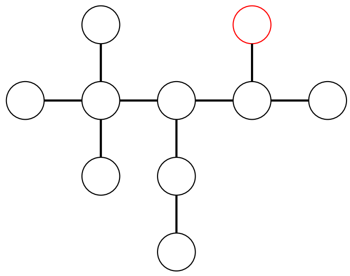


$U(T) = 3$

Example A.4. Examples where $U(T') = U(T)$:



$U(T) = 2$

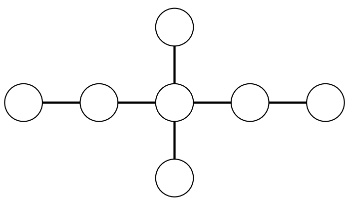


$U(T) = 2$

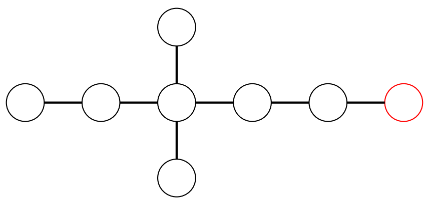
(3) At a pendent vertex:

3.1. The diameter increases:
$U(T)$ could increase, stay the same, or decrease.

Example A.5. Examples where $U(T')$ is greater than $U(T)$:

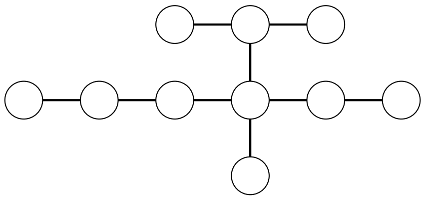


$d(T) = 5$
$U(T) = 3$

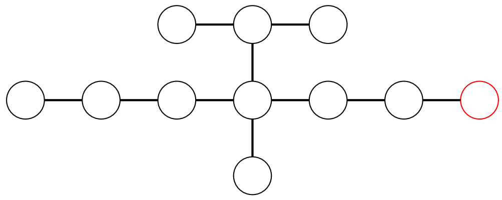


$d(T) = 6$
$U(T) = 4$

Example A.6. Examples where $U(T') = U(T)$:

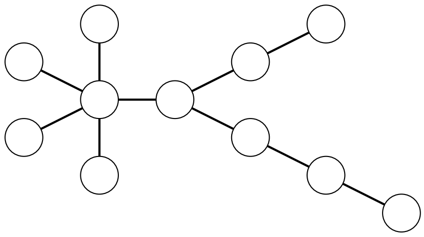


$d(T) = 6$
$U(T) = 3$

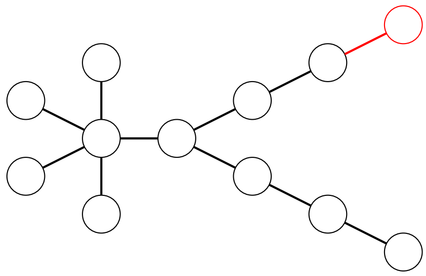


$d(T) = 7$
$U(T) = 3$

Example A.7. Examples where $U(T')$ is less than $U(T)$:



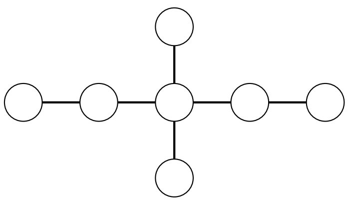
$d(T) = 6$
$U(T) = 3$



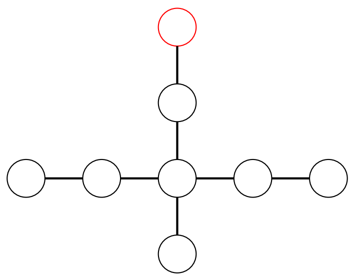
$d(T) = 7$
$U(T) = 2$

3.2. The diameter stays the same:
$U(T)$ could stay the same or decrease.

Example A.8. Examples where $U(T') = U(T)$:

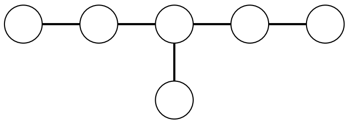


$d(T) = 5$
$U(T) = 3$

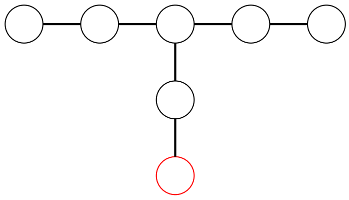


$d(T) = 5$
$U(T) = 3$

Example A.9. Examples where $U(T')$ is less than $U(T)$:



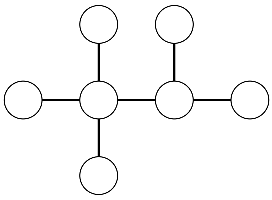
$d(T) = 5$
$U(T) = 4$



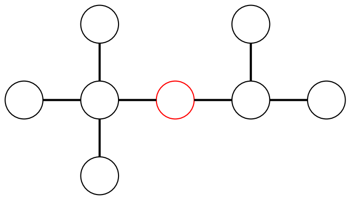
$d(T) = 5$
$U(T) = 3$

(4) Subdivision of an edge on the connecting path:
$U(T)$ could increase, stay the same, or decrease.

Example A.10. Examples where $U(T')$ is greater than $U(T)$:

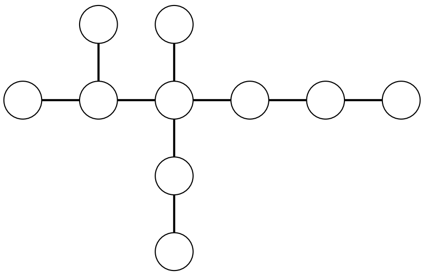


$d(T) = 4$
$U(T) = 2$

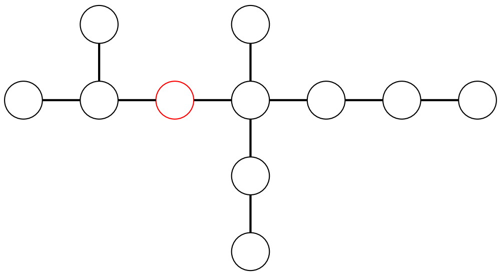


$d(T) = 5$
$U(T) = 3$

Example A.11. Examples where $U(T') = U(T)$:

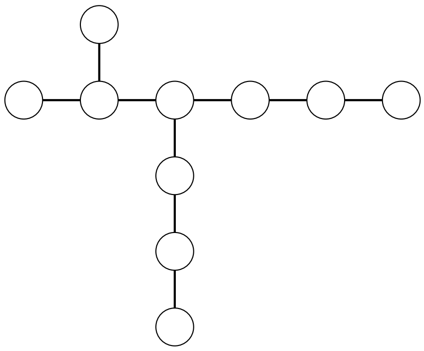


$d(T) = 6$
$U(T) = 3$

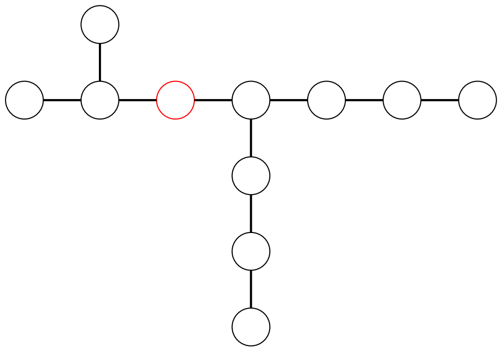


$d(T) = 7$
$U(T) = 3$

Example A.12. Examples where $U(T')$ is less than $U(T)$:



$d(T) = 7$
$U(T) = 4$



$d(T) = 7$
$U(T) = 3$

Bibliography

[JS] C.R. Johnson and C. Saiago, Eigenvalues, Multiplicities and Graphs, *Cambridge University Press*, (2018).

[JLW] C.R. Johnson, A. Li, and A. Walker, Ordered multiplicity lists for eigenvalues of symmetric matrices whose graph is a linear tree, *Discrete Mathematics*, **333** (2014), 39–55.

[JL-DS] C.R. Johnson, A. Leal Duarte, and C. Saiago, Inverse eigenvalue problems and lists of multiplicities of eigenvalues for matrices whose graph is a tree: the case of generalized stars and double generalized stars, *Linear Algebra Appl.*, **373** (2003), 311–330.

[In] M. Ingwersen, Notes on the computation of $U(T)$, *College of William & Mary, Research Experience for Undergraduate* (2018).

[JL-D] C.R. Johnson and A. Leal Duarte, The maximum multiplicity of an eigenvalue in a matrix whose graph is a tree, *Linear and Multilinear Algebra*, **46** (1999), 139-144.

[JL-DS03] C.R. Johnson, A. Leal Duarte, and C. Saiago, The Parter-Wiener theorem: refinement and generailzation, *SIAM J. Matrix Anal. Appl.*, **25:2** (2003), 352-361.

[JW] C.R. Johnson and T. Wakhare, The inverse eigenvalue problem for linear trees, *Discrete Mathematics*, submitted.

[JWW] C.R. Johnson, T. Wakhare and E. Wityk, The proportion of trees that are linear, *Discrete Mathematics*, under review.

[JM] C.R. Johnson and B. Mudrak. Multiplicity lists by diameter: All trees of diameter < 7, in progress.

[OEIS] The On-Line Encyclopedia of Integer Sequences® (OEIS®), oeis.org/.

[DIJ] W. Ding, M. Ingwersen, and C.R. Johnson. The Minimum Number of Multiplicity 1 Eigenvalues among Real Symmetric Matrices whose Graph is a 2-Linear Tree, *Linear and Multilinear Algebra*, submitted.