Properties of greedy trees

by

Valisoa Razanajatovo Misanantenaina

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics at Stellenbosch University

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof. Stephan Wagner

December 2014

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
Valisoa Razanajatovo Misanantenaina

Date: 2014/09/30 ....

Abstract

Properties of greedy trees

Valisoa Razanajatovo Misanantenaina

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MSc

December 2014

A greedy tree is constructed from a given degree sequence using a simple greedy algorithm that assigns the highest degree to the root, the second, the third, . . . , -highest degree to the root's neighbours, etc. This particular tree is the solution to numerous extremal problems among all trees with given degree sequence. In this thesis, we collect results for some distance-based graph invariants, the number of subtrees and the spectral radius in which greedy trees play a major role. We show that greedy trees are extremal for the aforementioned graph invariants by means of two different approaches, one using level greedy trees and majorization, while the other one is somewhat more direct. Finally, we prove some new results on greedy trees for additive parameters with specific toll functions.

Uittreksel

Eienskappe van die gulsige bome

Valisoa Razanajatovo Misanantenaina
Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MSc

Desember 2014

’n Gulsige boom word vanuit ’n gegewe graadry deur middel van ’n eenvoudige gulsige algoritme gebou, wat die hoogste graad aan die wortel toewys, die tweede-, derde-, . . . , -hoogste graad aan die wortel se bure, ens. Hierdie spesifieke boom is die oplossing van ’n groot aantal ekstremale probleme onder al die bome met gegewe graadry. In hierdie tesis beskou ons ’n versameling van resultate oor afstand-gebaseerde grafiekinvariante, die aantal subbome en die spektraalstraal waarin gulsige bome ’n belangrike rol speel. Ons bewys dat gulsige bome ekstremaal vir die bogenoemde grafiekinvariante is deur van twee verskillende tegnieke gebruik te maak: een met behulp van vlak-gulsige bome en majorering, en ’n ander metode wat effens meer direk is. Laastens bewys ons sommige nuwe resultate oor gulsige bome vir additiewe parameters met spesifieke tolfunksies.

Acknowledgements

I heartily express my profound gratitude to my supervisor Prof. Stephan Wagner, who continuously advised and guided me throughout the entire research period and during the writing of this thesis. It was kind of him to accept to translate the abstract of this thesis into Afrikaans.
My sincere thanks go to the Faculty of Science of Stellenbosch University and the African Institute for Mathematical Sciences (AIMS) for financially and materially supporting this project.
I am grateful to staff and colleagues in the Mathematics Division, especially Isaac who helped me a lot in improving my spoken and written English.
My appreciation also goes to my beloved Tsinjo for his love and support, and to my parents, my brothers, and all my family members for their constant encouragement and prayer.

"How great are God's riches! How deep are his wisdom and knowledge!...For all things were created by him, and all things exist through him and for him. To God be the glory for ever! Amen."
Romans, 11 : 33*a*/36

Dedications

To my parents on their 30^{th} *wedding anniversary,*

Contents

Declaration **i**

Abstract **ii**

Uittreksel **iii**

Acknowledgements **iv**

Dedications **v**

Contents **vi**

List of Figures **viii**

1 Introduction **1**

2 Basic notions **3**
2.1 Introduction 3
2.2 Greedy and level-greedy Trees 7

3 The majorization approach **13**
3.1 Preliminaries 13
3.2 Distance-based invariants 16
3.3 Number of subtrees 20

4 Direct approach **30**
4.1 The Wiener index 30
4.2 Spectral radius 39

5 Additive parameters **48**
5.1 Preliminaries 48

5.2 Additive parameters and level-greedy trees 50
5.3 Additive parameters and greedy trees 55

List of References **63**

List of Figures

2.1 A cycle C_4 and its subgraphs. 4
2.2 A star 5
2.3 A rooted tree T 6
2.4 A greedy tree. (Only the first eight vertices are labelled). 8
2.5 A rooted level-greedy tree. 9
2.6 An edge-rooted level-greedy tree. 9
2.7 A level-greedy forest. 12

3.1 Decomposition of T. 23
3.2 Example of the embedding from T_v to T_u for a rooted greedy tree. 26
3.3 A Volkmann tree $V_{17,3}$. 29

4.1 Labeling of a path and the components. 31
4.2 A star S_5 40
4.3 A greedy tree of order 12 with 4 leaves. 45
4.4 Trees of order 6. 47

5.1 T and T' in the proof of Lemma 5.2.6. 54
5.2 The tree T_H. 56
5.3 Rooted trees with degree sequence $(3, 3, 3, 2, 2, 1, 1, 1, 1, 1)$. 60

Chapter 1

Introduction

Part of extremal graph theory, one of the classical areas of graph theory, deals with the study of extremal (minimal or maximal) graphs in a given class of graphs with respect to certain graph invariants. In a nutshell, it studies how the structure of a graph influences its invariants. In this project, we particularly focus on the class of trees with given degree sequence, where the so-called "greedy trees" appear as the solution to numerous extremal problems.

In the literature, greedy trees first appeared as trees satisfying the breadth first search ordering [5, 18, 36]. In this thesis, we use a more recent definition by which greedy trees are constructed from a given degree sequence using a simple greedy algorithm that assigns the highest degree to the root, the second, the third, ..., -highest degree to the root's neighbor, etc [32]. Actually, this construction is a generalisation of "good trees" [28], also known as Volkmann trees, which are often extremal among trees of order n with maximum degree.

The main aim of this work is to collect and unify results involving greedy trees which have been obtained by several researchers, describe general proof techniques, and find new extremal properties of greedy trees. In order to achieve this goal, we will first define some basic notions of graph theory, and review in particular some properties of trees, which are the basic objects of our study. Moreover, we mention properties of greedy trees resulting from their definition. We also introduce the notion of rooted level greedy and edge-rooted level greedy trees which are in some sense the equivalent of greedy trees among trees with given level degree sequence. A core component of Chapter 2 is a direct proof of a theorem found in [24] that links greedy trees to level greedy trees. As we will see, this theorem will be useful in proving our results presented in subsequent chapters.

In Chapter 3, we prove the extremality of greedy trees by using the majorization approach on trees with given level degree sequence. We then use the theorem stated in Chapter 2 to extend the results to greedy trees. To illustrate this method, we will prove that greedy trees minimize the Wiener index, the hyper-Wiener index and the generalized Wiener index [24] and maximize the Harary index [30] and the number of subtrees [2]. For the latter graph invariant, we also compare trees with different degree sequences to rederive known results on trees of given order n, trees with given maximum degree and trees with given number of leaves.

In Chapter 4, we use a more direct approach to prove that greedy trees minimize the Wiener index [32] and maximize the spectral radius [5]. More specifically, this approach involves rearranging the edges in a tree and studying influence of this modification on the aforementioned graph invariants. Again, as for the number of subtrees, we can extend the results by comparing trees with different degree sequences.

Finally, in Chapter 5 we prove some new results on greedy trees. As it turns out, greedy trees are extremal for additive parameters with specific toll functions. More precisely, if the additive parameter has a toll function that depends only on the order of the tree, and which is increasing concave (resp. decreasing convex), then greedy trees minimize (resp. maximize) it. To this end, we combine the two methods, i.e., we use majorization on the order of the rooted subtrees of the trees to prove the extremality of level greedy trees and then rearrange the edges to see the effect on the parameters to conclude that the greedy trees are indeed extremal.

Chapter 2

Basic notions

2.1 Introduction

A graph is defined as a pair of sets $(V(G), E(G))$, where the elements of $V(G)$ are called vertices of G, and the elements of $E(G)$, which are two-element subsets of $V(G)$, are called edges of G. $|V(G)|$ is the order of G and $|E(G)|$ its size. For simplicity an edge $\{u, v\}$ will be denoted by uv.
In a graphical representation, each vertex is indicated by a point and each edge by a line joining the points which represent its edges, see Figure 2.1. A *simple undirected graph* does not contain loops, i.e., edges that join a vertex to itself, nor multiple edges between two vertices, and its edges do not have a direction. All graphs considered throughout the whole thesis will be simple and undirected graphs.
Two graphs G and G' are identical (written $G = G'$), if $V(G) = V(G')$ and $E(G) = E(G')$. However, there are also instances of graphs which are so similar that they could be represented by the same diagram. In such a situation, the graphs are said to be isomorphic, formally defined as follows:

Definition 2.1.1. Two graphs G and G' are *isomorphic* ($G \cong G'$) if there is a bijection $\ell : V(G) \to V(G')$ such that $uv \in E(G) \iff \ell(u)\ell(v) \in E(G')$.

Definition 2.1.2. A graph G' is a *subgraph* of G if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. If in particular $V(G) = V(G')$, then G' is called a spanning subgraph of G.

Remark 2.1.3. Let $\{v_1, v_2, \ldots, v_k\}$ be a subset of the vertices of a graph G.

The subgraph of G which results from removing the vertices $v_1, v_2, \ldots, v_k$ and the edges containing them from G will be denoted by $G - \{v_1, v_2, \ldots, v_k\}$. The subgraph of G which results from removing the edges $v_1w_1, \ldots, v_kw_k$ will be denoted by $G - \{v_1w_1, \ldots, v_kw_k\}$.

Definition 2.1.4. A graph (V, E), where $V = \{v_1, v_2, \ldots, v_n\}$, $E = \{v_1v_2, v_2v_3, \ldots, v_{n-1}v_n\}$, with $n \geq 1$, and where $v_i \neq v_j$ if $i \neq j$, is called a path and denoted by P_n.
Let $n \geq 3$ and P_n be a path, with the same notation as before. The graph $(V(P_n), E(P_n) \cup \{v_1v_n\})$ is called a cycle and denoted C_n.
The length of a path (cycle) is the number of edges in the path (cycle).

Example 2.1.5. The first two pictures are isomorphic representations of a cycle of length 4, and the last two show some of its subgraphs.

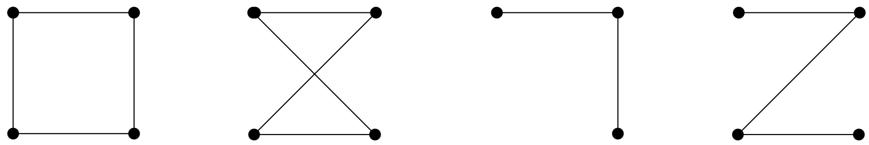


Figure 2.1: A cycle C_4 and its subgraphs.

Now that the notions of paths and cycles have been introduced, we can talk about connectivity and acyclic graphs.

Definition 2.1.6. A graph G is *connected* if every pair of its vertices is joined by a path. Otherwise, G is disconnected.
A *forest* is an *acyclic* graph, i.e., it does not contain a cycle. A *tree* is a connected forest.

Definition 2.1.7. For any vertex $v \in G$, the set $\{u \in V(G) : uv \in E(G)\}$ is called the *neighbourhood* of v and denoted by $N(v)$. The cardinality of $N(v)$, denoted by $d_G(v)$ (we write $d(v)$ if there is no ambiguity), is the *degree* of v. A vertex of degree 1 is called a *leaf*.
We denote by $(d_1, d_2, \ldots, d_n)$ the degree sequence of a graph G, where d_i stands for the degree of the i-th vertex and we may assume that $d_1 \geq d_2 \cdots \geq d_n$.

The following proposition links the degree sequence of a graph and its number of edges.

Proposition 2.1.8. For a graph with m edges, we have

$$\sum_{v \in V(G)} d(v) = 2m.$$

Proof. An edge contributes to the degree of two vertices. Thus when we take the sum of all degrees, each edge is counted exactly twice. Therefore the statement holds. □

Definition 2.1.9. The distance $d(u, v)$ between two vertices u and v is the length of the shortest path between u and v in G.

Example 2.1.10. This is an example of a star, which is a tree of order n with $n - 1$ leaves. Considering Figure 2.2, we can see that this star has 8 leaves and one vertex r called center such that $d(r) = 8$. Its degree sequence is therefore $(8, 1, 1, 1, 1, 1, 1, 1, 1)$. The sum of its degrees is $16 = 2 * 8$ which confirms Proposition 2.1.8. Moreover, the distance between the center r and one of the leaves is 1, while the distance between two leaves is 2.

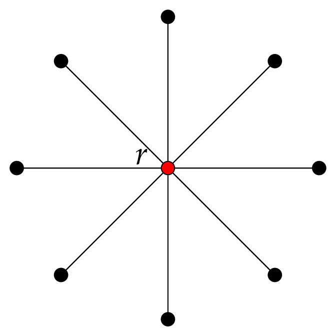


Figure 2.2: A star

In this thesis, we specifically focus on trees. We first collect some well-known facts and properties of trees (see for example [4, 6]), which will be useful later on. It is easy to see that any two vertices of a tree T are connected by a unique path. This unique path connecting two vertices u, v is denoted by $P_T(u, v)$ ($P(u, v)$ if it is clear what tree is considered). A *rooted tree* (T, r) is obtained by specifying one vertex r as the root. The height of a vertex v of a rooted tree T with root r, denoted $h_T(v)$, is the distance between r and v, while the height of T is just the greatest height of its vertices. For any two different vertices u, v in a rooted tree (T, r), we say that v is a successor of u (or u is an ancestor of v) if $P_T(r,u) \subset P_T(r,v)$. If in addition u and v are adjacent to each other, we say that v is a child of u (or u is a parent of v). Two vertices u, v are siblings of each other if they share the same parent. We say that a vertex v is at level i if $d_T(r,v) = i$.

Example 2.1.11. This is an example of a rooted tree T of height 3. As we can see in Figure 2.3, the vertex v has 3 children v_1, v_2, v_3.

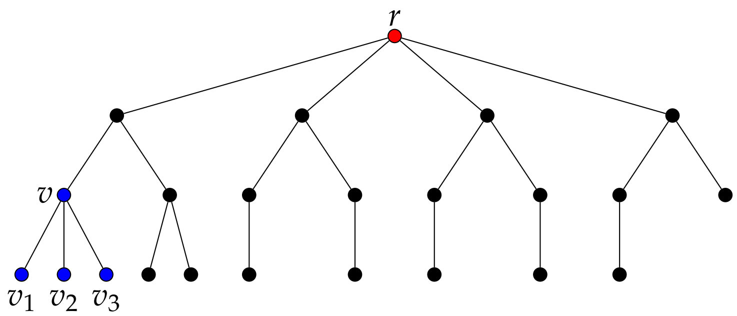


Figure 2.3: A rooted tree T

Theorem 2.1.12. *In a tree T, the number of edges m is equal to $n-1$, where n is the order of T.*

Proof. Let us prove the statement by induction on n. If $n = 1$, T is a single vertex, so $m = 0$.
Suppose the theorem is true for all trees on $n < \ell$ vertices, and let T be a tree on ℓ vertices. Let $uv \in E(T)$. Then $T - uv$ is disconnected and we obtain two components T_1 and T_2, which are trees on fewer vertices than ℓ. Therefore $m = |E(T_1)| + |E(T_2)| + 1 = |V(T_1)| - 1 + |V(T_2)| - 1 + 1 = n - 1$. □

Proposition 2.1.13. A sequence $(d_1, d_2, \ldots, d_n)$ of positive integers is a degree sequence of an n-vertex tree if and only if

$$\sum_{i=1}^{n} d_i = 2(n-1) \tag{2.1.1}$$

Proof. If $(d_1, d_2, \ldots, d_n)$ is a degree sequence of an n-vertex tree, then by using Theorem 2.1.12 and Proposition 2.1.8, we have $\sum_{i=1}^{n} d_i = 2(n-1)$.

Let us prove by induction that if $(d_1, d_2, \ldots, d_n)$ is a degree sequence of positive integers satisfying (2.1.1), then it is a degree sequence of an n-vertex tree. For $n = 2$, the only possible sequence satisfying (2.1.1) is $(1, 1)$, which is a degree sequence of a tree with two vertices and a single edge. Suppose that the statement is true for $n = \ell$, we have to prove that it holds for $n = \ell + 1$. Note that if $(d_1, d_2, \ldots, d_{\ell+1})$ is a sequence satisfying (2.1.1), then there exist i, j such that $d_i > 1$ and $d_j = 1$ (for otherwise the sum $d_1 + d_2 + \cdots + d_{\ell+1}$ would be at least $2(\ell + 1) > 2\ell$ or at most $\ell + 1 < 2\ell$). From this sequence, we can obtain a sequence of ℓ positive integers $(d_1, d_2, \ldots, d_i - 1, d_{i+1}, \ldots, d_{j-1}, d_{j+1}, \ldots, d_{\ell+1})$ satisfying (2.1.1). However by the induction hypothesis, this sequence of ℓ integers corresponds to the degree sequence of an ℓ-vertex tree, and the operation to get back to the original degree sequence is to attach a pendant edge to a vertex u with $d(u) = d_j - 1$. Thus the statement holds. □

2.2 Greedy and level-greedy Trees

In this section, we introduce greedy trees, which are the main subject of the thesis. We also state some properties and theorems which are useful in the next chapters. The greedy tree has been defined in various equivalent ways in previous papers [2, 24, 36]. In this thesis, we define the greedy tree as follows:

Definition 2.2.1. Given a sequence satisfying Proposition 2.1.13, the greedy tree is constructed by the following "greedy algorithm":

(i) Label the vertex with the largest degree v (the root);

(ii) Label the neighbours of v as $v_1, v_2, \ldots$, and assign the largest degrees available to them such that $d(v_1) \geq d(v_2) \geq \cdots$;

(iii) Label the neighbours of v_1 (except v) as $v_{11}, v_{12}, \ldots$, and then do the same for $v_2, v_3, \ldots$;

(iv) Repeat (ii) and (iii) for all the newly labeled vertices. Always start with the neighbours of the labeled vertex with the largest degree whose neighbours are not labeled yet.

Example 2.2.2. Let us construct a greedy tree with the following degree sequence: $(4, 4, 3, 3, 3, 3, 3, 3, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$. We choose the vertex v such that $d(v) = 4$ on level 0. Then v is adjacent to four vertices v_1, v_2, v_3, v_4 on level 1 whose degrees are $4, 3, 3, 3$ respectively. Then v_1, v_2, v_3, v_4 are adjacent to three, two, two and two vertices respectively on level 2. These vertices are denoted by $v_{11}, v_{12}, v_{13}, v_{21}, \ldots, v_{42}$, and their degrees are $3, 3, 3, 2, 2, 2, 1, 1, 1$ respectively. Finally $v_{11}, v_{12}, v_{13}, v_{21}, v_{22}, v_{31}$ are adjacent to two, two, two, one, one and one vertex respectively on level 3. The degrees of these vertices are all 1.

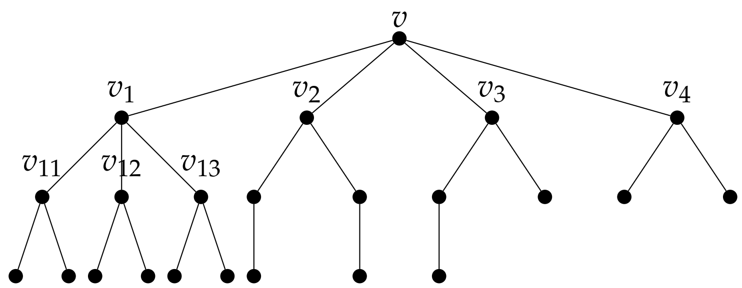


Figure 2.4: A greedy tree. (Only the first eight vertices are labelled).

From the definition of the greedy tree, we get:

Lemma 2.2.3 ([31, 32])**.** A rooted tree T with a given degree sequence is a greedy tree with the canonical root if:

(i) the root r has the largest degree;

(ii) the heights of any two leaves differ by at most 1;

(iii) for any two vertices u and v, if $h_T(u) < h_T(v)$, then $d(u) \geq d(v)$;

(iv) for any two vertices u and v of the same height, $d(u) > d(v)$ implies that $d(u') \geq d(v')$ for any successors u' of u and v' of v of the same height;

(v) for any two vertices u and v of the same height, $d(u) > d(v)$ implies that $d(u') \geq d(v')$ for any siblings u' of u and v' of v of the same height.

Definition 2.2.4. For $i = 0, 1, \ldots, H$, let multisets $\{a_{i,1}, a_{i,2}, \ldots, a_{i,\ell_i}\}$ of non-negative numbers be given such that $\ell_0 = 1$ (resp. $\ell_0 = 2$) and

$$\ell_{i+1} = \sum_{j=1}^{\ell_i} a_{i,j}.$$

Assume that the elements of each multiset are sorted, i.e. $a_{i,1} \geq a_{i,2} \geq \cdots \geq a_{i,\ell_i}$. The level-greedy tree (with height H) corresponding to this sequence of multisets is the rooted (resp. edge-rooted) tree whose j-th vertex at level i has outdegree $a_{i,j}$.

Example 2.2.5. Figure 2.5 shows a rooted level-greedy tree with the following outdegree sequences on each level : $\{(3), (3, 2, 2), (3, 2, 2, 1, 1, 1)\}$.

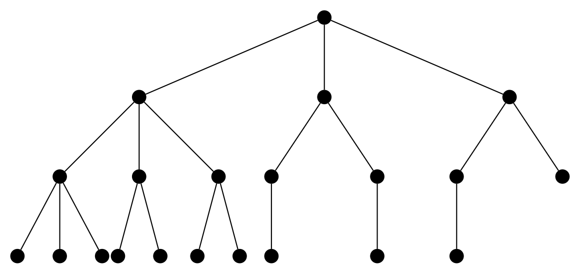


Figure 2.5: A rooted level-greedy tree.

Example 2.2.6. Figure 2.6 shows an edge-rooted tree with the following outdegree sequences: $\{(2, 1), (3, 2, 2), (3, 2, 2, 1, 1, 1)\}$.

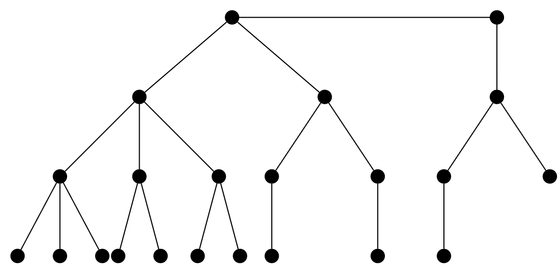


Figure 2.6: An edge-rooted level-greedy tree.

The following theorem found in [24] relates greedy trees and level-greedy trees. The proof is based on the "semi-regular" property defined in [28]. Here, we present an alternative proof using a direct approach.

Theorem 2.2.7. *If a tree is level-greedy with respect to any possible choice of root or edge-root, then it is greedy.*

Proof. Let T be a level-greedy tree with respect to any root and edge-root. Let us choose the root r in the following way:

- Take r as the vertex with maximum degree.
- If there are several vertices of highest degree, we sum all the degrees of their neighbours and take the one which maximizes this sum.
- If there is still more than one vertex satisfying this condition, we continue summing the degrees of all the vertices at distance $2, 3, \ldots, k$. At a certain point, we end up with one unique vertex which satisfies all the conditions and take it as our root. Otherwise, we can take any of the vertices that are left.

By our hypothesis, we know that T is level-greedy with respect to r. For T to be greedy, it is left to show that the degree of a vertex on a higher level is always greater or equal to the degree of a vertex on a lower level. Without loss of generality, we can consider two vertices u, v on consecutive levels, such that u has the minimum degree on level k and v has the maximum degree on level $k+1$.

Now let us suppose that $d(u) < d(v)$. We consider the middle edge e of the path between u and v as our new root. The endpoints of e are r and r_1, which are respectively the ancestors of u and v. Moreover u and v are now on the same level k. The following cases can occur:

Case 1 : $d(r) \neq d(r_1)$

From the choice of r, we have $d(r) > d(r_1)$. Since by hypothesis T is level-greedy with respect to e, the vertices are ordered in such a way that successors of r have greater (or at least equal) degree than successors of r_1 at the same level. But u is a successor of r, v a successor of r_1, and $d(u) < d(v)$, i.e., the inequality goes in the opposite direction, which is a contradiction to T being level-greedy.

Case 2 : $d(r) = d(r_1)$

Let us consider the successors of r and r_1.

- If at some level $\ell < k$ the sum of the degrees of the successors of r does not equal the sum of the degrees of the successors of r_1, then by the choice of r, we know that the sum of the degrees of the successors of r is greater. This means that the successors of r have higher degrees. With a similar reasoning as in Case 1, this yields a contradiction.
- If at all levels $\ell < k$, we have equality of the sum of the degrees of the successors of r and r_1, then at level k, we have the same number of successors of r and r_1, denoted respectively by $v_1^r, \ldots, v_j^r$ and $v_1^{r_1}, \ldots, v_j^{r_1}$. However, by the choice of r, the sum of the degrees of the successors of r is greater or equal to the analogous sum for r_1. If it is greater, we can apply a similar reasoning as before. Otherwise, since we have $d(u) < d(v)$, there exists h such that $d(v_h^{r_1}) < d(v_h^r)$. This implies that the vertices at level k will not be ordered according to their degrees, which is a contradiction to T being level-greedy.

Therefore, we have $d(u) \geq d(v)$, which shows that T is a greedy tree.

□

For the rest of this chapter, we are going to generalize the notion of greedy trees to greedy forests, according to [2]. These definitions and concepts will be required in subsequent chapters.

Definition 2.2.8. Let F be a rooted forest where the maximum height of any component is k. The level degree sequence of F is the sequence

$$D = (V_0, \ldots, V_k), \tag{2.2.1}$$

where, for any $0 \leq i \leq k$, V_i is the non-increasing sequence formed by the degrees of the vertices of F on the i-th level.

Definition 2.2.9. Let $D = ((a_{0,1}, \ldots, a_{0,k_0}), (a_{1,1}, \ldots, a_{1,k_1}), \ldots, (a_{n,1}, \ldots, a_{n,k_n}))$ be the level degree sequence of a forest. In a similar way as for the level greedy tree, if we assume that the elements of each multiset are sorted, i.e. $a_{i,1} \geq a_{i,2} \geq \cdots \geq a_{i,k_i}$ for $i = 0, \ldots, n$, the level greedy forest is the rooted forest whose j-th vertex at level i has outdegree $a_{i,j}$.

Example 2.2.10. Figure 2.7 shows a level-greedy forest whose outdegree sequences on each level are: $\{(2,2,2),(3,2,2,1,1,1),(2,2,2,1,1,1,1,1,1)\}$.

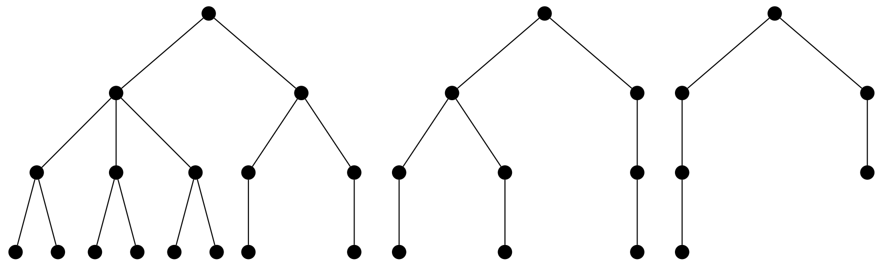


Figure 2.7: A level-greedy forest.

Remark 2.2.11. A connected level greedy forest is a level greedy tree. This definition coincides with Definition 2.2.4.

Let T_1 and T_2 be two rooted trees. For $j \in \{1,2\}$ and $\ell \geq 0$ let $\mathcal{V}_{\ell,j} = \{v^\ell_{j,1}, \ldots, v^\ell_{j,k_{\ell,j}}\}$ be the set of vertices at level ℓ of T_j. We write $T_1 \triangleright T_2$ if the height of T_1 is at least that of T_2 and for any $\ell \geq 0$, we have

$$\min\{d(v^\ell_{1,1}), \ldots, d(v^\ell_{1,k_{\ell,1}})\} \geq \max\{d(v^\ell_{2,1}), \ldots, d(v^\ell_{2,k_{\ell,2}})\}$$

if $\mathcal{V}_{\ell,2}$ is not empty. The relation $\triangleright$ is easily seen to be transitive.

Remark 2.2.12. Let F be a rooted forest. F is a level greedy forest if and only if its components can be labeled as $F_1, \ldots, F_t$ such that each of $F_1, \ldots, F_t$ is a level greedy tree and $F_1 \triangleright \cdots \triangleright F_t$. A tree T rooted at v is a level greedy tree if and only if $T - v$ is a level greedy forest.

Chapter 3

The majorization approach

For many graph invariants, such as the Wiener index, number of subtrees, and spectral moments, the path and the star attain the extremal values (minimum or maximum) among all trees with the same order. A lot of research [2, 18, 24, 32, 36] has been conducted to find such extremal trees with various restrictions such as fixed maximum degree, fixed number of leaves and fixed degree sequence. For the last restriction, it turns out that the greedy tree either minimizes or maximizes the aforementioned graph invariants. In this chapter, we prove this fact in several instances by applying the theory of majorization on rooted and edge-rooted level greedy trees, and consequently extending the result to greedy trees. We will focus on distance-based graph invariants and the number of subtrees. The majorization approach can be used as well with spectral moments (see [1]).

3.1 Preliminaries

We denote by S_n the set of all permutations of $\{1, \ldots, n\}$. Let $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ be sequences of nonnegative numbers. We say that B *majorizes* A if for all $1 \leq k \leq n$ we have

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i.$$

If for any $\sigma \in S_n$ the sequence B majorizes $(a_{\sigma(1)}, \ldots, a_{\sigma(n)})$, then we write

$$A \preceq B.$$

Remark 3.1.1. Note that majorization is transitive, i.e.,

$$A \preceq B, B \preceq C \Longrightarrow A \preceq C.$$

Remark 3.1.2. Let $\sigma \in S_n$ be such that $a_{\sigma(1)} \geq \cdots \geq a_{\sigma(n)}$. It is easy to see that $(a_{\sigma'(1)}, \ldots, a_{\sigma'(n)}) \preceq (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$ for any $\sigma' \in S_n$. Thus, $A \preceq B$ is equivalent to $(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) \preceq B$.

Remark 3.1.3. In some papers (see [20]), in the definition of majorization we have equality

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i,$$

for the last n.

Lemma 3.1.4 (Rearrangement inequality [11, Theorem 368])**.** For any two sequences of real numbers $a_1 \geq a_1 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ and for any $\sigma \in S_n$, we have

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \geq a_{\sigma(1)} b_1 + a_{\sigma(2)} b_2 + \cdots + a_{\sigma(n)} b_n.$$

Proof. Let us first consider the case when $n = 2$. Suppose $a_1 \geq a_2$ and $b_1 \geq b_2$, then

$$(a_1 - a_2)(b_1 - b_2) \geq 0,$$

which implies

$$a_1 b_1 + a_2 b_2 \geq a_2 b_1 + a_1 b_2.$$

For the general case, let us prove the statement by contradiction. Let $a_1 \geq a_1 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$. Suppose that there exists a permutation σ different from the identity such that the sum

$$a_{\sigma(1)} b_1 + a_{\sigma(2)} b_2 + \cdots + a_{\sigma(n)} b_n$$

is maximal (since there are only finitely many permutations, the maximum must exist). Then there exists $j \in \{1, ..., n-1\}$ such that $\sigma(j) \neq j$ and $\sigma(i) = i$ for all $i \in \{1, ..., j-1\}$. Hence $\sigma(j) > j$ and there exists $k \in \{j+1, ..., n\}$ with $\sigma(k) = j$. However, from the case $n = 2$, we have $a_j b_j + a_{\sigma(j)} b_k \geq a_{\sigma(j)} b_j + a_j b_k$. It implies that the sum of products can only increase if we instead pair a_j with b_j and $a_{\sigma(j)}$ with b_k (unless $a_j = a_{\sigma(j)}$ or $b_j = b_k$, in which case either we can interchange j and $\sigma(j)$ or j and k without affecting the sum). In any case, we obtain a new permutation σ' with $\sigma'(i) = i$ for $i \in \{1, 2, \ldots, j\}$, and the sum does not decrease. After a finite number of iterations, we always reach the identity permutation. $\square$

The following lemma, found in [24], is one of the keys to proving our theorems later on.

Lemma 3.1.5. Suppose that we have nonnegative sequences such that $(a'_1, \dots a'_n) \preceq (a_1, \dots, a_n)$ and $(b'_1, \dots, b'_n) \preceq (b_1, \dots, b_n)$. Then we have

$$a'_1 b'_1 + \dots + a'_n b'_n \leq a_1 b_1 + \dots + a_n b_n.$$

Proof. Suppose that the sequences $(a'_1, \dots a'_n)$ and $(b'_1, \dots, b'_n)$ satisfy the conditions in Lemma 3.1.5, and that the sum

$$a'_1 b'_1 + \dots + a'_n b'_n$$

is a maximum, which exists since the inequalities define a compact set. By Lemma 3.1.4, we may assume that $a'_1 \geq \dots \geq a'_n$ and $b'_1 \geq \dots \geq b'_n$. If $(a_1, \dots, a_n) \neq (a'_1, \dots, a'_n)$, let h be the smallest index such that $a_1 + \dots + a_h > a'_1 + \dots + a'_h$; h exists by the definition of majorization. Let $\varepsilon > 0$ be the difference between the two sides of the inequality. Replacing a'_h by $a'_h + \varepsilon$ and a'_{h+1} by $a'_{h+1} - \varepsilon$, we obtain a new $(2n)$-tuple of numbers satisfying the majorization, while the sum $a'_1 b'_1 + \dots + a'_n b'_n$ will not decrease, since it changes by $\varepsilon(b'_h - b'_{h+1}) \geq 0$. Note that during this process, the index h will increase. We can repeat the same argument until h reaches n and we end up with $a_1 = a'_1, \dots, a_n = a'_n, b_1 = b'_1, \dots, b_n = b'_n$. □

Example 3.1.6. Let us consider the following sequences $(4,3,3), (5,3,2), (3,2,1), (2,2,2)$. We have

$$(4,3,3) \preceq (5,2,3) \preceq (5,3,2),$$
$$(2,2,2) \preceq (3,1,2) \preceq (3,2,1),$$

and

$$4 \cdot 2 + 3 \cdot 2 + 3 \cdot 2 = 20,$$
$$5 \cdot 3 + 2 \cdot 2 + 3 \cdot 1 = 22,$$
$$5 \cdot 3 + 3 \cdot 1 + 2 \cdot 2 = 22,$$
$$5 \cdot 3 + 3 \cdot 2 + 2 \cdot 1 = 23.$$

This confirms Lemma 3.1.5.

3.2 Distance-based invariants

Mathematical chemists often use a class of graph invariants, known as topological indices in the chemical literature, to establish a connection between a compound's molecular graph and its characteristics. Among those topological indices, a number of distance-based invariants received great attention. One of the most classical and most thoroughly studied distance-based invariants is the *Wiener index* of a graph G, introduced by Wiener [33], which is the sum of the distances between all pairs of vertices, denoted by

$$W(G) = \sum_{\{v,w\} \subseteq V(G)} d(v,w), \tag{3.2.1}$$

where $d(v,w)$ is the distance between two vertices $v,w \in V(G)$.
Let us mention also some invariants generalizing the Wiener index. The *generalised Wiener index* (see [8]) is defined as follows:

$$W_\alpha(G) = \sum_{\{v,w\} \subseteq V(G)} d(v,w)^\alpha, \tag{3.2.2}$$

where α is some real number.
Let v and w be vertices of a tree, and denote by $n(v,w)$ the number of vertices u (including v itself) for which the path from u to w passes through v. Then the *hyper Wiener index* [16] is defined as

$$WW(G) = \sum_{\substack{\{v,w\} \subseteq V(G) \\ v \neq w}} n(v,w)n(w,v). \tag{3.2.3}$$

The hyper-Wiener index can also be expressed as follows:

$$WW(G) = \sum_{\{v,w\} \subseteq V(G)} \binom{d(v,w)+1}{2}. \tag{3.2.4}$$

In fact, $n(v,w)n(w,v)$ counts the number of paths $P(u,u')$ such that $P(u,u')$ contains $P(v,w)$. On the other hand, the number of paths contained in $P(u,u')$ is $\binom{d(u,u')+1}{2}$.
In order to characterize the tree which minimizes the aforementioned graph invariants, we study a graph invariant $p_k(T)$ introduced in [24].

Definition 3.2.1. We denote by $p_k(T)$ the number of pairs of vertices $\{u,v\}$ in T, with $u \neq v$, such that $d(u,v) \leq k$.

Example 3.2.2. Let us consider the cases $k = 1$ and $k = 2$:

- $p_1(T) = n - 1$, where T is a tree of order n. In fact, we observe that $d(u, v) \leq 1 \iff d(u, v) = 1$, and the number of pairs satisfying this property is the same as the number of edges in T.
- $p_2(T) = \frac{1}{2}\sum_{v\in V(T)} d(v)^2$. A pair $\{u, v\}$ satisfies $d(u, v) \leq 2$ if and only if it satisfies $d(u, v) = 2$ or $d(u, v) = 1$. Thus

$$\begin{aligned}
p_2(T) &= \frac{1}{2}\left(\sum_{w\in V(T)} \sum_{\substack{(u,v) \\ u\neq v, uw, wv\in E(T)}} 1 + \sum_{\substack{(u,v) \\ uv\in E(T)}} 1\right) \\
&= \frac{1}{2}\left(\sum_{w\in V(T)} d(w)(d(w) - 1) + \sum_{u} d(u)\right) \\
&= \frac{1}{2}\sum_{w\in V(T)} d(w)^2.
\end{aligned}$$

Theorem 3.2.3 ([24]). *Let $d_1 \geq d_2 \geq \ldots d_n$ be positive integers satisfying Proposition 2.1.13, and let k be another arbitrary positive integer. Among all trees with degree sequence $(d_1, d_2, \ldots, d_n)$, the greedy tree has the largest number $p_k(T)$.*

In order to prove Theorem 3.2.3, we will mention two lemmas which characterize the tree maximizing $p_k(T)$ among rooted and edge rooted trees. Let D be a level degree sequence of a rooted tree. We denote by $\mathbb{T}_r(D)$ the set of all trees that have D as a level degree sequence. In particular, $G(D)$ is the level greedy tree with level sequence D.

Lemma 3.2.4 ([24]). If $T \in \mathbb{T}_r(D)$, then $p_k(T) \leq p_k(G(D))$.

Proof. Let us compare the number of pairs $\{u, v\}$ satisfying $d(u, v) \leq k$ for a tree $T \in \mathbb{T}_r(D)$. Suppose u is at level i and v is at level j. Two cases can be considered.

Case 1: If $i + j \leq k$, then automatically $d(u, v) \leq k$. The number of pairs, at those levels, satisfying the condition is the same for any T.

Case 2: Otherwise, the condition is satisfied if and only if u and v have the same ancestor at level $r = \lceil (i + j - k)/2 \rceil$. Let us count those pairs. We denote by $w_1, w_2, \ldots, w_m$ the vertices at level r, and by $x_1, \ldots, x_m$ and

$y_1, \ldots, y_m$ the number of their respective successors at level i and level j. The number of pairs that we have to count is

$$x_1 y_1 + \cdots + x_m y_m$$

for $i \neq j$. If we count the number of pairs which will satisfy $d(u, v) \leq k$ for $i = j$, it is $\binom{x_1}{2} + \cdots + \binom{x_m}{2}$, and since we consider trees with the same outdegrees at each level, the sum $x_1 + \cdots + x_m$ is constant, so only the sum of the squared terms $x_1^2 + \cdots + x_m^2$ really matters.
Now, let $(x_1, \ldots, x_m)$, $(y_1, \ldots, y_m)$ be the number of successors at level i and j for $G(D)$. By the definition of $G(D)$, $(x_1, \ldots, x_m)$, $(y_1, \ldots, y_m)$ are non-decreasing sequences. Let $(x_1', \ldots, x_m')$, $(y_1', \ldots, y_m')$ be the corresponding sequences for any tree $T \in \mathbb{T}_r(D)$. Since T and $G(D)$ have the same level degree sequences, by Remark 3.1.2, we have $(x_1', \ldots x_m') \preceq (x_1, \ldots, x_m)$ and $(y_1', \ldots, y_m') \preceq (y_1, \ldots, y_m)$. Therefore, Lemma 3.1.5 gives us that

$$x_1 y_1 + \cdots + x_m y_m \geq x_1' y_1' + \cdots + x_m' y_m'. \tag{3.2.5}$$

Thus $G(D)$ indeed maximizes the number of pairs of vertices at levels i and j that have a common ancestor at level r, for every i and j. Hence $p_k(T) \leq p_k(G(D))$. □

Let D_e be a level degree sequence of an edge rooted tree. We denote by $\mathbb{T}_e(D)$ the set of all trees that have D_e as a level degree sequence, and $G(D_e)$ is the edge rooted level greedy tree with level sequence D_e. By an analogous proof as in Lemma 3.2.4, we have

Lemma 3.2.5 ([24])**.** If $T \in \mathbb{T}_e(D)$, then $p_k(T) \leq p_k(G(D_e))$.

Proof of Theorem 3.2.3. Let T be any tree that maximizes $p_k(T)$. Let $r \in V(T)$ be a root of T. Lemma 3.2.4 ensures that T can be chosen to be a level-greedy tree with the same outdegree sequence, such that $p_k(T)$ will not decrease. In a similar way, if we take $e \in E(T)$ as an edge-root of our tree, by Lemma 3.2.5, replacing T by a level-greedy tree with the same outdegree sequence will improve $p_k(T)$. We iterate the process for different vertex and edge roots in T. Note that each step of this process strictly increases $p_4(T)$: to see why, consider the first level ℓ where the outdegrees change when we replace our rooted (or edge-rooted) tree by a level-greedy tree. Then if $x_1, x_2, \ldots, x_m$ and $x_1', x_2', \ldots, x_m'$ are the number of successors at level

$\ell + 1$ of the vertices at level $\ell - 1$ in the greedy tree and the original tree respectively, we have $(x_1, x_2, \ldots, x_m) \neq (x'_1, x'_2, \ldots, x'_m)$ and the inequality (3.2.5) becomes strict. Since $p_4(T)$ has to remain bounded, this means that the process will stop. Hence we end up with a tree T that is level-greedy with respect to any root and any edge-root. Therefore by Theorem 2.2.7 in Chapter 2, T is greedy. $\square$

Theorem 3.2.3 is a strong tool to characterize extremal trees for distance-based graph invariants. The following corollary illustrates this statement, especially for the three graph invariants we mentioned earlier.

Corollary 3.2.6 ([24]). Let $f(x)$ be any nonnegative, nondecreasing function, defined for positive integers x. Then the graph invariant

$$W_f(T) = \sum_{\{v,w\} \subseteq V(G)} f(d(v,w))$$

is minimized by the greedy tree among all trees with given degree sequence.

Proof. Let $q_k(T)$ be the number of pairs (u,v) such that $d(u,v) > k$. Then $p_k(T) + q_k(T) = \binom{n}{2}$. Theorem 3.2.3 implies that the greedy tree minimizes $q_k(T)$. Note that

$$\begin{aligned} W_f(T) &= \sum_{k \geq 0} (f(k+1) - f(k)) |\{\{v,w\} \subseteq V(T) : d(v,w) > k\}| \\ &= \sum_{k \geq 0} (f(k+1) - f(k)) q_k(T), \end{aligned}$$

and by the definition of f, $f(k+1) - f(k)$ is nonnegative for all k (we set $f(0) = 0$). This implies the desired statement. $\square$

Remark 3.2.7. Note that $W_f(T)$ can only decrease when $p_k(T)$ increases.

By suitably choosing f in Corollary 3.2.6, we obtain the results for the classical Wiener index by taking $f(x) = x$, the hyper-Wiener index with $f(x) = \frac{x(x+1)}{2}$ and the generalized Wiener index with $f(x) = x^\alpha$.
The Harary index $H(G)$ of a graph G, introduced more recently [23], is defined as the "reciprocal analogue" of the Wiener index, namely

$$H(G) = \sum_{\{v,w\} \subseteq V(G)} \frac{1}{d(v,w)}.$$

Wagner et al. [30] found that the trees that minimize the Harary index are not necessary equal to those that maximize the Wiener index. However, the tree that minimizes the Wiener index corresponds to the one which maximizes the Harary index. This result is obtained by the following corollary, using an analogous proof to Corollary 3.2.6, stated as follows:

Corollary 3.2.8 ([30]). Let $f(x)$ be any nonnegative, nonincreasing function of x. Then the graph invariant

$$W_f(T) = \sum_{\{v,w\} \subseteq V(G)} f(d(v,w))$$

is maximized by the greedy tree among all trees with given degree sequence.

3.3 Number of subtrees

The number of subtrees of a tree plays an important role in phylogenetic reconstruction [17]. A lot of research has been conducted on the number of subtrees of a tree, and in particular on trees that maximize or minimize it. This includes the article [28], as well as [2, 37], which deal with the number of subtrees in trees with given degree sequence. In this section, we present the result obtained in [2] which states that the greedy tree not only maximizes the total number of subtrees, but the number of subtrees of any given order.

Theorem 3.3.1. *Among all trees T with degree sequence D, the number $n_k(T)$ of subtrees of order k attains its maximum when T is the greedy tree $G(D)$.*

As in Section 3.2, to prove Theorem 3.3.1 we characterize the tree maximizing the number of subtrees of any size, among rooted and edge rooted trees with the same level degree sequence.

To prove the statement for trees, let us first consider two-component forests with a given level degree sequence.

Let $\mathbb{F}_D$ (resp. $\mathbb{T}_D$) be the set of all rooted two-components forests (resp. trees) with level degree sequence D. We denote by $G_1(D)$ and $G_2(D)$ the two connected components of the level greedy forest $G(D)$, where we assume that $|V(G_1(D))| \geq |V(G_2(D))|$. Similarly, we write F_1 and F_2 for the components of a rooted forest $F \in \mathbb{F}_D$.

Definition 3.3.2. Let F be a rooted forest which has n levels of vertices. The *level sequence* of a subforest F' of F is the sequence $(s_0, \ldots, s_{n-1})$, where s_i is the number of vertices of F' at level i in F. We write $S^- = (s_1, \ldots, s_{n-1})$ for the sequence obtained from S by removing the first term.

We denote by $n_S(F)$ the number of subtrees in F with level sequence S. For any integer $k \geq 1$, the number $n_k(F)$ of subtrees of order k in F can be written as the sum of $n_S(F)$ over all possible sequences S that sum to k. Let us formulate a lemma which describes the behaviour of $n_S(F)$ for two-component forests.

Lemma 3.3.3 ([2])**.** Let D be a given level degree sequence of a two-component forest. For any level sequence $S = (s_0, s_1, \ldots, s_{L(D)})$ and for any $F \in \mathbb{F}_D$ we have

$$(n_S(F_1), n_S(F_2)) \preceq (n_S(G_1(D)), n_S(G_2(D))). \tag{3.3.1}$$

Now, let us relate n_S to rooted trees.

Lemma 3.3.4 ([2])**.** Let D be a level degree sequence of a rooted tree. For any $T \in \mathbb{T}_D$ and for any level sequence $S = (s_0, s_1, \ldots, s_{L(D)})$ we have

$$n_S(T) \leq n_S(G(D)). \tag{3.3.2}$$

Proof. Let T be a rooted tree with degree sequence D. To prove Lemma 3.3.4, we may consider two cases depending on whether the root is present or not:

Case 1: $s_0 = 0$. In this case, the set of subtrees consists of the subtrees of all connected components obtained by removing the root. Let $C(T)$ be the set of the connected components of $T - r(T)$. Suppose that there are elements H_1 and H_2 of $C(T)$ such that $H_1 \cup H_2$ is not a level greedy forest, and let B be the level degree sequence of $H_1 \cup H_2$. By Lemma 3.3.3, we have

$$n_{S^-}(H_1) + n_{S^-}(H_2) \leq n_{S^-}(G_1(B)) + n_{S^-}(G_2(B)).$$

By replacing H_1 and H_2 with $G_1(B)$ and $G_2(B)$ respectively, we obtain a new rooted tree T^1, with the same level degree sequence as T, satisfying

$$n_S(T) \leq n_S(T^1).$$

Note that during this process the number of two-component forests with components in $C(T^1)$ which are level greedy is increasing in comparison to $C(T)$. Thus, we may iterate the process K times to end up with a tree T^K such that

$$n_S(T) \leq n_S(T^K),$$

where any two elements in $T^K - r(T^K)$ form a level greedy forest. By Remark 2.2.12, such a situation is reached only when T^K is a level greedy tree.

Case 2: $s_0 = 1$. Let us prove this second case by an induction on the height of the tree $L(D)$.

The case where $L(D)$ is 0 or 1 is obvious, since the corresponding sets $\mathbb{T}_D$ contain only one element. Assume that the lemma is true whenever $L(D) \leq k$ for some $k \geq 2$. Now, we have to prove the statement for the case $L(D) = k + 1$.

Let us prove this statement by an induction on the degree of the root r.

If $d(r) = 0$, then $\mathbb{T}_D$ consists of one single vertex, thus the statement holds.

If $d(r) = 1$, then there are two subcases:

- If $s_1 = 0$, the only possible subtrees with such a level sequence consist of a root only, and these subtrees are present in every tree $T \in \mathbb{T}_D$. Similarly, if $s_1 \geq 2$, in view of the fact that $r = 1$, there are no subtrees satisfying this case.

- If $s_1 = 1$. In this case, the root does not play any role, and the number of subtrees satisfying the sequence is the same as the number of subtrees obtained by removing the root. Thus, we can apply the induction hypothesis on the height of the tree to get the result.

If $d(r) = 2$, then we have the following subcases:

- If $s_1 \geq 3$, then $n_S(T) = n_S(G(D)) = 0$ for any $T \in \mathbb{T}_D$.

- If $s_1 = 2$, let T' and U be the trees obtained from T and $G(D)$ respectively by merging the root and its neighbours. Let $S' = (s_1, s_3, \ldots, s_n)$. Then we can use the induction hypothesis on the height to get

$$n_S(T) = n_{S'}(T') \leq n_{S'}(U) = n_S(G(D)).$$

- If $s_1 = 1$, then using Lemma 3.3.3, we obtain

$$n_S(T) = n_S^-(H_1) + n_S^-(H_2) \leq n_{S^-}(H_1') + n_{S^-}(H_2') = n_S(G(D)),$$

where H_1, H_2 (resp. H_1', H_2') are the two components of $C(T)$ (resp. $C(G(D))$).

- The case $s_1 = 0$ corresponds to subtrees that consist only of a root.

Now, we assume that (3.3.2) holds for $r \leq \ell$ for some $\ell \geq 2$. We set $r = \ell + 1$. Let A, A' be subtrees of T as in Figure (3.1).

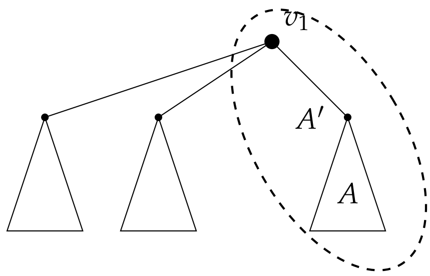


Figure 3.1: Decomposition of T.

Let $\mathbb{S}$ be the set of all possible level sequences whose first term is 1. Then we have

$$\begin{aligned} n_S(T) &= \sum_{\substack{S_1, S_2 \in \mathbb{S} \\ S_1^- + S_2^- = S^-}} n_{S_1}(A') n_{S_2}(T - A) \\ &= \sum_{\substack{S_1, S_2 \in \mathbb{S} \\ S_1^- + S_2^- = S^-}} n_{S_1^-}(A) n_{S_2}(T - A), \end{aligned}$$

where $S_1^- + S_2^-$ is obtained by summing S_1^- and S_2^- term by term. Let D_1 be the level degree sequence of A, and let D_2 be the level degree sequence of $(T - A)$. Since $L(D_1) \leq k$, and since the degree of the root of $(T - A)$ is equal to ℓ, we use the induction hypothesis to get

$$n_{S_1^-}(A) \leq n_{S_1^-}(G(D_1)), \tag{3.3.3}$$

$$n_{S_2}(T - A) \leq n_{S_2}(G(D_2)). \tag{3.3.4}$$

From equations (3.3.3) and (3.3.4), we obtain

$$n_{S_1^-}(G(D_1)) n_{S_2}(G(D_2)) \geq n_{S_1^-}(A) n_{S_2}(T - A). \tag{3.3.5}$$

Let T^1 be the rooted tree obtained by adding an edge joining the two roots of $G(D_2)$ and $G(D_1)$ and taking the root of $G(D_2)$ as root of T^1. Then

$$n_S(T^1) = \sum_{\substack{S_1,S_2 \in \mathbb{S} \\ S_1^- + S_2^- = S^-}} n_{S_1^-}(G(D_1)) n_{S_2}(G(D_2)).$$

Since (3.3.5) is valid for any S_1 and S_2 satisfying the relation $S^- = S_1^- + S_2^-$, it implies that

$$n_S(T^1) \geq n_S(T).$$

We iterate the process to obtain a sequence

$$n_S(T) \leq n_S(T^1) \leq \cdots \leq n_S(T^K).$$

This operation has to stop when there is no more branch A to replace. From the construction of T^K and Remark 2.2.12, we know that any $d(r) - 1$ elements of $C(T^K)$ form a level greedy forest.
Thus, since $d(r) \geq 3$, any two elements of $C(T^K) = T^K - r(T^K)$ form a level greedy forest, and using again Remark 2.2.12, T^K coincides with $G(D)$. Therefore, $n_S(T^K) \leq n_S(G(D))$ is true for any possible degree of the root r. Hence, the result follows. □

An analogous lemma holds also for edge-rooted trees:

Lemma 3.3.5 ([2])**.** Let D be the level degree sequence of an edge-rooted tree. For any $T \in \mathbb{T}_D$, we have $n_S(T) \leq n_S(G(D))$ for any level sequence $S = (s_1, s_2, \ldots, s_{L(D)})$.

Proof. If $s_1 \leq 1$, then the lemma follows clearly from Lemma 3.3.4 since the edge between the roots does not play any role. The case $s_1 = 2$ is obtained again from Lemma 3.3.4, by merging the two endpoints of the edge root to get a vertex rooted tree. For $s_1 \geq 3$, there are no subtrees with such a level sequence. □

Proof of Theorem 3.3.1. As in the proof of Theorem 3.2.3, if a rooted or edge rooted tree is transformed to a level-greedy tree with the same level degree sequence, then for some k, $p_k(T)$ from Section 3.2 strictly increases while all other $p_k(T)$ either stay the same or increase as well, which means that this process has to stop and we end up with a greedy tree (Theorem 2.2.7). Thus, Theorem 3.3.1 follows immediately as a consequence of Lemma 3.3.4 and 3.3.5. □

It is also interesting to compare the greedy trees associated to different degree sequences. It will allow us to characterize the extremal trees with respect to the number of subtrees in a variety of different classes of trees. The following proposition relates two different degree sequences of trees in terms of majorization.

Proposition 3.3.6 ([37]). Let $B = (b_0, \ldots, b_{n-1})$ and $B' = (b'_0, \ldots, b'_{n-1})$ be two nonincreasing degree sequences of trees. If $B \preceq B'$, then there exists a series of degree sequences $B_1, \ldots, B_k$ such that $B \preceq B_1 \preceq \cdots \preceq B_k \preceq B'$, where B_i and B_{i+1} differ at exactly two entries, say $b_j(b'_j)$ and $b_k(b'_k)$ of $B_i(B_{i+1})$, with $b'_j = b_j + 1, b'_k = b_k - 1$ and $j < k$.

Proof. Let B and B' be two nonincreasing degree sequences of trees such that $B \preceq B'$ and $B \neq B'$. Let ℓ be the smallest index for which $b_\ell > b'_\ell$, and let k be the largest index smaller than ℓ for which $b_k < b'_k$ (they exist by the definition of majorization). Then, we have $b_i = b'_i$ for $k < i < \ell$.
We set

$$\begin{aligned} B^1 &= (b^1_0, \ldots, b^1_{n-1}) \\ &= (b'_0, \ldots, b'_{k-1}, b'_k - 1, b'_{k+1}, \ldots, b'_{\ell-1}, b'_\ell + 1, b'_{\ell+1}, \ldots, b'_{n-1}). \end{aligned}$$

We may observe that $b'_k - 1 \geq b_k$ and $b'_\ell + 1 \leq b_\ell$, hence:

$$B \preceq B^1 \preceq B'.$$

Besides, by Proposition 2.1.13, B^1 is a degree sequence of a tree.
In a similar fashion, we can determine B^2 such that $B \preceq B^2 \preceq B^1$. We iterate the process to obtain the sequence

$$B \preceq B^k \preceq \cdots \preceq B^1 \preceq B'.$$

□

We denote by $n_k(T, v)$ the number of subtrees of order k that contain the vertex v. The following lemma compares the value of n_k for the vertices on the same level of a level-greedy tree.

Lemma 3.3.7 ([2]). Let $D = ((a_{0,1}), (a_{1,1}, \ldots, a_{1,k_1}), \ldots, (a_{n-1,1}, \ldots, a_{n-1,k_{n-1}}))$ (resp. $D = ((a_{0,1}, a_{0,2}), (a_{1,1}, \ldots, a_{1,k_1}), \ldots, (a_{n-1,1}, \ldots, a_{n-1,k_{n-1}}))$) be a level degree sequence in a nonincreasing order of a rooted (edge-rooted) tree. Then for all $0 \leq \ell \leq L(D)$ and $k \geq 1$, we have

$$n_k(G(D), g^\ell_1) \geq n_k(G(D), g^\ell_2) \geq \cdots \geq n_k(G(D), g^\ell_{k_\ell}),$$

where g_j^ℓ is the j-th vertex on level ℓ, such that $d(g_j^\ell) = a_{\ell,j}$.

Proof. Let $u = g_i^\ell$ and $v = g_j^\ell$ with $i < j$ be two vertices on the same level ℓ, and let w be their closest common ancestor. Let u' and v' be the children of w such that $u \in T_{u'}$ and $v \in T_{v'}$. By the definition of a rooted (resp. edge-rooted) level greedy tree, the degrees of the vertices on some level in $T_{u'}$ are greater or equal to those of all the vertices on the same level in $T_{v'}$. Therefore there is an isomorphic embedding Φ of $T_{v'}$ into $T_{u'}$ that maps v to u as depicted in Figure 3.2. Now, let us construct an injection that maps a subtree R of $G(D)$ that contains v to a subtree R' of $G(D)$ that contains u. We distinguish three different cases:

(i) If R contains both u and v, then we simply set $R' = R$.

(ii) If R neither contains u, nor w, then we set $R' = \Phi(R)$.

(iii) If R does not contain u, but contains w, then let x be the first vertex (closest to w) on the path from w to u that is not contained in R, and let y be the vertex on the path from w to v which is on the same level as x. Replace $R \cap T_y$ by $\Phi(R \cap T_y)$ to obtain R'.

It is clear that this is an injection that preserves the size of subtrees, so it follows that $n_k(T, u) \geq n_k(T, v)$.

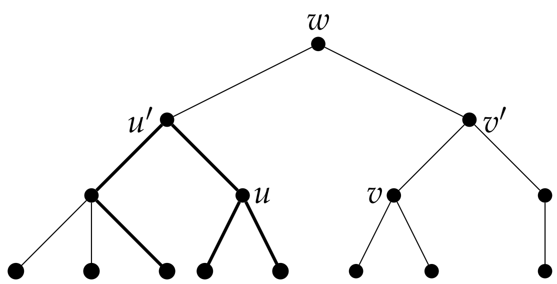


Figure 3.2: Example of the embedding from T_v to T_u for a rooted greedy tree.

□

Now, we are able to prove the following theorem:

Theorem 3.3.8 ([2])**.** *Let B and B' be the degree sequences of trees of the same order such that $B \preceq B'$. Then for any positive integer k, we have*

$$n_k(G(B)) \leq n_k(G(B')).$$

Proof. By Proposition 3.3.6, there exists a degree sequence B_1 with $B \preceq B_1 \preceq B'$ such that B and B_1 only differ in two places, i.e. $B = (b_0, b_1, \ldots, b_i, \ldots, b_j, \ldots, b_{n-1})$ and $B_1 = (b_0, b_1, \ldots, b_i + 1, \ldots, b_j - 1, \ldots, b_{n-1})$ with $i < j$. Consider two vertices u and v in $G(B)$ such that $d_{G(B)}(u) = b_i$ and $d_{G(B)}(v) = b_j$. If the length of the path in $G(B)$ joining u and v is even, let w be the middle vertex of this path; otherwise, let e be the middle edge. We know that $G(B)$ is level-greedy with respect to w (resp. e), and u and v are on the same level h. We have $u = g_k^h$ and $v = g_\ell^h$ for some $k < \ell$. Without loss of generality, we may assume that ℓ is the largest index such that $d_{G(B)}(g_\ell^h) = b_j$. Let x be a child of v, and let H be the branch rooted at x. Then $G(B) - H$ is still a level greedy tree.

Consider the tree $T = G(B) - vx + ux$, which has degree sequence B_1. Subtrees of $G(B)$ are still subtrees in T except for those that contain both v and x, but not u. On the other hand, we gain subtrees that contain u and x but not v.

Therefore, we have

$$n_k(T) - n_k(G(B)) = \sum_{k_1+k_2=k} n_{k_1}(H, x)(n_{k_2}(G(B) - H, u) - n_{k_2}(G(B) - H, v)), \tag{3.3.6}$$

but by Lemma 3.3.7, $n_{k_2}(G(B) - H, u) \geq n_{k_2}(G(B) - H, v)$, which implies that (3.3.6) is nonnegative. Thus,

$$n_k(G(B_1)) \geq n_k(T) \geq n_k(G(B)).$$

We apply the process repeatedly to the sequence in Proposition 3.3.6 to obtain the result:

$$n_k(G(B)) \leq n_k(G(B_1)) \leq \cdots \leq n_k(G(B_r)) \leq n_k(G(B')).$$

□

Let $\nu(T)$ be the total number of subtrees of T.

Example 3.3.9. Let us compute $\nu(P_n)$ and $\nu(S_n)$:

- $\nu(P_n) = \binom{n+1}{2}$. In fact the number of subtrees of P_n corresponds to the number of ways to choose 2 out of n vertices as the end-vertices for the subpath, allowing the 2 vertices to be identical.

- $\nu(S_n) = 2^{n-1} + n - 1$. There are two ways to choose a subtree in a star. Either it does not contain the center, so that the only possibilities are isolated vertices, and there are $n-1$ choices. Or it contains the center and we may choose any subset of the remaining vertices, for which there are 2^{n-1} choices.

Theorem 3.3.8 implies a number of corollaries, such as general bounds for the number of subtrees of a tree of order n (see [27]), the extremal tree among trees of order n with maximum degree Δ (see [15]), and a few more (see [2]).

Corollary 3.3.10. Let T be any tree of order n. Then

$$\nu(T) \leq 2^{n-1} + n - 1,$$

with equality if and only if T is the star S_n.

Proof. Let T be a tree of order n with degree sequence D. Let $B' = (n-1, 1, \ldots, 1)$ be a sequence with n terms. It is clear that B' is the degree sequence of S_n. Moreover $D \preceq B'$. Considering that $\nu(T)$ is the sum of all $n_k(T)$ over all $1 \leq k \leq n$ and that the star is greedy for B' (in fact, it is the only tree with degree sequence B'), the statement follows from Theorems 3.3.1 and 3.3.8. □

Let $\mathbb{T}_{n,\Delta}$ be the set of all trees of order n with maximum degree Δ.

Definition 3.3.11 ([10])**.** The Volkmann tree $V_{n,\Delta} \in \mathbb{T}_{n,\Delta}$ is constructed as follows:
If $n \in \{1, \ldots, \Delta+1\}$, then $V_{n,\Delta}$ is the star.
Let $n > \Delta + 1$. Define n_i as

$$n_i = 1 + \sum_{j=1}^{i} \Delta(\Delta-1)^{j-1}, \quad \text{for } i = 1, 2, \ldots$$

and choose k such that

$$n_{k-1} < n \leq n_k.$$

Then, calculate the parameters m and h from

$$m = \left\lfloor \frac{n - n_{k-1}}{\Delta - 1} \right\rfloor$$

and

$$h = n - n_{k-1} - (\Delta - 1)m.$$

The vertices of $V_{n,\Delta}$ are arranged into $k+1$ levels. At level 0, there is one vertex labeled $v_{0,1}$. At level i ($i = 1, 2, \ldots, k-1$), there are $\Delta(\Delta-1)^{i-1}$ vertices, labeled $v_{i,1}, v_{i,2}, \ldots, v_{i,\Delta(\Delta-1)^{i-1}}$. They are connected in that order to the vertices at level $i-1$. At level k, there are $n - n_{k-1}$ vertices, labeled by $v_{k,1}, v_{k,2}, \ldots, v_{k,n-n_{k-1}}$. These are connected to the vertices at level $k-1$, so that $\Delta - 1$ vertices from level k are adjacent to vertices $v_{k-1,1}, v_{k-1,2}, \ldots, v_{k-1,m}$. The remaining h vertices at level k (if any) are connected to the vertex $v_{k-1,m+1}$ at level $k-1$.

Example 3.3.12. Figure 4.3 shows a Volkmann tree $V_{17,3}$. We observe that $n_2 = 10 < 17 < 22 = n_3$, which means we have a tree on 4 levels, with parameters $m = 3$ and $h = 1$.

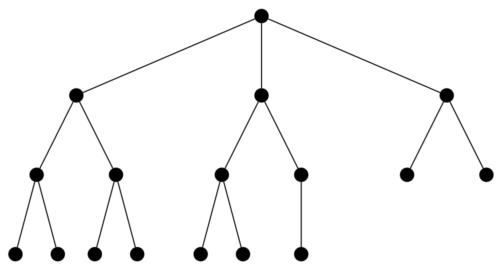


Figure 3.3: A Volkmann tree $V_{17,3}$.

Corollary 3.3.13. If $T \in \mathbb{T}_{n,\Delta}$, then $\nu(T) \leq \nu(V_{n,\Delta})$.

Proof. We can consider the Volkmann tree as a greedy tree with level sequence $B_\Delta = (\Delta, \ldots, \Delta, r, 1, \ldots, 1)$, where $0 \leq r < \Delta$. For any other sequence B of a tree in $\mathbb{T}_{n,\Delta}$, it is clear that $B \preceq B_\Delta$. Hence we get the result by applying again Theorems 3.3.1 and 3.3.8. □

Corollary 3.3.14 ([2])**.** Among trees of order n with s leaves, the greedy tree $G(D)$ corresponding to the sequence $D = (s, 2, \ldots, 2, 1, 1, \ldots, 1)$ maximizes the total number of subtrees.

Proof. Let T be a tree of order n with s leaves and degree sequence $D = (d_0, d_1, \ldots, d_{n-1})$. We know that $d_i > 1$ for $i = 1, \ldots, n-s-1$ and $d_i = 1$ for $i = n-s, \ldots, n-1$. Now it is easy to see that $(d_0, d_1, \ldots, d_{n-1}) \preceq (s, 2, \ldots, 2, 1, \ldots, 1)$. Thus, the result follows by using Theorems 3.3.1 and 3.3.8. □

Chapter 4

Direct approach

In this chapter, we review a direct approach to prove the extremality of greedy trees. For this purpose, we focus particularly on the Wiener index and the spectral radius of a graph. The proofs for other parameters such as the number of subtrees ([37]), spectral moments ([18]) or Laplacian spectral radii ([35]) can follow similar lines to the ones we consider here.

4.1 The Wiener index

As defined in Chapter 3, the Wiener index of a graph is the sum of the distances between all pairs of vertices in the graph. Recall that $\mathbb{T}_D$ is the set of all trees with degree sequence D, $G(D)$ is the greedy tree corresponding to D, and $W(T)$ denotes the Wiener index of T. The main task of this section is to prove the following theorem:

Theorem 4.1.1. *For $T \in \mathbb{T}_D$, $W(T) \geq W(G(D))$.*

For convenience, we call a tree T "optimal" if it minimizes the Wiener index among all trees with the same degree sequence. In order to prove Theorem 4.1.1, let us describe a decomposition of a tree T, found in [32]. Let x, y be two vertices of T. The path $P_T(x, y)$ from x to y can be written as $x_k x_{k-1} \ldots x_2 x_1 z y_1 y_2 \ldots y_{k-1} y_k$ when the length of $P_T(x, y)$ is even, or $x_k x_{k-1} \ldots x_2 x_1 y_1 y_2 \ldots y_{k-1} y_k$ when it is odd, where $x_k = x, y_k = y$.
Let G_1 be the graph resulting from T by deleting all edges in $P_T(x, y)$. We denote by X_i, Y_i and Z the components that contain respectively x_i, y_i, and z for $i = 1, \ldots, k$. We also denote by $X_{>k}$ (resp. $Y_{>k}$) the trees induced by the vertices in $V(X_{k+1}) \cup V(X_{k+2}) \cup \ldots$ (resp. $V(Y_{k+1}) \cup V(Y_{k+2}) \cup \ldots$). Figure

4.1 shows such a labeling with respect to a path of odd length (without z). Without loss of generality, assume that $|V(X_1)| \geq |V(Y_1)|$.

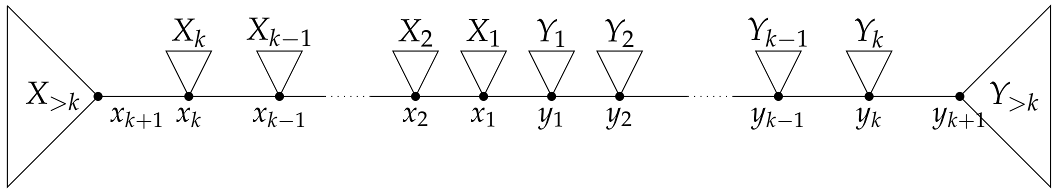


Figure 4.1: Labeling of a path and the components.

Lemma 4.1.2 ([31, 32])**.** In an optimal tree, if $|V(X_i)| \geq |V(Y_i)|$ for $i = 1, 2 \ldots, k-1$ and $|V(X_{>k-1})| \geq |V(Y_{>k-1})|$, then we can assume $d(x_k) \geq d(y_k)$.

Proof. Suppose $d(x_k) < d(y_k)$. Set $s = d(y_k) - d(x_k)$, and let $v_1, v_2, \ldots, v_s$ be neighbours of y_k other than y_{k-1} and y_{k+1}: delete the s edges $y_k v_i$ and add edges $x_k v_i$ $(i = 1, \ldots, s)$ instead. After this operation we will have $d(x_k) \geq d(y_k)$, and the degree sequence of the tree is preserved. We show that this operation will not increase the Wiener index. Let S be the set of vertices in the components of $T \backslash \{y_k v_1, y_k v_2, \ldots, y_k v_s\}$ that contain $v_1, v_2, \ldots, v_s$.
Note that in our operation, the lengths of the paths between vertices only change if exactly one of the end vertices is in S. We have four cases to consider.
Case 1: If the other end of the path is in X_i, for $i = 1, \ldots, k-1$, the distance decreases by $2i-1$ (resp. $2i$) if the length of the path is odd (resp. even). Thus the total contribution to the Wiener index decreases by: $\sum_{i=1}^{k-1}(2i-1)|V(X_i)||S|$ (resp. $\sum_{i=1}^{k-1}(2i)|V(X_i)||S|$).
Case 2: If the other end of the path is in Y_i, for $i = 1, \ldots, k-1$, the distance increases by $2i-1$ (resp. $2i$) if the length of the path is odd (resp. even). Thus the total contribution to the Wiener index increases by: $\sum_{i=1}^{k-1}(2i-1)|V(Y_i)||S|$ (resp. $\sum_{i=1}^{k-1}(2i)|V(Y_i)||S|$).
Case 3: If the other end of the path is in $X_{>k-1}$, the distance decreases by $2k-1$ (resp. $2k$) if the length of the path is odd (resp. even).Thus the total contribution to the Wiener index decreases by: $(2k-1)|V(X_{>k-1})||S|$ (resp. $(2k)|V(X_{>k-1})||S|$).
Case 4: If the other end of the path is in $Y_{>k-1} \backslash S$, the distance increases by $2k-1$ (resp. $2k$) if the length of the path is odd (resp. even). Thus the total

contribution to the Wiener index increases by: $(2k-1)(|V(Y_{>k-1})|-|S|)|S|$ (resp. $(2k)(|V(Y_{>k-1})|-|S|)|S|$).
Therefore, the total amount of change that occurs is

$$\sum_{i=1}^{k-1}(2i-1)(|V(Y_i)|-|V(X_i)|)|S| \\ +(2k-1)(|V(Y_{>k-1})|-|S|-|V(X_{>k-1})|)|S|$$

for a path without z, and

$$\sum_{i=1}^{k-1}(2i)(|V(Y_i)|-|V(X_i)|)|S|) \\ +(2k)(|V(X_{>k})|-(|V(Y_{>k-1})|-|S|-|V(X_{>k-1})|)|S|$$

for a path with z, which is nonpositive in view of our hypothesis. □

Remark 4.1.3. In Lemma 4.1.2, if at least one strict inequality holds in the conditions, then we can replace "can assume" by "must have" in the statement.

Lemma 4.1.4 ([31, 32])**.** Let P be a path of an optimal $T \in \mathbb{T}_D$ whose end vertices are leaves.
If the length of P is odd $(2m-1)$, then the vertices of P can be labeled as $u_m u_{m-1} \ldots u_1 w_1 w_2 \ldots w_m$, where U_i, W_i are the components that contain respectively u_i, w_i such that

$$|V(U_1)| \geq |V(W_1)| \geq |V(U_2)| \geq |V(W_2)| \geq \cdots \geq |V(U_m)| = |V(W_m)| = 1.$$

If the length of P is even $(2m)$, then the vertices of P can be labeled as $u_{m+1}u_m u_{m-1} \ldots u_1 w_1 w_2 \ldots w_m$ such that

$$|V(U_1)| \geq |V(W_1)| \geq |V(U_2)| \geq |V(W_2)| \geq \cdots \geq |V(W_m)| = |V(U_{m+1})| = 1,$$

To prove Lemma 4.1.4, we need to consider an equivalent expression for the Wiener index stated in the following proposition.

Proposition 4.1.5 ([8])**.** The Wiener index can also be written as follows:

$$W(T) = \sum_{uv \in E(T)} n(u)n(v),$$

where $n(u)$ (resp.$n(v)$) is the number of vertices in the component that contains u (resp. v) after removing uv.

Proof. The expression on the right hand side of the equation counts the number of pairs $\{u', v'\}$ such that $P(u', v')$ contains uv. On the other hand, if we have a path $P(u', v')$, then the number of edges that it contains is $d(u', v')$, which coincides with the previous definition of the Wiener index. □

Now, we need a theorem from Hardy, Littlewood and Pólya [11] that connects sequences and bilinear forms. For a sequence a of the form $a_{-n}, \ldots a_{-n+1}, \ldots, a_{-1}, a_0, a_1, \ldots\ a_{n-1}, a_n$, let a^+ be the sequence obtained by rearranging the elements so that

$$a_0^+ \geq a_1^+ \geq a_{-1}^+ \geq a_2^+ \geq a_{-2}^+ \geq \cdots \geq a_n^+ \geq a_{-n}^+.$$

Theorem 4.1.6 (Cf. [11, Theorem 371])**.** *Suppose that $c, x,$ and y are nonnegative sequences, and c is symmetrically increasing, so that*

$$c_0 \leq c_1 = c_{-1} \leq c_2 = c_{-2} \leq \cdots \leq c_{2k} = c_{-2k},$$

while the x and y are given except for their order. Then the bilinear form

$$\sum_{r=-k}^{k} \sum_{s=-k}^{k} c_{r-s} x_r y_s$$

attains its minimum when x is x^+ and y is y^+.

Proof of Lemma 4.1.4. We provide the proof for a path of odd length, the other case can be shown in a similar manner. Let $P = u_m u_{m-1} \ldots\ u_1 w_1 w_2 \ldots w_m$ be as described in Lemma 4.1.4. Let A be a sequence whose indices are from $-m$ to $m-1$ such that

$$A_{-m} = |V(U_m)|, A_{-m+1} = |V(U_{m-1})|, \ldots, A_{-1} = |V(U_1)|,$$
$$A_0 = |V(W_1)|, A_1 = |V(W_2)|, \ldots A_{m-1} = |V(W_m)|.$$

Now, let us consider the contribution C of the edges of the path P to the expression for the Wiener index in Proposition 4.1.5, since the contribution of all other edges remains the same. We have:

$$\begin{aligned} C = {} & A_{-m}(A_{-m+1} + \cdots + A_{m-1}) \\ & + (A_{-m} + A_{-m+1})(A_{-m+2+\cdots} + A_{m-1}) \\ & \vdots \\ & + (A_{-m} + A_{-m+1} + \cdots + A_{m-2})A_{m-1}. \end{aligned}$$

We observe that it can be written as a sum of terms of the form A_rA_s, so let us count how many there are of each. A product A_rA_s occurs if A_r and A_s are in different parentheses, which happens $k = r - s$ times. Now, as $A_rA_s = A_sA_r$, we set $c_{r-s} = k/2$ and $c_{s-r} = k/2$. Then, C can be written as

$$C = \sum_{r=-m}^{m-1} \sum_{s=-m}^{m-1} c_{r-s} A_r A_s,$$

where c is symmetrically increasing:

$$c_0 = 0 \leq c_1 = c_{-1} = 1/2 \leq \cdots \leq c_{2m-1} = c_{-2m+1} = m.$$

Hence, by using Theorem 4.1.6, S attains its minimum if A is A^+. Therefore,

$$|V(U_1)| \geq |V(W_1)| \geq |V(U_2)| \geq \cdots \geq |V(U_m)| = |V(W_m)|$$

for an optimal tree. □

Lemma 4.1.7 ([31, 32])**.** In an optimal tree, for a path with labelling as in Lemma 4.1.4, we have

$$d(u_1) \geq d(w_1) \geq d(u_2) \geq d(w_2) \geq \cdots \geq d(u_m) = d(w_m) = 1$$

if the path is of odd length $(2m - 1)$; and

$$d(u_1) \geq d(w_1) \geq d(u_2) \geq d(w_2) \geq \cdots \geq d(u_m) \geq d(w_m) = d(w_{m+1}) = 1$$

if the path is of even length $(2m)$.

Proof. We show the proof for a path of odd length, the other case is similar. First, we have

$$|V(U_1)| \geq |V(W_1)| \geq |V(U_2)| \geq |V(W_2)| \geq \cdots \geq |V(U_m)| = |V(W_m)| = 1.$$

By suitably choosing y_i and x_i in Lemma 4.1.2, we get the following results:
Case 1: Let $y_1 = u_{i+1}, y_2 = u_{i+2} \ldots; \quad x_1 = u_i, x_2 = u_{i-1}, \ldots, x_{i+1} = w_1, \ldots$. Since

$$|V(Y_{>1})| = \sum_{k=i+2}^{m} |V(U_k)| < \sum_{k=1}^{m} |V(W_k)| + \sum_{k=1}^{i-1} |V(U_k)| = |V(X_{>1})|,$$

then $d(u_{i+1}) = d(y_1) \leq d(x_1) = d(u_i)$. This applies to u_i, u_{i+1} for $i = 1, \ldots, m-1$, hence

$$d(u_1) \geq d(u_2) \geq \cdots \geq d(u_m).$$

Case 2: Let $y_1 = w_i, y_2 = w_{i-1} \ldots, y_{i+1} = u_1, \ldots; \quad x_1 = w_{i+1}, x_2 = w_{i+2}, \ldots$. In a similar way as in Case 1, $V(Y_{>1}) > V(X_{>1})$, so $d(w_i) = d(y_1) \geq d(x_1) = d(w_{i+1})$. Again, this applies for $i = 1, \ldots, m-1$, hence

$$d(w_1) \geq d(w_2) \geq \cdots \geq d(w_m).$$

Case 3: Let $x_i = u_i$ and $y_i = w_i$ for $i = 1, 2 \ldots, m$, then we obtain $d(u_i) \geq d(w_i)$ for $i = 1, 2 \ldots, m$.
Case 4: Let $z = u_1, y_i = u_{i+1}, x_i = w_i$, to get $d(w_i) \geq d(u_{i+1})$ for $i = 1, 2, \ldots, m-1$.
The result follows by combining all the cases. □

Let us write $g_T(u) = \sum_{v \in V(T)} d(u, v)$. Note that $W(T) = \frac{1}{2} \sum_{v \in V(T)} g_T(v)$.

Lemma 4.1.8 ([13])**.** For any tree T, $g_T(u)$ is minimized at one or two adjacent vertices in the whole tree. We call such a vertex "centroid".

Proof of Theorem 4.1.1. On any path of an optimal tree labeled as in Lemmas 4.1.4 and 4.1.7, we have

$$g_T(v) \geq g_T(u_1), \tag{4.1.1}$$

for any other vertex v in the path, where the degree $d(u_1)$ and the size $|V(U_1)|$ of the associated branch U_1 are maximal among vertices of the path. In fact, we observe that for smaller k, the number of vertices w satisfying $d(u_1, w) \leq k$ is greater than the number of vertices satisfying $d(u_r, w) \leq k$ for $r \neq 1$.
In view of Lemma 4.1.8, two cases can be considered:

- If there is only one vertex in the centroid, we label it as v.
- Otherwise, we simply choose either one as v and the other as v_1.

We only show the theorem for the first case, the second one is analogous. In an optimal tree T considered as a tree rooted at v, we know from (4.1.1) that v has the largest degree, so (i) in Lemma 2.2.3 is satisfied.
Now, consider any path starting at a leaf u, passing through v, and ending at a leaf w, so that the only common ancestor of u and w is v. By Lemma

4.1.7, we have $|d(u,v) - d(w,v)| = 0$ if the length of the path is even and $|d(u,v) - d(w,v)| = 1$ otherwise. Hence, the heights of any two leaves differ by at most 1, which shows that (ii) holds in Lemma 2.2.3.

Furthermore, for a vertex x of height i and a vertex y of height $j(i < j)$, if y is a successor of x, consider the same path from a leaf u to a leaf w passing through y; then by Lemma 4.1.7 $d(x) \geq d(y)$. Otherwise, consider the path that passes through y', y, u, x, x', where y', x' are leaf successors of (or equal to) y, x respectively, and u their first common ancestor. We have $u_1 = u$ by (4.1.1) and Lemma 4.1.4, and $x = u_{k+1}, y = w_\ell$ or $x = w_k, y = u_{\ell+1}$, where $k = i - h_T(u), \ell = j - h_T(u)$, so we can see $k + 1 \leq \ell$. Again by Lemma 4.1.7, $d(x) \geq d(y)$, so (iii) in Lemma 2.2.3 holds.

Now, for two non-leaves x and y on the same level i such that $d(x) > d(y)$, let x', y' on the same level j be the successors of x, y respectively. Consider the longest path that passes through y', y, u, x, x', where u is the first common ancestor of x and y. We have $u_1 = u$ by (4.1.1) and Lemma 4.1.4, so $x = w_k, x' = w_\ell, y = u_{k+1}, y' = u_{\ell+1}$, since $d(x) > d(y)$, where $k = i - h_T(u), l = j - h_T(u)$. Using Lemma 4.1.7, $d(x') \geq d(y')$, hence (iv) in Lemma 2.2.3 is satisfied.

Finally, let x_0 (x') and y_0 (y') be the parents (siblings) of x and y described above respectively, and let x'' and y'' on level j be successors of x' and y' respectively. The conclusion of (iv) implies that

$$|V(T_{x_0} \backslash T_{x'})| > |V(T_{y_0} \backslash T_{y'})|. \tag{4.1.2}$$

Now, consider the longest path passing through y'', y', u, x', x'', where u is the common ancestor of x and y that is on the path $P(x', y')$. By (4.1.1) and Lemma 4.1.4, we have $u_1 = u$, so $x' = w_k, x'' = w_\ell, y' = u_{k+1}, y'' = u_{\ell+1}$ by (4.1.2), where $k = i - h_T(u), \ell = j - h_T(u)$. Thus from Lemma 4.1.7 we have $d(x') \geq d(y')$ and $d(x'') \geq d(y'')$, which satisfies (v) in Lemma 2.2.3.

In summary, by Lemma 2.2.3, the optimal tree is the greedy tree. □

As for the number of subtrees in Section 3.3, let us compare trees with different degree sequences.

Lemma 4.1.9 ([36]). Let $G(D)$ be the greedy tree corresponding to the degree sequence D. Let x and y be two vertices in $V(G(D))$ such that $d_{G(D)}(x) \geq d_{G(D)}(y) \geq 2$. Let x' be a child of y and let T' be the tree obtained from $G(D)$ by deleting the edge yx' and adding the edge xx'. Then

(i) $T' \in \mathbb{T}_{D'}$, with $D \preceq D'$.

(ii) $W(G(D)) > W(T')$.

Proof. For (i), it is clear that $d_{T'}(x) = d_{G(D)}(x) + 1, d_{T'}(y) = d_{G(D)}(y) - 1$ and the other vertex degrees stay the same. Since $d_{G(D)}(x) \geq d_{G(D)}(y)$, we have $D \preceq D'$.
Now, for (ii), consider the longest path that passes through x, w, y where w is the common ancestor of x and y. We may assume that the length of the path is odd, the case when it is even is similar. Using the labelling of Lemma 4.1.4, we have $u_1 = w, x = u_k, y = w_\ell$, where $k \leq \ell$ since $d(x) \geq d(y)$. Let X' be the branch formed by x' and its successors. Note that the distance between two vertices changes only if one of them is in X'. If one of the end vertices is in U_i, the amount of the change is given by

$$\sum_{i=1}^{k}(2i - 1 + \ell - k)|V(U_i)||X'| + \sum_{i=k+1}^{m}(\ell + k - 1)|V(U_i)||X'|.$$

On the other hand, if one of the end vertices is in W_i but not in X', the amount of the change is

$$\begin{aligned}&\sum_{i=1}^{\ell-1}(-2i + 1 + \ell - k)|V(W_i)||X'| \\ &- (\ell + k - 1)(|V(W_\ell)| - |X'|)|X'| - \sum_{i=\ell+1}^{m}(\ell + k - 1)|V(W_i)||X'|.\end{aligned}$$

Thus,

$$\begin{aligned}&W(G(D)) - W(T') \\ =& \sum_{i=1}^{k}(2i-1)(|V(U_i)| - |V(W_i)|)|X'| + \sum_{i=1}^{k}(\ell - k)(|V(W_i)| + |V(U_i)|)|X'| \\ &+ \sum_{i=k+1}^{\ell-1}(\ell + k - 1)(|V(U_i)| - |V(W_i)|)|X'| + \sum_{i=k}^{\ell-1}(2\ell - 2i)|V(W_i)||X'| \\ &+ (\ell + k - 1)(|V(U_\ell)| - |V(W_\ell)| + |X'|)|X'| \\ &+ \sum_{i=\ell+1}^{m}(\ell + k - 1)(|V(U_i)| - |V(W_i)|)|X'|.\end{aligned}$$

By Lemma 4.1.4, $|V(U_i)| \geq |V(W_i)|$, so each term in the summation is nonnegative. Moreover, since $|X'| > 0$, the change is strictly positive due to the term $(\ell + k - 1)|X'||X'|$. Thus $W(G(D)) > W(T')$. □

Theorem 4.1.10 ([36]). *Let B and B' be the degree sequences of trees of the same order such that $B \preceq B'$. Then we have*

$$W(G(B)) \geq W(G(B')).$$

Proof. By Proposition 3.3.6, there exists a degree sequence B_1 with $B \preceq B_1 \preceq B'$ such that B and B_1 only differ in two places, i.e. $B = (b_0, b_1, \ldots, b_i, \ldots, b_j, \ldots, b_{n-1})$ and $B_1 = (b_0, b_1, \ldots, b_i + 1, \ldots, b_j - 1 , \ldots, b_{n-1})$ with $i < j$. Let us consider two vertices u, v in the greedy tree $G(B)$, such that $d_{G(B)}(u) = b_i$ and $d_{G(B)}(v) = b_j$. By the definition of majorization, we may assume $b_i \geq b_j \geq 2$. Now, let x be a child of v, and T' be the tree obtained by removing vx and adding ux. Lemma 4.1.9 provides that $T' \in \mathbb{T}_{B_1}$, and $W(T') < W(G(B))$. Furthermore, by Theorem 4.1.1, we have $W(T') \geq W(G(B_1))$. Thus

$$W(G(B_1)) \leq W(T') \leq W(G(B)).$$

By repeating the same process, the theorem follows. □

Now, it is not surprising that the trees maximizing the number of subtrees correspond to those which minimize the Wiener index. Let us mention some corollaries similar to those for the total number of subtrees.

Corollary 4.1.11 ([8]). Let T be any tree of order n. Then,

$$(n-1)^2 \leq W(T),$$

with equality if and only if T is the star S_n.

Proof. The proof is similar to the proof of Corollary 3.3.10 but using Theorems 4.1.1 and 4.1.10. Now let us compute the the Wiener index of the star S_n. We know that the distance between the center and any other vertex is one, while the distance between two non-center vertices is two, so we have

$$W(S_n) = (n-1) + 2\binom{n-1}{2} = (n-1)^2.$$

□

Corollary 4.1.12 ([10]). Among trees of order n with maximum degree Δ, the Volkmann tree $V_{n,\Delta}$ minimizes the Wiener index.

Corollary 4.1.13 ([36])**.** Among trees of order n with s leaves, the greedy tree $G(D)$ corresponding to the sequence $D = (s, 2, \ldots, 2, 1, 1, \ldots, 1)$ minimizes the Wiener index.

4.2 Spectral radius

Definition 4.2.1. Let G be a graph of order n. Let us label the vertices of G by $v_1, v_2, \ldots, v_n$. The *adjacency matrix* of G is the square matrix $A(G) = (a_{i,j})_{1 \le i,j \le n}$ where:

$$a_{i,j} = \begin{cases} 1 & \text{if} \quad v_i v_j \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

In particular, for all i and j in $\{1, \ldots, n\}$, we have $a_{i,i} = 0$ and $a_{i,j} = a_{j,i}$, meaning that $A(G)$ is a nonnegative symmetric matrix.
The *characteristic polynomial* of an n-vertex graph G is defined by $\det(xI_n - A(G))$, where I_n is the identity matrix of order n, and will be denoted $\phi_G(x)$.
The *eigenvalues* of the graph G are the zeros of $\phi_G(x)$.

Remark 4.2.2. Since $A(G)$ is symmetric, all the roots of $\phi_G(x)$ are real.

Definition 4.2.3. The *spectral radius* of a graph G is the largest eigenvalue of $A(G)$.

Definition 4.2.4 ([7])**.** An $n \times n$ matrix A is *reducible* if we can partition $1, \ldots, n$ into two non-empty subsets E, F such that $a_{ij} = 0$ if $i \in E$ and $j \in F$.
A matrix that is not reducible, is *irreducible*.

Proposition 4.2.5 ([7])**.** The adjacency matrix of graph $A(G)$ is irreducible if and only if G is connected.

Since we are dealing with trees, which are connected graphs, the following result applies.

Theorem 4.2.6 (Perron-Frobenius Theorem, [12])**.** *If $A(G)$ is irreducible, then its spectral radius $\lambda(G)$ is simple (an eigenvalue of multiplicity one) and positive. It corresponds to the unique positive unit eigenvector f, that we refer to as Perron vector of G.*

Example 4.2.7. Consider the star S_5 pictured in Figure 4.2.

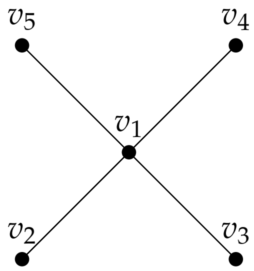


Figure 4.2: A star S_5

Its adjacency matrix is given by

$$A(S_5) = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then, its characteristic polynomial is

$$\phi_{S_5}(x) = \begin{vmatrix} x & -1 & -1 & -1 & -1 \\ -1 & x & 0 & 0 & 0 \\ -1 & 0 & x & 0 & 0 \\ -1 & 0 & 0 & x & 0 \\ -1 & 0 & 0 & 0 & x \end{vmatrix} = x^5 - 4x^3 = x^3(x-2)(x+2).$$

It is easy to see that the eigenvalues of S_5 are $2, -2, 0$. Hence the spectral radius of S_5 is 2, which corresponds to the Perron vector

$$f = \begin{pmatrix} f(v_1) \\ f(v_2) \\ f(v_3) \\ f(v_4) \\ f(v_5) \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Let us mention a property of the characteristic polynomial that we will need later on.

Lemma 4.2.8 ([7])**.** Let T be a tree, let u be a leaf and v the vertex which is adjacent to u. Then $\Phi_T(x) = x\Phi_{T-u}(x) - \Phi_{T-u-v}(x)$.

A lot of research has been done on the spectral radius of graphs, especially upper and lower bounds ([19, 25]). In this section, we characterize the trees that maximize the spectral radius among all trees with the same degree sequence. The main result can be stated as follows:

Theorem 4.2.9 ([5])**.** *If $T \in \mathbb{T}_D$, then $\lambda(T) \leq \lambda(G(D))$, where $G(D)$ is the greedy tree with degree sequence D.*

Let T be a tree of order n. For a vertex v in T, and a unit vector f, set $N_f(v) = \sum_{uv \in E} f(u)$. By the definition of the adjacency matrix $A = A(T)$ we observe that $(Af)(v) = N_f(v)$. The Rayleigh quotient of A on vectors f is defined by

$$\mathcal{R}_T(f) = \frac{\langle Af, f\rangle}{\langle f, f\rangle} = \frac{\left(\sum_{v\in V} f(v)\right)\left(\sum_{uv\in E} f(u)\right)}{\sum_{v\in V} f(v)^2} = \frac{2\sum_{uv\in E} f(u)f(v)}{\sum_{v\in V} f(v)^2}, \tag{4.2.1}$$

where $\langle , \rangle$ is the scalar product.

Proposition 4.2.10 ([12])**.** Let $\mathcal{S}$ be the set of unit vectors on V. Then

$$\lambda(T) = \max_{f\in\mathcal{S}} \mathcal{R}_T(f) = 2\max_{f\in\mathcal{S}} \sum_{uv\in E} f(u)f(v).$$

Moreover, if $\mathcal{R}_T(f) = \lambda(T)$ for a positive function $f \in \mathcal{S}$, then f is an eigenvector corresponding to the largest eigenvalue $\lambda(T)$ of $A(T)$, i.e., it is a Perron vector.

The technique used to prove Theorem 4.2.9 is by rearranging the edges, and considering the behaviour of the spectral radius λ. For convenience, let us again call a tree which maximizes the spectral radius an "optimal tree".

Lemma 4.2.11 ([5])**.** Let T be an optimal tree and f be the Perron vector of T. For vertices $u, v \in T$, if $d(u) > d(v)$, then $f(u) > f(v)$.

Proof. Let $s = d(u) - d(v) > 0$ and suppose $f(u) \leq f(v)$. Let T' be a tree obtained from T by removing s edges of the form uw_k, and adding the s edges vw_k for $k = 1, \ldots, s$, where w_k are neighbours of u such that $P(w_k, v) \supset P(u, v)$. It is clear that T and T' have the same degree sequence.

Moreover, if E and E' denote the edge sets of T and T' respectively, we have

$$\begin{aligned}\mathcal{R}_{T'}(f) - \mathcal{R}_T(f) &= \langle A(T')f, f\rangle - \langle A(T)f, f\rangle \\ &= 2\left(\sum_{xy\in E'\setminus E} f(x)f(y) - \sum_{x'y'\in E\setminus E'} f(x')f(y')\right) \\ &= 2\left(\sum_{k=1}^{s}(f(v)-f(u))f(w_k)\right) \geq 0.\end{aligned}$$

Hence by Proposition 4.2.10, $\lambda(T') \geq \mathcal{R}_{T'}(f) \geq \mathcal{R}_T(f) = \lambda(T)$. However, $\lambda(T') = \lambda(T)$ if and only if f is also an eigenvector corresponding to $\lambda(T')$ on T' so that

$$\lambda(T')f(v) = \sum_{i=1}^{s} f(w_k) + \sum_{xv\in E} f(x) > \sum_{xv\in E} f(x) = \lambda(T)f(v), \qquad (4.2.2)$$

which is a contradiction. Therefore, $\lambda(T') > \lambda(T)$, which is a contradiction of T being optimal. □

Remark 4.2.12. From equation (4.2.2), we can see that in an optimal tree, $f(u) = f(v)$ can only occur if $d(u) = d(v)$.

Lemma 4.2.13 ([5])**.** Let T be an optimal tree and f be the Perron vector of T. Suppose there are two vertices u, v such that $f(u) \geq f(v)$. If u', v' are neighbours of u, v respectively with $P(u', v') \subset P(u, v)$ or $P(u', v') \supset P(u, v)$ such that $f(u) \geq f(u')$ and $f(v) \geq f(v')$, then $f(u') \geq f(v')$.

Proof. Suppose that $f(u') < f(v')$. Let T' be a tree obtained from T by deleting the edges uu' and vv' and replacing them by edges uv' and vu'. Clearly, T and T' have the same degree sequence. Moreover, if E, E' denote the edge sets of T and T' we have

$$\begin{aligned}\mathcal{R}_{T'}(f) - \mathcal{R}_T(f) &= \langle A(T')f, f\rangle - \langle A(T)f, f\rangle \\ &= 2\left(\sum_{xy\in E'\setminus E} f(x)f(y) - \sum_{x'y'\in E\setminus E'} f(x')f(y')\right) \\ &= 2[f(u)f(v') + f(v)f(u')] - 2[f(u)f(u') + f(v)f(v')] \\ &= 2(f(u) - f(v))(f(v') - f(u')) \geq 0.\end{aligned}$$

Again, by Proposition 4.2.10, $\lambda(T') \geq \mathcal{R}_{T'}(f) \geq \mathcal{R}_T(f) = \lambda(G)$. However, $\lambda(T') = \lambda(T)$ if and only if f is also an eigenvector corresponding to $\lambda(T')$

on T' so that

$$\lambda(T)f(u) = (A(T)f)(u) = f(u') + \sum_{wu \in E \cap E'} f(w)$$
$$= \lambda(T')f(u) = (A(T')f)(u) = f(v') + \sum_{wu \in E \cap E'} f(w),$$

which would imply that $f(u') = f(v')$. Since we have a strict inequality $\lambda(T') > \lambda(T)$, this contradicts the assumption that T is optimal. □

Proof of Theorem 4.2.9. Let T be an optimal tree among all trees with the same degree sequence. Let f be a Perron vector of T associated to $\lambda(T)$. At level 0, let us choose the vertex v_0 as the root for which $f(v_0)$ is a maximum. By Lemma 4.2.11, $d(v_0) \geq d(v)$ for any $v \in T$. Now at level 1, we denote the neighbours of v_0 by $v_1, \ldots, v_{k_0}$. From Lemma 4.2.13, by taking $u = v_0$ and v another vertex of T, we know that the neighbours v_k of v_0 have greater Perron vector entries $f(v_k)$ than other vertices. We may assume without loss of generality that $f(v_1) \geq f(v_2) \geq \cdots \geq f(v_{k_0})$. From Lemma 4.2.11, we know that $d(v_1) \geq d(v_2) \geq \cdots \geq d(v_{k_0})$. Let us now consider the vertices at level 2. If v'_i is a child of v_i and v'_j a child of v_j, with $i < j$, by Lemma 4.2.13 $f(v'_i) \geq f(v'_j)$. If $f(v'_i) > f(v'_j)$, we apply Lemma 4.2.11 to obtain $d(v'_i) > d(v'_j)$, otherwise by Remark 4.2.12 $d(v'_i) = d(v'_j)$. Hence the vertices at level 2 are ordered. Moreover, by applying again Lemma 4.2.13 $f(v'_{k_0}) \geq f(v)$ for any v at level greater than 2. We continue to apply Lemma 4.2.13 and Lemma 4.2.11 for higher levels, and we obtain that at each level ℓ, $d(v_1^\ell) \geq d(v_2^\ell) \geq \cdots \geq d(v_{k_\ell}^\ell)$ and $d(v_{k_\ell}^\ell) \geq d(v_1^{\ell+1})$. This construction corresponds to the greedy algorithm, thus T is the greedy tree. □

As it is for the other graph invariants mentioned earlier, it is interesting to compare the spectral radius for trees having different degree sequences.

Theorem 4.2.14 ([5]). *Let B and B' be the degree sequences of trees of the same order such that $B \preceq B'$. Then we have*

$$\lambda(G(B)) \leq \lambda(G(B')),$$

Equality holds if and only if $B = B'$.

Proof. By Proposition 3.3.6, if $B \neq B'$, there exists a degree sequence B_1 with $B \preceq B_1 \preceq B'$ such that B and B_1 only differ in two places, i.e.,

$B = (b_0, b_1, \ldots, b_i, \ldots, b_j, \ldots, b_{n-1})$ and $B_1 = (b_0, b_1, \ldots, b_i + 1, \ldots, b_j - 1, \ldots, b_{n-1})$ with $i < j$. Let us consider two vertices u, v in the greedy tree $G(B)$ such that $d_{G(B)}(u) = b_i$ and $d_{G(B)}(v) = b_j$. By the definition of majorization, we may assume $b_i \geq b_j \geq 2$. Now, let x be a child of v, and T' be the tree obtained by removing vx and adding ux. Then clearly $T' \in \mathbb{T}(B_1)$. Using Lemma 4.2.11, we have $f(u) \geq f(v)$, where f is the Perron vector of $G(B)$. If A and A' are the adjacency matrices of $G(B)$ and T' respectively, and E and E' their edge sets, then we get

$$\begin{aligned}
\mathcal{R}_{T'}(f) - \mathcal{R}_{G(B)}(f) &= \langle A'f, f\rangle - \langle Af, f\rangle \\
&= 2\left(\sum_{xy \in E' \setminus E} f(x)f(y) - \sum_{uv \in E \setminus E'} f(u)f(v)\right) \\
&= 2(f(u)f(x) - f(v)f(x)) \\
&= 2(f(u) - f(v))f(x) \geq 0.
\end{aligned}$$

Thus, $\lambda(T') \geq \lambda(G(B))$. Moreover, by Theorem 4.2.9, $\lambda(T') \leq \lambda(G(B_1))$. Therefore

$$\lambda(G(B)) \leq \lambda(G(B_1)).$$

We iterate this process to obtain the result.
Furthermore, equality holds if and only if f is also an eigenvector corresponding to $\lambda(T')$ on T' so that

$$\lambda(T')f(u) = f(x) + \sum_{wu \in E} f(w) > \sum_{wu \in E \cap E'} f(w) = \lambda(G(B))f(u),$$

which is impossible. □

Theorems 4.2.9 and 4.2.14 allow us to find bounds for the spectral radius of different classes of trees.

Corollary 4.2.15 ([34])**.** Among trees of order n with s leaves, the greedy tree $G(D(s))$ corresponding to the sequence $D(s) = (s, 2, \ldots, 2, 1, 1, \ldots, 1)$ maximizes the spectral radius.

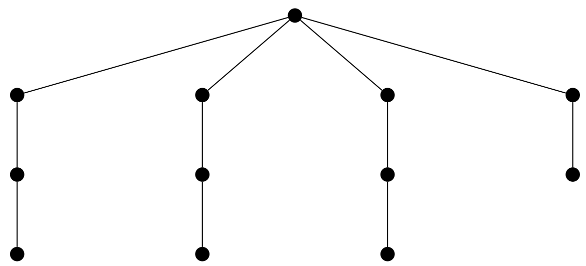


Figure 4.3: A greedy tree of order 12 with 4 leaves.

Proof. The proof is similar to the proof of Corollary 3.3.14, but using Theorems 4.2.9 and 4.2.14. □

Remark 4.2.16. Wu et al. ([34]) established that the characteristic polynomial of $G(D(s))$ can be expressed in terms of Chebyshev polynomials, and they could compute for some s the corresponding spectral radius. More precisely, they found that for $\lfloor \frac{n}{2} \rfloor \leq s \leq n-1$, one has

$$\lambda(G(D(s))) = \sqrt{\frac{s+1+\sqrt{(s+1)^2-4(2s-n+1)}}{2}}.$$

Let $\mathbb{T}_{n,\Delta}$ be the set of all trees of order n with maximum degree Δ, $\Delta \geq 3$.

Corollary 4.2.17 ([26])**.** If $T \in \mathbb{T}_{n,\Delta}$, then

(i) if $n = \Delta + 1$, then T is the star S_n and $\lambda(T) = \sqrt{n-1}$.

(ii) if $\Delta + 1 \leq n \leq 2\Delta$, then $\lambda(T) \leq \sqrt{\frac{n-1+\sqrt{(n-2\Delta)^2+2n-3}}{2}}$, with equality if and only if T is the Volkmann tree $V_{n,\Delta}$.

(iii) if $2\Delta < n \leq \Delta^2 + 1$, then $\lambda(T) \leq \sqrt{2\Delta - 1}$, with equality if and only if T is a complete Δ-ary tree with 3 levels.

(iv) if $n > \Delta^2 + 1$, then $\lambda(T) < 2\sqrt{\Delta - 1}\cos\frac{\pi}{2k+1}$, where k is the height of $V_{n,\Delta}$.

Proof. Let $T \in \mathbb{T}_{n,\Delta}$ with degree sequence D. It is clear that $D \preceq (\Delta, \Delta, \ldots, r, 1, \ldots, 1)$, where $0 \leq r < \Delta$, which is the degree sequence of the Volkmann tree. Hence, by Theorems 4.2.9 and 4.2.14, $\lambda(T) \leq \lambda(V_{n,\Delta})$. Now, let us consider all the cases.

For the first case (i), the only possible tree is the n-vertex star, so the result holds.

For (ii), we have $\Delta + 1 \leq n \leq 2\Delta$. For this case the Volkmann tree V_1 has 3 levels, moreover all vertices are leaves except for the root and another vertex a such that a is attached to $s = n - \Delta - 1$ leaves. Let $b_1, \ldots, b_s$ be the leaves attached to a. By using Lemma 4.2.8 repeatedly, we obtain

$$\begin{aligned}
\Phi_{V_1}(x) &= x\Phi_{V_1-b_1}(x) - \Phi_{V_1-a-b_1}(x) = x\Phi_{V_1-b_1}(x) - x^{s-1}\Phi_{S_\Delta}(x) \\
&= x\Phi_{V_1-b_1}(x) - x^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&= x(x\Phi_{V_1-b_1-b_2}(x) - x^{s-2}(x^\Delta - (\Delta-1)x)) - x^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&= x^2\Phi_{V_1-b_1-b_2}(x) - 2x^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&\ \ \vdots \\
&= x^s\Phi_{V_1-b_1-b_2-\cdots-b_s}(x) - sx^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&= x^s\Phi_{S_{\Delta+1}}(x) - sx^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&= x^s(x^{\Delta+1} - \Delta x^{\Delta-1}) - sx^{s-1}(x^\Delta - (\Delta-1)x^{\Delta-2}) \\
&= x^{s+\Delta-1}(x^2 - \Delta) - sx^{s+\Delta-3}(x^2 - (\Delta-1)) \\
&= x^{s+\Delta-3}(x^4 - (\Delta+s)x^2 + s(\Delta-1)).
\end{aligned}$$

We know that $\lambda(V_1) \neq 0$, which implies it is the largest solution of the equation $x^4 - (\Delta+s)x^2 + s(\Delta-1) = 0$. Then, $x^2 = \frac{\Delta+s+\sqrt{(\Delta+s)^2-4s(\Delta-1)}}{2}$. Replacing s by $n - \Delta - 1$ we get $x^2 = \frac{n-1+\sqrt{(n-2\Delta)^2+2n-3}}{2}$ and therefore $\lambda(V_1) = \sqrt{\frac{n-1+\sqrt{(n-2\Delta)^2+2n-3}}{2}}$.
Now, consider case (iii) where $2\Delta < n \leq \Delta^2 + 1$. We let V_2 be the Volkmann tree corresponding to the degree sequence of T, so that V_2 has 3 levels. We have already proved that $\lambda(T) \leq \lambda(V_2)$. V_2 is a subgraph (possibly equal to) of the complete Δ-ary tree $V*$ with 3 levels, so $\lambda(T) \leq \lambda(V_2) \leq \lambda(V^*)$. Moreover, $\lambda(V^*) = \sqrt{2\Delta - 1}$ (see [22]). Hence, the result holds.
Finally, consider case (iv) where $n > \Delta^2 + 1$. Denote by V_3 the Volkmann tree corresponding to the degree sequence of T. Now the height of V_3 is greater than three. Once again, we have $\lambda(T) \leq \lambda(V_3)$. With a similar reasoning as in (ii), V_3 is a subgraph of a complete Δ-ary tree V^* with the same height as V_3. By some matrix manipulations, Song et al ([26]) proved that $\lambda(V^*) < 2\sqrt{\Delta-1}\cos\frac{\pi}{2k+1}$. Thus the result holds. □

To end this chapter, let us give an example of a partial ordering of trees and see the behaviour of the Wiener index and the spectral radius.

Example 4.2.18. Figure 4.4 shows all the trees of order 6, with their Wiener index and spectral radius.

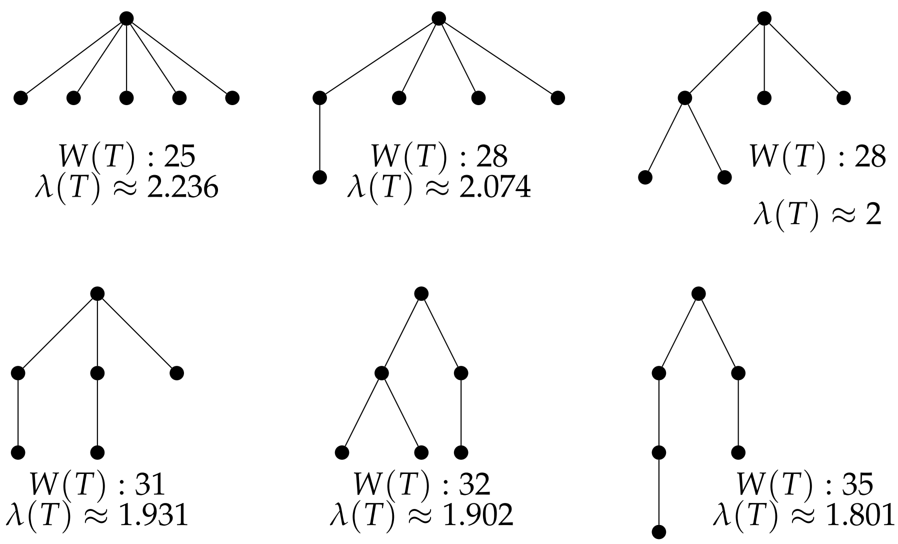


Figure 4.4: Trees of order 6.

Chapter 5

Additive parameters

Additive parameters occur frequently, especially in computer science in the analysis of divide-and-conquer algorithms. A lot of research has been done to compute the distribution of such parameters (see [9, 21, 29]). In this chapter, we present a slightly different point of view in characterizing extremal trees for graph invariants which are "additive". This includes several natural examples, as can be seen in the following.

5.1 Preliminaries

Let T be a rooted tree with root r. Let us denote the set of branches attached to r by $\{T_1, \ldots, T_k\}$.

Definition 5.1.1. A parameter $A(T)$ is called *additive* if

$$A(T) = A(T_1) + A(T_2) + \cdots + A(T_k) + f(T),$$

where f is called a *toll function*. We assume that for the single-vertex tree, we have $f(\bullet) = A(\bullet)$.

Example 5.1.2. (i) The number of leaves is an additive parameter, in fact:

$$\ell(T) = \ell(T_1) + \cdots + \ell(T_k) + f(T),$$

where

$$f(T) = \begin{cases} 1 & \text{if } |T| = 1, \\ 0 & \text{otherwise.} \end{cases}.$$

We assume the root is never considered as a leaf unless $|T| = 1$.

(ii) The total path length, which is the sum of distances from the root is also additive:

$$p(T) = p(T_1) + \cdots + p(T_k) + |T| - 1.$$

Definition 5.1.3. For each vertex $u \in V(T)$, recall that T_u is the subtree of the rooted tree T induced by u and its all successors. In particular if u is the root, then $T_u = T$.
Let us write $\alpha_T(u) = |V(T_u)|$ and $\alpha(T) = (\alpha_T(u), u \in V(T))$, where we may assume that $\alpha_T(u)$ are ordered in a non-increasing way in the sequence $\alpha(T)$.

Proposition 5.1.4. Let A be an additive parameter such that

$$A(T) = A(T_1) + \cdots + A(T_k) + f(|T|).$$

Then we have

$$A(T) = \sum_{u \in V(T)} f(\alpha_T(u)).$$

Proof. Let us prove the statement by induction on the height of the tree. (Recall that the height h is the maximum distance between the root and a vertex in T)
Since $f(\bullet) = A(\bullet)$, the statement is true for height 0. Suppose it is true for any trees of height less or equal to h. Let us compute A for a tree T of height $h+1$. If $T_1, T_2, \ldots, T_k$ are the branches attached to the root, we obtain

$$A(T) = A(T_1) + \cdots + A(T_k) + f(|T|),$$

but $T_1, \ldots, T_k$ are of height less or equal to h, so by the induction hypothesis we get

$$A(T) = \sum_{u \in V(T_1)} f(\alpha_T(u)) + \cdots + \sum_{u \in V(T_k)} f(\alpha_T(u)) + f(|T|) = \sum_{u \in V(T)} f(\alpha_T(u)).$$

□

Let us present some further simple properties [36] of majorization which will be used later.

Proposition 5.1.5. Let $A = (a_1, \ldots, a_k, b_1, \ldots, b_k)$ and $B = (a_1 + b, \ldots, a_k + b, b_1 - b, \ldots, b_k - b)$ be two nonnegative integer sequences with $b > 0$. If $a_i \geq b_i$ for $i = 1, \ldots, k$, then $A \preceq B$.

Proposition 5.1.6. Let $A = (a_1, \ldots, a_n)$ and $A' = (a'_1, \ldots, a'_m)$ be two sequences. We denote by (A, A') the sequence obtained by combining the two, i.e., $(A, A') = (a_1, \ldots, a_n, a'_1, \ldots, a'_m)$. If $A \preceq B$ and $A' \preceq B'$, then $(A, A') \preceq (B, B')$.

Definition 5.1.7. Let $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ be two sequences. We write $A \leq B$ if $a_i \leq b_i$ for $i = 1, \ldots, n$. If at least one of the inequalities is strict, then we write $A < B$.

5.2 Additive parameters and level-greedy trees

Let us characterize first the extremal trees among trees with given level degree sequence. Then, we will extend this result to trees with fixed degree sequence. Let $\mathbb{T}_r(D)$ be the set of all rooted trees with level degree sequence D. The two following lemmas that can be found in [36] show the behaviour of $\alpha(T)$ under modifications of the tree T.

Lemma 5.2.1. Let $T \in \mathbb{T}_r(D)$. Suppose that u and v are at the same level, that they are successors of w and that there are two internally disjoint paths $P(u, w) = (u, u_1, \ldots, u_k, w)$ and $P(v, w) = (v, v_1, \ldots, v_k, w)$ such that $\alpha_T(u) < \alpha_T(v)$, and $\alpha_T(u_i) \geq \alpha_T(v_i)$ for $i = 1, \ldots, k$. Let T' be the tree obtained from T by deleting the edges u_1u and v_1v and adding the edges u_1v and v_1u. Then $T' \in \mathbb{T}_r(D)$ and $\alpha(T) \preceq \alpha(T')$.

Proof. Note that all the degrees are preserved by this operation, hence T and T' have the same outdegrees at each level, in other words $T' \in \mathbb{T}_r(D)$. Set $b = \alpha_T(v) - \alpha_T(u) > 0$. Clearly, $\alpha_{T'}(v_i) = \alpha_T(v_i) - b$ and $\alpha_{T'}(u_i) = \alpha_T(u_i) + b$ for $i = 1, \ldots, k$, so by Proposition 5.1.5,

$$\begin{aligned}
&(\alpha_T(u_1), \ldots, \alpha_T(u_k), \alpha_T(v_1), \ldots, \alpha_T(v_k)) \\
&\preceq (\alpha_T(u_1) + b, \ldots, \alpha_T(u_k) + b, \alpha_T(v_1) - b, \ldots, \alpha_T(v_k) - b) \\
&= (\alpha_{T'}(u_1), \ldots, \alpha_{T'}(u_k), \alpha_{T'}(v_1), \ldots, \alpha_{T'}(v_k)).
\end{aligned}$$

Note also that for any vertex $y \in V(T) \backslash \{u_1, \ldots, u_k, v_1, \ldots, v_k\}$, we have $\alpha_{T'}(y) = \alpha_T(y)$. Therefore by Proposition 5.1.6, $\alpha(T) \preceq \alpha(T')$. $\square$

Lemma 5.2.2. Let $T \in \mathbb{T}_r(D)$. Suppose that u and v are at the same level, that they are successors of w and that there are two internally disjoint paths $P(u,w) = (u, u_1, \ldots, u_k, w)$ and $P(v,w) = (v, v_1, \ldots, v_k, w)$ such that $\alpha_T(u) \geq \alpha_T(v)$, and $\alpha_T(u_i) \geq \alpha_T(v_i)$ for $i = 1, \ldots, k$. If $d_T(u) < d_T(v)$ (recall that $d_T(u)$ is the degree of u in T), set $s = d_T(v) - d_T(u) > 0$, and let T' be the tree obtained from T by deleting the s edges vx_i and adding s edges ux_i, $i = 1, \ldots, s$, where $x_1, \ldots, x_s$ are children of v. Then $T' \in \mathbb{T}_r(D)$ and $\alpha(T) \preceq \alpha(T')$.

Proof. We can see that $d_{T'}(u) = d_T(v)$ and $d_{T'}(v) = d_T(u)$, while the degree of the other vertices is preserved, so $T' \in \mathbb{T}_r(D)$. Set $b = \sum_{i=1}^{s} \alpha_T(x_i) > 0$. We can see that $\alpha_{T'}(v) = \alpha_T(v) - b$, $\alpha_{T'}(v_i) = \alpha_T(v_i) - b$ and $\alpha_{T'}(u) = \alpha_T(u) + b$, $\alpha_{T'}(u_i) = \alpha_T(u_i) + b$ for $i = 1, \ldots, k$. Hence by Proposition 5.1.5, we have

$$\begin{aligned}&(\alpha_T(u), \alpha_T(u_1), \ldots, \alpha_T(u_k), \alpha_T(v), \alpha_T(v_1), \ldots, \alpha_T(v_k))\\ &\preceq (\alpha_{T'}(u), \alpha_{T'}(u_1), \ldots, \alpha_{T'}(u_k), \alpha_{T'}(v), \alpha_{T'}(v_1), \ldots, \alpha_{T'}(v_k)).\end{aligned}$$

Moreover, for any vertex $y \in V(T) \backslash \{u_1, \ldots, u_k, v_1, \ldots, v_k\}$, we have $\alpha_{T'}(y) = \alpha_T(y)$. Therefore by Proposition 5.1.6, $\alpha(T) \preceq \alpha(T')$. □

Lemma 5.2.3. For any $T \in \mathbb{T}_r(D)$, we have

$$\alpha(T) \preceq \alpha(G(T)),$$

where $G(T)$ is the rooted level greedy tree corresponding to D.

Proof. Let $T \in \mathbb{T}_r(D)$. Let us denote by $u_{1,1}, u_{1,2}, \ldots, u_{1,\ell_1}$ the vertices of T at level 1, where without loss of generality we may assume that $d_T(u_{1,1}) \geq d_T(u_{1,2}) \geq \cdots \geq d_T(u_{1,\ell_1})$. Suppose there exist i, j with $i < j$ such that $\alpha_T(u_{1,i}) < \alpha_T(u_{1,j})$, but $d_T(u_{1,i}) > d_T(u_{1,j})$. We apply Lemma 5.2.2 to obtain a new tree T' such that $d_{T'}(u_{1,j}) > d_{T'}(u_{1,i})$ and $\alpha_{T'}(u_{1,j}) \geq \alpha_{T'}(u_{1,i})$. Moreover, $\alpha(T) \preceq \alpha(T')$. We iterate this process to any possible i, j to end up with a tree T_1, such that if $u^1_{1,1}, u^1_{1,2}, \ldots, u^1_{1,\ell_1}$ are the vertices of T_1 at level 1, then $d_{T^1}(u^1_{1,1}) \geq d_{T^1}(u^1_{1,2}) \geq \cdots \geq d_{T^1}(u^1_{1,\ell_1})$ and $\alpha_{T^1}(u^1_{1,1}) \geq \alpha_{T^1}(u^1_{1,2}) \geq \cdots \geq \alpha_{T^1}(u^1_{1,\ell_1})$. Furthermore $\alpha(T) \preceq \alpha(T_1)$.
Now let us consider level 2. Suppose u is a child of $u^1_{1,i}$ and v is a child of $u^1_{1,j}$, with $i < j$ such that $\alpha_{T_1}(u) < \alpha_{T_1}(v)$. We apply Lemma 5.2.1 to obtain a new tree T'_1 such that $\alpha_{T'_1}(u) > \alpha_{T'_1}(v)$ and $\alpha(T_1) \preceq \alpha(T'_1)$. Now

if $d_{T_1'}(u) < d_{T_1'}(v)$, we apply again Lemma 5.2.2 to obtain a tree T_1'' such that $d_{T_1''}(u) \geq d_{T_1''}(v)$ and $\alpha_{T_1''}(u) \geq \alpha_{T_1''}(v)$. Moreover, $\alpha(T_1') \preceq \alpha(T_1'')$. By repeating this operation for all vertices at level 2, we end up with a new tree T_2 such that the vertices at level 2 are ordered in a nonincreasing order according to their degrees and $\alpha(T_1) \preceq \alpha(T_2)$.
We continue to apply Lemma 5.2.1 and 5.2.2 for the next levels, and we end up with a tree T_H such that the vertices at all levels are ordered, which means $T_H \cong G(T)$. Furthermore, $\alpha(T) \preceq \alpha(T_1) \preceq \alpha(T_2) \ldots \preceq \alpha(T_H)$. Since majorization is transitive, the statement follows. □

The main part of this section is to prove that for specific toll functions, the level greedy tree either minimizes or maximizes the additive parameter.

Theorem 5.2.4. *If a parameter is additive and the toll function is of the form $t(T) = f(|T|)$, where f is a convex (concave) function then it is maximised (minimised) by the rooted level-greedy tree among all trees with the same outdegree level sequence.*

In order to prove Theorem 5.2.4, we need the following inequality, which is known as Karamata's inequality.

Lemma 5.2.5 ([14])**.** Let $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ be two sequences of real numbers such that $a_1 + \cdots + a_n = b_1 + \cdots + b_n$. If the sequence B majorizes A, and f is a convex funtion, then the inequality

$$\sum_{i=1}^{n} f(a_i) \leq \sum_{i=1}^{n} f(b_i)$$

holds.

Proof. Let $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ be two sequences such that $A \preceq B$, and let f be a convex function. Without loss of generality, assume that A and B are in a non-increasing order and f is continuously differentiable. f is convex on an interval I if and only if f lies above all its tangents, i.e.:

$$f(x) - f(y) \geq (x - y) f'(y),$$

for all x, y in I. Thus $f(a_i) - f(b_i) \geq (a_i - b_i)f'(b_i)$ holds for all $i = 1, \ldots, n$. Let us compute the difference between $\sum_{i=1}^{n} f(a_i)$ and $\sum_{i=1}^{n} f(b_i)$.

$$\sum_{i=1}^{n} f(a_i) - \sum_{i=1}^{n} f(b_i) = \sum_{i=1}^{n} f(a_i) - f(b_i) \geq \sum_{i=1}^{n} (a_i - b_i)f'(b_i)$$
$$= (a_1 - b_1)(f'(b_1) - f'(b_2)) + (a_1 + a_2 - b_1 - b_2)(f'(b_2) - f'(b_3))$$
$$+ (a_1 + a_2 + a_3 - b_1 - b_2 - b_3)(f'(b_3) - f'(b_4)) + \cdots$$
$$+ (\sum_{i=1}^{n-1} a_i - \sum_{i=1}^{n-1} b_i)(f'(b_{n-1}) - f'(b_n)) + (\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i)f'(b_n).$$

Note that for all $k = 1, \ldots, n-1$, $\sum_{i+1}^{k} a_i - \sum_{i=1}^{k} b_i \leq 0$ by the definition of the majorization, and $f'(b_k) - f'(b_{k+1}) \leq 0$ by the convexity of f. Moreover, the last term is 0 by our assumptions. Thus this difference is nonnegative. □

Proof of Theorem 5.2.4. Let A be an additive parameter with a convex toll function f. Then Proposition 5.1.4 provides that

$$A(T) = \sum_{u \in V(T)} f(\alpha_T(u)),$$

and Lemma 5.2.3 tells us that for any tree $T \in \mathbb{T}_r(D)$,

$$\alpha(T) \preceq \alpha(G(T)).$$

Using Lemma 5.2.5, we obtain that

$$A(T) = \sum_{u \in V(T)} f(\alpha_T(u)) \leq A(G(T)) = \sum_{u \in V(G(T))} f(\alpha_{G(T)}(u)).$$

If f is concave, we use the fact that $-f$ is convex to obtain the result. □

Let us study the behaviour of $\alpha(T)$ when we consider it as an edge rooted tree. Let T be a rooted tree with a root u and v one of its neighbours. Now, we consider the tree T as an edge rooted tree with root $e = uv$ whose level degree sequence is given by D.
Let $\mathbb{T}_e(D)$ be the set of all edge rooted trees with level degree sequence D.

Lemma 5.2.6. If $T_e \in \mathbb{T}_e(D)$, then:

$$(\alpha_T(w) | w \neq v) \preceq (\alpha_{G(T_e)}(w) | w \neq v),$$

where $G(T_e)$ is the edge-rooted level greedy tree corresponding to D.

Proof. Let T' be the tree obtained by merging u and v to obtain a vertex rooted tree with root r (as seen in Figure 5.1). We can easily see that

$$(\alpha_T(w)|w \in (V(T)\backslash\{u,v\})) = (\alpha_{T'}(w)|w \in V(T')\backslash r)$$

and

$$(\alpha_{G(T_e)}w|w \in V(G(T_e))\backslash\{u,v\}) = (\alpha_{G(T')}(w)|w \in V(G(T'))\backslash r).$$

Using Lemma 5.2.3, we know that $\alpha(T') \preceq \alpha(G(T')$ and of course $\alpha_{T'}(r) = \alpha_{G(T')}(r)$, so $(\alpha_{T'}(w)|w \in V(T')\backslash r)) \preceq (\alpha_{G(T')}(w)|w \in V(G(T'))\backslash r)$. Thus

$$(\alpha_T(w)|w \in (V(T)\backslash\{u,v\})) \preceq (\alpha_{G(T_e)}(w)|w \in V(G(T_e))\backslash\{u,v\}).$$

Moreover $\alpha_T(u) = \alpha_{G(T_e)}(u)$, so using Proposition 5.1.6, we have $(\alpha_T(w)|w \neq v) \preceq (\alpha_{G(T_e)}(w)|w \neq v)$.

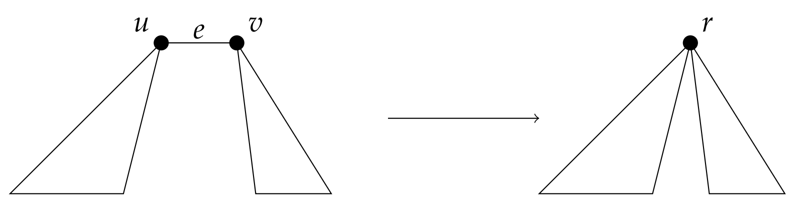


Figure 5.1: T and T' in the proof of Lemma 5.2.6.

□

Remark 5.2.7. Let us discuss the value of $\alpha_T(v)$. Two cases can be considered:

Case 1: if $d_T(u) \geq d_T(v)$, then after reshuffling the vertices at each level to get a level greedy tree, the vertices are ordered in such a way that successors of u have greater or equal degree than successors of v. It implies that the degree of the successors of v in the edge-rooted level greedy tree is smaller than for any other trees with the same degree sequence. Thus $\alpha_T(v) \geq \alpha_{G(T_e)}(v)$.

Case 2: if $d_T(u) < d_T(v)$, then after reshuffling the vertices, they are now ordered in such a way that the successors of v have greater or equal degree than successors of u. Hence $\alpha_T(v) \leq \alpha_{G(T_e)}(v)$.

5.3 Additive parameters and greedy trees

For our purposes, let us consider the following definition of a greedy tree as given in [2].

Definition 5.3.1. If a tree is level greedy with k levels and its level degree sequence satisfies

$$\min(a_{i,1}, \ldots, a_{i,\ell_i}) \geq \max(a_{i+1,1}, \ldots, a_{i+1,\ell_{i+1}}),$$

for any level i, $0 \leq i \leq k-1$, then it is a greedy tree.

Let $\mathbb{T}(D)$ be the set of all rooted trees (with root r) with given degree sequence $D = (d_0, d_1, \ldots, d_n)$. We denote by $G(D)$ the greedy tree corresponding to D. We have already seen in the previous section how $\alpha(T)$ behaves when we rearrange vertices on the same level. Now, we would like to know the effect on $\alpha(T)$ caused by rearranging vertices at different levels. In the spirit of Lemma 3.5 in [36], we have

Lemma 5.3.2. Let $T \in \mathbb{T}(D)$. Suppose that u is a successor of v, that there is a path $P(u,v) = (u, u_1, \ldots, u_k, v)$ and that $d_T(u) > d_T(v)$. Set $s = d_T(u) - d_T(v) > 0$, and let T' be a tree obtained from T by deleting the s edges ux_i, and adding s edges $vx_i, i = 1, \ldots, s$, where $\{x_1, \ldots, x_s\}$ are children of u. Then $T' \in \mathbb{T}(D)$ and $\alpha(T') < \alpha(T)$.

Proof. We can see that $d_T(u) = d_{T'}(v), d_T(v) = d_{T'}(u)$, and the degrees of the other vertices stay the same, hence $T' \in \mathbb{T}_r(D)$. Set $b = \sum_{i=1}^{s} \alpha_T(x_i) > 0$. Then $\alpha_{T'}(u) = \alpha_T(u) - b < \alpha_T(u)$ and $\alpha_{T'}(u_i) = \alpha_T(u_i) - b < \alpha_T(u_i)$ for $j = 1, \ldots, k$. Moreover for any $y \in V \backslash \{u, u_1, \ldots, u_k\}$, we have $\alpha_{T'}(y) = \alpha_T(y)$. Hence the statement holds. □

Now, we are ready to prove our main theorem, stated as follows:

Theorem 5.3.3. *If a parameter is additive and the toll function is of the form $t(T) = f(|T|)$, where f is a decreasing and convex (increasing and concave) function then it is maximised (minimised) by the greedy tree among all rooted trees with the same degree sequence.*

We prove Theorem 5.3.3 following the same construction as given in [2, Theorem 3].

Proof. Let A be an additive parameter with a strictly decreasing and convex toll function f. Then Proposition 5.1.4 shows that

$$A(T) = \sum_{u \in V(T)} f(\alpha_T(u)).$$

Let T_H be a tree in $\mathbb{T}(D)$ that satisfies $A(T_H) \geq A(T)$ for any rooted tree T with the same degree sequence D.

We observe from Lemma 5.3.2 that we can choose T_H in such a way that the root has maximum degree and the degrees of the vertices decrease as we move away from the root following a path. This decreases α, and since f is decreasing, A will increase as a result. Therefore T_H can be considered as a rooted tree satisfying this property. Hence, if $D = (d_0, d_1, \ldots, d_n)$ is a degree sequence in non-increasing order, then $d(r) = d_0$ (we write $d(u)$ for $d_{T_H}(u)$) and there exists a neighbour v of r with $d(v) = d_1$.

By Theorem 5.2.4, T_H can be chosen to be a level-greedy tree.

Let e be the edge rv and T'_H the component of $T_H - e$ that contains the root r, as shown in Figure 5.2.

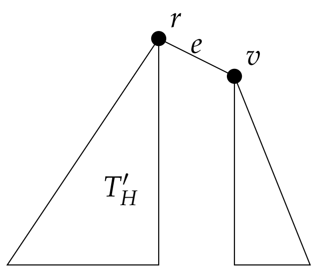


Figure 5.2: The tree T_H.

Let us consider T_H as an edge rooted tree with e as the root. By Lemma 5.2.6 we have

$$(\alpha_{T_H}(w) | w \neq v) \preceq (\alpha_{G(T_H)}(w) | w \neq v). \tag{5.3.1}$$

Moreover by our construction, $d(v) \leq d(r)$, so following the first case in Remark 5.2.6, we have

$$\alpha_{T_H}(v) \geq \alpha_{G(T_H)}(v). \tag{5.3.2}$$

We notice that the equality in equation (5.3.2) holds if and only if T_H is already level greedy with respect to e. Thus if we reshuffle the branches in

T_H to become a level greedy tree with edge root e, and apply Karamata's inequality to equation (5.3.1), we obtain

$$\sum_{w \neq v} f(\alpha_{T_H}(w)) \leq \sum_{w \neq v} f(\alpha_{G(T_H)}(w)).$$

Now using the fact that f is decreasing together with equation (5.3.2), we obtain

$$f(\alpha_{T_H}(v)) \leq f(\alpha_{G(T_H)}(v)).$$

Therefore the reshuffling process will increase A. Since all processes we described will strictly increase some additive parameter A', at some point it has to stop and we end up with a tree T_H that is level greedy with respect to r and also level greedy with respect to the edge e.

For contradiction, assume T_H (vertex rooted) is not isomorphic to the rooted greedy tree $G(D)$. Then for some $i \geq 2$, there exist vertices u_i and u_{i+1} at levels i and $i+1$ respectively such that $d(u_i) < d(u_{i+1})$. Now we distinguish four different cases. In each of them, we make use of the fact that T_H is a rooted level greedy tree, and w_i and w_{i+1} always denote vertices at level i and $i+1$ respectively.

Case 1: If $u_i, u_{i+1} \in V(T'_H)$, then there exists a vertex $w_{i+1} \in V(T_H - T'_H)$ such that $d(u_i) < d(u_{i+1}) \leq d(w_{i+1})$.

Case 2: If $u_i, u_{i+1} \in V(T_H - T'_H)$, then there exists a vertex $w_i \in V(T'_H)$ such that $d(w_i) \leq d(u_i) < d(u_{i+1})$. If level i of T'_H is already empty, we set $d(w_i) = 0$, and the argument that follows is still valid.

Case 3: If $u_i \in V(T_H - T'_H)$ and $u_{i+1} \in V(T'_H)$, then there exist vertices $w_i \in V(T'_H)$ and $w_{i+1} \in V(T_H - T'_H)$ such that

$$d(w_i) \leq d(u_i) < d(u_{i+1}) \leq d(w_{i+1}).$$

The case that level i of T'_H is empty is treated in the same way as before. Hence all three cases above can be reduced to the following fourth case:

Case 4: $u_i \in V(T'_H)$ and $u_{i+1} \in V(T_H - T'_H)$, but this contradicts the fact that T_H is level greedy as edge rooted tree with root e.

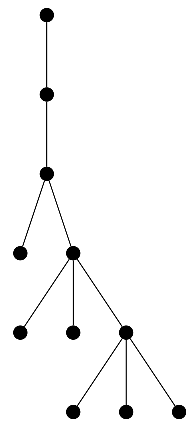
We obtain the case when f is strictly increasing and concave in a similar way. □

It is also interesting to find trees "opposite" to greedy trees in maximizing or minimizing additive parameters.

Definition 5.3.4 ([24])**.** A *caterpillar* is a tree with the property that a path remains if all leaves are deleted. In our case, we will call a caterpillar any rooted tree where removing all leaves produces a rooted tree with precisely one leaf.

Definition 5.3.5. A *reverse greedy caterpillar* is a rooted caterpillar for which the degrees of its internal vertices increase from the root to a leaf.

Example 5.3.6. A reverse greedy caterpillar with the following degree sequence: $(4, 4, 3, 2, 1, 1, 1, 1, 1, 1, 1)$.



Lemma 5.3.7. Let $T \in \mathbb{T}(D)$. Suppose u and v are successors of w and there are two internally disjoint paths $P(u, w) = (u, u_1, \ldots, u_k, w)$ and $P(v, w) = vw$ such that $u_k \neq v$, $k \geq 1$ and $d_T(v) > d_T(u) = 1$. Set $s = d_T(v) - 1 > 0$, and let T' be a tree obtained from T by deleting the s edges vx_i and adding s edges ux_i, $i = 1, \ldots, s$, where $x_1, \ldots, x_s$ are children of v. Then T' is a rooted tree with the same degree sequence and $\alpha(T) < \alpha(T')$.

Proof. Since $d_{T'}(u) = d_T(v)$ and $d_{T'}(v) = d_T(u)$, T' and T have the same degree sequence. Set $b = \sum_{i=1}^{s} \alpha_T(x_i) > 0$. We see that v becomes a leaf and $\alpha_{T'}(u) = \alpha_T(v), \alpha_{T'}(v) = \alpha_T(u)$, $\alpha_{T'}(u_i) = \alpha_T(u_i) + b > \alpha_T(u_i)$ for $i = 1, \ldots, k$. Moreover, for any vertex $y \in V(T) \setminus \{u, u_1, \ldots, u_k, v\}$, we have $\alpha_{T'}(y) = \alpha_T(y)$. Therefore $\alpha(T) < \alpha(T')$. □

Theorem 5.3.8. *If a parameter is additive and the toll function is of the form $t(T) = f(|T|)$, where f is a strictly increasing (decreasing) function then it is maximised (minimised) by the reverse greedy caterpillar among all rooted trees with the same degree sequence.*

Proof. As in Theorem 5.3.3, let A be an additive parameter with a strictly increasing toll function f. Then Proposition 5.1.4 shows that

$$A(T) = \sum_{u \in V(T)} f(\alpha_T(u)).$$

Let T_H be a tree in $\mathbb{T}_r(D)$ that satisfies $A(T_H) \geq A(T)$ for any rooted tree T with the same degree sequence D. Suppose that there is more than one non-leaf vertex in T_H. Assume there are two non-leaf vertices u_i, v_i at level i. At level 1, we apply Lemma 5.3.7 to v_1 and a leaf successor of u_1 to obtain a new tree T'_H such that $\alpha(T_H) < \alpha(T'_H)$. Since f is increasing, this process makes A increase. Repeating this process, we end up with a tree that has only one non-leaf vertex at level 1. A increases with each application of Lemma 5.3.7. We iterate this idea for the next levels to obtain a tree that has only one non-leaf vertex at each level.
By Lemma 5.3.2, we can choose T_H such that the root has minimum degree and the degrees of the vertices increase as we move away from the root following a path (except for a leaf). The associated sequence α increases, and since f is increasing, A will increase.
In view of Definition 5.3.5, we see that T_H is the reverse greedy caterpillar. □

Remark 5.3.9. Note that in Theorem 5.3.8, the toll function does not have to be convex or concave.

Now let us illustrate Theorems 5.3.3 and 5.3.8 by some examples.

Corollary 5.3.10. The total path length $P(T)$ described in Example 5.1.2 (ii) is maximized by the reverse greedy caterpillar, and the greedy tree is one of the trees that minimize it.

Proof. The total path length is additive with toll function $t(T) = |T| - 1$, which is increasing. Hence by Theorem 5.3.8, it is maximized by the reverse greedy caterpillar. The toll function t is also convex, but not strictly, so the greedy tree $G(D)$ minimizes $P(T)$, but it is not the only such tree: any tree with the same level degree sequence as $G(D)$ also minimizes it. □

Example 5.3.11. Let us consider the rooted trees with degree sequence $D = (3, 3, 3, 2, 2, 1, 1, 1, 1, 1)$ depicted in Figure 5.3.

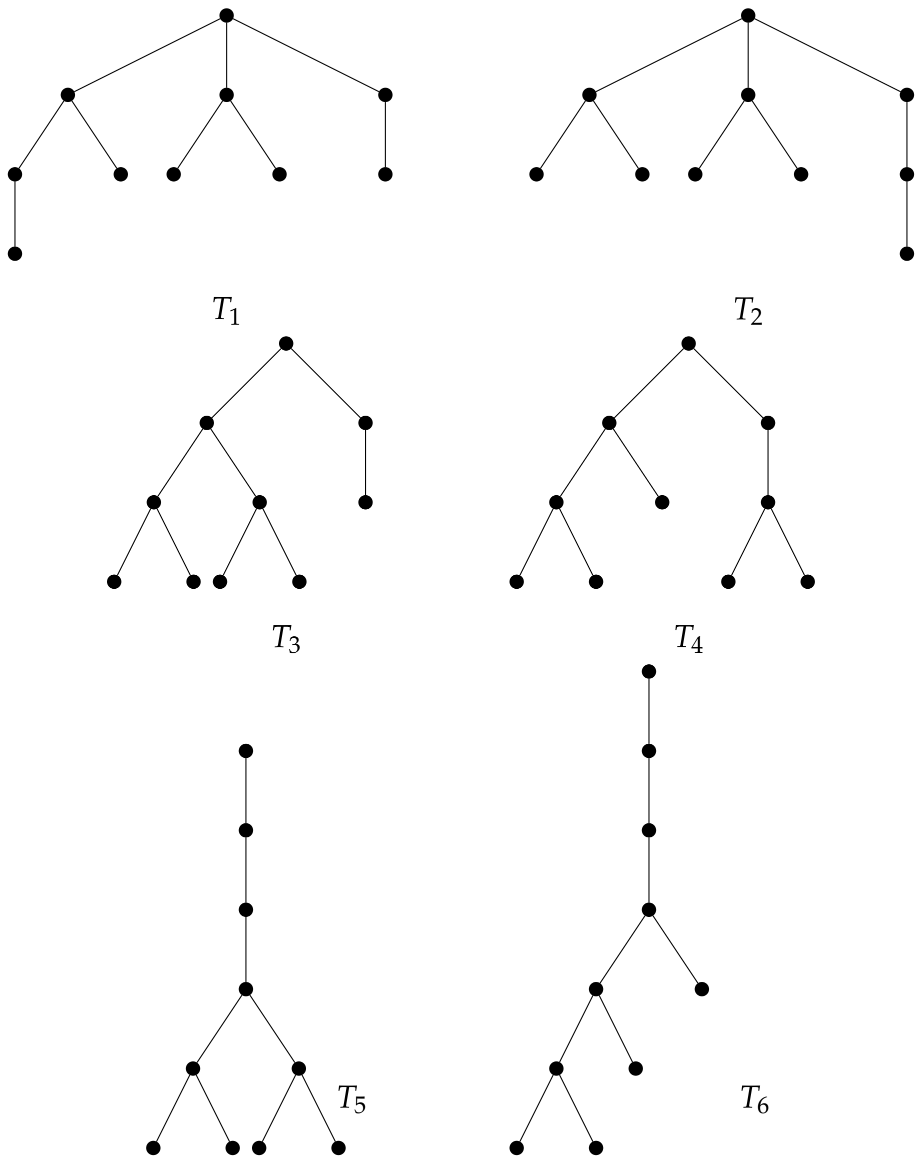


Figure 5.3: Rooted trees with degree sequence $(3, 3, 3, 2, 2, 1, 1, 1, 1, 1)$.

We compute their total path length to obtain:

$$P(T_1) = P(T_2) = 1 \times 3 + 2 \times 5 + 3 \times 1 = 16$$
$$P(T_3) = P(T_4) = 1 \times 2 + 2 \times 3 + 3 \times 4 = 20$$
$$P(T_5) = 1 \times 1 + 2 \times 1 + 3 \times 1 + 4 \times 2 + 5 \times 4 = 34$$
$$P(T_6) = 1 \times 1 + 2 \times 1 + 3 \times 1 + 4 \times 2 + 5 \times 2 + 6 \times 2 = 36.$$

The results confirm Corollary 5.3.10. Note that T_1 has the same level degree sequence as T_2, the same for T_3 and T_4. However since the toll function is not strictly convex, nor concave, we cannot distinguish their total path lengths from one another.

Definition 5.3.12 ([3])**.** The *factorial* of a tree is a graph invariant defined recursively as follows:

$$T! = n \prod_{i=1}^{k} T_i!, \tag{5.3.3}$$

where T is a rooted tree of order n, and $\{T_1, \ldots, T_k\}$ is the set of branches attached to the root r. We set $\bullet ! = 1$.

For a combinatorial aspect, we remark that the factorial of a tree counts the number of functions $V \to V$ such that each vertex is mapped to a successor (or itself).

Example 5.3.13. Let us compute the factorial of the path P_n rooted at one of its end vertices and the star S_n rooted at its center.

$$P_n! = nP_{n-1}! = n(n-1)P_{n-2}! = \cdots = n!\bullet ! = n!,$$

which gives us the ordinary factorial. For the star, we get

$$S_n! = n \prod_{i=1}^{n-1} \bullet ! = n.$$

Corollary 5.3.14. The tree factorial is minimized by the greedy tree and maximized by the reverse greedy caterpillar among all trees with the same degree sequence.

Proof. Note that

$$\begin{aligned} T! &= |T| \prod_{i=1}^{k} T_i! \\ &= \exp\left(\log(|T| \prod_{i=1}^{k} T_i!) \right) \\ &= \exp\left(\sum_{i=1}^{k} \log(T_i!) + \log(|T|) \right) = \exp(\ell(T)). \end{aligned}$$

As we can see, $\ell(T)$ is an additive parameter, whose toll function is $t(T) = \log(|T|)$. Since log is an increasing concave function, by Theorem 5.3.3 and 5.3.8, $\ell(T)$ is minimized by the greedy tree and maximized by the reverse greedy caterpillar among all rooted trees with the same degree sequence. Moreover exp is increasing, so $\ell(T)$ and $T!$ have the same extremal trees. Thus, the statement follows. □

Example 5.3.15. Let us consider again the trees with degree sequence $D = (3,3,3,2,2,1,1,1,1,1)$ depicted in Figure 5.3, and compute their factorials.

$$\begin{aligned}
T_1! &= 10(4(2) \times 3 \times 2) = 480 \\
T_2! &= 10(3 \times 3 \times 3(2)) = 540 \\
T_3! &= 10(7(3)(3) \times 2) = 1260 \\
T_4! &= 10(5(3) \times 4(3)) = 1800 \\
T_5! &= 10(9)(8)(7(3 \times 3)) = 45360 \\
T_6! &= 10(9)(8)(7(5(3))) = 75600.
\end{aligned}$$

The results obtained confirm Corollary 5.3.14. Moreover, note that T_1 and T_2 (as well as T_3 and T_4) have the same level sequence but T_1 and T_3 are level greedy. The fact that $T_1! < T_2!$ and $T_3! < T_4!$ follows from Theorem 5.2.4. These examples show that the convexity of the function is important.

It is worth mentioning that an additive parameter $L(T)$ with toll function $\log(|T|)$ corresponds also to the so-called "shape parameter" ([21]), it has been shown that $\exp(L(T))$ provides a measure of the shape of T. This parameter has the same extrema as the tree factorial.

List of References

[1] E. O. D. Andriantiana and S. Wagner. Spectral moments of trees with given degree sequence. *Linear Algebra Appl.*, 439(12):3980–4002, 2013.

[2] E. O. D. Andriantiana, S. Wagner, and H. Wang. Greedy trees, subtrees and antichains. *Electron. J. Combin.*, 20(3):P28, 2013.

[3] S. T. Belinschi, M. Bożejko, F. Lehner, and R. Speicher. The normal distribution is infinitely divisible. *Adv. Math.*, 226(4):3677–3698, 2011.

[4] N. Biggs. *Algebraic Graph Theory*. Cambridge Mathematical Library. Cambridge University Press, 1993.

[5] T. Bıyıkoglu and J. Leydold. Graphs with given degree sequence and maximal spectral radius. *Electron. J. Combin.*, 15:R119, 2008.

[6] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*. Macmillan Press, New York, 1976.

[7] D.M. Cvetković, M. Doob, and H. Sachs. *Spectra of graphs: theory and application*. Pure Appl. Math. Academic Press, 1980.

[8] A. A. Dobrynin, R. Entringer, and I. Gutman. Wiener index of trees: theory and applications. *Acta Appl. Math.*, 66(3):211–249, 2001.

[9] J. A. Fill and N. Kapur. Limiting distributions for additive functionals on catalan trees. *Theor. Comput. Sci.*, 326(1):69–102, 2004.

[10] M. Fischermann, A. Hoffmann, D. Rautenbach, L. Székely, and L. Volkmann. Wiener index versus maximum degree in trees. *Discrete Appl. Math.*, 122(1):127–137, 2002.

[11] G. H. Hardy, John E. Littlewood, and G. Pólya. *Inequalities*. Cambridge university press, 1952.

[12] R. A. Horn and C. R. Johnson. *Matrix analysis. Reprinted with corrections*. Cambridge university press, 1990.

[13] C. Jordan. Sur les assemblages de lignes. *J. Reine Angrew. Math.*, 70:185–190, 1869.

[14] Z. Kadelburg, D. Dukic, M. Lukic, and I. Matic. Inequalities of Karamata, Schur and Muirhead, and some applications. *Teach. Math.*, 8(1):31–45, 2005.

[15] R. Kirk and H. Wang. Largest number of subtrees of trees with a given maximum degree. *SIAM J. Discrete Math.*, 22(3):985–995, 2008.

[16] D. J. Klein, I. Lukovits, and I. Gutman. On the definition of the hyper-Wiener index for cycle-containing structures. *J. Chem. Inf. Comp. Sci.*, 35(1):50–52, 1995.

[17] B. Knudsen. Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree. In *Algorithms in Bioinformatics*, pages 433–446. Springer, 2003.

[18] S. Li and Y. Song. On the spectral moment of trees with given degree sequences. *arXiv preprint arXiv:1209.2188*, 2012.

[19] B. Liu, J. Shen, and X. Wang. On the largest eigenvalue of non-regular graphs. *J. Comb. Theory, Ser. B*, 97:1010–1018, 2007.

[20] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications: Theory of Majorization and Its Applications*. Springer, 2010.

[21] A. Meir and J. W. Moon. On the log-product of the subtree-sizes of random trees. *Random Struct. Algor.*, 12(2):197–212, 1998.

[22] R. Oscar and S. Ricardo. The spectra of the adjacency matrix and Laplacian matrix for balanced trees. *Linear Algebra Appl.*, (403):97–117, 2005.

[23] D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić. On the Harary index for the characterization of chemical graphs. *J. Math. Chem.*, 12(1):235–250, 1993.

[24] N. S. Schmuck, S. G. Wagner, and H. Wang. Greedy Trees, Caterpillars, and Wiener-type Graph Invariants. *MATCH Commun. Math. Comput. Chem.*, 68:273–292, 2012.

[25] J. Shu and Y. Wu. Sharp upper bounds on the spectral radius of graphs. *Linear Algebra Appl.*, 377:241–248, 2004.

[26] H. Song, Q. Wang, and L. Tian. New upper bounds on the spectral radius of trees with the given number of vertices and maximum degree. *Linear Algebra Appl.*, 439(9):2527–2541, 2013.

[27] L. A. Székely and H. Wang. On subtrees of trees. *Adv. Appl. Math.*, 34(1):138–155, 2005.

[28] L. A. Székely, H. Wang, and T. Wu. The sum of the distances between the leaves of a tree and the semi-regular property. *Discrete Math.*, 311(13):1197–1203, 2011.

[29] S. Wagner. Additive tree functionals with small toll functions and subtrees of random trees. *DMTCS Proceedings*, (AQ):67–80, 2012.

[30] S. Wagner, H. Wang, and X-D. Zhang. Distance-based graph invariants of trees and the Harary index. *Filomat*, 27(1):41–50, 2013.

[31] H. Wang. Corrigendum: The extremal values of the Wiener index of a tree with given degree sequence. *Discrete Appl. Math.*, 157:3754, 2009.

[32] H. Wang. The extremal values of the Wiener index of a tree with given degree sequence. *Discrete Appl. Math.*, 156:2647–2654, 2009.

[33] H. Wiener. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.*, 69:17–20, 1947.

[34] B. Wu, E. Xiao, and Y. Hong. The spectral radius of trees on k pendant vertices. *Linear Algebra Appl.*, 395:343–349, 2005.

[35] X-D. Zhang. The Laplacian spectral radii of trees with given degree sequences. *Discrete Math.*, 308(15):3143–3150, 2008.

[36] X-D. Zhang, Q-Y. Xiang, L-Q. Xu, and R-Y. Pan. The Wiener index of trees with given degree sequences. *MATCH Commun. Math. Comput. Chem*, 60:623–644, 2008.

[37] X-M. Zhang, X-D. Zhang, D. Gray, and H. Wang. The number of subtrees of trees with given degree sequence. *J. Graph Theory*, 73(3):280–295, 2013.