Spring 4-21-2011

EXTREMAL TREES AND RECONSTRUCTION

Andrew Ray
University of Nebraska-Lincoln, ray.andrew@gmail.com

EXTREMAL TREES AND RECONSTRUCTION

by

Andrew Benjamin Ray

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Mathematics

Under the Supervision of Professor Jamie Radcliffe

Lincoln, Nebraska

May, 2011

EXTREMAL TREES AND RECONSTRUCTION

Andrew Benjamin Ray, Ph. D.

University of Nebraska, 2011

Adviser: Jamie Radcliffe

Problems in two areas of graph theory will be considered.

First, I will consider extremal problems for trees. In these questions we examine the trees that maximize or minimize various invariants. For instance the number of independent sets, the number of matchings, the number of subtrees, the sum of pairwise distances, the spectral radius, and the number of homomorphisms to a fixed graph. I have two general approaches to these problems. To find the extremal trees in the collection of trees on n vertices with a fixed degree bound I use the certificate method. The certificate is a branch invariant, related to, but not the same as, the original invariant. We exploit the recursive structure of the problem. The second approach is geared towards finding the trees with given degree sequence that are extremal. I have a common approach involving labelings of the vertices corresponding to each invariant; the canonical example of which is labeling the vertices by the components of the leading eigenvector. This approach yields strictly stronger results when combined with a majorization result.

Second, I will consider two problems in graphs reconstruction. For these problems we are given limited information about a graph and decide whether the graph is uniquely determined by this data. The first problem is reconstruction of trees from their k-subtree matrix; a generalization of the Wiener matrix. This includes the problem of reconstruction from the Wiener matrix which was an open problem. Two vertices are adjacent if the corresponding entry is the largest in either its row or

its column. The second problem is reconstructing graphs from metric balls of their vertices. I give a solution to the conjecture that every graph with no pendant vertices and girth at least $2r+3$ can be reconstructed from its metric balls of radius r. We do so by examining the intersections of metric balls and their sizes.

ACKNOWLEDGMENTS

I would like to thank my advisor Jamie Radcliffe for all his guidance. I am grateful to have had the opportunity to work with him on many interesting problems for the last few years. I also would like thank the rest of my committee: Stephen Hartke, Christine Kelly, and Jitender Deogun for helping make this happen. Finally, I would like to thank my wonderful wife Elizabeth for putting up with all the late nights and short weekends over the last five years.

Contents

Chapter 1

Introduction

1.1 Maximizing and Minimizing Graph Invariants

Graph invariants distill the structure of a graph down to a number independent of the graph's representation. Thus understanding an invariant gives us information about the structure of a graph. Many graph invariants have been developed, each telling us something different about the structure of the graph.

One graph invariant that we will consider, in various contexts, in this thesis is the Wiener index. The Wiener index of a graph is defined to be the sum of the distances between each pair of vertices in the graph. Thus the Wiener index of K_n is $\binom{n}{2}$, and that of the P_n is (counting the number of pairs at distance k) $\sum_{k=1}^{n-1}(n-k)k = \binom{n+2}{3}$. A particularly interesting class of graphs to examine, for this and many invariants, is the class of trees on n vertices. In Figure 1.1 we list every tree (up to isomorphism) on 7 vertices together with its Wiener index and number of independent sets.

It is a well known fact that for trees the Wiener index is maximized by the path and minimized by the star. Roughly speaking, the less branched the tree is, the higher the Wiener index. A first observation is that trees with the same maximum degree

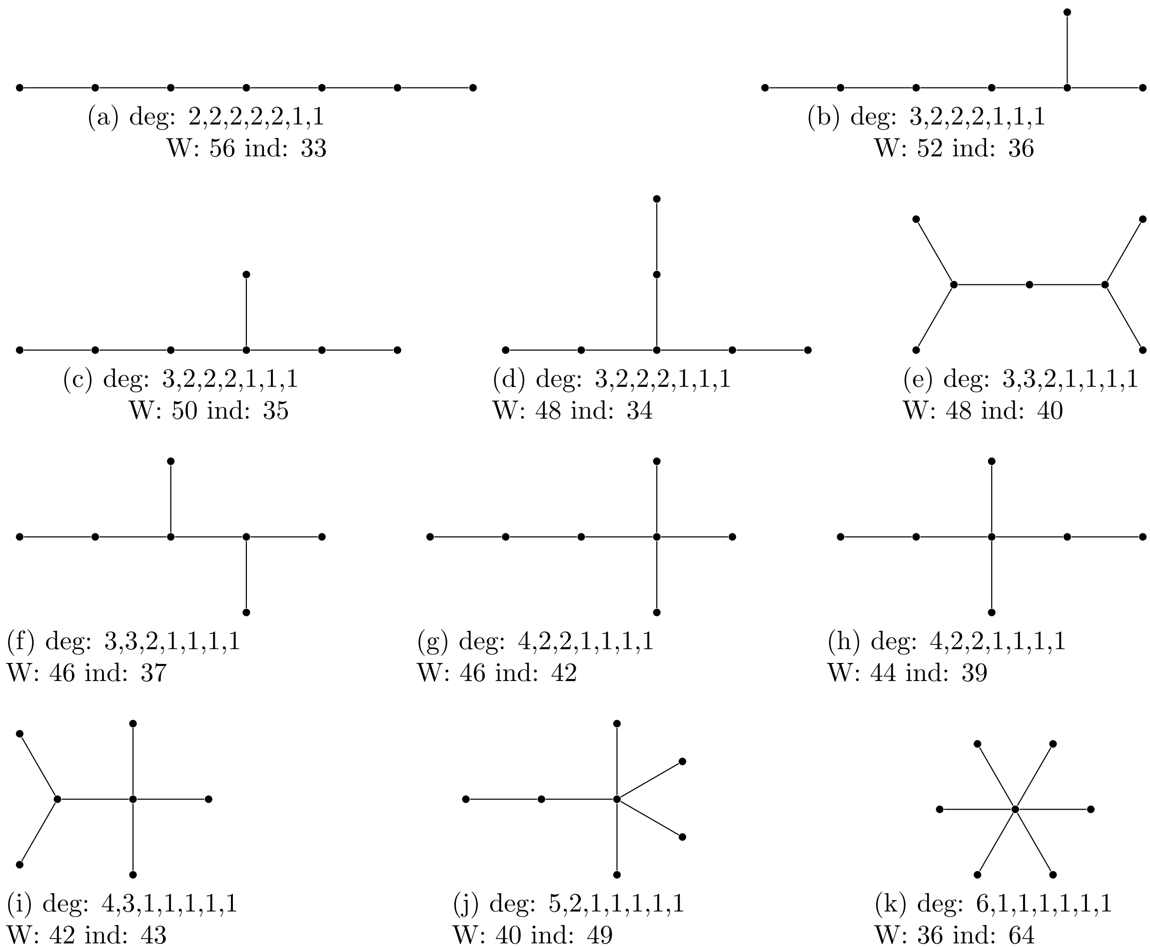


Figure 1.1: All trees on 7 vertices with their degree sequence (deg), Wiener index (W), and number of independent sets (ind).

appear to be clustered together in this order. One may wonder which trees have the largest or smallest Wiener index under the restriction of maximum degree. Since the path has the maximum Wiener index and it has maximum degree less than any other tree, the only meaningful question is which tree or trees with a given maximum degree *minimize* the Wiener index. Fischermann, Hoffmann, Rautenbach, Székely, and Volkmann characterized such trees [4]. They proved that the tree with maximum degree at most $d+1$ is the ball $B_{n,d}$. This is the tree where every vertex (except for at most one) has degree $d+1$ or 1 and vertices are packed as close together as possible.

A second similar observation would be that trees with the same degree sequence are clustered together. In his thesis, Jelen [10] addressed this question showing that the tree with a given degree sequence that minimizes the Wiener index is the again the ball (suitably reinterpreted) and that the trees maximizing the Wiener index are caterpillars.

It is an interesting fact that many pairs of graph invariants are strongly correlated for trees (for an explicit result of these correlations see [20]). One such correlation is the Wiener index and the number of independent sets. In Figure 1.1 we list the values of both for all trees on 7 vertices. The ordering of trees by the number of independent sets is similar to the reverse of the ordering by Wiener index; it is maximized by the star and minimized by the path. But even taking this reversal into account, it is not the case that the ordering by the Wiener index and the ordering by the number of independent sets is the same; they are simply similar. Thus, it should be no surprise that the same questions (which tree minimizes or maximizes this number in some family of trees?) when asked about the number of independent sets give slightly different answers. Heuberger and Wagner showed that the festoon maximized the number of independent sets in the class of trees with a given maximum degree [6]. This tree, the festoon, was completely new at the time and its structure was not completely understood, preventing them from providing an elegant definition. We will rectify this in Chapter 3 as well as solve the problem of maximizing the number of independent sets for trees with a given degree sequence. We in fact prove quite a bit more. We give a result for a larger class of invariants.

A natural class of graph invariants comes from considering the number of homomorphisms to a fixed graph H. The number of independent sets is actually one of these invariants. It counts the number of homomorphisms to the graph H_{ind} given in Figure 1.2. The only restriction of a homomorphism to H_{ind} is that we cannot send

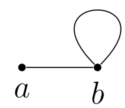


Figure 1.2: H_{ind}

adjacent vertices to vertex a. Hence if we have an independent set in a graph G we can send every vertex of the independent set to a and every other vertex to b and have a homomorphism of G to H_{ind}. Conversely if we have a homomorphism from G to H_{ind} the inverse image of a is an independent set in G. Counting homomorphisms of trees to graphs is fairly easy. Start with an arbitrary choice of root r for a given tree T, and for each vertex v in the target graph H, consider the number of homomorphisms that send r to v. For each branch of r we can send its root to any neighbor of v. All of these choices can be made independently and recursively for each branch. For concreteness, let's consider the problem of counting the number of independent sets in T. We keep track, for each branch of T of how many independent sets in that branch contain the root (i.e., send the root to $a \in V(H_{ind})$) and how many do not. We work our way up from the leafs. The number of independent sets of a branch containing the root is the product of the number of independent sets of each of its branches not containing the root. The number of independent sets of a branch not containing the root is the product of the number of independent sets of each of its branches. This is illustrated in Figure 1.3 where each vertex is labeled with the number of independent sets of the branch at that vertex containing the root and not containing the root. In the end the total number of independent sets is just the sum of the number containing and not containing the root. In Figure 1.3 this is $96 + 210 = 306$.

The number of colorings of a graph with k colors is another example of counting homomorphisms. The number of homomorphisms to the complete graph on k vertices

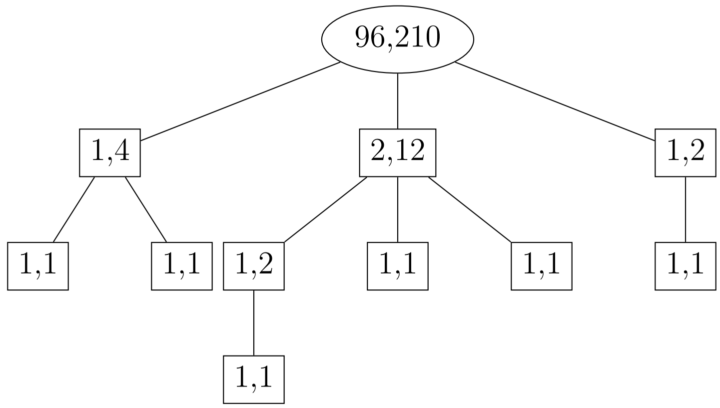


Figure 1.3: Example of counting independent sets.

is the number of k-colorings; for a given homomorphism the inverse images of vertices correspond to the color classes in a coloring. Using the above method we can count the number of homomorphisms of a tree on n vertices to the complete graph on k vertices. There are k choices for the root and since the degree of each vertex in the complete graph is $k-1$, there are $k-1$ choices for the root of each branch. The same is true for every branch and so the total number of homomorphisms or k-colorings is simply $k(k-1)^{n-1}$. The number of k colorings of T is the evaluation of the chromatic polynomial of T at k. But every tree on n vertices has the same number of k-colorings for all k so every tree on n vertices has the same chromatic polynomial $t(t-1)^{n-1}$. (This is of course a well known result.)

For non-regular target graphs in general the number of homomorphisms from trees on n vertices is not constant. As we will see, many different target graphs have the same trees that maximize the number of homomorphisms. In particular, for strongly biregular graphs the extremal tree is the ball or the festoon. This is one of many applications in Chapter 3.

1.2 Graph Reconstruction Problems

Reconstruction is another interesting topic in graph theory. In general it asks: is a graph determined by some collection of partial information about it? The most famous of such problems is the Kelly-Ulam-Reconstruction Conjecture that claims that each graph (on 3 or more vertices) is determined up to isomorphism by the isomorphism classes of its induced proper subgraphs, or equivalently, that a graph is determined up to isomorphism by the multiset of isomorphism classes of single vertex deleted subgraphs. A more precise statement of this problem is that G is Kelly-Ulam-reconstructable if $G \simeq G'$ whenever G' is a graph on the same vertex set V with $G - v \simeq G' - v$ for all $v \in V$. Kelly proved the conjecture for all trees [11] and since then much work has been done on the subject. Unfortunately this thesis will not solve this conjecture.

There are many other reconstruction questions. One such question is whether one can reconstruct a tree from its Wiener matrix. Randić, Guo, Oxley, Krishnapriyan, and Naylor conjectured that for trees an entry in the Wiener Matrix is the largest in its row or column if and only if the corresponding vertices are adjacent [16]. The Wiener matrix for a tree is the matrix with entries equaling the number of paths containing the corresponding pair of vertices. A generalization of this problem is the k-subtree matrix which has entries equaling the number of subtrees with maximum degree at most k containing the corresponding pair of vertices. The Wiener matrix is a special case of this when $k = 2$. We will show in Chapter 4 that we can reconstruct a tree from any of these matrices.

Another such question is whether a graph can be reconstructed from the collection of metric balls of radius r about each vertex. A metric ball of radius r about a vertex is simply the set of vertices within distance r. It is easy to construct examples where reconstruction from these metric balls fails. With no pendant vertices and large enough girth we can always reconstruct form metric balls. However, there is an open question of how large the girth must be to reconstruct from metric balls of radius r. Levenshtein [14] proved that graphs with no pendant vertices and girth at least $2r + 2\lceil (r-1)/4 \rceil + 1$ can be reconstructed from metric balls of radius r. He conjectured that that any graph with no pendant vertices and girth at least $2r + 3$ can be reconstructed from metric balls of radius r. In Chapter 4, we will prove this conjecture. This is the best possible result.

Chapter 2

Notation and Definitions

A *graph* is an ordered pair $G = (V(G), E(G))$ of *vertices* $V(G)$ which is some finite set and *edges* $E(G)$ representing a connection between two vertices. When possible we will write $V = V(G)$ and $E = E(G)$. For graphs the edges will be subsets of size two of the vertex set corresponding to the two *endpoints* of the edge. A vertex and an edge are said to be *incident* if the vertex is one of the endpoints of the edge. Two vertices are *adjacent* if there is an edge containing them. The *neighborhood* of a vertex v, written $N_G(v)$ (or $N(v)$ when clear), is the set of vertices adjacent to v. For *looped* graphs the edges can also have both endpoints be the same vertex. For convenience, instead of writing $u, v \in E$ for an edge of a graph G we may write $u \sim_G v$ (or $u \sim v$ when clear) or simply uv. For a non-edge we will write $u \not\sim v$. The order of a graph G, written $|G|$, is the size of the vertex set. Often we will simply write n when $|G| = n$ and the graph is clear by context.

A *directed graph* D is a graph with an orientation on each edge; one endpoint is the start and the other is the end of the edge. In figures this is denoted with an arrow from the start to the end. When speaking about a directed edge starting at u and ending at v we will write $u \to v$ or simply $\overrightarrow{uw}$.

A *path* between two vertices u and v is a list of distinct vertices (except possibly the first and last vertex) starting with u and ending with v such that each adjacent pair is an edge. A *directed path* from u to v is a path where the edge for each adjacent pair is directed towards the second vertex. A *cycle* is a path that starts and ends at the same vertex. A *forest* is a graph with no cycles. A graph is *connected* if there is a path between any two vertices. A *tree* is a connected graph with no cycles. The *length* of a path is one less than the number of vertices in the list, i.e., the number of edges. The *distance* between two vertices is the length of the shortest path between them.

The *degree* of a vertex is the number of edges it is incident to. The *degree sequence* of a graph is a list of the degrees of each vertex, usually sorted in non-increasing or non-decreasing order. A *leaf* of a tree is a vertex of degree one.

A *homomorphism* between graphs G and H is a function $f : V(G) \to V(H)$ such that $f(u)f(v) \in E(H)$ for every $uv \in E(G)$. Two graphs G and H are *isomorphic* if there is an invertible homomorphism whose inverse is also a homomorphism, which we denote by $G \simeq H$. We also write $\mathrm{Hom}(G, H)$ for the set of homomorphisms from G to H and $\hom(G, H) = |\mathrm{Hom}(G, H)|$. When we speak about uniqueness of graphs or unlabeled graphs we mean unique up to isomorphism.

A (real) graph *invariant* is a function with domain the set of all graphs and range $\mathbb{R}$ such that isomorphic graphs have the same value. Put another way, an invariant is some property of the graph that does not depend on the current representation of the graph.

Let $\mathcal{T}_{n,d}$ be the collection of trees on n vertices and maximum degree at most $d+1$. Let $\mathcal{T}_\pi$ be the collection of trees with degree sequence π. Note that a degree sequence π of length n is that of a tree if and only if $\pi(i) \geq 1$ for all i and $\sum \pi(i) = 2n - 2$.

A *subgraph* of $G = (V, E)$ is a graph $G' = (V', E')$ that consists of $V' \subseteq V$ and

$E' \subseteq E$, we will write $G' \subseteq G$. A subgraph $G' \subseteq G$ is *proper* if $G' \neq G$. A subgraph $G' \subset G$ is *induced* by the set $A \subset V$, written $G' = G[A]$, if its vertex set is A and it contains every edge in G between two vertices in A. A *subtree* is a subgraph that is a tree.

A *rooted tree* is simply a tree with a vertex or edge designated to be the *root*. If T is a tree and $v \in V(T)$ we write T^v for T rooted at v. If T is a tree and $vw \in E(T)$ we write T^{vw} for T rooted at the edge vw. If T is a rooted tree we write $\mathrm{root}(T)$ for its root. A *birooted forest* is a forest with two distinguished vertices: the left and right root. We write $T^{v,w}$ for the birooted forest T with left root v and right root w.

A *branch* of a tree is a rooted subtree induced by the vertices on one side of an edge. More formally, if T is a rooted tree with $\mathrm{root}(T) = v$ and $u \in V(T) \setminus \{v\}$ then the *branch of* T *at* u, denoted T_u^v (or T_u when clear by context), is the subgraph of T consisting of all vertices that are separated from v by u. In other words $w \in T_u$ if and only if u is in the unique path from w to v. Similarly, if T is a rooted tree with $\mathrm{root}(T) = vw$ and $u \in V(T)$ then the *branch of* T *at* u, denoted T_u^{vw} (or T_u when clear by context), is the subgraph of T consisting of all vertices that are separated from both v and w by u. In both cases, we will consider the branch T_u to be rooted at u. The *branches of* T^v will be the collection of branches of T at the neighbors of v, that is $\{T_u^v : u \in \mathrm{N}(v)\}$.

Example 1. The branches of T^f, where T is the tree in Figure 2.1, are $T[e, a, d, h]$, $T[b, c, g]$, and $T[i]$. If T is rooted at the edge ef then the branch at f , i.e. T_f^{ef}, is $T[f, b, c, g, i]$, and the branch at b, i.e. T_b^{ef}, is $T[b, c, g]$.

Let T be a rooted tree. The *depth* of a vertex in T^v is the distance to v, and, in the case of T^{vw}, the minimum of the distances to v and w. The *height* of a rooted tree is the greatest depth of a leaf.

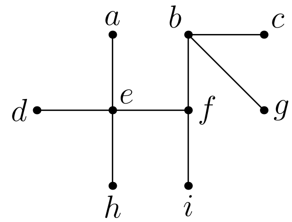


Figure 2.1: Example

The *children* of u with respect to v in tree T, denoted $N_T^v(u)$ (or $N^v(u)$ when clear), are $N(u) \cap T_u^v$. If T is a rooted tree and $u \in V(T)$ is not the root then the *predecessor* of u is the unique neighbor of u that is closer to the root. The predecessor of u is denoted u'. In the case T^{vw} of a tree rooted at an edge, neither v' nor w' are defined.

One of the simplest rooted trees that we will encounter is the complete d-ary tree. We denote the *complete d-ary tree* on with n levels by C_n. It is defined inductively as follows. C_1 is the rooted tree with one vertex. For $n > 1$, C_n is the rooted tree having d branches all equal to C_{n-1}. For convenience we write C_0 for the empty rooted tree. We can consider C_0 to be a branch of any vertex of a tree when needed, specifically when we need d branches of a vertex we will make up any deficit with copies of C_0.

The *path* on n vertices is the graph P_n consisting of a single path of length $n-1$. The *star* on n vertices is the graph $K_{1,n-1}$ with one vertex adjacent to all others.

Let $[n] = \{1, 2, \ldots, n\}$.

Chapter 3

Extremal Trees

3.1 Discussion

Extremal graph theory is one of the cornerstones of graph theory. In general it asks: given a class of graphs, what are the extremal values of some invariant? Further we would like to know what graphs achieve these extremal values, that is, what are the extremal graphs? For the set of all trees on n vertices and many interesting invariants (e.g. diameter, number of leafs, number of independent sets, number of matchings, largest eigenvalue, Wiener index, and more), the extremal trees are the star and path. These questions, however, become substantially more interesting if we restrict to trees with bounded degrees or a fixed degree sequence.

A motivating example is the class of chemical trees, that is, the class of trees with maximum degree at most four. As the name suggests this class has applications in chemistry. Alkanes are chemical compounds that consist of only carbon and hydrogen atoms linked by single bonds in an acyclic manor. Each carbon atom must have four bonds to other atoms and each hydrogen must be joined to a carbon atom, the resulting structure of the carbon atoms is a tree of maximum degree four. The relevance to chemistry is that various graph theoretical invariants on chemical trees correlate well to physical properties of the corresponding alkane, such as their boiling points [21].

The problem of maximizing or minimizing an invariant on trees of bounded maximum degree has been addressed for independent sets, matchings, energy, Wiener index, spectral radius, Laplacian spectral radius, and number of subtrees [6, 7, 12, 19, 4, 9, 18, 22]. For all of these invariants it is the case that the optimal trees without degree restrictions are the path and the star. In $\mathcal{T}_{n,d}$ the path is still an extremal tree, the *trivial extremal tree*; we will concern ourselves with the other extremal trees from here on. For all of these invariants the other extremal tree is either the ball or the festoon (for detailed descriptions of these trees see Section 3.2.3). Unfortunately the previous methods used to determine the extremal trees for these various invariants are, for the most part, all different, and somewhat cumbersome. The goal of Section 3.2 is to present one proof that can be applied to all of these invariants and to make it clear why we get two different trees, the ball and the festoon, depending on the invariant.

For trees with a fixed degree sequence very little is known. Jelen addressed maximizing and minimizing the Wiener index in his 2002 dissertation using the superdominance order [10]. More recently in 2008 Bıyıkoğlu and Leydold addressed maximizing the spectral radius [3]. Their approach also applies to the Laplacian spectral radius. The tree which minimizes the Wiener index and maximizes the spectral radius is the ball. The goal of Section 3.3 is to give a general result in the spirit of Bıyıkoğlu and Leydold that can be used to show that other invariants are maximized or minimized in $\mathcal{T}_\pi$ by the ball or festoon. Additionally, if a *majorization* result can be shown for the invariant this will imply the corresponding result for bounded degree.

Our techniques allow us to extend the known results in a variety of directions.

For instance consider the problem of maximizing the (weighted) number of homomorphisms from a tree to a given target graph. This gives us an invariant for each weighted target graph. We will show that if the target graph satisfies a simple condition then we can use these new methods to find the tree with bounded degree or a fixed degree sequence that maximizes this quantity. In addition, we prove the interesting fact that the optimal tree we get for this invariant is the ball or festoon depending on the target graph, independent of the weighting.

The method for bounded degree uses a *certificate* that will be specific to each invariant. The certificate is a branch invariant that satisfies specific properties that will be key to the proof. There are two types of certificates, increasing and decreasing, and the type determines which tree is optimal, the ball or the festoon. We will prove the result using a generic certificate and show that to apply it to any invariant we simply need to present a certificate for that invariant. The main result of Section 3.2 is as follows.

Main Theorem 1. *Let σ be an invariant of trees with a certificate ρ. If ρ is increasing then σ is optimized in $\mathcal{T}_{n,d}$ by the ball $B_{n,d}$. If ρ is decreasing then σ is optimized in $\mathcal{T}_{n,d}$ by the festoon $F_{n,d}$.*

The method for a fixed degree sequence is more complex. An invariant has an associated *labeling*. If the labeling refines the degree and is direct for a maximal (or minimal) tree then the tree is a ball. If the labeling refines the degree and is alternating for a maximal (or minimal) tree then the tree is a festoon. The main result of 3.3 is as follows.

Main Theorem 2. *Let $T \in \mathcal{T}_\pi$ and f a labeling of T that is a refinement of the degree. If f is a direct labeling of T then $T \simeq B_\pi$. If f is an alternating labeling of T then $T \simeq F_\pi$.*

In Section 3.2.6 we look at applications of the results of Section 3.2 for bounded degree trees to the following invariants: number of independent sets, value of matching generating polynomial (at $x > 0$), number of weighted homomorphisms to a strongly biregular graph, number of subtrees, Wiener index, spectral radius, and Laplacian spectral radius. In Section 3.3.5 we look at applications of the results of Section 3.3 for trees with a fixed degree sequence to the number of homomorphisms to a strongly biregular graph, the number of matchings, and the number of subtrees.

We close the chapter in Section 3.4 with some open problems.

3.2 Bounded Degree

Let d be some fixed integer. We would like to find the tree that maximizes or minimizes an invariant in $\mathcal{T}_{n,d}$.

3.2.1 Certificates

To better understand the invariant we are trying to optimize it is useful to think of computing a related invariant for the tree's branches.

Definition 1. A *branch invariant* is an invariant of possibly empty rooted trees.

A certificate is a special branch invariant, one that satisfies a certain branch exchange property. This property is the key to the proof of the main theorem. Roughly it says that in an extremal tree, branches with small values appear together, and those with large values appear together. To be more precise, looking at the certificates values for the branches of two vertices, we have the smaller values adjacent to one vertex and the larger adjacent to the other vertex.

Branch Exchange Property. Let $\mathcal{S} \subseteq \mathcal{T}_{n,d}$ be a set of trees, and ρ a branch invariant. Then ρ satisfies the branch exchange property on $\mathcal{S}$ if for each $T \in \mathcal{S}$ and $l \neq r$ vertices of T we have the following. Let $\{L_i\}_{i=1}^d$ be the d branches of T_l^r and $\{R_i\}_{i=1}^d$ the d branches of T_r^l, (see Figure 3.1). Then either $\max_i(\rho(L_i)) \leq \min_i(\rho(R_i))$ or $\min_i(\rho(L_i)) \geq \max_i(\rho(R_i))$.

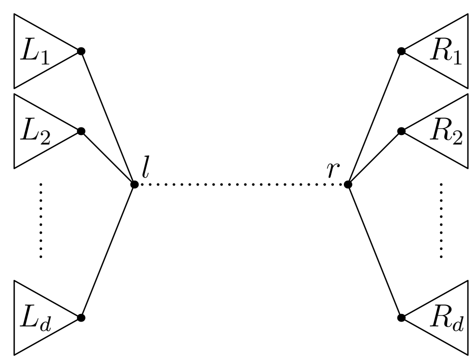


Figure 3.1: The Branch Exchange Property

In our applications the branch invariant ρ mentioned in the branch exchange property is *not* the tree invariant σ we are optimizing. The two are related only by the seemingly weak connection that ρ should satisfy the branch exchange property on the class of σ-extremal trees.

Definition 2. Let σ be an invariant to be optimized. A *certificate* for σ is a branch invariant ρ satisfying.

1. The values of ρ are in $I = (0, 1]$ and $\rho(C_0) = 1$, $\rho(C_1) < 1$.

2. There exists a continuous symmetric function $f : I^d \to I$ that is either strictly increasing or strictly decreasing, such that if T is a rooted tree with (possibly empty) branches $T_1, \ldots, T_d$ then

$$\rho(T) = f(\rho(T_1), \ldots, \rho(T_d)).$$

The function f will be referred to as the certificate's *recursive definition*.

3. ρ satisfies the Branch Exchange Property (BEP) on the set of σ-extremal trees.

The certificate is said to be *increasing* if f is increasing and *decreasing* if f is decreasing.

Sometimes it is more natural to work with a branch invariant that takes its values on $[0, \infty)$, which just requires a slight modification of the above.

Definition 3. A *wide certificate* for σ is a branch invariant μ satisfying properties 2 & 3 of Definition 2 but instead of values in $(0, 1]$, it takes values in $I = [0, \infty)$ with $\mu(C_0) = 0$, $\mu(C_1) > 0$.

Proposition 1. *If σ has a wide certificate μ then it has a certificate ρ. Moreover ρ is increasing if and only if μ is.*

Proof. Let μ be a wide certificate and let g be its associated function. Define

$$\rho(T) = \frac{1}{1 + \mu(T)}$$

and

$$f(x_1, \ldots, x_d) = \frac{1}{1 + g(\frac{1}{x_1} - 1, \ldots, \frac{1}{x_d} - 1)}.$$

Observe that since $\mu(T) \in [0, \infty)$ we have $\rho(T) \in (0, 1]$. Also $\rho(C_0) = (1 + 0)^{-1} = 1$ and $\rho(C_1) = (1 + \mu(C_1))^{-1} < 1$ since $\mu(C_1) > 0$. It is clear that if g is a continuous symmetric monotonic function then so is f. Clearly, if $\mu(T_1) < \mu(T_2)$ then $\rho(T_1) > \rho(T_2)$. It is straightforward to check that if μ satisfies the branch exchange property for σ-extremal trees then so does ρ. Finally, since f is the composition of g with two other strictly decreasing functions, it is clear that f is strictly increasing when g is, and f is strictly decreasing when g is. □

3.2.2 Outline

Given a extremal tree T rooted at v and a certificate ρ it will be useful to think of labeling each vertex u with the value $\rho(T_u)$. When we have an increasing certificate we would like to find a choice of root so that there is a plane tree drawing such that at every level the values are increasing from left to right, and the largest value at any level is at most the smallest value at the next level. This will show that the leafs of the extremal tree must all be on the last two levels. By results in this section we see that all but at most one of the branches of the root is complete, and the possibly incomplete branch satisfies this condition inductively. This is the defining property of the ball, so the extremal tree for an invariant with an increasing certificate is the ball.

When we have a decreasing certificate we would like to find a choice of root so that there is a plane tree drawing such that at every level the values are alternately increasing and decreasing from left to right, and the values on any level are more 'central' than on the next level. This will show that the leafs of the extremal tree must all be on the last three levels. Using basic facts from this section we show that all but possibly one of the branches of the root are complete, i.e. C_h, C_{h-1}, C_{h-2} where h is the height of the tree. The possibly remaining incomplete branch also satisfies this property, but we will see that if its height is h we can only have branches of C_h or C_{h-2} and possibly one incomplete branch that also satisfies the same property. This is the defining property of the festoon. Thus, if an invariant has a decreasing certificate, its extremal tree is the festoon.

Observe that a plane tree drawing is equivalent to giving an ordering of the vertices at each level, the extra structure that we described above makes an order on all of the vertices except the root. This is what we will call a strong ordering of the vertices.

In general strong orderings are hard to find; for most extremal trees there is only one choice of root such that the remaining vertices have a strong order. To assist us in finding the strong order we will start with what we call a weak ordering, which we can find for any choice of root in an extremal tree.

Definition 4. Let T be a tree rooted at vertex v or an edge v_1v_2. A *certificate ordering* of T with respect to certificate ρ is a partial order, $\preceq$, of the vertices, excluding v in the vertex rooted case, such that if u and w are vertices at the same depth then the following conditions are satisfied.

1. u and w are comparable. Moreover, if $\rho(T_u) < \rho(T_w)$ then $u \prec w$.

2. If u and w are not adjacent to v or equal to v_1 or v_2, and $u \prec w$ then,

   a) if f is increasing then $u' \preceq w'$,

   b) if f is decreasing then $w' \preceq u'$.

Remark. There is a difference between edge rooted and vertex rooted trees. If T is rooted at vertex v the value of $\rho(T^v)$ will be irrelevant. Instead we will care about the value of $\rho(T_v^u)$ for $u \in N(v)$. If T is rooted at edge vw then we will care about the values of $\rho(T_v^{vw})$ and $\rho(T_w^{vw})$. This is why we exclude the vertex root from the certificate order and have the specific exclusions in condition 2 above.

We will use the following terminology to refer to some special elements of a tree with a certificate order.

Definition 5. We say c is the *center* of a decreasing certificate with recursive definition f if

$$f(c, c, \ldots, c) = c.$$

We will show in Section 3.2.5 that every decreasing certificate has a unique center.

Definition 6. Fix a certificate order $\preceq$ on T. The *first vertex at depth k* is the minimal element at depth k. The *last vertex at depth k* is the maximal element at depth k. In the case of a decreasing certificate we also have the following. A vertex u is *before the center* if $\rho(T_u) < c$ and is *after the center* if $\rho(T_u) > c$. We will refer to, for instance, the last vertex at depth k before the center, etc.

The formal definitions of a weak and strong order follow.

Definition 7. Let T be a tree rooted at vertex v or an edge v_1v_2. A *weak ordering* of T with respect to the certificate ρ is a certificate ordering such that vertices at different depths are not comparable.

Definition 8. Let T be a tree rooted at vertex v or an edge v_1v_2. A *strong ordering* of T with respect to the certificate ρ is a certificate ordering satisfying the following additional conditions.

1. It is a total order.
2. If $\rho(T_u) < \rho(T_w)$ then $u \prec w$ (no depth restriction).
3. If f is increasing and u, w are two vertices (besides the root in the vertex rooted case) such that $u \prec w$, then the depth of u is less than or equal to that of w.
4. If f is decreasing and u, w are two vertices (besides the root in the vertex rooted case) such that $u \prec w$, then the following holds. If w is before the center then the depth of w is less than or equal to that of u. If u is after the center then the depth of u is less than or equal to that of w.

Now we prove some basic lemmas that will be used later.

Lemma 2. *Assume f is a increasing recursive definition for ρ. Let T be any nonempty rooted tree having d branches (possibly empty), with all degrees at most $d+1$, then $\rho(T) \leq \rho(C_1)$. Furthermore for $T \neq C_1$ this is strict, i.e. $\rho(T) < \rho(C_1)$.*

Proof. Notice that $\rho(C_1) = f(1, \ldots, 1)$ is the maximum value of f, and thus since we can compute $\rho(T)$ for any nonempty T using f we must have $\rho(T) \leq \rho(C_1)$ and furthermore any tree T with a nonempty branch T_0 will give less then this maximum since $\rho(T_0) < 1$ and f is increasing. □

Remark. In this situation it is easy to establish that $1 = \rho(C_0) > \rho(C_1) > \rho(C_2) > \cdots > 0$ by induction with the same argument as above. However, we do not use this fact.

Lemma 3. *Assume f is a decreasing recursive definition for ρ. Let T be any nonempty rooted tree having d branches (possibly empty), with all degrees at most $d+1$, then $\rho(C_1) \leq \rho(T) \leq \rho(C_2)$. Furthermore if $T \neq C_1$ then $\rho(T) > \rho(C_1)$, and if $T \neq C_2$ then $\rho(T) < \rho(C_2)$.*

Proof. Notice that $\rho(C_1) = f(1, \ldots, 1)$ is the minimum value of f, and thus since we can compute $\rho(T)$ for any nonempty T using f we must have $\rho(T) \geq \rho(C_1)$. Furthermore, any tree T with a nonempty branch T_0 will give more than this minimum since $\rho(T_0) < 1$ and f is decreasing. Now since we have $\rho(T) \geq \rho(C_1)$ for *all* rooted trees (even the empty one), then in particular, if T is a nonempty tree with branches $T_1, \ldots, T_d$ we have $\rho(T_i) \geq \rho(C_1)$ for all i. Thus, since f is decreasing, $\rho(T) = f(\rho(T_1), \ldots, \rho(T_d) \leq f(\rho(C_1), \ldots, \rho(C_1)) = \rho(C_2)$, and if any one of the branches is not C_1 by the above we have that this is strict. □

Remark. In this situation it is easy to establish that $0 < \rho(C_1) < \rho(C_3) < \rho(C_5) < \cdots < \rho(C_4) < \rho(C_2) < \rho(C_0) = 1$ by induction with the same argument as above.

However we do not use this fact.

Lemma 4. *If T is a σ-extremal tree and ρ is a certificate with recursive definition f, then using the notation of the branch exchange property from Section 3.2.1:*

- *If f is increasing then, $\rho(T_l^r) \leq \rho(T_r^l)$ if and only if $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$, and $\rho(T_l^r) \geq \rho(T_r^l)$ if and only if $\min\{\rho(L_i)\} \geq \max\{\rho(R_i)\}$.*
- *If f is decreasing then, $\rho(T_l^r) \geq \rho(T_r^l)$ if and only if $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$, and $\rho(T_l^r) \leq \rho(T_r^l)$ if and only if $\min\{\rho(L_i)\} \geq \max\{\rho(R_i)\}$.*

Proof. Recall, $\rho(T_l^r) = f(\rho(L_1), \ldots, \rho(L_d))$ and $\rho(T_r^l) = f(\rho(R_1), \ldots, \rho(R_d))$. □

Corollary 5. *Let T be an extremal tree with respect to an invariant with a certificate ρ. Then every vertex except possibly one has degree 1 or $d+1$.*

Proof. For a contradiction suppose there is more than one vertex of degree not 1 or $d+1$, say u and w are two of them. Applying the branch exchange property to u and w yields a contradiction since they will each have C_0 as a branch and also some other nonempty branch. □

Corollary 6. *Let T be an extremal tree with respect to an invariant with a certificate ρ. If $l \neq r$ and $\rho(T_l^r) = \rho(T_r^l)$ then $T_l^r = T_r^l = C_h$ for some h.*

Proof. By the previous lemma all of the branches of these two trees will have equal ρ-values. For the same reason the branches of these branches will all have equal ρ-values, and so on. At some depth we have a leaf, and since all the ρ-values at that depth are equal by Lemma 2 or 3, everything at that depth is a leaf. Hence they are both C_h for some h. □

Lemma 7 (Standard Weak Ordering)**.** *If T is extremal then with respect to any choice of vertex or edge as a root, there is a weak ordering of T, unique up to interchange of isomorphic branches.*

Proof. It is trivial to order the vertices at depth one – simply order the neighbors of the root in increasing order of their ρ values. If T is rooted at an edge then it is also trivial to order the two vertices incident to the edge. Assume that we have a weak ordering up to some depth k. Applying the branch exchange lemma to every pair of vertices at depth k, we see that their neighbors at depth $k+1$ have ρ values that are not interlaced. Further notice that by Lemma 4, the fact that f is monotonic forces the neighbors at depth $k+1$ to have ρ values in the same or opposite order as their predecessors respectively, so there is an ordering of the neighbors at depth $k+1$ satisfying all the required properties. Hence we have a weak ordering. It is clear by Corollary 6 that this is unique up to interchange of isomorphic branches. □

Remark. A weak ordering is unique up to interchange of isomorphic branches of a vertex, i.e. root preserving automorphisms. Hence we refer to the one given above as *the weak ordering.*

Our goal is to show that when T is extremal then there is always a choice of root where there is a strong ordering of the vertices, and prove that this forces the tree to be one of the two optimal trees, the ball or the festoon, depending only on whether f is increasing or decreasing.

3.2.3 The Optimal Trees

The two nontrivial optimal trees that we encounter are the ball and the festoon. The ball can be thought of as a rooted tree with $d+1$ branches all equal to C_h (the complete d-ary tree with h levels) plus one extra partial level of leafs added in order to the last level. Thus the ball contains all vertices within distance h of the root, and all vertices are within distance $h+1$ of the root. The vertices at distance $h+1$ are clustered as closely together as can be. In chemistry the ball is also known as a dendrimer. A technical definition follows.

Definition 9 (Ball)**.** Let $\mathcal{B}_{h,\delta}$ for $\delta \in \{d, d+1\}$ be the collection of rooted trees with δ branches, $\delta - 1$ of which are each equal to C_h or C_{h-1} and the last branch is an element of $\mathcal{B}_{h-1,d}$, where $\mathcal{B}_{0,\delta} = \{C_1\}$. The *ball* on n vertices and bounded degree $d+1$, denoted $B_{n,d}$, is the unique tree of $\mathcal{B} = \bigcup_{h=0}^{\infty} \mathcal{B}_{h,d+1}$ with n vertices.

A festoon can be thought of in a similar manner as a rooted tree with $d+1$ branches all from C_h or C_{h+1} plus one extra level of C_2's added in order to the second to last level. Thus the festoon contains all vertices within distance h of the root, and all vertices are within distance $h+2$ of the root. The vertices at distance $h+2$ are clustered as closely together as can be. A technical definition follows.

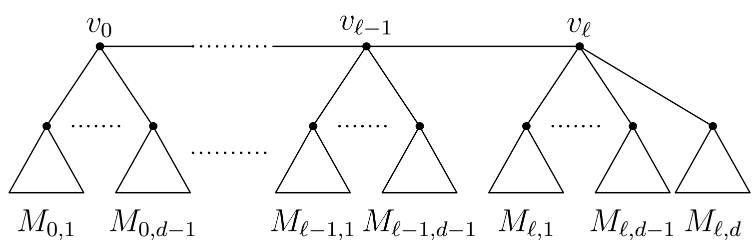
Definition 10 (Festoon)**.** Let $\mathcal{F}_h$ be the collection of rooted trees with d branches, $d-1$ of which are each equal to C_h or C_{h-2} and the last branch is an element of $\mathcal{F}_{h-1}$, where $\mathcal{F}_0 = \{C_1\}$. Let $\mathcal{F}_h^*$ be the collection of rooted trees with $d+1$ branches, d of which are equal to C_h, C_{h-1}, or C_{h-2} and the last an element of $\mathcal{F}_{h-1}$. The *festoon* on n vertices and bounded degree $d+1$, denoted $F_{n,d}$, is the unique tree of $\mathcal{F} = \bigcup_{h=0}^{\infty} \mathcal{F}_h^*$ with n vertices. In this definition we exclude choice of C_h when $h < 0$ since they are undefined.

The festoon tree was first defined by Heuberger and Wagner in [6] as the solution for two optimization problems. The first problem was to maximize the number of independent sets a tree of fixed size and bounded degree could have. The second was the similar problem to minimize the number of matchings that a tree of fixed size and bounded degree could have. The number of independent sets and the number of matchings are both examples of invariants that are a measure of branching, that is to say that for a fixed tree size the maximum and minimum examples of these invariants are the star and the path and 'small' changes to the graph produce small changes in the invariants.

Later in [7] it was also shown that the festoon tree also minimizes the energy of a tree of fixed size and bounded degree. The energy of a graph is the sum of the absolute value of its eigenvalues but it can also be computed by the Coulson integral in terms of the matching generating polynomial. It turns out that the festoon tree minimizes the matching generating polynomial for all positive values of x and therefore it minimizes the energy.

The original definition given by Heuberger and Wagner in [6] can be paraphrased as the following:

Definition 11 (Old Festoon)**.** There is a unique tree $H_{n,d}$ with n vertices and bounded degree $d+1$ that can be decomposed as



with $M_{k,1}, \dots, M_{k,d-1} \in \{C_k, C_{k+2}\}$ for $0 \le k < \ell$ and either $M_{\ell,1} = \cdots = M_{\ell,d} = C_{\ell-1}$ or $M_{\ell,1} = \cdots = M_{\ell,d} = C_\ell$ or $M_{\ell,1}, \dots, M_{\ell,d} \in \{C_\ell, C_{\ell+1}, C_{\ell+2}\}$, where at least two of $M_{\ell,1}, \dots, M_{\ell,d}$ equal $C_{\ell+1}$.

This definition can lead to some trees that at first glance look rather strange. For example in [8] we are given the following example of $F_{69,2}$ shown in Figure 3.2. For

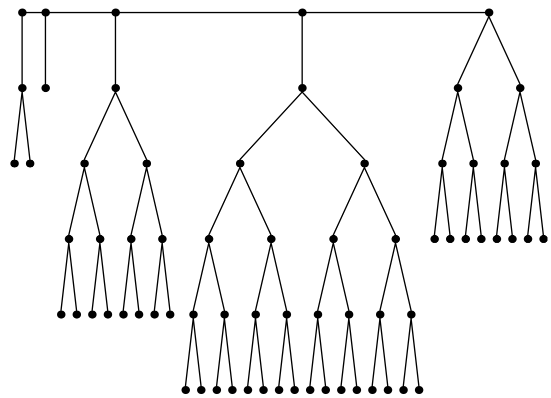


Figure 3.2: $F_{69,2}$ in the form of the old definition

more examples see Figure 9 in [8]. Furthermore, the algorithm given for creating these trees is equally hard to understand (see Algorithm 1 and 2 in [8]). We would like to show that Definitions 10 and 11 are equivalent. The essential change in the definition is just the requirement on the last set of branches $M_{\ell,1}, \ldots, M_{\ell,d}$.

To see that our definitions are equivalent consider the following. If $M_{\ell,1} = \cdots = M_{\ell,d} = C_{\ell-1}$ then combining these branches together we have one branch of C_ℓ at v_ℓ and so if we root at $v_{\ell-1}$ we will have one branch of C_ℓ, $d-1$ branches of $C_{\ell-1}$ and $C_{\ell+1}$ and the last branch follows the same recursive definition. This is consistent with the new definition when $h = \ell + 1$. If $M_{\ell,1} = \cdots = M_{\ell,d} = C_\ell$ then combining these branches together we have one branch of $C_{\ell+1}$ at v_ℓ and so if we root at $v_{\ell-1}$ we will have one branch of $C_{\ell+1}$, $d-1$ branches of $C_{\ell-1}$ and $C_{\ell+1}$ and the last branch follows the same recursive definition. This is consistent with the new definition when $h = \ell + 1$. If $M_{\ell,1}, \ldots, M_{\ell,d} \in \{C_\ell, C_{\ell+1}, C_{\ell+2}\}$, where at least two of $M_{\ell,1}, \ldots, M_{\ell,d}$ equal $C_{\ell+1}$ then rooting at v_ℓ we are consistent with the new definition when $h = \ell+2$. In a similar fashion we can go in the other direction so the definitions are equivalent.

Below are easy algorithms to build the ball and festoon. When implementing them it is fastest to keep track of the lowest numbered vertex that is not full degree.

Algorithm Festoon(n, d): returns a labeled tree that is the festoon on n vertices with bounded degree $d+1$.

1. Input n and d.

2. If $n \leq d+1$ then return the star $K_{1,n-1}$ and label the vertices $1, \ldots, n$ with the center vertex labeled as 1.

3. If $n \geq d+2$ then let $T = Festoon(n-d-1, d)$ and let k be the smallest numbered vertex in T having degree d or less. To T we will add a star $K_{1,d}$ with the center labeled $n-d$ and leafs $n-d+1, \ldots, n$, we will return this tree with an edge added between vertex k and $n-d$.

One should compare this to the algorithm for creating a ball on n vertices with bounded degree $d+1$.

Algorithm Ball(n, d): returns a labeled tree that is the ball on n vertices with bounded degree $d+1$.

1. Input n and d.

2. If $n = 1$ then return the single vertex graph with its vertex labeled as 1.

3. If $n \geq 2$ then let $T = Ball(n-1, d)$ and let k be the smallest numbered vertex in T having degree d or less. To T we will add a leaf with label n adjacent to the vertex labeled k and return this tree.

3.2.4 Increasing Certificate Solution

We will first consider the case where we have an increasing certificate ρ for the invariant σ. Then in the next section we will consider the case where we have a decreasing certificate. Given an increasing certificate ρ for the invariant σ we will do the following. First we give a technical lemma saying that for an extremal tree predecessor of the last vertex at any given depth at least 2 is not the double predecessor of the first vertex at the next depth. Then we show that an extremal tree has a strong order and prove that we must have a ball.

Lemma 8. *Let $k \geq 3$. If T is extremal and has a non-leaf root and no leafs at depth $k-2$ or less, then with respect to a weak ordering, if u is the first at depth k and w is the last at depth $k-1$, then $u'' \neq w'$. That is, the branches $T_{u'}$ and $T_{w'}$ are disjoint.*

Proof. First we will show that if T has no leafs at depth $k-1$ ($k \geq 2$) or less then $u' \neq w$. If it is the case that $w = u'$, then there is no other vertex at depth $k-1$, for if $x \neq w$ is at this depth, by assumption $x \prec w$ and so since we have a weak ordering $x_0 \prec u$ for any x_0 such that $x_0' = x$, a contradiction to the choice of u. But this is a contradiction since now x must be a leaf, but by assumption there is no leaf at depth $k-1$. Now since there is no other vertex at depth $k-1$ and no leafs at lesser depths we must have that the root is a leaf, a contradiction. Therefore $w \neq u'$. If there is a leaf at depth $k-1$ then by definition if w is the maximum with respect to the weak ordering at depth $k-1$ then it must be a leaf and thus w neighbors no vertex at depth k. Therefore if T has a non-leaf root and there are no leafs at depth $k-2$ or less then $u' \neq w$. Observe that u' is the first at depth $k-1$ and if there are no leafs at depth $k-2$ then w' is the last at depth $k-2$ so applying the above argument again we have that $u'' \neq w'$ as long as $k \geq 3$. □

Lemma 9. *If T is extremal on $n \geq 3$ vertices, then there is a choice of root, v, that is not a leaf, such that the following hold.*

- *If $u \in N(v)$ then $\rho(T_v^u) \leq \rho(T_u^v)$.*

- *If* $1 \leq \mathrm{depth}(w) < \mathrm{depth}(u)$ *then* $\rho(T_w) \leq \rho(T_u)$.

Proof. Start by orienting all edges uw such that $\overrightarrow{uw}$ if $\rho(T_u^w) > \rho(T_w^u)$. (In case of equality pick an arbitrary orientation.) Let v be any non-leaf sink in this orientation of T. Such a choice exists since every leaf is just C_1 which has the largest ρ-value of any nonempty rooted tree by assumption. This choice clearly satisfies the first condition of the lemma. Now we will show that branches at lesser depths have lower ρ-values. For the case of branches at depth 1, let w be maximum in the weak ordering at depth 1 (so $w' = v$) and u be minimum in the weak ordering at depth 2 (if there is no vertex at depth 2 then the result is trivial). By the previous lemma's proof, $u' \neq w$ and thus since $\rho(T_v^{u'}) \leq \rho(T_{u'}^v)$ (by first claim) we have by the branch exchange property that $\rho(T_w) \leq \rho(T_u)$ as desired. Now for the inductive step ($k \geq 3$), let u be at depth k and minimum with respect to the weak ordering (if none exists we are done), and let w be at depth $k-1$ and maximum with respect to the weak ordering. We would like to show that $\rho(T_w) \leq \rho(T_u)$. If $w' \neq u''$ then this is a consequence of the branch exchange property since we have $\rho(T_{w'}) \leq \rho(T_{u'})$ by the inductive hypothesis. If it is the case that $w' = u''$, then by the previous lemma there is a leaf at depth $k-2$ or less. But then by the inductive hypothesis every vertex at depth $k-1$ is a leaf, contradicting the existence of u. □

Lemma 10. *If T is extremal then there exists a vertex v such that T^v has a strong order.*

Proof. Start with choice of root provided by Lemma 9 and the standard weak ordering. If u is the minimum with respect to the weak ordering at depth k, and w is the maximum with respect to the weak ordering at depth $k-1$, we can take $w \prec u$ for each $k \geq 2$ to give us a total ordering of the vertices. By Lemma 9 this is clearly a strong ordering. □

Theorem 11. *Let σ be a invariant of trees with an increasing certificate ρ. Then σ is optimized by the ball $B_{n,d}$.*

Proof. We begin with the choice of root v provided by the previous lemma. Let h be the height of the tree, that is the maximum distance from the root to a leaf. We need to show that the minimum distance to a leaf is at least $h-1$ and that all but at most one branch is complete and that the one possibly incomplete branch satisfies this condition recursively, that is all but one of its branches are complete, etc. To show this we need only consider where all the leafs of the tree are in the strong ordering provided by the previous lemma. Since $\rho(C_1) > \rho(T)$ for nonempty $T \neq C_1$ we have $u \prec w$ for any non-leaf u and leaf w. Thus, by the conditions of strong ordering, any non-leaf vertex has depth at most that of any leaf, so the minimum distance to a leaf must be at least $h-1$. To see that at most one branch is not complete, let $\{T_i\}$ be the branches of v and suppose that T_{i_0} is incomplete. We claim that if $\rho(T_i) > \rho(T_{i_0})$ then $T_i = C_{h-1}$, if $\rho(T_i) < \rho(T_{i_0})$ then $T_i = C_h$, and if $\rho(T_i) = \rho(T_{i_0})$ then $i = i_0$. By Corollary 6 we cannot have any branch with the same ρ-value as T_{i_0}. If $\rho(T_i) > \rho(T_{i_0})$ then by inductively applying the branch exchange property to each branch we have that at every depth the vertices at depth k in T_{i_0} precede those at depth k in T_i and thus since T_{i_0} has a leaf at depth h (from v), we must have $T_i = C_h$. The other case is similar. Then we can consider the incomplete branch inductively with its induced strong order. Therefore σ is optimized by the ball $B_{n,d}$. □

This is the tree of bounded degree that has previously been shown to maximize the leading eigenvalue (both standard and Laplacian), minimize the number of subtrees, and minimize the Wiener index, that is the sum of distances between every pair of vertices [4, 9, 12, 18, 22].

3.2.5 Decreasing Certificate Solution

Given a decreasing certificate ρ for the invariant σ we have the following.

Lemma 12. *There is a unique solution c to the equation $x = f(x, \ldots, x)$ such that $0 < c < 1$. This root has the property that if $T_1, T_2, \ldots, T_d$ are the branches of a tree T and $\rho(T_j) < c$ for all j then $\rho(T) > c$. Conversely, if $\rho(T_j) > c$ for all j then $\rho(T) < c$.*

Proof. Let $\hat{f}(x) = f(x, x, \ldots, x)$. Note that $\hat{f} : [0, 1] \to [0, 1]$ is a strictly decreasing function hence there is a unique solution c to the equation $x = \hat{f}(x)$ such that $0 < c < 1$. Hence c is the unique solution to the equation $x = f(x, \ldots, x)$. Now suppose we have some tree T with d branches $T_1, \ldots, T_d$ all with $\rho(T_i) < c$. Since f is decreasing we have $\rho(T) = f(\rho(T_1), \ldots, \rho(T_d)) > f(c, \ldots, c) = c$. The other inequality is similar. $\square$

Remark. For a vertex u we will say $u < c$ to mean $\rho(T_u) < c$ and $u > c$ to mean $\rho(T_u) > c$.

Definition 12. A vertex rooted tree T is *alternating* if it has d possibly empty branches $T_1, T_2, \ldots, T_d$ such that,

1. $\rho(T) \neq c$
2. If $\rho(T) < c$ then $\rho(T_i) > c$ for all i,
3. If $\rho(T) > c$ then $\rho(T_i) < c$ for all i,
4. Each branch T_i is itself alternating.

Definition 13. We say that branches $T_1, \ldots, T_r$ (or their corresponding root vertices) are on the *same side* of c to mean that either $\rho(T_i) > c$ for all i or $\rho(T_i) < c$ for all

i. If a collection of branches is not on the same side of c then they are said to be on *both sides* of c. A branch T' is on the *c-side* of T if T' is on the same side of T as c, that is $(\rho(T) - \rho(T'))(\rho(T) - c) \geq 0$. A *split edge* is an oriented edge $\vec{uw}$ such that T_w^u has branches on both sides of c.

Lemma 13. *If $\vec{uw}$ is a split edge then T_u^w is alternating and T_w^u is on the c-side of T_u^w.*

Proof. Take any vertex in T_u^w and compare it to w using the branch exchange property. Since w has branches on both sides of c the other vertex can only have branches on one side of c. Since this is true for all choices in T_u^w, it is alternating. Furthermore, applying the branch exchange property to u and w we see that T_w^u is on the c-side of T_u^w. □

Lemma 14. *If T is extremal on $n \geq 3$ vertices, then there is a choice of root that is not a leaf, such that all of its branches are alternating.*

Proof. Start with an arbitrary choice of internal vertex as a root. If all of the tree's branches are alternating with respect to this root, we are done. Otherwise, let v be the vertex at greatest depth such that T_v is not alternating, note that this cannot be a leaf. The tree T_v must have branches with ρ-values on both sides of c. This is because none of its branches can have ρ-values of c, otherwise v was not chosen at maximum depth, and if all of the branches of T_v are on one side of c then T_v is on the other side, hence T_v is alternating, a contradiction. Let u be the predecessor of v relative to this root, then we have shown $\vec{uv}$ is a split edge. By the preceding lemma we have that T_u^v is alternating, and by choice of v all the other branches of v are alternating. Therefore if we root at v all of our tree's branches will be alternating. □

For many cases this is the proper choice of root, however, there are some special cases where we need to modify our choice.

Lemma 15. *If T is extremal on $n \geq 3$ vertices and v is a non-leaf root for which all branches are alternating and has $d(v) \leq d$, then either every nonempty branch of v is C_1 or C_2, or there is a choice of non-leaf root that has degree $d+1$ and all of its branches are alternating.*

Proof. Suppose $d(v) \leq d$. If every nonempty branch of v is C_1 or C_2 then we are done. Suppose not, then since v is the only vertex of degree not equal to 1 or $d+1$ we have that there must be a C_2 at depth at least 2 from v. Let $u \neq v$ be the predecessor of this C_2, and apply the branch exchange property to v and u. Since $d(v) \leq d$ at least one of the d branches of T_v^u is C_0. Thus since one of the branches of T_u^v is C_2, we must have that all of the branches of T_v^u are C_0 or C_2. Thus if v' is the predecessor of v with respect to u, we have that v' has degree $d+1$ and all of its branches are alternating. □

Lemma 16. *If T is extremal on $n \geq 3$ vertices, then one of the following is true.*

- *There is a choice of vertex root v that is not a leaf, such that for $u \in N(v)$, T_v^u is on the c-side of T_u^v, T_u^v is alternating, and T_u^v is not on the same side of c for all choices of u.*
- *There is a choice of edge root vw, such that its branches T_v^w and T_w^v are not on the same side of c, are both alternating, and for any $u \in N(v) \cup N(w) - \{v, w\}$, T_v^w and T_w^v are on the c-side of T_u^{vw} .*

Proof. We start with the choice of root v provided by the previous lemma, and note that if we are in the special case where all of its branches are C_1 or C_2 then the result is trivial.

Case 1. Suppose that all the branches have ρ-values on one side of c. Let $u, v_1, v_2, \ldots, v_d$ be the neighbors of v. Without loss of generality suppose that $\max\{\rho(T^v_{v_i})\} \leq \rho(T^v_u) < c$. Let $u_1, u_2, \ldots, u_d$ be the neighbors of u not equal to v. By the assumption that T_u is alternating we have that $c < \rho(T_{u_i})$, and since $\max\{\rho(T^v_{v_i})\} < c$ by the recursive definition $c < \rho(T^u_v)$. If $c < \rho(T^u_v) \leq \min\{\rho(T^u_{u_i})\}$ then we can root at vu and we are done. If not then there is some i such that $c < \rho(T^u_{u_i}) < \rho(T^u_v)$. Change the root to $v := u_i$ and start this case again from the beginning, noting that all the branches are alternating and on one side of c. Note that we will never go back to a previous choice of root since we are getting ρ-values closer to c, thus this will terminate. Note that this argument will also work if the degree of v is not $d+1$, taking the remaining branches to be empty.

Case 2. Suppose that all but one branch is on one side of c. Let $u, v_1, v_2, \ldots, v_d$ be the neighbors of v. Without loss of generality suppose that $\rho(T_u) < c < \min\{\rho(T_{v_i})\}$. If $\rho(T^v_u) \leq \rho(T^u_v) < c$ then the choice of v satisfies the lemma and we are done. If $\rho(T^u_v) < \rho(T^v_u) < c$ then the choice of u satisfies the lemma and note that all of its branches are alternating.

Case 3. Suppose there is at least two branches of v on each side of c. Let $u \in N(v)$, notice $\overrightarrow{uv}$ is a split edge so T^u_v is on the c-side of T^v_u. □

Lemma 17. *If T is extremal on $n \geq 3$ vertices, then there is a choice of root that is not a leaf, such that all of its branches are alternating and we have a strong ordering with respect to this choice.*

Proof. Let T be rooted at the choice of edge or vertex given by the previous lemma. First consider the case where we are rooted at an edge v_1v_2, and without loss of generality $v_1 < c < v_2$ in which case $v_1 \prec w_2$ in the weak ordering. By the previous lemma for any vertex u at depth 1, if $u < c$ then $\rho(T_u) \leq \rho(T_{v_1})$ and so we can take

$u \prec v_1$, and if $u > c$ then $\rho(T_{v_1}) \leq \rho(T_u)$ and so we can take $v_1 \prec u$. This gives a strong ordering up to depth 1 in the edge rooted case and in the vertex rooted case we already have a strong order up to depth 1. In either case we have vertices on both sides of c in this order. Suppose we have a strong ordering of all vertices up to depth $m \geq 1$. Let w be maximum with respect to the weak ordering at depth $m+1$ such that $w < c$ and let u be minimum with respect to the strong order up to depth m, by construction u is at depth m and $u < c$, consequently u and w are not in the same branch of T. By the inductive hypothesis (or in the vertex rooted case when $m = 1$, the previous lemma) we have that $c < \rho(T_{u'}) \leq \rho(T_{w'})$ and so by the branch exchange property we have that $\rho(T_w) \leq \rho(T_u) < c$, so we can take $w \prec u$. If w is minimum with respect to the weak order at depth $m+1$ such that $w > c$ and u is maximum with respect to the strong order up to depth m, a similar argument shows we can take $u \prec w$. □

Theorem 18. *Let σ be a invariant of trees with a decreasing certificate ρ. Then σ is optimized by the festoon $F_{n,d}$.*

Proof. We begin with the choice of root provided by the previous lemma. Let h be the height of the tree, that is the maximum distance from the root to a leaf (the distance from an edge being the minimum of the distances from its adjacent vertices). We need to show that the minimum distance to a leaf is at least $h-2$, that all but at most one branch is complete, and that the one possibly incomplete branch satisfies this condition recursively, that is all but one of its branches are complete, etc. To show this we need only consider where all the leafs of the tree are in the strong ordering provided by the previous lemma. Since $\rho(C_1) < \rho(T)$ for nonempty $T \neq C_1$ we have $w \prec v$ for any non-leaf v and leaf w. Thus by the conditions of strong ordering, any non-leaf vertex has depth at most one more then that of any leaf, so the minimum distance to a leaf must be at least $h-2$. To see that at most one branch is not complete, let $\{T_i\}$ be the branches of the root and suppose that T_{i_0} is incomplete. Observe that since it is alternating, T_{i_0} has leafs only at depth h and $h-2$. We claim that if $T_i (i \neq i_0)$ is on the c-side of T_{i_0} and they are on the same side of c, then $T_i = C_h$, and if T_i is not on the c-side of T_{i_0} (they still are on the same side of c though) then $T_i = C_{h-2}$, and if T_i is not on the same side of c as T_{i_0} then $T_i = C_{h-1}$. By Corollary 6 we cannot have any branch with the same ρ-value as T_{i_0}. If $T_i (i \neq i_0)$ is on the c-side of T_{i_0} and they are on the same side of c, then by inductively applying the branch exchange property to each branch we have $c < \rho(T_u) < \rho(T_w)$ or $\rho(T_w) < \rho(T_u) < c$ for all u in $\rho(T_i)$ and w in T_{i_0} both at depth k and thus since T_{i_0} has a leaf at depth h from the root, we must have $T_i = C_h$. The other cases are similar. Then we can consider the incomplete branch inductively with its induced strong order. Notice that in the edge rooted case one of the two branches of the root is C_{h-1} so a vertex root choice of the vertex adjacent to the edge root which is not part of this C_{h-1} will be consistent with the definition of the festoon. Also note that since all the branches of the root are alternating for the recursive part we do not get any C_{h-1}'s. Therefore σ is optimized by the festoon $F_{n,d}$. □

In the collection $\mathcal{T}_{n,d}$, this tree has previously been shown to maximize the number of independent sets, minimize the number of matchings and minimize energy [6, 7].

3.2.6 Applications

Number of Independent Sets

For the first application of this method we will maximize the number of independent sets of a tree in the collection $\mathcal{T}_{n,d}$. We will take $\rho(T)$ to be the ratio $\sigma_0(T)/\sigma(T)$ where $\sigma(T)$ is the total number of independent sets and $\sigma_0(T)$ is the total number of independent sets not containing the root. We will also use the quantity $\sigma_1(T) = \sigma(T) - \sigma_0(T)$.

Definition 14. Let $f : (0,1]^d \to (0,1]$ be defined as,

$$f(x_1, \dots, x_d) = \frac{1}{1 + \prod_{i=1}^{d} x_i}.$$

Notice that f is a continuous symmetric decreasing function. We have the following.

Lemma 19.

$$\begin{aligned} \sigma_0(T) &= \prod_{i=1}^{r} \sigma(T_i) \\ \sigma_1(T) &= \prod_{i=1}^{r} \sigma_0(T_i) \\ \rho(T) &= f\left(\rho(T_1), \dots, \rho(T_d)\right) \end{aligned}$$

Proof. The first two are clear and the last is a simple manipulation of the first two,

$$\rho(T) = \frac{\sigma_0(T)}{\sigma_0(T) + \sigma_1(T)} = \frac{1}{1 + \frac{\sigma_1(T)}{\sigma_0(T)}} = \frac{1}{1 + \prod_{i=1}^{d} \rho(T_i)}$$

□

Lemma 20 (Branch Exchange)**.** *Let T be a tree in $\mathcal{T}_{n,d}$ with $\sigma(T)$ maximum and let $l \neq r$ be vertices of T. Let $\{L_i\}$ be the d branches of T_l^r and $\{R_i\}$ the d branches of T_r^l. We have* $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$ *or* $\min\{\rho(L_i)\} \geq \max\{\rho(R_i)\}$.

Proof. Let T' be the maximal subtree of T having l and r as leafs. Let $\mathcal{I}$ be the collection of independent sets of T' and define the following quantities,

$$\sigma_{00}(T') = \#\{A \in \mathcal{I} : l \notin A, r \notin A\}$$
$$\sigma_{01}(T') = \#\{A \in \mathcal{I} : l \notin A, r \in A\}$$
$$\sigma_{10}(T') = \#\{A \in \mathcal{I} : l \in A, r \notin A\}$$
$$\sigma_{11}(T') = \#\{A \in \mathcal{I} : l \in A, r \in A\}.$$

We have that the total number of independent sets of T is,

$$\begin{aligned}
\sigma(T) =& \sigma_{00}(T')\sigma_0(L)\sigma_0(R) + \sigma_{01}(T')\sigma_0(L)\sigma_1(R) \\
&+ \sigma_{10}(T')\sigma_1(L)\sigma_0(R) + \sigma_{11}(T')\sigma_1(L)\sigma_1(R) \\
=& \sigma_0(L)\sigma_0(R)\Bigg(\sigma_{00}(T') + \sigma_{01}(T')\frac{\sigma_1(R)}{\sigma_0(R)} \\
&+ \sigma_{10}(T')\frac{\sigma_1(L)}{\sigma_0(L)} + \sigma_{11}(T')\frac{\sigma_1(L)}{\sigma_0(L)}\frac{\sigma_1(R)}{\sigma_0(R)}\Bigg) \\
=& \sigma_0(L)\sigma_0(R)\Bigg(\sigma_{00}(T') + \sigma_{01}(T')\prod_{i=1}^{d}\rho(R_i) \\
&+ \sigma_{10}(T')\prod_{i=1}^{d}\rho(L_i) + \sigma_{11}(T')\prod_{i=1}^{d}\rho(L_i)\prod_{i=1}^{d}\rho(R_i)\Bigg)
\end{aligned}$$

Consider permuting the branches $\{L_1, L_2, \dots, L_d, R_1, R_2, \dots, R_d\}$ by a permutation π. Notice that this preserves the required structure of T; it still has n vertices and degree at most $d+1$. We would like to consider how one of these permutations changes the total number of independent sets. By assumption T maximizes $\sigma(T)$, so a permutation of the branches can only reduce this quantity. Observe that $\sigma_0(L)\sigma_0(R) = \prod_{i=1}^{r}\sigma(L_i)\prod_{i=1}^{r}\sigma(R_i)$ is invariant under π. Also, $\prod_{i=1}^{d}\rho(L_i)\prod_{i=1}^{d}\rho(R_i)$ is invariant under π, thus we need only maximize,

$$\sigma_{01}(T')\prod_{i=1}^{d}\rho(R_i)+\sigma_{10}(T')\prod_{i=1}^{d}\rho(L_i).$$

Thus if $\sigma_{01}(T') \geq \sigma_{10}(T')$ this is maximized when $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$ and if $\sigma_{01}(T') \leq \sigma_{10}(T')$ this is maximized when $\min\{\rho(L_i)\} \geq \max\{\rho(R_i)\}$. $\square$

Theorem 21. *The number of independent sets of a tree in $\mathcal{T}_{n,d}$ is maximized by the festoon $F_{n,d}$.*

Proof. We have a decreasing certificate ρ, thus we are done. $\square$

This result was previously shown in [6] using a different method that also relied on the branch exchange property.

Matching Generating Polynomial

A similar example is minimizing the number of matchings of a tree in the collection $\mathcal{T}_{n,d}$. A generalization of this problem is minimizing the weighted number of matchings where the weight of a matching I is defined to be $\lambda^{|I|}$ for some $\lambda \in \mathbb{R}$ and the weighted number of matchings is just the sum of these weights over all matchings. This quantity is the definition of the matching generating polynomial evaluated at λ. Explicitly, the matching generating polynomial for a graph G is the polynomial $M(G,\lambda) = \sum_{k\geq 0} m(G,k)\lambda^k$ where $m(G,k)$ is the number of matchings in G of exactly k edges.

We will show that the matching generating polynomial is minimized for all $\lambda > 0$ by the festoon. As before we will need some auxiliary quantities for our computations. For a rooted tree T, let $m_1(T,k)$ to be the number of matchings of k edges covering the root and let $m_0(T,k)$ be the number not covering the root. Let $M_j(T,\lambda) =$

$\sum_{k\geq 0} m_j(G,k)\lambda^k$ for $j \in \{0,1\}$. We will fix an arbitrary choice of $\lambda > 0$ and for a rooted tree T define

$$\mu(T) = \frac{M_0(T,\lambda)}{M(T,\lambda)}.$$

In Lemma 3.1 of [7] it is shown that μ satisfies the branch exchange property for extremal (minimum matching polynomial at λ) trees. Furthermore, $\mu(T)$ has a recursive definition, $\mu(T) = g(\mu(T_1), \ldots, \mu(T_d))$, where

$$g(x_1, \ldots, x_d) = \frac{1}{1 + \lambda \sum_{i=1}^{d} x_i}.$$

It is clear that g is decreasing and so we have a decreasing certificate ρ (where $\rho(T) = (1 + \mu(T))^{-1}$) and so the matching generating polynomial at λ is minimized by the festoon $F_{n,d}$.

Weighted Homomorphisms

Given the previous application it is natural to think about doing the same thing for independent sets. In fact we can do even better.

Definition 15. A *homomorphism* from a graph G to another graph H is a map $f : V(G) \to V(H)$ such that if $u \sim v$ then $f(u) \sim f(v)$. Let $\mathrm{Hom}(G,H)$ be the collection of homomorphisms from G to H, and let $\mathrm{hom}(G,H)$ be the number of such homomorphisms.

Remark. An independent set can be identified with a homomorphism to a graph on two vertices connected by an edge, one of the vertices having a loop.

We can also assign weights $w : V(H) \to (0,\infty)$ to an image graph to make the following generalization of the number of homomorphisms from G to H, for details see [5].

Definition 16. Let $w : V(H) \to (0, \infty)$ be an assignment of weights. Then the *weight of a homomorphism* f from G to H is

$$wt(f) = \prod_{u \in G} w(f(u)) = \prod_{v \in H} w(v)^{|f^{-1}(v)|}$$

and the weighted number of the homomorphisms from G to H is

$$\hom_w(G, H) = \sum_{f \in \mathrm{Hom}(G,H)} wt(f).$$

Notice that when all the weights are one then this agrees with $\hom(G, H)$.

Remark. The weighted number of independent sets of any graph G is the weighted number of homomorphisms to a graph in Figure 3.3.

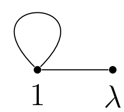


Figure 3.3: Image graph for counting the weighted number of independent sets.

Definition 17. A graph G is *strongly biregular* if there exists a partition of the vertices $A \cup B = V(G)$ and constants w, x, y, z such that for every $a \in A$ and $b \in B$ we have $w = d_A(a), x = d_B(a), y = d_A(b), z = d_B(b)$.

For the rest of the section let us consider H to be a fixed strongly biregular graph. We will assume that H is connected, since we could consider each connected component separately. We will assume that H is not regular, for if H is regular of degree r then the number of homomorphisms from a tree on n vertices to H is exactly $|H|(r-1)^{n-1}$, a constant. Let $A \cap B$ be the strongly biregular partition of the vertices of H, and let $a \in A$ and $b \in B$ be representatives of the two partitions.

We will assume that the vertices in each partition have equal non-zero weights, say $\mu \neq 0$ is the weight of a and $\nu \neq 0$ the weight of b. We will make the following abbreviations: $\alpha = \mu d_A(a), \beta = \nu d_B(a), \gamma = \mu d_A(b), \delta = \nu d_B(b)$. Without loss of generality we will assume that $\alpha + \beta < \gamma + \delta$, else we can switch A and B.

Definition 18. Let $f : (0,1]^d \to (0,1]$ be defined as

$$f(x_1, \ldots, x_d) = \frac{\alpha \prod_{i=1}^r x_i + \beta}{\gamma \prod_{i=1}^r x_i + \delta}.$$

Definition 19. For a rooted tree T, let $\mu\,\mathrm{h}_a(T)$ be the weighted number of homomorphisms from T to H sending the root to a, and $\nu\,\mathrm{h}_b(T)$ the weighted number sending the root to b. Let $\rho(T)$ be the ratio

$$\rho(T) = \frac{\alpha\,\mathrm{h}_a(T) + \beta\,\mathrm{h}_b(T)}{\gamma\,\mathrm{h}_a(T) + \delta\,\mathrm{h}_b(T)}.$$

Lemma 22. *Let T be rooted with branches $T_1, \ldots, T_r$. Then*

$$\begin{aligned}
\mathrm{h}_a(T) &= \prod_{i=1}^{r} \big(\alpha\,\mathrm{h}_a(T_i) + \beta\,\mathrm{h}_b(T_i)\big) \\
\mathrm{h}_b(T) &= \prod_{i=1}^{r} \big(\gamma\,\mathrm{h}_a(T_i) + \delta\,\mathrm{h}_b(T_i)\big) \\
\rho(T) &= f\left(\rho(T_1), \ldots, \rho(T_d)\right)
\end{aligned}$$

Proof. The first two equations are clear. The last equation is a simple manipulation of the first two,

$$\rho(T) = \frac{\alpha\,\mathrm{h}_a(T) + \beta\,\mathrm{h}_b(T)}{\gamma\,\mathrm{h}_a(T) + \delta\,\mathrm{h}_b(T)} = \frac{\alpha\frac{\mathrm{h}_a(T)}{\mathrm{h}_b(T)} + \beta}{\gamma\frac{\mathrm{h}_a(T)}{\mathrm{h}_b(T)} + \delta} = \frac{\alpha \prod_{i=1}^r \rho(T_i) + \beta}{\gamma \prod_{i=1}^r \rho(T_i) + \delta} = f\left(\rho(T_1), \ldots, \rho(T_d)\right).$$

$\square$

Remark. We have that the weighted number of homomorphisms from T to H is,

$$\hom_w(T,H) = |A|\mu\, \mathrm{h}_a(T) + |B|\nu\, \mathrm{h}_b(T).$$

Remark. Its easy to verify that f is increasing when $d_B(a)d_A(b) < d_A(a)d_B(b)$, decreasing when $d_B(a)d_A(b) > d_A(a)d_B(b)$, and constant when they are equal. Also $f(x)$ is continuous symmetric.

The following shows that ρ satisfies the branch exchange property for $\hom(\,\cdot\,, H)$.

Lemma 23 (Branch Exchange)**.** *Let T be a tree in $\mathcal{T}_{n,d}$ with $\hom_w(T,H)$ maximum and let $l \neq r$ be vertices of T. Let $\{L_i\}$ be the d branches of T_l^r and $\{R_i\}$ the d branches of T_r^l. We have $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$ or $\min\{\rho(L_i)\} \geq \max\{\rho(R_i)\}$.*

Proof. Let T' be the maximal subtree of T having l and r as leafs. For $x, y \in \{a,b\}$, define

$$\mathrm{h}_{xy}(T') = \sum_{\substack{f \in \mathrm{Hom}(T',H) \\ f(l)=x, f(r)=y}} wt(f).$$

We have that the weighted number of homomorphisms from T to H is,

$$\begin{aligned}
\hom_w(T,H) &= \mathrm{h}_{AA}(T')\,\mathrm{h}_a(L)\,\mathrm{h}_a(R) + \mathrm{h}_{AB}(T')\,\mathrm{h}_a(L)\,\mathrm{h}_b(R) \\
&\qquad + \mathrm{h}_{BA}(T')\,\mathrm{h}_b(L)\,\mathrm{h}_a(R) + \mathrm{h}_{BB}(T')\,\mathrm{h}_b(L)\,\mathrm{h}_b(R) \\
&= \mathrm{h}_b(L)\,\mathrm{h}_b(R)\left(\mathrm{h}_{AA}(T')\frac{\mathrm{h}_a(L)\,\mathrm{h}_a(R)}{\mathrm{h}_b(L)\,\mathrm{h}_b(R)} + \mathrm{h}_{AB}(T')\frac{\mathrm{h}_a(L)}{\mathrm{h}_b(L)}\right. \\
&\qquad \left. + \mathrm{h}_{BA}(T')\frac{\mathrm{h}_a(R)}{\mathrm{h}_b(R)} + \mathrm{h}_{BB}(T')\right) \\
&= \mathrm{h}_b(L)\,\mathrm{h}_b(R)\left(\mathrm{h}_{AA}(T')\prod_{i=1}^{d}\rho(L_i)\rho(R_i) + \mathrm{h}_{AB}(T')\prod_{i=1}^{d}\rho(L_i)\right. \\
&\qquad \left. + \mathrm{h}_{BA}(T')\prod_{i=1}^{d}\rho(R_i) + \mathrm{h}_{BB}(T')\right).
\end{aligned}$$

Consider permuting the branches $\{L_1, L_2, \ldots, L_d, R_1, R_2, \ldots, R_d\}$ by a permutation π, notice that this preserves the required structure of T, it still has n vertices and degree at most $d+1$. We would like to consider how one of these permutations changes the weighted number of homomorphisms. By assumption T maximizes $\hom_w(T,H)$, so a permutation of the branches can only reduce the weighted number of homomorphisms. Observe that $\mathrm{h}_b(L)\,\mathrm{h}_b(R) = \prod_{i=1}^{r}\big(\gamma\,\mathrm{h}_a(L_i)+\delta\,\mathrm{h}_b(L_i)\big)\big(\gamma\,\mathrm{h}_a(R_i)+\delta\,\mathrm{h}_b(R_i)\big)$ is invariant under π. Also, $\prod_{i=1}^{d}\rho(L_i)\rho(R_i)$ is invariant under π, thus we need only maximize,

$$\mathrm{h}_{AB}(T')\prod_{i=1}^{d}\rho(L_i) + \mathrm{h}_{BA}(T')\prod_{i=1}^{d}\rho(R_i).$$

Thus, if $\mathrm{h}_{AB}(T') \geq \mathrm{h}_{BA}(T')$ then this is maximized when $\min\{L_i\} \geq \max\{R_i\}$, and if $\mathrm{h}_{AB}(T') \leq \mathrm{h}_{BA}(T')$ then this is maximized when $\max\{\rho(L_i)\} \leq \min\{\rho(R_i)\}$. $\square$

Remark. If f is constant then, by argument of the previous lemma, $\hom_w(T,H)$ is invariant under permutations of any of its branches and thus is constant for all trees on n vertices. Note that it is fair to consider $\rho(C_0)$ to be this constant value since it is still compatible with the recursive definition.

Theorem 24. *The weighted number of homomorphisms of a tree T in $\mathcal{T}_{n,d}$ to a strongly biregular graph H with partition A and B and representatives a and b respectively,* $\hom_w(T,H)$*, is maximized by the festoon $F_{n,d}$ when $d_A(a)d_B(b) - d_B(a)d_A(b) < 0$ and by the ball $B_{n,d}$ when $d_A(a)d_B(b) - d_B(a)d_A(b) > 0$.*

Proof. In the first case we have a decreasing certificate ρ and in the second case we have an increasing certificate ρ, thus we are done. □

Remark. Notice that the weights do not matter. If T maximizes $\hom(T,H)$ then it maximizes $\hom_w(T,H)$ for any assignment of non-zero weights to the vertices of H.

Example 2. An interesting application of this result is the *Widom-Rowlinson model* in statistical physics. Simply stated it is $hom(T, P_2^\circ)$, where P_2° is the path of length 2 with loops at every vertex (see Figure 3.4). It is easy to see that P_2° is strongly biregular with partition $A = \{a, c\}$ and $B = \{b\}$. We have that $d_A(a)d_B(b) - d_B(a)d_A(b) = 1 \cdot 1 - 1 \cdot 2 = -1 < 0$. So in $\mathcal{T}_{n,d}$, $hom(T, P_2^\circ)$ is uniquely maximized by the festoon $F_{n,d}$.

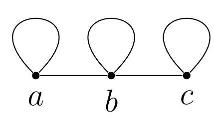


Figure 3.4: The graph P_2°.

Number of Subtrees

We will show that the ball maximizes the number of subtrees of a tree with bounded degree. A subtree is just a connected subgraph of a tree. Let $\sigma(T)$ be the number of subtrees of T, $\mu(T)$ the number containing the root of T. Note that

$$\mu(T) = \prod_{i=1}^{d} (1 + \mu(T_i))$$

So if $g(x_1, \dots, x_d) = \prod_{i=1}^{d}(1 + x_i)$, then $\mu(T) = g(\mu(T_1), \dots, \mu(T_d))$ and g is increasing. With the standard certificate of T for the branch exchange property, let $\sigma_{00}(T')$ be the number of subtrees of T' not containing l or r, let $\sigma_{10}(T')$ be the number containing l but not r and so on. Then we have the following

$$\begin{aligned}
\sigma(T) &= \sigma_{11}(T')\mu(L)\mu(R) + \sigma_{10}(T')\mu(L) + \sigma_{01}(T')\mu(R) + \sigma_{00}(T') + \sigma_0(L) + \sigma_0(R) \\
&= \sigma_{11}(T') \prod_{i=1}^{d} (1 + \mu(L_i)) (1 + \mu(R_i)) + \sigma_{10}(T') \prod_{i=1}^{d} (1 + \mu(L_i)) \\
&\quad + \sigma_{01}(T') \prod_{i=1}^{d} (1 + \mu(R_i)) + \sigma_{00}(T') + \sum_{i=1}^{d} \sigma(L_i) + \sum_{i=1}^{d} \sigma(R_i).
\end{aligned}$$

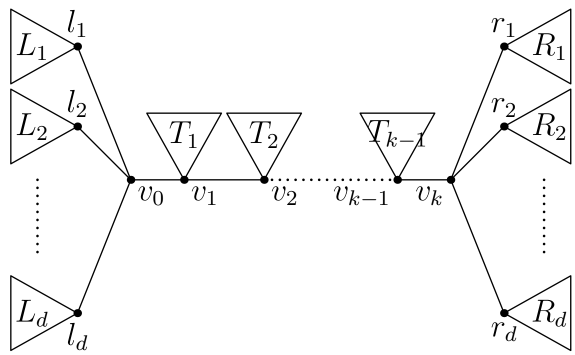
This is maximized when $\max\{\mu(L_i)\} \leq \min\{\mu(R_i)\}$ or $\min\{\mu(L_i)\} \geq \max\{\mu(R_i)\}$. Therefore we have a increasing certificate ρ, and so the number of subtrees is maximized when T is the ball $B_{n,d}$. This result was previously shown in [12, 19] using a different method.

Wiener Index

We will show that the ball minimizes the Wiener index of a tree with bounded degree. The Wiener index is the sum of the distances between every pair of vertices. The Wiener index $W(T)$ of a tree can also be computed using the following formula:

$$W(T) = \sum_{uv \in E} |T_u^v||T_v^u|$$

This simplifies the computation of whether switching two branches will increase or decrease the Wiener index, as the only terms in the above sum that change are the ones for edges on the path between the roots of these branches. Let T be decomposed as seen below.



Let $a_i = |L_i|$, $b_i = |R_i|$, $n_i = |T_i|$, and for ease of notation $n_0 = 1 + \sum a_i$, $n_k = 1 + \sum b_i$, and $n = |T|$. Then we have

$$W(T) = \sum_{\substack{uv \in E \\ uv \neq l_i v_0,\, uv \neq v_k r_i \\ uv \neq v_i v_{i+1}}} |T_u^v||T_v^u| + \sum_{i=1}^{k} \left(\sum_{j=0}^{i-1} n_j \right) \left(\sum_{j=i}^{k} n_j \right) + \sum_{i=1}^{d} a_i \left(n - a_i \right) + \sum_{i=1}^{d} b_i \left(n - b_i \right) \tag{3.1}$$

Note that the first term is invariant under permutation of the branches $\{L_i, B_i \mid i = 1, \ldots, d\}$. The second term simplifies as follows.

$$\begin{aligned}
\sum_{i=1}^{k}\left(\sum_{j=0}^{i-1} n_j\right)\left(\sum_{j=i}^{k} n_j\right) &= \sum_{i=1}^{k}\left(n_0 + \sum_{j=1}^{i-1} n_j\right)\left(n_k + \sum_{j=i}^{k-1} n_j\right) \\
&= \sum_{i=1}^{k}\left(n_0 n_k + n_0 \sum_{j=i}^{k-1} n_j + n_k \sum_{j=1}^{i-1} n_j + \left(\sum_{j=1}^{i-1} n_j\right)\left(\sum_{j=i}^{k-1} n_j\right)\right) \\
&= k n_0 n_k + n_0 \sum_{i=1}^{k}\sum_{j=i}^{k-1} n_j + n_k \sum_{i=1}^{k}\sum_{j=1}^{i-1} n_j \\
&\quad + \sum_{i=1}^{k}\left(\sum_{j=1}^{i-1} n_j\right)\left(\sum_{j=i}^{k-1} n_j\right) \\
&= k\left(1 + \sum_{i=i}^{d} a_i\right)\left(1 + \sum_{i=i}^{d} b_i\right) + \left(1 + \sum_{i=i}^{d} a_i\right)\sum_{i=1}^{k}\sum_{j=i}^{k-1} n_j \\
&\quad + \left(1 + \sum_{i=i}^{d} b_i\right)\sum_{i=1}^{k}\sum_{j=1}^{i-1} n_j + \sum_{i=1}^{k}\left(\sum_{j=1}^{i-1} n_j\right)\left(\sum_{j=i}^{k-1} n_j\right)
\end{aligned} \tag{3.2}$$

So (3.2) is minimized when the a_i are all greater than or equal to the b_i or vice versa.

$$\sum_{i=1}^{d} a_i (n - a_i) + \sum_{i=1}^{d} b_i (n - b_i) = \left(\sum_{i=1}^{d} a_i + \sum_{i=1}^{d} b_i\right) n - \sum_{i=1}^{d} a_i^2 - \sum_{i=1}^{d} b_i^2$$

This shows that the last two terms of (3.1) are invariant. Therefore, $\mu(T) = |T|$ satisfies the branch exchange property for the Wiener index and has $\rho(C_0) = 0$ and $\rho(C_1) = 1$. Let $g(x_1, \ldots, x_d) = 1 + \sum_{i=1}^{d} x_i$, then g is a recursive definition of $\mu(T)$ and is increasing. So we have an increasing certificate ρ and hence $W(T)$ is minimized by the ball. This result was previously shown in [4] and later in [9] using different methods.

Average path length is also minimized by the ball since the Wiener index is the sum of the lengths of every path and there is a fixed number of paths in a tree of size n.

Largest Eigenvalue

Definition 20. The characteristic polynomial of a graph G is the polynomial $\phi_G(x) = \det(xI - A)$ where A is the adjacency matrix of G. The roots of this polynomial are the eigenvalues which are all real since A is symmetric. The largest eigenvalue will be λ_G. Note $\lambda_G > 0$ for any connected graph with at least one edge.

It is well known that if H is a subgraph of G then $\lambda_H \leq \lambda_G$ and for $x > \lambda_G$ we have $\phi_G(x) > 0$. A well known recursive formula for a graph G with a bridge edge uv, that is an edge not in any cycle, is $\phi_G = \phi_{G-uv} - \phi_{G-u-v}$. We would like to show that the ball has the largest eigenvalue for trees with bounded degree. To do this we will start with an extremal tree T_0, and for $x > \lambda_{T_0}$ we will show that T_0 minimizes the value of the characteristic polynomial at x. Further we will show for a tree T (a subgraph of T_0) rooted at r that $\mu(T) = \phi_{T-r}(x)/\phi_T(x)$ corresponds to an increasing certificate ρ of $\sigma(T_0) = \phi_{T_0}(x)$, and therefore T_0 is the ball. Note that $\phi_{C_0} = 1$, by convention we will have $\phi_{C_0-r} = 0$ where r is the fake root of C_0. A recursive definition for the characteristic polynomial of a tree is [15],

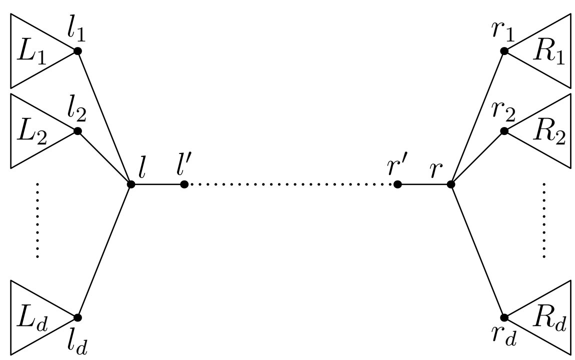
$$\phi_T(x) = \phi_{T-r}(x)\left(x - \sum_{i=1}^{r} \frac{\phi_{T_i-r_i}(x)}{\phi_{T_i}(x)}\right),$$

where T has root r and branches T_i with roots r_i. Note that the convention of $\phi_{C_0-r} = 0$ is consistent with this recursive definition. It is important to note that since we will only be considering values of $x > \lambda_{T_0}$, we will have that $\phi_T(x) > 0$ for all trees T of bounded degree $d+1$ and n or fewer vertices. The above formula yields a recursive definition of μ,

$$\mu(T) = \frac{1}{x - \sum_{i=1}^{d} \mu(T_i)}.$$

The corresponding symmetric function is $g(x_1, \ldots, x_d) = (x - \sum x_i)^{-1}$ which is increasing.

To establish the branch exchange property decompose a tree T as shown in the diagram below where $l \neq r$ are two arbitrary vertices and l' is the predecessor of l relative to r and r' is the predecessor of r relative to l. The graph $T - ll' - r'r$ has at most 3 components. Denote the one containing l by L, containing r by R, and the one containing l' and r' (if such a component exists) by T'.



Using the preceding formulas we can write the characteristic polynomial of T at x (assuming $l' \neq r$ and $l' \neq r'$, although it is an easy exercise to see that these cases are the same) as:

$$\begin{aligned}
\phi_T(x) &= \phi_{T-ll'-r'r}(x) - \phi_{T-ll'-r'-r}(x) - \phi_{T-l-l'-r'r}(x) + \phi_{T-l-l'-r'-r}(x) \\
&= \phi_L(x)\phi_{T'}(x)\phi_R(x) - \phi_L(x)\phi_{T'-r'}(x)\phi_{R-r}(x) \\
&\quad - \phi_{L-l}(x)\phi_{T'-l'}(x)\phi_R(x) + \phi_{L-l}(x)\phi_{T'-l'-r'}(x)\phi_{R-r}(x) \\
&= \phi_{L-l}(x)\phi_{R-r}(x)\left[\phi_{T'}(x)\frac{1}{\mu(L)}\frac{1}{\mu(R)} - \phi_{T'-r'}(x)\frac{1}{\mu(L)}\right. \\
&\quad \left. -\phi_{T'-l'}(x)\frac{1}{\mu(R)} + \phi_{T'-l'-r'}(x)\right] \\
&= \prod_{i=1}^{d}\phi_{L_i}(x)\phi_{R_i}(x)\left[\phi_{T'}(x)\left(x - \sum_{i=1}^{d}\mu(L_i)\right)\left(x - \sum_{i=1}^{d}\mu(R_i)\right)\right. \\
&\quad \left. -\phi_{T'-r'}(x)\left(x - \sum_{i=1}^{d}\mu(L_i)\right) - \phi_{T'-l'}(x)\left(x - \sum_{i=1}^{d}\mu(R_i)\right) + \phi_{T'-l'-r'}(x)\right]
\end{aligned}$$

Up to a positive constant factor the part of the above expression that is not constant under permutation of the branches is:

$$\phi_{T'}(x)\left(\sum_{i=1}^{d}\mu(L_i)\right)\left(\sum_{i=1}^{d}\mu(R_i)\right) - \phi_{T'-r'}(x)\left(\sum_{i=1}^{d}\mu(L_i)\right) - \phi_{T'-l'}(x)\left(\sum_{i=1}^{d}\mu(R_i)\right)$$

This quantity is minimized when $\max\{\mu(L_i)\} \leq \min\{\mu(R_i)\}$ or $\min\{\mu(L_i)\} \geq \max\{\mu(R_i)\}$. Therefore, we have an increasing certificate ρ and so the characteristic polynomial is minimized at $x > \lambda_{T_0}$ by the ball $B_{n,d}$, hence the ball has the largest eigenvector in $\mathcal{T}_{n,d}$. This result was previously shown in [18] using a method that depended on the entries of the principal eigenvector.

Largest Laplacian Eigenvalue

Using a similar method as in the previous section we can show that the ball maximizes the largest Laplacian eigenvalue. This is the largest eigenvalue of the Laplacian matrix $L = D - A$ where A is the adjacency matrix and D is the diagonal matrix of vertex degrees.

Definition 21. The Laplacian polynomial of a graph G is the polynomial $\psi_G(x) = \det(xI - L)$. The roots of this polynomial are the eigenvalues which are all real since L is symmetric. The largest eigenvalue will be λ_G, note $\lambda_G > 0$ for any connected graph with at least one edge.

Definition 22. Let H be a subgraph of G. The modified Laplacian polynomial of H is the polynomial $\psi'_H(x) = \det(xI - L_H)$, where L is the Laplacian matrix of G and L_H is the restriction of the matrix to the vertices of H. The largest root of this polynomial will be denoted as λ_H.

As before it is easy to check that for $x > \lambda_G$ we have $\psi_G(x) > 0$ and $\psi'_H(x) > 0$ for all subgraphs H. A recursive formula for a graph G with a bridge edge uv, that is an edge not in any cycle, is

$$\psi_G = \psi'_{G-uv} - \psi'_{G-u-v}.$$

We would like to show that the ball has the largest eigenvalue for trees with bounded degree. To do this we will start with an extremal tree T_0, and for $x > \lambda_{T_0}$ we will show that T_0 minimizes the value of the Laplacian polynomial at x. Further we will show for a tree T (a subgraph of T_0) rooted at r that $\mu(T) = \psi'_{T-r}(x)/\psi'_T(x) + 1$ (and $\mu(C_0) = 0$) corresponds to an increasing certificate ρ of $\sigma(T_0) = \psi'_{T_0}(x)$, and therefore T_0 is the ball. Note that $\psi'_{C_0} = 1$ as it is the determinant of the 0×0 matrix. A recursive definition for the modified Laplacian polynomial of a tree is,

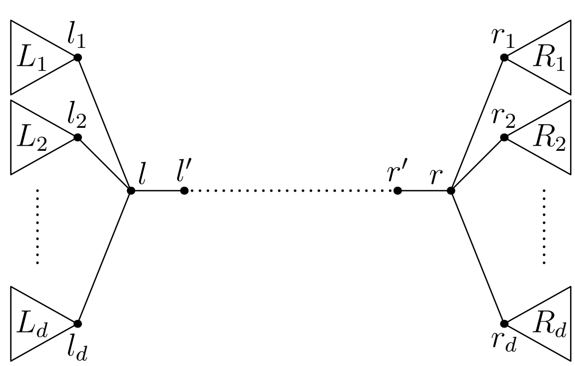
$$\psi'_T(x) = \psi'_{T-r}(x)\left(x - 1 - \sum_{i=1}^{r}\left(\frac{\psi'_{T_i - r_i}(x)}{\psi'_{T_i}(x)} + 1\right)\right),$$

where T has root r and branches T_i with roots r_i. Note that the convention of $\psi_{C_0 - r} = 0$ is consistent with this recursive definition. It is important to note that since we will only be considering values of $x > \lambda_{T_0}$, we will have that $\psi'_T(x) > 0$ for all trees T of bounded degree $d+1$ and n or fewer vertices. The above formula yields a recursive definition of μ,

$$\mu(T) = \frac{1}{x - 1 - \sum_{i=1}^{d} \mu(T_i)} + 1.$$

The corresponding symmetric function is $g(x_1, \ldots, x_d) = (x - \sum x_i)^{-1}$ which is increasing.

To establish the branch exchange property decompose a tree T as shown in the diagram below where $l \neq r$ are two arbitrary vertices and l' is the predecessor of l relative to r and r' is the predecessor of r relative to l. The graph $T - ll' - r'r$ has at most 3 components denote the one containing l by L, containing r by R, and containing l' and r' (if such a component exists) by T'.



Using the preceding formulas we can write the characteristic polynomial of T at

x (assuming $l' \neq r$ and $l' \neq r'$, although it is an easy exercise to see that these cases are the same) as:

$$\begin{aligned}
\psi_T(x) &= \psi'_{T-ll'-r'r}(x) - \psi'_{T-ll'-r'-r}(x) - \psi'_{T-l-l'-r'r}(x) + \psi'_{T-l-l'-r'-r}(x) \\
&= \psi'_L(x)\psi'_{T'}(x)\psi'_R(x) - \psi'_L(x)\psi'_{T'-r'}(x)\psi'_{R-r}(x) \\
&\quad - \psi'_{L-l}(x)\psi'_{T'-l'}(x)\psi'_R(x) + \psi'_{L-l}(x)\psi'_{T'-l'-r'}(x)\psi'_{R-r}(x) \\
&= \psi'_{L-l}(x)\psi'_{R-r}(x)\left[\psi'_{T'}(x)\frac{1}{\mu(L)-1}\frac{1}{\mu(R)-1} - \psi'_{T'-r'}(x)\frac{1}{\mu(L)-1}\right. \\
&\quad \left. -\psi'_{T'-l'}(x)\frac{1}{\mu(R)-1} + \psi'_{T'-l'-r'}(x)\right] \\
&= \prod_{i=1}^{d}\psi'_{L_i}(x)\psi'_{R_i}(x)\left[\psi'_{T'}(x)\left(x-1-\sum_{i=1}^{d}\mu(L_i)\right)\left(x-1-\sum_{i=1}^{d}\mu(R_i)\right)\right. \\
&\quad -\psi'_{T'-r'}(x)\left(x-1-\sum_{i=1}^{d}\mu(L_i)\right) - \psi'_{T'-l'}(x)\left(x-1-\sum_{i=1}^{d}\mu(R_i)\right) \\
&\quad \left. +\psi'_{T'-l'-r'}(x)\right]
\end{aligned}$$

Up to a positive constant factor the part of the above expression that is not constant under permutation of the branches is:

$$\psi'_{T'}(x)\left(\sum_{i=1}^{d}\mu(L_i)\right)\left(\sum_{i=1}^{d}\mu(R_i)\right) - \psi'_{T'-r'}(x)\left(\sum_{i=1}^{d}\mu(L_i)\right) - \psi'_{T'-l'}(x)\left(\sum_{i=1}^{d}\mu(R_i)\right)$$

This quantity is minimized when $\max\{\mu(L_i)\} \leq \min\{\mu(R_i)\}$ or $\min\{\mu(L_i)\} \geq \max\{\mu(R_i)\}$. Therefore, we have an increasing certificate ρ and so the Laplacian polynomial is minimized at $x > \lambda_{T_0}$ by the ball $B_{n,d}$, hence the ball has the largest Laplacian eigenvector in $\mathcal{T}_{n,d}$. This result was previously shown in [22] using a method that depended on the entries of the principal eigenvector.

3.3 Fixed Degree Sequence

3.3.1 Direct and Alternating Labelings

Definition 23. A *labeling* of a graph G is a function $f : V(G) \to \mathbb{R}$.

Definition 24. Let f and g be labelings of a graph G. We say f is a *refinement* of g if $f(u) \leq f(v)$ implies $g(u) \leq g(v)$ for all $u, v \in V(G)$.

Definition 25. A labeling f of graph G is *complete* if $f(u) \neq f(v)$ for all $u, v \in V(G)$ such that $u \neq v$.

It is clear that if we have a labeling that is complete we can make a labeling on $[n]$ by numbering the vertices smallest through largest with $1, 2, \ldots, n$.

Definition 26. Let f be a labeling of T. We say f is *direct* if

$$f(l) \leq f(r) \implies f(l_0) \leq f(r_0) \qquad \forall l \neq r, l_0 \in N^r(l), r_0 \in N^l(r).$$

We say f is *alternating* if

$$f(l) \leq f(r) \implies f(l_0) \geq f(r_0) \qquad \forall l \neq r, l_0 \in N^r(l), r_0 \in N^l(r).$$

Proposition 25. *If g is a refinement of f and g is direct (respectively, alternating) then f is direct (respectively, alternating).*

Proof. Clear by definition of refinement. □

Lemma 26. *If g is a direct (respectively, alternating) labeling of a tree T that is a refinement of the degree then there exists a complete refinement f of g that is also a direct (respectively, alternating) labeling of the vertices.*

Proof. The cases of $n = 1, 2$ are trivial. We will suppose g is not a complete labeling otherwise the result is also trivial. Note that since g is a direct or alternating order,

$$g(l) = g(r) \implies g(l_0) = g(r_0) \qquad \forall l \neq r, l_0 \in N^r(l), r_0 \in N^l(r). \tag{3.3}$$

First note that we cannot have a path on 3 vertices in T with each vertex having the same label; else by (3.3) the labeling is constant and thus not a refinement of degree since $n \geq 3$. Let $V_x = \{v \in V(T) : g(v) = x\}$. We cannot have distinct $u, v, w \in V_x$ such that w is in the path from u to v. If we did then since $d(w) \geq 2$, every vertex in V_x also has the same degree of at least two. Applying (3.3) to each pair of u, v, w yields that every neighbor of w has a common g value, say y, and so does every $u_0 \in N^v(u)$ and $v_0 \in N^u(v)$. Picking one such $u_0 \in N^v(u)$ and $v_0 \in N^u(v)$ as well as a neighbor w_0 of w on the path between u and v, yields a triple of the same kind but with distance 2 greater between u_0 and v_0. But this can be repeated without limit, a contradiction. Therefore, the minimal subtree T_x of T containing V_x is such that every vertex of V_x is a leaf and in this subtree the predecessor of each element of V_x is uniquely defined.

Now we will show that if there is $v \in V(T_x)$ with $g(v) = y$ then $T_y \subset T_x$. Suppose not then there is some $w \notin V(T_x)$ with $G(v) = y$. Since v is not a leaf of T_x there is a path $v = v_0, v_1, \ldots, v_k = u$ to a leaf u of T_x (i.e. $g(u) = x$) that does not intersect the path from v to w. Now let $w_0 = w$ and $w_{i+1} \in N^{v_i}(w_i)$ for $i = 0, 1, \ldots, k-1$ (such vertices exist since the labeling is a refinement of the degree and the degrees of $v_0, v_1, \ldots, v_{k-1}$ are at least 2). For $i = 0, 1, \ldots, k-1$ since $g(v_i) = g(w_i)$ by property (3.3) we have $g(v_{i+1}) = g(w_{i+1})$. But then $g(w_k) = g(v_k) = x$ and T_x contains the whole path including w, a contradiction.

By the previous claim we can pick an x such that T_x is minimal with respect to inclusion for x with $|V_x| \geq 2$. The values for the predecessors of V_x in T_x are distinct, with the exception that multiple elements of V_x can have the same predecessor. Let $v \in V_x$ be such that $g(v')$ is maximum among all predecessors of V_x in T_x. Let $h : V(T) \to \mathbb{R}$ be such that $h(w) = g(w)$ for all $w \neq v$ and $h(v) = g(v) + \frac{1}{2}\min\{g(a) - g(b) : a \neq b\}$ if g was direct or $h(v) = g(v) - \frac{1}{2}\min\{g(a) - g(b) : a \neq b\}$ if g was alternating. The labeling h by construction is clearly direct (respectively, alternating) by construction and we have reduced the number of pairs of vertices with equal labels. Thus, there is a complete labeling f refining g that is also direct (respectively, alternating). □

3.3.2 The Ball and Festoon

Definition 27. Let π be indexed in non decreasing order and $\sum_{i=1}^{n} \pi(i) = 2n - 2$. The *directed ball* with degree sequence π, denoted $\vec{B}_\pi$, is the directed graph with vertex set $[n]$ and edges $l \to k$ if $l \in S_k^B$, where S_k^B is defined below.

$$S_k^B = \begin{cases} \mathbb{Z} \cap \left(\sum_{i=1}^{k-1}(\pi(i) - 1), \sum_{i=1}^{k}(\pi(i) - 1)\right] & , k \neq n \\ \mathbb{Z} \cap \left(\sum_{i=1}^{n-1}(\pi(i) - 1), n - 1\right] & , k = n \end{cases}$$

The *ball*, B_π, is the undirected version of this graph.

Example 3. Let $\pi = (1,1,1,1,1,1,1,1,2,2,3,3,4,4)$, then

$$\begin{aligned} S_1^B = \cdots = S_8^B &= \{\} \\ S_9^B &= \{1\} \\ S_{10}^B &= \{2\} \\ S_{11}^B &= \{3,4\} \\ S_{12}^B &= \{5,6\} \\ S_{13}^B &= \{7,8,9\} \\ S_{14}^B &= \{10,11,12,13\} \end{aligned}$$

and B_π is the tree in Figure 3.5.

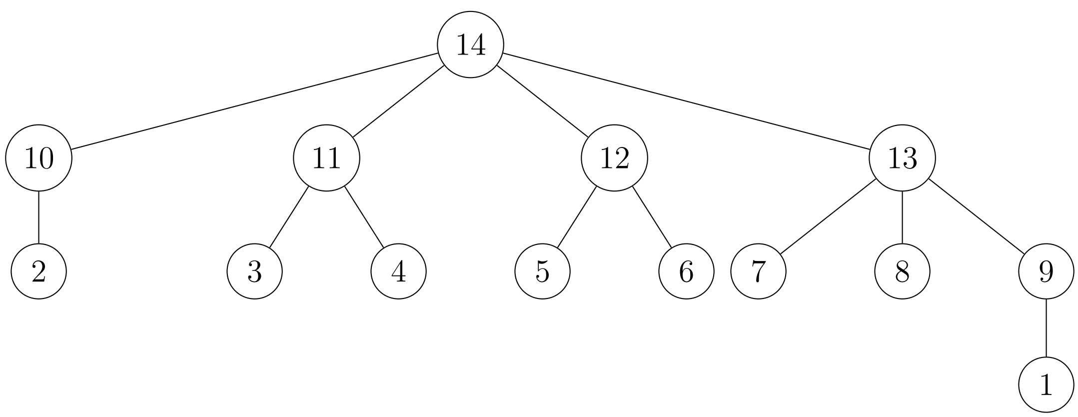


Figure 3.5: B_π for $\pi = (1,1,1,1,1,1,1,1,2,2,3,3,4,4)$.

Theorem 27. *B_π is a tree.*

Proof. Notice that $\{S_k^B\}_{k=1}^n$ partitions $[n-1]$, $l < k$ for all $l \in S_k$ (i.e., $l \to k \Rightarrow l < k$), $|S_k^B| = \pi(k) - 1$ for all $k \neq n$, and $|S_n^B| = \pi(n)$. In $\overrightarrow{B_\pi}$ every vertex except n has exactly one outgoing edge, the vertex at the end of this edge is a strictly larger integer, and thus every vertex has a directed path to n. Therefore, B_π is connected and has $n-1$ edges and thus is a tree. Furthermore, its degree sequence is π. $\square$

Remark. B_π has Prüfer Code

$$\underbrace{1,1,\ldots,1}_{\pi(1)-1},\underbrace{2,2,\ldots,2}_{\pi(2)-1},\ldots,\underbrace{n,n,\ldots,n}_{\pi(n)-1}.$$

For example, if

$$\pi = (1,1,1,1,1,1,1,1,2,2,3,3,4,4)$$

then B_π has Prüfer Code

$$9,10,11,11,12,12,13,13,13,14,14,14.$$

Definition 28. Let π be indexed in non decreasing order and $\sum \pi(i) = 2n-2$. The *directed festoon* with degree sequence π, denoted $\vec{F}_\pi$, is the directed graph with vertex set $[n]$ and edges $l \to k$ if $l \in S_k^F$, where S_k^F is defined below. Let $c = \min\left\{k : k > \sum_{i=k+1}^n (\pi(i)-1)\right\}$.

$$S_k^F = \begin{cases} \mathbb{Z} \cap \left(2+\sum_{i=k+1}^n(\pi(i)-1), 2+\sum_{i=k}^n(\pi(i)-1)\right] & ,1 \le k < c \\ \mathbb{Z} \cap \left(\sum_{i=k+1}^n(\pi(i)-1), 2+\sum_{i=k}^n(\pi(i)-1)\right] \setminus \{c\} & ,k = c \\ \mathbb{Z} \cap \left(\sum_{i=k+1}^n(\pi(i)-1), \sum_{i=k}^n(\pi(i)-1)\right] & ,c < k \le n. \end{cases}$$

The *festoon*, F_π, is the undirected version of this graph.

Example 4. Let $\pi = (1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 3, 3, 4, 4)$, then $c = 11$ and

$$\begin{aligned} S_1^F = \cdots = S_8^F &= \{\} \\ S_9^F &= \{14\} \\ S_{10}^F &= \{13\} \\ S_{11}^F &= \{9, 10, 12\} \\ S_{12}^F &= \{7, 8\} \\ S_{13}^F &= \{4, 5, 6\} \\ S_{14}^F &= \{1, 2, 3\} \end{aligned}$$

and F_π is the tree in Figure 3.6.

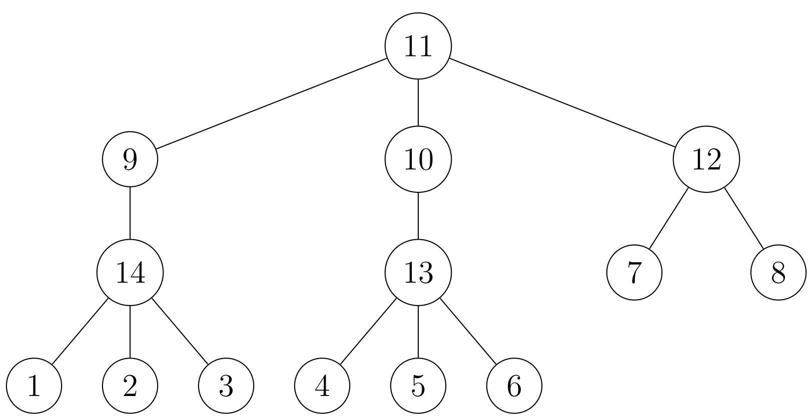


Figure 3.6: F_π for $\pi = (1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 3, 3, 4, 4)$.

Theorem 28. *F_π is a tree.*

Proof. Notice that $\{S_k^F\}_{k=1}^n$ partition $[n] - c$ and so in $\overrightarrow{F_\pi}$ every vertex but c has exactly one outgoing edge (and c has none). We have no loops because $k \notin S_k^F$ for all k. F_π has degree sequence π since $|S_k^F| = \pi(k) - 1$ for all $k \neq c$ and $|S_c^F| = \pi(c)$.

Define $g(x) : [n] \to [n]$ by $g(x) = y$ if $x \to y$ and $g(c) = c$. This is well defined since each vertex can have only one outgoing edge except c which has none. By the definitions of the S_k^F's we have that this function is decreasing. The composition of decreasing functions is increasing so $g \circ g$ is increasing. This shows that along any directed path the subsequence consisting of every other vertex in the path is monotonic (strictly since the values are distinct). In particular, since c is the only fixed point of g every directed path leads to c. Thus, F_π is connected and on $n-1$ edges so it is a tree. □

3.3.3 Uniqueness

Theorem 29. *The ball B_π is the unique tree in $\mathcal{T}_\pi$ where the degree is a direct labeling of the vertices.*

Proof. Let T be the ball as defined in Definition 27. It has the canonical labeling $f(k) = k$ which we will now show is a direct labeling. Consider two vertices k and l with $k < l$ and $i \in N^l(k)$, $j \in N^k(l)$. We would like to show that $i < j$. The following argument refers to the root orientation. In the case where $k \to i$ then the path from l to k must be directed and hence $l < k$, a contradiction. In the case where $i \to k$ and $l \to j$ then $i < k < l < j$. In the case where $i \to k$ and $j \to l$ then since $x < y$ for all $x \in S_k$ and $y \in S_l$ we have $i < j$. Since the labeling f is a refinement of the degree, the degree is a direct labeling of the vertices.

For the remainder we will let T be the undirected version of the above tree. This is the only tree where the degree is a direct labeling of the vertices. Let T' be a tree in $\mathcal{T}_\pi$ where the degree is a direct labeling of the vertices. By Lemma 26 there exists f' a complete refinement of the degree sequence that is a direct labeling of the vertices of T'. Without loss of generality we may assume f' has range $[n]$ and we may further assume the vertex set for T' is $[n]$ and this is a canonical labeling. We would like to show the identity map $\phi : V(T') \to V(T)$ is an isomorphism. Since f and f' are both refinements of π, ϕ preserves degrees. We will show $N^n_T(k) = N^n_{T'}(k)$ for every vertex $k \in [n]$. Once we have shown this for a vertex we will say the vertex is *verified* (and if not we say it is *unverified*). Let $A \subseteq [n]$ be the set of verified vertices; we start with $A = \emptyset$. At each stage we will try to verify the smallest unverified vertex, say k. It will always be the case that $N^c_T(k) \subseteq A$ (the children of k in T are always smaller). Suppose $N^n_T(k) \neq N^n_{T'}(k)$ then there exists $i \in N^n_{T'}(k) \setminus N^n_T(k)$ and $j \in N^n_T(k) \setminus N^n_{T'}(k)$. Consider the path between i and j in T'. Then k is the vertex adjacent to i in this path. We let l be the vertex adjacent to j in this path. At this point l cannot be verified because its child in T', j, has a different predecessor in T. Since k is the smallest unverified vertex we have $k < l$. Also $i > j$ since per the construction of T the children of k in T are the smallest vertices without full degree in $T[A]$. This prevents f' from being a direct labeling of T'. Thus, the children of k must be the same in T and T'. We need not verify n since its neighborhood is implicit once everything else is verified. Therefore, $T \simeq T'$.

□

Theorem 30. *The festoon F_π is the unique tree in $\mathcal{T}_\pi$ where the degree is an alternating labeling of the vertices.*

Proof. Let T be the festoon as defined in Definition 28. This tree has the canonical labeling $f(k) = k$ which we will now show is an alternating labeling. Consider two vertices k and l with $k < l$ and $i \in N^l(k)$, $j \in N^k(l)$. We would like to show that $i > j$. The following argument refers to the root orientation. In the case where $i \to k$ and $j \to l$, since g is decreasing we must have $i > j$. The case where $k \to i$ and $l \to j$ can not happen. Consider the path P from k to l (which does not contain i and j), both ends of this path must be oriented out. But then some vertex in the path has two outgoing edges, a contradiction. In the case where $i \to k$ and $l \to j$, the path from i to j is directed. The vertices along this path alternate between two sequences; one is decreasing to c and the other is increasing to c. Thus, if the distance between k and l is even then they are both larger/smaller than c then i and j are on the other side and in the opposite order to that of k and l, i.e. $i > j$. If the distance is odd then k and l are on different sides of c and so are i and j but opposite of k and l, i.e. $i > j$. The case where $k \to i$ and $j \to l$ is identical.

This is the only tree where the degree is an alternating labeling of the vertices. Let T' be a tree in $\mathcal{T}_\pi$ where the degree is an alternating labeling of the vertices. By Lemma 26 there exists f' a complete refinement of the degree sequence that is an alternating labeling of the vertices of T'. Without loss of generality we may assume f' has range $[n]$ and we may further assume the vertex set for T' is $[n]$ and this is a canonical labeling. We would like to show that the identity map $\phi : V(T') \to V(T)$ is an isomorphism. Since f and f' are both refinements of π, ϕ preserves degrees. We will show $N^c_T(k) = N^c_{T'}(k)$ for every vertex $k \in [n]$. Once we have shown this for a vertex we will say it is *verified* (and if not we say it is *unverified*). Let $A \subseteq [n]$ be the set of verified vertices; we start with $A = \emptyset$. At each stage we will try to verify either the largest unverified vertex greater than c or smallest unverified vertex less than c, say k. There will always be a choice such that $N^c_T(k) \subseteq A$ (if not all of the children of both the smallest and largest unverified vertex are verified then they could not be the largest and smallest by the construction of T). Suppose $N^c_T(k) \neq N^c_{T'}(k)$ then there exists $i \in N^c_{T'}(k) \setminus N^c_T(k)$ and $j \in N^c_T(k) \setminus N^c_{T'}(k)$. Consider the path between i and j in T'. Then k is the vertex adjacent to i in this path. We let l be the vertex adjacent to j in this path. At this point l cannot be verified because its child in T', j, has a different predecessor in T. In the case where k is the largest unverified vertex greater than c, we have $k > l$; also $i > j$ since per the construction of T the children of k in T are the smallest vertices without full degree in $T[A]$. In the case where k is the smallest unverified vertex less than c, we have $k < l$; also $i < j$ since per the construction of T the children of k in T are the largest vertices without full degree in $T[A]$. In both cases this prevents f' from being an alternating labeling of T'. Thus, the children of k must be the same in T and T'. We need not verify c since its neighborhood is implicit once everything else is verified. Therefore, $T \simeq T'$.

□

3.3.4 Conclusion

This result will be used as follows. For an invariant σ we will show that the extremal tree that maximizes σ in $\mathcal{T}_\pi$ will have a direct or an alternating labeling that is a refinement of the degree. Then by the following two immediate corollaries the extremal tree must be the ball or the festoon.

Corollary 31. *If a tree T with degree sequence π has a direct labeling f that is a refinement of the degree then $T \simeq B_\pi$.*

Corollary 32. *If a tree T with degree sequence π has an alternating labeling f that is a refinement of the degree then $T \simeq F_\pi$.*

To get an even stronger result consider the following definition.

Definition 29. Let π and π' be degree sequences written in non increasing order. We say π *majorizes* π' written $\pi \rhd \pi'$ if $\sum_{i=1}^{k} \pi(i) \geq \sum_{i=1}^{k} \pi'(i)$ for $k = 1, 2, \ldots, n-1$ and $\sum_{i=1}^{n} \pi(i) = \sum_{i=1}^{n} \pi'(i)$.

This defines a partial order on degree sequences. We say σ has a majorization result if whenever $\pi \rhd \pi'$, the extremal tree for π gives a larger value for σ than the extremal tree for π' (or suitably modified when we want to minimize σ).

In $\mathcal{T}_n$, all trees on n vertices, every degree sequence is majorized by

$$\pi_1 = (n-1, 1, 1, \ldots, 1).$$

Both the ball and festoon with this degree sequence are the star $K_{1,n-1}$.

In $\mathcal{T}^*_{n,k}$, trees on n vertices and exactly k leafs, every degree sequence is majorized by

$$\pi_2 = (k, \underbrace{2, 2, \ldots, 2}_{n-k-1}, \underbrace{1, 1, \ldots, 1}_{k}).$$

The ball with this degree sequence is the almost equally subdivided star with k leafs. The festoon with this degree sequence is the broom.

In $\mathcal{T}_{n,d}$, trees on n vertices and degree at most $d+1$, every degree sequence is majorized by

$$\pi_3 = (\underbrace{d+1, d+1, \ldots, d+1}_{\lfloor \frac{n-2}{d} \rfloor}, r, 1, 1, \ldots, 1)$$

Here we have $B_{\pi_3} = B_{n,d}$ and $F_{\pi_3} = F_{n,d}$.

Therefore, given an extremal tree result for an invariant σ in $\mathcal{T}_\pi$ and a majorization result (this is the case for every invariant we consider) we can determine which tree any any of the above families of trees maximize σ. Hence, these results are strictly stronger than all previous results for these families. The same is true for any family of trees that is determined by its degree sequence and has a unique maximal element in the majorization order.

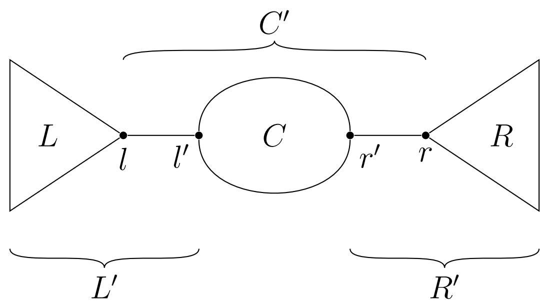


Figure 3.7: Decomposition of a tree T

3.3.5 Applications

In all of the following examples let T be a tree and define the following subtrees. Let l' and r' be distinct vertices of T and let $l \in N_T^{r'}(l')$ and $r \in N_T^{l'}(r')$. There are three components of $T - ll' - rr'$; let L be the one containing l, R be the one containing r, and C the remaining component which will contain l' and r' (see Figure 3.7). Let $L' = L + ll'$, $R' = R + rr'$, and $C' = C + ll' + rr'$. The tree $T_{L\leftrightarrow R} = T - ll' - rr' + lr' + rl'$ is the tree T with the branches L and R *switched*. The tree $T_{L\to r'} = T - ll' + lr'$ is the tree T with the branch L *shifted* to r'. Shifting can be thought of as switching with an empty branch. It is important to note that switching preserves all degrees and shifting preserves all degrees but that of l' and r' for which $d_{T_{L\to r'}}(l') = d_T(l') - 1$ and $d_{T_{L\to r'}}(r') = d_T(r') + 1$.

For each example below we wish to maximize some invariant σ of trees. For each invariant we will give a labeling f_T of each tree T. This labeling will be related to σ in such a way that when $T \in \mathcal{T}_\pi$ has $\sigma(T)$ maximum the labeling f_T of T is direct or alternating. We will show this by comparing $\sigma(T)$ and $\sigma(T_{L\leftrightarrow R})$; by assumption $\sigma(T) - \sigma(T_{L\leftrightarrow R}) \geq 0$.

Next, at the same time we will show that the labeling is a refinement of the degree and also give a majorization result. We do this by comparing $\sigma(T)$ and $\sigma(T_{L\to r'})$ for an arbitrary tree T. In general we will show that if $f_T(l') \leq f_T(r')$ then $\sigma(T_{L\to r'}) > \sigma(T)$

and moreover $f_{T_{L\to r'}}(l') < f_{T_{L\to r'}}(r')$. Hence, if $d(l') > d(r')$ then we could shift $d(l') - d(r')$ branches of l' (relative to r') to r' and have a tree with the same degree sequence and a larger value of σ; so f_T is a refinement of degree for extremal trees T. Furthermore, this shows that if $\pi \triangleright \pi'$ then the extremal tree for σ with degree sequence π has larger σ value than the extremal tree for π'.

Therefore, by the previous section if $T \in \mathcal{T}_\pi$ maximizes σ then $T \simeq B_\pi$ or $T \simeq F_\pi$ and if $T \in \mathcal{T}_{n,d}$ maximizes σ then $T \simeq B_{n,d}$ or $T \simeq F_{n,d}$. The argument is similar for minimizing σ, however, in general only one argument will apply per invariant.

Homomorphisms

Let H be a strongly biregular graph with partition A, B; representatives $a \in A, b \in B$; and degrees $\alpha = d_A(a), \beta = d_B(a), \gamma = d_A(b), \delta = d_B(b)$. Without loss of generality we will assume that $\alpha + \beta < \gamma + \delta$, else we can switch A and B. Define the quantities

$$\begin{aligned}
h^T_{XY}(u,v) &= |\{f \in \mathrm{Hom}(T,H) : f(u) \in X, f(v) \in Y\}| \\
h^T_{xy}(u,v) &= |\{f \in \mathrm{Hom}(L,H) : f(u) = x, f(v) = y\}| \\
h^T_X(v) &= |\{f \in \mathrm{Hom}(T,H) : f(v) \in X\}| \\
h^T_x(v) &= |\{f \in \mathrm{Hom}(L,H) : f(v) = x\}|
\end{aligned}$$

Lemma 33. *For any tree T and vertex v we have $h^T_a(v) \le h^T_b(v)$.*

Proof. We will prove this by induction on the height of T. In the base case of height zero then $h^T_a(v) = h^T_b(v) = 1$. Now suppose T has height k and the claim is true for all trees of height less than k. Let $v_1, v_2, \ldots, v_d$ be the neighbors of v and let $T_i = T^v_{v_i}$.

Then,

$$\frac{h_a^T(v)}{h_b^T(v)} = \frac{\prod_{i=1}^d(\alpha h_a^{T_i}(v_i) + \beta h_b^{T_i}(v_i))}{\prod_{i=1}^d(\gamma h_a^{T_i}(v_i) + \delta h_b^{T_i}(v_i))}$$
$$= \prod_{i=1}^d \frac{\alpha \frac{h_a^{T_i}(v_i)}{h_b^{T_i}(v_i)} + \beta}{\gamma \frac{h_a^{T_i}(v_i)}{h_b^{T_i}(v_i)} + \delta}.$$

Let $g(x) = \frac{\alpha x+\beta}{\gamma x+\delta}$. It is easy to verify that $g((0,1]) \subset (0,1)$; clearly $g(x) > 0$ for $x \in (0,1]$ and $g(x) < 1$ for $x \in (\frac{-\delta}{\gamma}, \frac{\delta-\beta}{\alpha-\gamma})$ a superset of $(0,1]$. By the induction hypothesis $\frac{h_a^{T_i}(v_i)}{h_b^{T_i}(v_i)} \in (0,1]$ thus $\frac{h_a^T(v)}{h_b^T(v)} = \prod_{i=1}^d g\left(\frac{h_a^{T_i}(v_i)}{h_b^{T_i}(v_i)}\right) \in (0,1)$. Therefore, $h_a^T(v) \leq h_b^T(v)$ and strict inequality if T is not a single vertex. □

Theorem 34. *Suppose that T maximizes* $\hom(\cdot, H)$ *in $\mathcal{T}_\pi$. If $\alpha\delta - \beta\gamma > 0$ then h_B^T is a direct labeling of T refining the degree. If $\alpha\delta - \beta\gamma < 0$ then h_B^T is an alternating labeling of T refining the degree. Furthermore, we have a majorization result for* $\hom(\cdot, H)$.

Proof. We will use the notation of Section 3.3.5. With the additional convention that $D' = L \cup R$ and $D = L' \cup R'$. Suppose $T \in \mathcal{T}_\pi$ maximizes $\hom(\cdot, H)$. Consider the following differences and easily verifiable factored forms.

$$\begin{aligned} \hom(T,H) - \hom(T_{L\leftrightarrow R}, H) = &\left(h_{BA}^C(l',r') - h_{AB}^C(l',r')\right) \\ &\cdot \left(h_{ba}^D(l',r') - h_{ab}^D(l',r')\right) \end{aligned} \tag{3.4}$$

$$h_B^T(l') - h_B^T(r') = h_{BA}^C(l',r')\, h_{ba}^D(l',r') - h_{AB}^C(l',r')\, h_{ab}^D(l',r') \tag{3.5}$$

$$h_{BA}^{C'}(l,r) - h_{AB}^{C'}(l,r) = \left(h_{BA}^C(l',r') - h_{AB}^C(l',r')\right)(\alpha\delta - \beta\gamma) \tag{3.6}$$

$$h_{ba}^D(l',r') - h_{ab}^D(l',r') = \left(h_{ba}^{D'}(l,r) - h_{ab}^{D'}(l,r)\right)(\alpha\delta - \beta\gamma) \tag{3.7}$$

$$h_B^T(l) - h_B^T(r) = h_{BA}^{C'}(l,r)\, h_{ba}^{D'}(l,r) - h_{AB}^{C'}(l,r)\, h_{ab}^{D'}(l,r) \tag{3.8}$$

For instance, in (3.6) we condition on what set l' and r' are sent to in each homomorphism to get:

$$\begin{aligned} h_{BA}^{C'}(l,r) - h_{AB}^{C'}(l,r) &= \beta\alpha h_{AA}^{C}(l',r') + \beta\gamma h_{AB}^{C}(l',r') \\ &\quad + \delta\alpha h_{BA}^{C}(l',r') + \delta\gamma h_{BB}^{C}(l',r') \\ &\quad - \alpha\beta h_{AA}^{C}(l',r') - \alpha\delta h_{AB}^{C}(l',r') \\ &\quad - \gamma\beta h_{BA}^{C}(l',r') - \gamma\delta h_{BB}^{C}(l',r') \\ &= \left(h_{BA}^{C}(l',r') - h_{AB}^{C}(l',r')\right)(\alpha\delta - \beta\gamma). \end{aligned}$$

Since $T_{L\leftrightarrow R} \in \mathcal{T}_\pi$ and T maximizes $\hom(\cdot, H)$ in $\mathcal{T}_\pi$ we have that (3.4) is non negative. Suppose (3.5) is non negative, then we must have (3.6) and (3.7) non negative. Therefore, (3.8) is non negative if $\alpha\delta - \beta\gamma > 0$ and non positive if $\alpha\delta - \beta\gamma < 0$. In the former case h_B^T is a direct labeling of T, in the latter it is an alternating labeling of T.

We would now like to show that if $h_B^T(l') \leq h_B^T(r')$ then $d(l') \leq d(r')$. We will do so by showing that we can shift L to r' to increase $\hom(\cdot, H)$.

$$\begin{aligned} \hom(T,H) - \hom(T_{L\to r'},H) = {}& (h_a^{L'}(l') - h_b^{L'}(l')) \\ & \cdot (h_{AB}^{T-L}(l',r') - h_{BA}^{T-L}(l',r')) \end{aligned} \tag{3.9}$$

$$h_B^T(l') - h_B^T(r') = h_b^{L'}(l') h_{BA}^{T-L}(l',r') - h_a^{L'}(l')(h_{AB}^{T-L}(l',r') \tag{3.10}$$

If $\hom(T,H) \geq \hom(T_{L\to r'},H)$ then (3.9) is non negative. By Lemma 33, $h_a^{L'}(l') < h_b^{L'}(l')$ and hence $h_{AB}^{T-L}(l',r') \leq h_{BA}^{T-L}(l',r')$. But then (3.10) is strictly positive, a contradiction. Therefore $\hom(T,H) < \hom(T_{L\to r'},H)$. Note that

$$h_B^{T_{L\to r'}}(l') - h_B^{T_{L\to r'}}(r') = h_a^{L'}(l')h_{BA}^{T-L}(l',r') \\ - h_b^{L'}(l')(h_{AB}^{T-L}(l',r')$$

is strictly negative since $h_a^{L'}(l') < h_b^{L'}(l')$ and $h_{AB}^{T-L}(l',r') > h_{BA}^{T-L}(l',r')$ by (3.9). Therefore $\hom(\cdot,H)$ is a refinement of the degree and we get a majorization result for free. $\square$

Number of Independent Sets

The number of independent sets of T is $\hom(T,H)$ where H is given in Figure 3.8

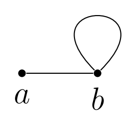


Figure 3.8: The graph H

Note that $\alpha\delta - \beta\gamma = 0\cdot 1 - 1\cdot 1 < 0$ and so by Theorem 34 $i_0(v)$, the number of independent sets of T not containing v, is an alternating labeling refining the degree of the tree that maximizes the number of independent sets in $\mathcal{T}_\pi$. Therefore, the number of independent sets for trees in $\mathcal{T}_\pi$ is maximized by the festoon F_π.

Matching Generating Polynomial

The matching generating polynomial for a graph G is the polynomial

$$M(G,\lambda) = \sum_{k\geq 0} m(G,k)\lambda^k$$

where $m(G,k)$ is the number of matchings in G of exactly k edges. We will show that in $\mathcal{T}_\pi$ the matching generating polynomial is minimized for all $\lambda > 0$ by the festoon

F_π. Note that $m(G, 0) = 1$ for all graphs so $M(G, \lambda) > 0$ for all $\lambda > 0$.

Theorem 35. *Suppose that T minimizes $M(\cdot, \lambda)$ in $\mathcal{T}_\pi$ for some $\lambda > 0$. Then $f_T(v) = M(T, \lambda) - M(T - v, \lambda)$ (counting matchings that contain v) is an alternating labeling of T refining the degree. Furthermore, we have a majorization result for $M(\cdot, \lambda)$.*

Proof. We will use the notation of Section 3.3.5. With the additional convention that $D' = L \cup R$. Fix $\lambda > 0$ and suppose $T \in \mathcal{T}_\pi$ minimizes $M(\cdot, \lambda)$. Consider the following differences and easily verifiable factored forms.

$$\begin{aligned} M(T, \lambda) - M(T_{L\leftrightarrow R}, \lambda) &= \lambda\big(M(D' - l, \lambda) - M(D' - r, \lambda)\big) \\ &\quad \cdot \big(M(C - l', \lambda) - M(C - r', \lambda)\big) \end{aligned} \tag{3.11}$$

$$\begin{aligned} M(T - l', \lambda) - M(T - r', \lambda) &= \lambda M(C - l' - r', \lambda)\big(M(D' - r, \lambda) - M(D' - l, \lambda)\big) \\ &\quad + M(D', \lambda)\big(M(C - l', \lambda) - M(C - r', \lambda)\big) \end{aligned} \tag{3.12}$$

$$\begin{aligned} M(T - l, \lambda) - M(T - r, \lambda) &= \lambda M(D' - l - r, \lambda)\big(M(C - r', \lambda) - M(C - l', \lambda)\big) \\ &\quad + M(C, \lambda)\big(M(D' - l, \lambda) - M(D' - r, \lambda)\big) \end{aligned} \tag{3.13}$$

Since $T_{L\leftrightarrow R} \in \mathcal{T}_\pi$ and T minimizes $M(\cdot, \lambda)$ in $\mathcal{T}_\pi$ we have that (3.11) is less than or equal to zero (so one of its factors is non negative and the other is non positive). Suppose $f_T(l') \leq f_T(r')$, then (3.12) is non negative. Therefore, (3.13) is non positive and we have $f_T(l) \geq f_T(r)$.

We would now like to show that if $f_T(l') \leq f_T(r')$ then $d(l') \leq d(r')$. We will do so by showing that we can shift L to r' to increase $M(\cdot,\lambda)$.

$$M(T,\lambda) - M(T_{L\to r'},\lambda) = \lambda M(L-l,\lambda)\big(M(T-L-l',\lambda) - M(T-L-r',\lambda)\big) \tag{3.14}$$

$$\begin{aligned} M(T-l',\lambda) - M(T-r',\lambda) = M(L,\lambda)\big(&M(T-L-l',\lambda) - M(T-L-r',\lambda)\big) \\ &- \lambda M(L-l,\lambda)M(T-L-l'-r',\lambda) \end{aligned} \tag{3.15}$$

Suppose $f_T(l') \leq f_T(r')$, then (3.15) is non negative. Therefore, $M(T-L-l',\lambda) - M(T-L-r',\lambda) > 0$ and hence $M(T,\lambda) > M(T_{L\to r'},\lambda)$. Note that

$$\begin{aligned} M(T_{L\to r'}-l',\lambda) - M(T_{L\to r'}-r',\lambda) = M(L,\lambda)\big(&M(T-L-l',\lambda) - M(T-L-r',\lambda)\big) \\ &+ \lambda M(L-l,\lambda)M(T-L-l'-r',\lambda) \end{aligned}$$

is strictly positive since (3.15) is non negative. Hence, $f_{T_{L\to r'}}(l') < f_{T_{L\to r'}}(r')$. Therefore, f_T is a refinement of the degree and we get a majorization result for free. $\square$

Energy

The energy of a graph is the sum of the magnitudes of its eigenvalues. Formally, if the spectrum of G (the spectrum of its adjacency matrix) is $\{\lambda_1,\ldots,\lambda_n\}$ then the energy of G is

$$E(G) = \sum_{i=1}^{n} |\lambda_i|.$$

For trees we can compute the energy with the Coulson integral

$$E(T) = \frac{2}{\pi}\int_0^\infty x^{-2}\log\Big(M\big(T,x^2\big)\Big)\,dx.$$

Therefore, since in $\mathcal{T}_\pi$ the festoon F_π minimizes $M(\cdot,\lambda)$ for all $\lambda > 0$ it must also minimize energy. It is also clear that we have a majorization result.

Number of Subtrees

Let $s(T)$ be the number of subtrees of T. Define the auxiliary quantity $s(T,v)$ to be the number of subtrees of T that contain the vertex v.

Theorem 36. *If T maximizes s in $\mathcal{T}_\pi$ then $s(T,\cdot)$ is a direct labeling of T.*

Proof. We will use the notation of Section 3.3.5. Suppose $T \in \mathcal{T}_\pi$ maximizes s. Consider the following differences and easily verifiable alternate forms.

$$s(T) - s(T_{L\leftrightarrow R}) = (s(L,l) - s(R,r))(s(C,l') - s(C,r')) \tag{3.16}$$

$$s(T,l') - s(T,r') = s(L',l')s(C-r',l') - s(R',r')s(C-l',r') \tag{3.17}$$

$$s(C-r',l') - s(C-l',r') = s(C,l') - s(C,r') = s(C'-r,l) - s(C'-l,r) \tag{3.18}$$

$$s(L',l') - s(R',r') = s(L,l) - s(R,r) \tag{3.19}$$

$$s(T,l) - s(T,r) = s(L,l)s(C'-r,l) - s(R,r)s(C'-l,r) \tag{3.20}$$

Since $T_{L\leftrightarrow R} \in \mathcal{T}_\pi$ and T maximizes s in $\mathcal{T}_\pi$ we have that (3.16) is non negative. Suppose (3.17) is non negative, then we must have (3.18) and (3.19) non negative. Therefore, equation (3.20) is non negative and $s(T,\cdot)$ is a direct labeling of T.

We would now like to show that if $s(T,l') \le s(T,r')$ then $d(l') \le d(r')$. We will do so by showing that we can shift L to r' to increase s.

$$s(T) - s(T_{L\to r'}) = s(L,l)(s(T-L,l') - s(T-L,r')) \tag{3.21}$$

$$s(T,l') - s(t,r') = s(L',l')s(T-L,l') - s(L,l)s(T-L,\{r',l'\}) - s(T-L,r') \tag{3.22}$$

If $s(T) \geq s(T_{L\to r'})$ then (3.21) is non negative and hence $s(T-L,l') \geq s(T-L,r')$. Notice that (3.22) is greater than $s(L',l')(s(T-L,l')-s(T-L,r'))$. But then (3.22) is strictly positive, a contradiction. Therefore $s(T) < s(T_{L\to r'})$. Note that

$$\begin{aligned} s(T_{L\to r'},l') - s(T_{L\to r'},r') &= s(T-L,l') + s(T-L,\{l',r'\})s(L,l) - s(T-L,r')s(L',l') \\ &< s(L',l')(s(T-L,l')-s(T-L,r')) \end{aligned}$$

Since (3.21) is negative so is $s(T_{L\to r'},l') - s(T_{L\to r'},r')$ i.e., $s(T_{L\to r'},l') < s(T_{L\to r'},r')$. Therefore, s is a refinement of the degree and we get a majorization result for free. $\square$

Wiener Index

The Wiener index is

$$W(T) = \sum_{u\sim v} d(u,v).$$

Define the auxiliary quantities

$$d(T,v) = \sum_{u\in V(T)} d(u,v)$$

and note that

$$W(T) = \frac{1}{2}\sum_{v\in V(T)} d(T,v).$$

We will show that $f_T(v) = W(T) - d(T,v)$ is a direct labeling of the vertices of the tree $T \in \mathcal{T}_\pi$ which minimizes $W(T)$. Furthermore, f_T refines the degree and we have a majorization result for W. We will do so with the analogous arguments for $d(T,\cdot)$.

Theorem 37. *Suppose that T has minimum Wiener index in $\mathcal{T}_\pi$. Let l' and r' be distinct vertices in T and $l \in N^{r'}(l')$ and $r \in N^{l'}(r')$. If $d(T,l') \geq d(T,r')$ then*

$d(T,l) \geq d(T,r)$ and $d(l') \leq d(r')$. Furthermore, we have a majorization result for W.

Proof. We will use the notation of Section 3.3.5. Consider the quantity $W(T) - W(T_{L\leftrightarrow R})$. The distances between pairs of vertices in $L \cup R$ are the same in both trees so they cancel. Similarly, the distances between pairs of vertices in C cancel. Therefore, counting the number of edges in paths between other pairs we have

$$\begin{aligned} W(T) - W(T_{L\leftrightarrow R}) &= |C|d(L',l') + |L|d(C,l') + |C|d(R',r') + |R|d(C,r') \\ &\quad - |C|d(L',l') - |L|d(C,r') - |C|d(R',r') - |R|d(C,r') \\ &= (d(C,l') - d(C,r'))(|L| - |R|). \end{aligned} \tag{3.23}$$

Note that $W(T) - W(T_{L\leftrightarrow R}) \leq 0$ since $W(T)$ is minimum. Therefore, either $d(C,l') \geq d(C,r')$ and $|R| \geq |L|$ or $d(C,l') \leq d(C,r')$ and $|R| \leq |L|$. Additionally, consider the quantities

$$\begin{aligned} d(T,l') - d(T,r') &= d(L',l') + d(C,l') + d(l',r')|R| + d(R',r') \\ &\quad - d(R',r') - d(C,r') - d(l',r')|L| - d(L',l') \\ &= d(C,l') - d(C,r') + d(l',r')(|R| - |L|) \end{aligned} \tag{3.24}$$

$$d(T,l) - d(T,l') = |C| + |R| - |L| \tag{3.25}$$

$$d(T,r) - d(T,r') = |C| + |L| - |R| \tag{3.26}$$

$$d(T,l) - d(T,r) = d(T,l') - d(T,r') + 2(|R| - |L|). \tag{3.27}$$

By assumption (3.24) is non negative and thus by the above we have $d(C,l') \geq d(C,r')$ and $|R| \geq |L|$. Therefore, (3.27) is non negative and $d(T,\cdot)$ is a direct labeling of the vertices of T. By a simple transformation so is $W(T) - d(T,\cdot)$.

We would now like to show that if $d(T,l') \geq d(T,r')$ then $d(l') \leq d(r')$. We will do so by showing that we can shift L to r' to decrease the Wiener index.

$$\begin{aligned} W(T) - W(T_{L\to r'}) &= |T-L|d(L',l') + |L|d(T-L,l') \\ &\quad - |T-L|d(L',l') - |L|d(T-L,r') \\ &= (d(T-L,l') - d(T-L,r'))|L| \end{aligned} \tag{3.28}$$

$$\begin{aligned} d(T,l') - d(T,r') &= d(L',l') + d(T-L,l') \\ &\quad - d(T-L,r') - d(l',r')|L| - d(L',l') \\ &= d(T-L,l') - d(T-L,r') - d(l',r')|L| \end{aligned} \tag{3.29}$$

If $W(T_{L\to r'}) \geq W(T)$ then (3.28) and hence (3.29) are non positive; but this contradicts the assumption that $d(T,l') \geq d(T,r')$, therefore $W(T_{L\to r'}) < W(T)$. Note that

$$\begin{aligned} d(T_{L\to r'},l') - d(T_{L\to r'},r') &= d(L',l') + d(l',r')|L| + d(T-L,l') \\ &\quad - d(T-L,r') - d(L',l') \\ &= d(T-L,l') - d(T-L,r') + d(l',r')|L| \end{aligned} \tag{3.30}$$

is strictly positive. Therefore $W(T) - d(T,\cdot)$ is a refinement of the degree and we get a majorization result for free. □

3.4 Open Problems

Knowing what trees in $\mathcal{T}_{n,d}$ and $\mathcal{T}_\pi$ maximize the number of homomorphisms to any strongly biregular graph, a natural question is what happens with more complex graphs. For instance consider the graphs H_1 and H_2 (see Figure 3.9). The number of homomorphisms to the graph H_1 appears to be maximized by the festoon (checked by computer for all trees of 75 vertices or less). However, the number of homomorphisms to the similar graph H_2 is not always maximized by the ball or festoon. Can we characterize the graphs that have the number of homomorphisms to them maximized by the festoon or the ball?

(a) H_1 (b) H_2

Figure 3.9: Two graphs

Chapter 4

Reconstruction

4.1 Discussion

A reconstruction problem asks us to recover a combinatorial object from partial information about it. We will explore two such problems in this chapter. First we will consider reconstructing trees from their Wiener matrix and other generalizations. Second we will consider reconstructing a graph of girth at least $2r+3$ from metric balls of radius r.

4.2 Trees from Matrices

4.2.1 Introduction

The Wiener index of a graph, introduced by Wiener in 1947, is the total of the distances between every pair of vertices. Wiener showed that for certain types of molecules his index correlated well with their physical properties [21]. An easy method of computing the Wiener index for a tree is to take a sum $\sum_e w_e$ over all edges e, where w_e is the product of the sizes of the components of $T - e$. It is easy to check that w_e equals the number of paths in T containing e.

Definition 30. Let T be a tree on n vertices. The *Wiener matrix* of T is the $n \times n$ matrix $W = (w_{ij})$ such that

$$w_{ij} = \begin{cases} \text{number of paths in } T \text{ containing } i \text{ and } j & i \neq j \\ 0 & i = j. \end{cases}$$

There is a strong connection between the Wiener index and the Wiener matrix, the Wiener index is half the sum of the entries of W corresponding to adjacent vertices. The hyper-Wiener index (WW) is half the sum of all entries of W.

Conjecture 38 (Randić, Guo, Oxley, Krishnapriyan, and Naylor [16])**.** *An entry in W is the largest in its row or column if and only if the corresponding vertices are adjacent.*

In other words a tree can be reconstructed from its Wiener matrix. We will prove this conjecture.

A natural generalization of this is to subtrees of bounded degree. Fix $k \geq 2$. Let $\mathcal{S}_k(T)$ be the collection of subtrees of T with maximum degree at most k.

Definition 31. Let T be a tree on n vertices. The k-subtree matrix of T is the $n \times n$ matrix $S_k = (s_{ij})$ such that

$$s_{ij} = \begin{cases} |\{T' \in \mathcal{S}_k(T) : i, j \in V(T')\}| & i \neq j \\ 0 & i = j. \end{cases}$$

Remark. When $k = 2$ this is the Wiener matrix.

We will show that we can use the same method to reconstruct from the the k-subtree matrix.

Theorem 39. *An entry in S_k is the largest in its row or column if and only if the corresponding vertices are adjacent.*

4.2.2 Result

We begin by defining the following quantity.

Definition 32. For a tree T with root r, let

$$\mu_k(T) = |\{T' \in \mathcal{S}_k(T) : r \in V(T'), d_{T'}(r) \leq k-1\}|.$$

Remark. When $k = 2$ this is the number of paths in T starting at r so $\mu_2(T) = |T|$. It may be helpful to the reader to first consider this case when reading.

Lemma 40. *If $uv \in E(T)$ consider the components of $T - uv$. Let $U = T_u^v$ be the component containing u and $V = T_v^u$ the component containing v. Then,*

$$s_{uv} = \mu_k(U)\mu_k(V).$$

Proof. Any subtree in $\mathcal{S}_k(T)$ can be broken up into a subtree in $S_k(U)$ and a subtree in $S_k(V)$ each with the degree at the root strictly less than k. In reverse any such pair of subtrees can be extended (with the edge uv) to be a subtree in $\mathcal{S}_k(T)$. □

Theorem 41. *An entry in S_k is the largest in its row or column if and only if the corresponding vertices are adjacent.*

Proof. For ease of notation we will write $\mu(A) = \mu_k(A)$.

($\Rightarrow$) Suppose s_{uv} is the largest entry in its row or column and $uv \notin E(T)$. Let w be any vertex on the path between u and v. Everything in $\mathcal{S}_k$ that contains both u and v also must contain w. Furthermore, the path between u and w is in $\mathcal{S}_k$ but

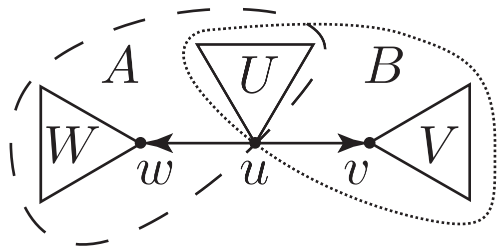


Figure 4.1: Case 1.

does not contain v. Therefore $s_{uw} > s_{uv}$ and similarly $s_{wv} > s_{uv}$, a contradiction, so $uv \in E(T)$.

($\Leftarrow$) Suppose $uv \in E(T)$. We would like to show that s_{uv} is the largest in its row or column. Direct the edges of T as follows. For an edge ab consider the components of $T - ab$. Let $A = T_a^b$ be the component containing a and $B = T_b^a$ the component containing b. Direct,

$$a \to b \text{ if } \mu(A) \leq \mu(B),$$

$$a \leftarrow b \text{ if } \mu(A) \geq \mu(B).$$

Without loss of generality, suppose $u \to v$. We will show s_{uv} is the largest entry in its row. (If $v \to u$ then s_{uv} is the largest entry in its column.)

Suppose there was some vertex $w \neq v$ such that $u \to w$. Then we have $w \leftarrow u \to v$ as in Figure 4.1. Using the labeling of the figure we have the following inequalities.

$$\mu(V) \geq \mu(A) > \mu(W)$$

$$\mu(W) \geq \mu(B) > \mu(V)$$

But this is a contradiction.

Now we would like to show s_{uv} is the largest entry in its row. We need only

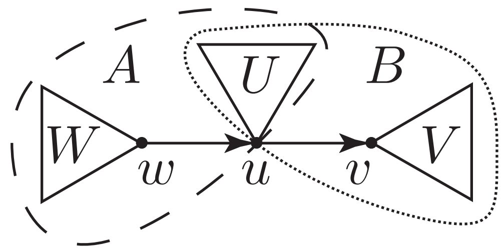


Figure 4.2: Case 2.

consider the neighbors of u by ($\Rightarrow$). Let $w \in N(u) - v$. By the claim we have $w \to u \to v$ as in Figure 4.2. Conditioning on whether a subtree contains w we get the following counts.

$$s_{uv} = s_{wv} + \mu(U)\mu(V)$$
$$s_{uw} = s_{wv} + \mu(U)\mu(W)$$

But $\mu(W) < \mu(A) \leq \mu(V)$, so $s_{uv} > s_{uw}$. $\square$

The sinks of this orientation are a generalization of the centroid.

Definition 33. A vertex is in the *k-subtree centroid* of a tree T if it is a sink of the orientation in Theorem 41. Alternatively, a vertex is in the k-subtree centroid of a tree T if it minimizes the maximum μ-value of its branches.

Remark. The 2-subtree centroid is the centroid. We will discuss this topic and the interpretation of the k-subtree centroid for $k \geq \Delta(T)$ in Section 4.2.3.

It is easy to construct trees where the k-subtree centroid and the centroid are different for $k \geq 3$. The k-subtree centroid takes into consideration branching whereas the centroid does not. See Figure 4.3 for an example.

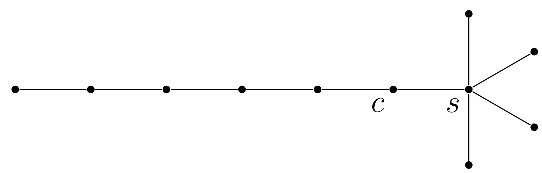


Figure 4.3: In this graph c is the centroid and s is the k-subtree centroid for $k \geq 3$.

Theorem 42. *The k-subtree centroid of a tree is either one vertex or two adjacent vertices.*

Proof. Suppose that there are two non adjacent members of the k-subtree centroid. Then on the path between these two vertices is another vertex u with two arrows directed out. But this case is specifically excluded in the proof of Theorem 41. Therefore, the k-subtree centroid is either one vertex or two adjacent vertices. Furthermore, this proves that all edges are oriented towards the k-subtree centroid. □

4.2.3 Special Cases

Wiener Matrix

When $k = 2$ the k-subtree matrix is the Wiener matrix and the sink(s) of the orientation in Theorem 41 are the vertices in the centroid.

Definition 34. The *centroid* of a tree T is the set of vertices v such that the components of $T - v$ all have size at most $n/2$, where $n = |T|$. Alternatively, the centroid is the set of vertices v that minimize the maximum size of a component of $T - v$.

A well known fact about centroids is they have only one or two adjacent vertices [17]. We include an independent proof of this fact for completeness.

Theorem 43. *The centroid of a tree is either one vertex or two adjacent vertices.*

Proof. Start with an arbitrary choice of vertex v. If v is not in the centroid then one of the components of $T-v$ has size strictly larger than $n/2$. Let v' be the neighbor of v in this component. One of the components of $T-v'$ is all but the largest component of $T-v$ along with the vertex v, and must have size at most $n/2$ since the remaining component of $T-v$ has size strictly larger than $n/2$. The remaining components of $T-v'$ are subsets of the largest component of $T-v$ and so are strictly smaller. If v' is not in the centroid we continue in this fashion, at each stage the size of the largest component is strictly less so eventually we have a vertex in the centroid.

Suppose the centroid of T contains two vertices v, w. Then by definition, the component C of $T-v$ containing w has size at most $n/2$ and the component D of $T-w$ containing v has size at most $n/2$. But $C \cup D = T$ so $|T| = n - |C \cap D|$. Thus $C \cap D$ must be empty and therefore $v \sim w$. □

Recall that $\mu_2(A) = |A|$ and so the the orientation of vw in Theorem 41 is towards the larger branch component of $T - vw$.

Corollary 44. *The centroid is the* 2*-subtree centroid.*

Proof. The size of the components of $T-vw$ for any edge vw add up to $n = |T|$. Thus, one component has size at least $n/2$ and the other has size at most $n/2$. First let us consider the case where these sizes are not the same and thus the inequalities are strict. Without loss of generality suppose the component $W = T_w^v$ of $T - vw$ containing w has size greater than $n/2$. We would like to show that the size of the largest component of $T-w$ is at less than the size of the largest component of $T-v$. One of the components of $T-v$ is W and thus the sum of the sizes of the remaining components of $T-v$ is $n-|W|-1$ which is less than $n/2-1$. One of the components of $T-w$ is $T-W$ which has size $n-|W|$ which is less than $n/2$, and the remaining components are strict subsets of W and so have size less than $|W|$. In the case where the sizes of the components of $T - vw$ are equal (to $n/2$) the size of the largest component of both $T - v$ and $T - w$ is $n/2$.

Therefore the orientation provided in Theorem 41 ($v \to w$ if $|T_v^w| \leq |T_w^v|$), of which the 2-centroid is the sink, is the same as the orientation: "$v \to w$ if the size of the largest component of $T - v$ is at most the size of the largest component of $T - w$," the sinks of which are clearly the centroid. Therefore, they have the same sink(s) and the 2-subtree centroid is the centroid. □

Subtree Matrix

When $k > \Delta(T)$ we are simply counting subtrees with no bound on maximum degree. Therefore, call the corresponding matrix the subtree matrix. In the orientation from Theorem 41 the sink(s) are the vertices that are in the maximum number of subtrees. To see this first consider the following lemma. In this section write μ in place of μ_k for $k > \Delta(T)$.

Lemma 45. *Let T be a tree and $a \sim b$ adjacent vertices. Let T^a denote T rooted at a, T^b denote T rooted at b, and set $A = T_a^b$, and $B = T_b^a$. Then*

$$\mu(T^a) > \mu(T^b) \iff \mu(A) > \mu(B)$$
$$\mu(T^a) < \mu(T^b) \iff \mu(A) < \mu(B)$$
$$\mu(T^a) = \mu(T^b) \iff \mu(A) = \mu(B)$$

Proof. Observe that we can compute $\mu(T^a)$ and $\mu(T^b)$ as follows. The quantity $\mu(T^a)$ can be broken down to two quantities: subtrees which contain a but not b, of which there are $\mu(A)$, and subtrees containing both a and b, of which there are $\mu(A)\mu(B)$.

We can do a symmetrical decomposition for $\mu(T^b)$ and so we have:

$$\mu(T^a) = \mu(A) + \mu(A)\mu(B)$$
$$\mu(T^b) = \mu(B) + \mu(A)\mu(B)$$

So the result is clear. □

Remark. There is no analog of this theorem for paths or trees with bounded maximum degree. For example, the number of paths containing a is not $|A| + |A||B|$ since this does not count paths in A through a neither of whose endpoints is a.

Definition 35. A vertex is in the *subtree centroid* of a tree T if it is in at least as many subtrees as any other vertex.

Corollary 46. *The subtree centroid of T is the k-subtree centroid when $k > \Delta(T)$.*

Proof. Direct the edges of the tree as follows: $v \to w$ if w is in at least as many subtrees as v. By Lemma 45 this is the same orientation as in Theorem 41. □

Remark. Moreover the number of subtrees that contain a vertex is strictly increasing as we move closer to the subtree centroid along a directed path.

4.2.4 Open Problems

What other generalizations of this problems can we solve? Can we say anything about the Wiener Matrix of a graph?

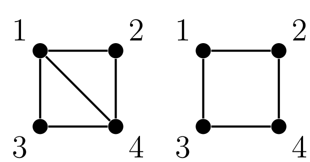
4.3 Graphs from Balls

4.3.1 Introduction

Recently Levenshtein, Konstantinova, Konstantinov, and Molodtsov [13] raised the question of whether a graph can be reconstructed from the function $B_r : V(G) \to \mathcal{P}(V(G))$ mapping $v \mapsto \{w \in V(G) : d_G(v, w) \leq r\}$ which we will call *the metric ball of radius r about v* or simply an *r-ball*. Clearly this will not be possible for all graphs when r is at least 2. A trivial case where reconstruction is impossible would be a graph of small diameter; any graph of diameter at most r will have $B_r(v) = V(G)$ for all vertices v (see Figure 4.4).

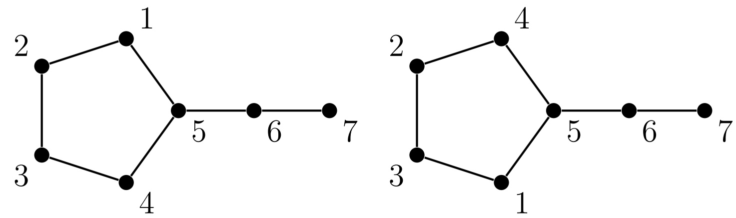
However, large diameter is insufficient if we are allowed small girth. Consider a cycle of length at most $2r+1$ attached to the end of a long path. The vertices on the cycle at the same distance from the path will have the same r-balls so their labels can be swapped to create another graph with the same r-balls (see Figure 4.5 for an example when $r = 2$).

Leafs are another problematic area. When $r \geq 4$ we can not always reconstruct (see Figure 4.6). However, leafs are trivial to recognize from balls in graphs of girth at least $2r+2$ since they are strict subsets of some other ball. Recursively removing leafs we can get to the 2-*core* of the graph and its associated r-balls for which we can then ask the question of reconstruction.



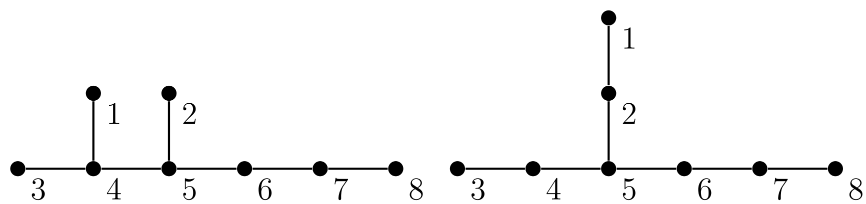
(a) G_1 (b) G_2

Figure 4.4: Both G_1 and G_2 have diameter 2 and hence the same 2-balls



(a) H_1 (b) H_2

Figure 4.5: Both H_1 and H_2 have the same 2-balls; note the labels of 1 and 4 are switched.



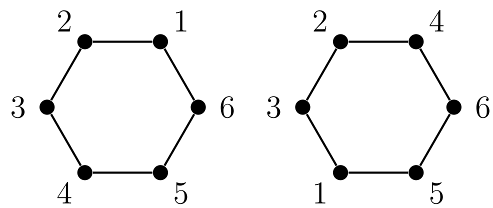
(a) T_1 (b) T_2

Figure 4.6: Both T_1 and T_2 have the same 4-balls

Definition 36. A graph $G \in \mathcal{F}$ is *reconstructable* (in $\mathcal{F}$) from its r-balls if for any $G' \in \mathcal{F}$ with the same family of r-balls as G we have $G = G'$. Similarly, $G \in \mathcal{F}$ is *reconstructable up to isomorphism* (in $\mathcal{F}$) if for any $G' \in \mathcal{F}$ with the same family of r-balls as G we have $G \simeq G'$.

Example 5. C_{2r+2}, the cycle on $2r+2$ vertices, is *not* reconstructible from its r-balls—the same graph with the labels of two antipodal vertices switched has the same r-balls. However, it *is* reconstructable up to isomorphism from its r-balls. (See Figure 4.7.)

Of particular interest is the value of $t(r)$ defined to be the minimum number t such that every graph $G \in \mathcal{F}$ is reconstructible (in $\mathcal{F}$) from its r-balls where $\mathcal{F}$ is the collection of simple connected graphs with no pendant vertices and girth at least t. The first result was from Levenshtein et al. [13] where they proved that if a graph



(a) C_6 (b) C_6'

Figure 4.7: Two different cycles on 6 vertices. C_6' has the labels of 1 and 4 switched.

G has girth at least 7 and a path of length 4 passing through any vertex then G can be reconstructed from its metric balls of radius 2, so as a corollary $t(2) = 7$. Shortly after, Levenshtein [14] gave a more general result, that $t(r) \leq 2r + 2\lceil (r-1)/4 \rceil + 1$. Furthermore, he conjectured that $t(r) = 2r + 3$. He proved the following theorem that gives us a path to proving his conjecture.

Theorem 47. *Suppose that for any simple connected graph G without pendant vertices with girth at least $2r + 3$ one can determine at least one edge of G using its r-balls. Then $t(r) = 2r + 3$.*

He did so by considering what he called dense covers of metric balls, showing that if a dense cover for the metric ball about a vertex v contains at least one neighbor of v then it must be exactly $N(v)$. This result is also implied by work of Adamaszek and Adamaszek [2]. For this class of graphs they gave a simple formula for finding the neighborhood of vertex v from the r-balls when one neighbor is already known. In the next section we will prove Levenshtein's conjecture by finding one edge of any graph in this class.

4.3.2 Result

An important fact about metric balls in graphs with large girth is the following.

Lemma 48. *If G is a graph with girth at least $2r+3$ and there is a path between vertices x and y of length $r+1$ or $r+2$ then $x \notin B_r(y)$.*

Proof. If $x \in B_r(y)$ then there is a path of length at most r between x and y. Combining this path with a path of length $r+1$ or $r+2$ between x and y creates a cycle of length at most $2r+2$, a contradiction. □

To prove Levenshtein's conjecture we will consider the following quantity.

Definition 37. Given $x, y \in V(G)$ we define $I_r(x,y) = B_r(x) \cap B_r(y)$ and $i_r(x,y) = |I_r(x,y)|$.

We will show that if we fix a vertex x and consider all other vertices in $B_r(x)$ this quantity is maximized only by vertices adjacent to x. Thus we can find at least one edge incident to x.

Theorem 49. *Let G be a graph with girth at least $2r+3$ and no pendant vertices. Consider some $x \in V(G)$. If $y \in B_r(x) \setminus \{x\}$ is such that $i_r(x,y) = \max_{z \in B_r(x) \setminus \{x\}} i_r(x,z)$ then $xy \in E(G)$.*

Proof. Suppose $y \in B_r(x) \setminus \{x\}$ but $xy \notin E(G)$. Consider the tree $T = G[B_r(x)]$. There is a one-to-one correspondence between $N(x)$ and components of $T - x$ since each component contains exactly one neighbor of x. We call the components of $T - x$ *branches*. Let x_0 be the unique neighbor of x that is in the same branch as y. We claim that $i_r(x,x_0) > i_r(x,y)$. To this end, define N to be the set of vertices at distance r from x and not in the same branch as y. By Lemma 48, it is clear that $I_r(x,x_0) = B_r(x) \setminus N$. To prove the claim we will show that $|B_r(y) \cap N| < |B_r(y)^c \cap I_r(x,x_0)|$, that is to say, what we gain from switching from y to x_0 is more than we lose. Explicitly,

$$
\begin{aligned}
I_r(x,y) - I_r(x,x_0) &= (B_r(x) \cap B_r(y)) \cap (B_r(x) \cap B_r(x_0))^c \\
&= B_r(y) \cap B_r(x) \cap B_r(x_0)^c \\
&= B_r(y) \cap N, \\
I_r(x,x_0) - I_r(x,y) &= (B_r(x) \cap B_r(x_0)) \cap (B_r(x) \cap B_r(y))^c \\
&= B_r(x) \cap B_r(x_0) \cap B_r(y)^c \\
&= B_r(y)^c \cap I_r(x,x_0).
\end{aligned}
$$

If $B_r(y) \cap N = \emptyset$ then any vertex at distance $r-1$ from x not in the same branch as y is in $B_r(y)^c \cap I_r(x,x_0)$. Such a vertex exists because there are no pendant vertices and $y \neq x_0$.

For $n \in B_r(y) \cap N$ with $d_G(y,n) = k \leq r$, let n' be the unique vertex $r+1-k$ steps above n in the tree. Note that there is a path of length $r+1$ from y to n' through n so $n' \notin B_r(y)$ by Lemma 48. Now for each distinct $n_1, n_2 \in B_r(y) \cap N$ we have that n_1' and n_2' are distinct else the symmetric difference of the path of length $r+1$ from y to n_1' and the path of length $r+1$ from y to n_2' is a cycle of length at most $2r+2$, a contradiction. So if $N_y = \{n' : n \in B_r(y) \cap N\}$ then $|N_y| = |B_r(y) \cap N|$ and $N_y \subseteq I_r(x,x_0) \setminus B_r(y)$. To find one more thing that we gain, let $n_0' \in N_y$ be of minimum distance to x. If n_0'' is the neighbor of n_0' one step closer to x ($n_0' \neq x$ by Lemma 48), we see that $n_0'' \notin N_y$ and there is a path of length $r+2$ between y and n_0'' so $n_0'' \notin B_r(y)$ by Lemma 48. This proves the claim. □

Corollary 50. *Every graph with no pendant vertices and girth at least $2r+3$ can be reconstructed from its r-balls. Moreover $t(r) = 2r+3$.*

Proof. Immediate from Theorem 47 and Theorem 49. □

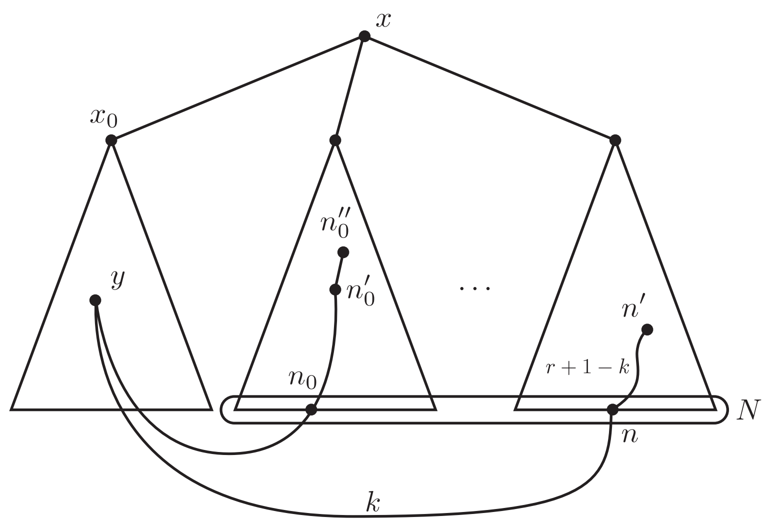


Figure 4.8: The tree T

Moreover, per the discussion in the introduction we can reconstruct the 2-core of any graph with girth at least $2r+3$ from its r-balls. For the vertices that are not in the 2-core, however, at least when $r \geq 4$, we can only in general ascertain what the closest vertex in the 2-core is and the distance form this vertex. When $r = 3$ it turns out we can reconstruct up to isomorphism if the diameter is at least 4.

4.3.3 Open Problems

This result settles the problem of reconstruction of graphs with large girth and no pendant vertices from metric balls. It is natural to consider the generalization of reconstruction up to isomorphism. Just as the cycle on $2r+2$ vertices is reconstructible up to isomorphism from balls of radius r, We conjecture that this is indeed the case for all graphs with girth $2r+2$ and no pendant vertices.

Conjecture 51. *Every graph with no pendant vertices and girth at least $2r+2$ can be reconstructed up to isomorphism from its r-balls.*

This conjecture has been proven for $r = 2$ by Adamaszek and Adamaszek [1] using a novel trick to build an isomorphism between two reconstructions. Unfortunately it is not clear how to extend this to larger values of r.

Another question is how much of a graph can be obtained from the matrix corresponding to sizes of intersections of balls. We conjecture that we can recover the whole graph from such a matrix.

Conjecture 52. *Every graph with no pendant vertices and girth at least* $2r+3$ *can be reconstructed from the matrix* $M = (m_{xy})$ *where*

$$m_{xy} = \begin{cases} i_r(x,y) & d(x,y) \le r \\ 0 & d(x,y) > r. \end{cases}$$

Bibliography

[1] A. Adamaszek and M. Adamaszek, *Uniqueness of graph square roots of girth six*, Arxiv preprint arXiv:0912.2291 (2009).

[2] ______, *Large-girth roots of graphs*, SIAM Journal on Discrete Mathematics **24** (2010), no. 4, 1501–1514.

[3] Türker Bıyıkoğlu and Josef Leydold, *Graphs with given degree sequence and maximal spectral radius*, Electron. J. Combin. **15** (2008), no. 1, Research Paper 119, 9. MR 2443134 (2009f:05163)

[4] Miranca Fischermann, Arne Hoffmann, Dieter Rautenbach, László Székely, and Lutz Volkmann, *Wiener index versus maximum degree in trees*, Discrete Appl. Math. **122** (2002), no. 1-3, 127–137. MR MR1907827 (2003d:05061)

[5] David Galvin and Prasad Tetali, *On weighted graph homomorphisms*, Graphs, morphisms and statistical physics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 63, Amer. Math. Soc., Providence, RI, 2004, pp. 97–104. MR MR2056231 (2005c:05015)

[6] Clemens Heuberger and Stephan G. Wagner, *Maximizing the number of independent subsets over trees with bounded degree*, J. Graph Theory **58** (2008), no. 1, 49–68. MR MR2404041 (2009i:05178)

[7] ———, *Chemical trees minimizing energy and Hosoya index*, J. Math. Chem. **46** (2009), no. 1, 214–230.

[8] ———, *On a class of extremal trees for various indices*, MATCH Commun. Math. Comput. Chem. **62** (2009), no. 3, 437–464. MR MR2568730

[9] F. Jelen and E. Triesch, *Superdominance order and distance of trees with bounded maximum degree*, Discrete Appl. Math. **125** (2003), no. 2-3, 225–233. MR MR1943114 (2003f:05038)

[10] Frank Jelen, *Superdominance order and distance of trees*, Ph.D. thesis, Rheinisch-Westflischen Technischen Hochschule Aachen, 2002.

[11] Paul J. Kelly, *A congruence theorem for trees*, Pac. J. Math. **7** (1957), 961–968.

[12] Russell Kirk and Hua Wang, *Largest number of subtrees of trees with a given maximum degree*, SIAM J. Discrete Math. **22** (2008), no. 3, 985–995. MR MR2424834 (2009e:05077)

[13] V. Levenshtein, E. Konstantinova, E. Konstantinov, and S. Molodtsov, *Reconstruction of a graph from 2-vicinities of its vertices*, Discrete Applied Mathematics **156** (2008), no. 9, 1399–1406.

[14] V.I. Levenshtein, *A conjecture on the reconstruction of graphs from metric balls of their vertices*, Discrete Mathematics **308** (2008), no. 5-6, 993–998.

[15] Bojan Mohar, *Computing the characteristic polynomial of a tree*, J. Math. Chem. **3** (1989), no. 4, 403–406. MR MR1039972 (91a:05034)

[16] M. Randic, X. Guo, T. Oxley, H. Krishnapriyan, and L. Naylor, *Wiener matrix invariants*, Journal of Chemical Information and Computer Sciences **34** (1994), no. 2, 361–367.

[17] KB Reid, *Centroids to centers in trees*, Networks **21** (1991), no. 1, 11–17.

[18] Slobodan K. Simić and Dejan V. Tošić, *The index of trees with specified maximum degree*, MATCH Commun. Math. Comput. Chem. **54** (2005), no. 2, 351–362. MR MR2201589 (2006i:05105)

[19] L.A. Székely and Hua Wang, *On subtrees of trees*, Advances in Applied Mathematics **34** (2005), no. 1, 138–155.

[20] Stephan G. Wagner, *Correlation of graph-theoretical indices*, SIAM J. Discrete Math. **21** (2007), no. 1, 33–46 (electronic). MR MR2299692 (2008d:05044)

[21] Harry Wiener, *Structural determination of paraffin boiling points*, Journal of the American Chemical Society **69** (1947), no. 1, 17–20.

[22] Aimei Yu and Mei Lu, *Laplacian spectral radius of trees with given maximum degree*, Linear Algebra Appl. **429** (2008), no. 8-9, 1962–1969. MR MR2446633 (2009f:05177)