Graphs with Equal Chromatic Symmetric Functions

Rosa Orellana
Dartmouth College
Mathematics Department
6188 Kemeny Hall
Hanover, NH 03755, USA.
Rosa.C.Orellana@dartmouth.edu

Geoffrey Scott
University of Michigan
Mathematics department
4080 East Hall, 530 Church Street
Ann Arbor, MI 48109, USA.
gsscott@umich.edu

May 11, 2014

Abstract

Stanley [9] introduced the chromatic symmetric function $X_G$ associated to a simple graph $G$ as a generalization of the chromatic polynomial of $G$. In this paper we present a novel technique to write $X_G$ as a linear combination of chromatic symmetric functions of smaller graphs. We use this technique to give a sufficient condition for two graphs to have the same chromatic symmetric function. We then construct an infinite family of pairs of unicyclic graphs with the same chromatic symmetric function, answering the question posed by Martin, Morin, and Wagner [7] of whether such a pair exists. Finally, we approach the problem of whether it is possible to determine a tree from its chromatic symmetric function. Working towards an answer to this question, we give a classification theorem for single-centroid trees in terms of data closely related to its chromatic symmetric function.

Keywords: chromatic symmetric function; graph coloring; unicyclic graphs; trees

Introduction

In 1995, Stanley [9] introduced a symmetric function $X_G = X_G(x_1, x_2, \ldots)$ associated with any simple graph $G$ (see Section 1 for a precise definition) called the chromatic symmetric function of $G$. $X_G$ has the property that when we specialize the variables to $x_1 = \cdots = x_k = 1$ and $x_i = 0$ for all $i > k$ then $X_G$ gives the number of ways to properly color the vertices of $G$ with $k$ colors. Hence $X_G(1, 1, \ldots, 1, 0, \ldots) = \chi_G(k)$, where $\chi_G$ is the chromatic polynomial of $G$.

One of the first questions posed by Stanley was whether $X_G$ determines $G$. As expected this is not the case, and Stanley provides the example of the kite and the bowtie as nonisomorphic graphs with the same $X_G$ [9, Fig. 1]. Although two nonisomorphic graphs may share the same chromatic symmetric function, Stanley conjectured that two nonisomorphic trees must have distinct chromatic symmetric functions. This conjecture is claimed to be true for trees with fewer than 23 vertices. This claim is found in the introduction of [7] and they cite Li-Yang Tang; however, the website containing this information is no longer available. Evidence that Stanley’s conjecture is true has been found by Morin [8] and Fougere [5] who showed that some families of trees are determined by the chromatic symmetric function. Martin, Morin and Wagner [7] showed that the degree sequence
and path sequence of a tree, $T$, can be obtained from $X_T$. They also show that some families of trees, called caterpillars and spiders, can be determined from their chromatic symmetric function.

A fundamental property of the chromatic polynomial is the deletion-contraction property, which allows us to write $\chi_G(k)$ as a linear combination of the chromatic polynomial of graphs with fewer edges. This property is the basis for inductive proofs of many other properties of the chromatic polynomial. Unfortunately $X_G$ does not satisfy a deletion-contraction law which makes it difficult to apply the useful technique of induction. Gebhard and Sagan [4] introduced a non-commutative version of $X_G$ that satisfies the deletion-contraction property and is a complete invariant of graphs. One of our results is a novel technique to decompose $X_G$ as a linear combination of chromatic symmetric functions of other graphs. And in the case when $G$ has a triangle we can write $X_G$ as a linear combination of chromatic symmetric functions of graphs with fewer edges than $G$.

There are many properties of $G$ that can be recovered from $X_G$. These include the number of vertices, the number of connected components, the number of matchings, and the girth. We have found that the number of triangles in $G$ can also be recovered from $X_G$. In the case that the graph is a tree, $T$, a lot more can be recovered from $X_T$; for example, the degree sequence can be recovered from $X_T$ [7]. This is no longer true for general graphs, we provide an example of a pair of non-isomorphic graphs with the same $X_G$ but different degree sequences. Although, the degree sequence can no longer be recovered from $X_G$ for arbitrary $G$, we show that the sum of the squares of the degrees can be recovered from $X_G$. This is a generalization of a result in [5] that shows the analogous result for trees.

In [7] the authors showed that $X_G$ is an complete invariant for two special families of unicyclic graphs and ask whether there exists a pair of unicyclic graphs with the same $X_G$. We answer this question in the affirmative by giving a pair of unicyclic graphs with the same chromatic symmetric function. In fact, our Theorem 4.2 gives a sufficient condition for two graphs to have the same chromatic symmetric function. We apply this theorem to construct infinitely many pairs of unicyclic graphs with the same $X_G$. The same technique can also be used to construct pairs of general graphs with the same $X_G$. We have also studied trees and we give a classification theorem for trees with one centroid. This classification arose from our study of the chromatic symmetric function of a tree when written in the power-sum symmetric basis.

Our paper is organized as follows. In Section 1 we review background information, set up notation, and define the chromatic symmetric function. In Section 2 we look at properties of $G$ that are determined by $X_G$ for general graphs. In particular, we show that the sum of the squares of the degrees as well as the number of triangles in a graph can be recovered from the chromatic symmetric function. In Section 3 we show how the chromatic symmetric function of a graph can be written as a linear combination of other chromatic symmetric functions. In Section 4 we focus our attention on unicyclic graphs. We also prove a sufficient condition for two graphs to have the same chromatic symmetric function and show how to construct pairs of graphs with the same $X_G$. In our last section, Section 5, we prove a classification theorem for trees with a single centroid that is closely related to the coefficients of the chromatic symmetric function when written in the power-sum symmetric basis.

1 Preliminaries

We assume that the reader is familiar with the basic facts about graphs found in any introductory graph theory book (see e.g., [2, 3, 6]). In this section we establish notation that will be used throughout the paper. A graph $G$ is an ordered pair $(V, E)$, where $V = V(G)$ is the vertex set and $E = E(G)$ is the edge set. All our graphs are simple, i.e., we do not allow loops or multiple edges.
The number of vertices \( \# V(G) \) is called the order of the graph. We will write \( uv \) for the edge joining the vertices \( u, v \in V(G) \) if such an edge exists. We say that \( u \) and \( v \) are endpoints of \( uv \), that \( uv \) is incident to \( u \) and \( v \), and that \( u \) is adjacent to \( v \). If two edges have no endpoints in common, they are disjoint. The degree \( d(v) \) of a vertex \( v \) is the number of edges incident to \( v \). The degree sequence of a graph \( G \) is the sequence \((d(v))_{v \in V(G)}\). An isolated vertex is a vertex of degree 0. A leaf is a vertex of degree 1. The girth of a graph is the number of distinct vertices in a shortest cycle in the graph. An acyclic graph has infinite girth.

A subgraph \( G' \subseteq G \) of a graph \( G = (V(G), E(G)) \) is a graph \( G' = (V'(G), E'(G)) \) such that \( V'(G) \subseteq V(G) \) and \( E'(G) \subseteq E(G) \). A subgraph is said to be induced by the vertex set \( V'(G) \) if every edge in \( E(G) \) having endpoints in \( V'(G) \) is also in \( E'(G) \). A subgraph \( H \) is a spanning subgraph of \( G \) if it has the same vertex set as \( G \). A subgraph is said to be a matching of size \( k \) if it consists of \( k \) disjoint edges on \( 2k \) vertices.

In this paper we are interested in certain classes of simple graphs. A graph is called unicyclic if it contains exactly one cycle, a forest if it contains no cycles, and a tree if it is a connected forest. Notice that a connected unicyclic graph with \( n \) vertices has \( n \) edges.

In the following proposition we summarize some well-known facts about trees. The reader may refer to [2, 3, 6] or any other introductory graph theory textbook for proofs of these facts.

**Proposition 1.1** ([2], pp. 99-100).

1. In a tree, any two vertices are connected by exactly one path.
2. Every tree on \( n \) vertices has \( n - 1 \) edges. In general, a forest on \( n \) vertices with \( c \) connected components has \( n - c \) edges.
3. Every nontrivial tree has at least two leaves. In general, if a forest contains \( c \) connected nontrivial components, then it contains at least \( 2c \) leaves.

We now give two definitions that are not as standard as the others we have given so far. We will use these definitions in Section 5. For further reading on these concepts see [6].

**Definition 1.2.** The weight of a vertex \( v \) of a tree \( T \) is the maximal number of edges in any subtree of \( T \) containing \( v \) as a leaf.

**Definition 1.3.** The centroid of a tree \( T \) is the set of all vertices of \( T \) having minimum weight.

An example of the weights of vertices of a tree is shown in Figure 1. In that graph, the vertex with weight 8 is the centroid of the tree.

![Figure 1: A tree with one centroid](image)

**Proposition 1.4** ([2], pp. 99). Every tree has a centroid consisting of either one vertex or two adjacent vertices.
A proper coloring of the vertices of a graph $G$ with $k$ colors is a function $\kappa : V(G) \to \{1, 2, \ldots, k\}$ such that $\kappa(v) \not= \kappa(w)$ for adjacent vertices $v, w \in V(G)$. Let $\mathbb{N}$ denote the positive integers. The chromatic polynomial of $G$ is a polynomial $\chi_G : \mathbb{N} \to \mathbb{N} \cup \{0\}$ having the property that $\chi_G(k)$ is the number of proper $k$-colorings of $G$.

For more properties of the chromatic polynomial, including the fact that it is a polynomial, see [6]. The graph $G$ in Figure 3 has chromatic polynomial $\chi_G(k) = k(k-1)^6$. One remarkable feature of the chromatic polynomial is that for any fixed $n > 0$, all trees with $n$ vertices have the same chromatic polynomial.

### 1.1 The chromatic symmetric function

This paper will focus on the chromatic symmetric function of a graph. Before defining this function, we review basic facts of symmetric functions.

A partition is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ of positive integers such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l$. The $\lambda_i$ are called the parts of $\lambda$. Furthermore, we say that $\lambda$ is a partition of $n$, written $\lambda \vdash n$, if $\sum \lambda_i = n$.

Let $x_1, x_2, \ldots$ be a countably infinite set of commuting indeterminates. For any positive integer $k$, define the power-sum symmetric function as

$$p_k = \sum_{i \geq 1} x_i^k$$

and for a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ we define

$$p_\lambda = p_{\lambda_1}p_{\lambda_2} \cdots p_{\lambda_l}$$

It is a well-known fact that $\{p_\lambda \mid \lambda \vdash n\}$ is a basis for the $\mathbb{Q}$-vector space $\Lambda_n$ of all symmetric functions that are homogeneous of degree $n$. For more details about symmetric functions see [11] Chap. 7.

Let $G$ be a simple graph. Stanley [9], see also [11] pp. 462-464, defined the chromatic symmetric function of $G$ as

$$X_G = X_G(x_1, x_2, \ldots) = \sum_\kappa \prod_{v \in V(G)} x_{\kappa(v)}$$

(1)
where the sum is over all proper colorings $\kappa$ and the $x_1, x_2, \ldots$ are a countably infinite set of commuting indeterminates. Since a coloring of a graph is invariant under permutation of the colors, $X_G$ is a symmetric, homogeneous function of degree $\#V(G)$. $X_G$ is a generalization of the well-known single variable chromatic polynomial of a graph $\chi_G(k)$. In fact, Stanley showed that if we set $x_1 = x_2 = \ldots = x_k = 1$ and $x_i = 0$ for all $i > k$, then

$$X_G(1, 1, \ldots, 1, 0, \ldots) = \chi_G(k).$$

Often when working with a symmetric function, it is helpful to expand it in terms of one of the many bases for the space of symmetric functions.

**Theorem 1.5.** [St1] 

$$X_G = \sum_{S \subseteq E(G)} (-1)^{\#S} p_{\pi(S)},$$

where $\pi(S)$ is the partition whose parts are the orders of the connected components of the subgraphs of $G$ induced by $S$. $\pi(S)$ is called the type of $S$, (see Figure 4 for an example of $\pi(S)$).

![Figure 4](image.png)

Figure 4: $S = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ is an edge set of type $\pi(S) = (4, 3, 2, 1, 1)$

There are several properties of a graph that we can determine from its chromatic symmetric function. In the following sections, we explore these properties.

## 2 Properties preserved by the chromatic symmetric function

We begin by studying the properties of a simple graph that are determined by its chromatic symmetric function. We review some known results and make two new contributions. In particular, we show that the number of triangles and the sum of the squares of the vertex degrees can be recovered from $X_G$. The results in this section will be applicable to later sections, when we restrict our attention to unicyclic graphs and trees.

**Remark 2.1.** If two graphs have the same chromatic symmetric function, then they have

1. the same number of edges,
2. the same number of vertices, and
3. the same number of matchings of $k$ edges (for any natural number $k$).

From Equation (2) the number of vertices can be recovered from the coefficient of $p_{(1,1,\ldots,1)}$, the number of edges from the coefficient of $p_{(2,1,\ldots,1)}$ and the number of $k$-matchings from the coefficient of $p_{(2,2,\ldots,2,1,\ldots,1)}$ (containing $k$ 2’s).

Fougere [5, Theorem 3.3.1] proved that the sum of the squared vertex degrees of a tree was determined by its chromatic symmetric function. This result was strengthened by Martin, Morin, and Wagner when they showed that the degree and path sequences of a tree can be recovered from its
chromatic symmetric function [7, Corollary 5]. In the following proposition, we show that Fougere’s result is true even for general graphs. However, the strengthened result in [7, Corollary 5] is not true for a general graph. An example of two graphs with the same chromatic symmetric function yet differing degree sequences is given in Figure 9.

**Proposition 2.2.** The sum of the squared vertex degrees of a graph $G$

$$\sum_{v \in V(G)} d(v)^2$$

can be obtained from $X_G$.

**Proof.** Let $S^{(2,2)}_G$ denote the number of spanning subgraphs of $G$ consisting of two disjoint edges and $\#V(G) - 4$ isolated vertices. Since there are no other spanning subgraphs having vertex partition $(2, 2, 1, \ldots, 1)$, it follows that $S^{(2,2)}_G$ is exactly the coefficient of $p_{(2,2,1,\ldots,1)}$ in $X_G$. Therefore, if $G$ and $H$ have the same chromatic symmetric function, then $S^{(2,2)}_G = S^{(2,2)}_H$. Next, let $S^{(3)}_G$ be the number of spanning subgraphs of $G$ consisting of two non-disjoint edges and $\#V(G) - 3$ isolated vertices. Since all 2-edge spanning subgraphs of $G$ consist of those tallied in $S^{(2,2)}_G$ or $S^{(3)}_G$, it follows that

$$\left( \frac{\#E(G)}{2} \right) = S^{(2,2)}_G + S^{(3)}_G.$$

For graphs $G$ and $H$ having the same chromatic symmetric function $\#E(G) = \#E(H)$ by Remark 2.1 and $S^{(2,2)}_G = S^{(2,2)}_H$ for the reasons above. Therefore, $S^{(3)}_G = S^{(3)}_H$. We can count $S^{(3)}_G$ by noting that any pair of non-disjoint edges is uniquely determined by its central vertex and a choice of two edges incident to that vertex. Using the fact that $G$ and $H$ contain the same number of edges (from Remark 2.1) and therefore must have the same degree sum, we calculate

$$\sum_{v \in V(G)} \binom{d(v)}{2} = \sum_{v \in V(H)} \binom{d(v)}{2}$$

$$\sum_{v \in V(G)} (d(v)^2 - d(v)) = \sum_{v \in V(H)} (d(v)^2 - d(v))$$

$$\sum_{v \in V(G)} d(v)^2 = \sum_{v \in V(H)} d(v)^2$$

\[\square\]

**Corollary 2.3.** The number of triangles $T_G$ in a graph $G$ can be obtained from $X_G$.

**Proof.** Note that the coefficient of $p_{(3,1,1,\ldots,1)}$ in $X_G$ consists of $S^{(3)}_G - T_G$. In the proof of Proposition 2.2 we saw that $S^{(3)}_G$ can be obtained from $X_G$. Therefore, $T_G$ can also be obtained from $X_G$. \[\square\]

Chow [1] proved that the planarity of a graph is not determined by its chromatic symmetric function. Also, in [7, Proposition 3], the authors show that the chromatic symmetric function of a graph determines its girth.
3 Decomposition techniques for graphs

A fundamental property of chromatic polynomials, $\chi_G(k)$, is the deletion-contraction property. If $G \setminus e$ denotes $G$ with edge $e$ deleted and $G/e$ denotes $G$ with edge $e$ contracted to a point, then

$$\chi_G(k) = \chi_{G \setminus e}(k) - \chi_{G/e}(k).$$

This property is often used to prove other properties of $\chi_G$ using induction on the number of edges. Unfortunately, there is no analogous deletion-contraction property for $X_G$. In this section we present a novel technique for writing the chromatic symmetric function of a graph as a linear combination of the chromatic symmetric function of other graphs. In the case that $G$ has girth three we are able to write $X_G$ as a linear combination of chromatic symmetric functions of graphs with fewer edges.

**Theorem 3.1.** Let $G$ be a graph where $e_1, e_2, e_3 \in E(G)$ form a triangle. Furthermore, define

- $G_{2,3} = (V(G), E(G) - \{e_1\})$
- $G_{1,3} = (V(G), E(G) - \{e_2\})$
- $G_3 = (V(G), E(G) - \{e_1, e_2\})$

Then

$$X_G = X_{G_{2,3}} + X_{G_{1,3}} - X_{G_3}.$$  

**Proof.** Consider the following partition of the set of spanning subgraphs of $G$:

- $G^1 = \{S \subseteq E(G) : e_1, e_2, e_3 \in S\}$
- $G^2 = \{S \subseteq E(G) : e_1, e_2 \in S, e_3 \notin S\}$
- $G^3 = \{S \subseteq E(G) : e_1, e_3 \in S, e_2 \notin S\}$
- $G^4 = \{S \subseteq E(G) : e_2, e_3 \in S, e_1 \notin S\}$
- $G^5 = \{S \subseteq E(G) : e_1, e_2, e_3 \notin S\}$
- $G^6 = \{S \subseteq E(G) : e_2 \in S, e_1, e_3 \notin S\}$
- $G^7 = \{S \subseteq E(G) : e_3 \in S, e_1, e_2 \notin S\}$
- $G^8 = \{S \subseteq E(G) : e_1, e_2, e_3 \notin S\}$

Then by Equation (2)

$$X_G = \sum_{S \subseteq E(G)} (-1)^{|S|} p_{\pi(S)} = \sum_{i=1}^{8} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)}$$

$$= \sum_{i=\{4,6,7,8\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)} + \sum_{i=\{3,5,7,8\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)}$$

$$- \sum_{i=\{7,8\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)} + \sum_{i=\{1,2\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)}$$

$$= X_{G_{2,3}} + X_{G_{1,3}} - X_{G_3} + \sum_{i=\{1,2\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)}.$$
It suffices to prove that the final term is equal to zero. Note, however, that for every subgraph in $G^2$, one can add the edge $e_3$ to get a corresponding subgraph in $G^1$. The addition of $e_3$ will not change the vertex partition of the subgraph in $G^2$, although it will add one edge. Thus, for every subgraph in $G^1$, there is a corresponding subgraph in $G^2$ with an opposite contribution to the chromatic symmetric function. Therefore,

$$\sum_{i \in \{1, 2\}} \sum_{S \in G^i} (-1)^{|S|} p_{\pi(S)} = 0$$

and the proof is complete. \[\square\]

To apply Theorem 3.1, we introduce the equivalence relation $\sim_X$ on linear combinations of graphs. Let $\{G_i\}_{i \leq p}$ and $\{H_i\}_{i \leq k}$ be sets of graphs, and let $\{c_i\}_{i \leq p}$ and $\{d_i\}_{i \leq k}$ be real numbers. We say that

$$\sum_{i \leq k} c_i G_i \sim_X \sum_{i \leq k} d_i H_i$$

if

$$\sum_{i \leq k} c_i X G_i = \sum_{i \leq k} d_i X H_i.$$

The following example illustrates how Theorem 3.1 can be applied to two graphs. The vertices of the graphs are labeled to correspond to the labels used in Theorem 3.1.

![Figure 5: Decomposition of Two Graphs](image)

Notice that the forests in the top expression are isomorphic to the forests in the bottom expression. Therefore, the unicyclic graph on the top left in Figure 5 has the same chromatic symmetric function as the unicyclic graph on the bottom left of Figure 5. This example shows that the degree sequence of a graph is not determined by its chromatic symmetric function.

However, this method has limitations. At this point in our discussion, we cannot apply it to graphs having no triangles. Also, even if we do manage to decompose the chromatic symmetric function of two graphs into linear combinations of chromatic symmetric functions of forests, we cannot always determine whether these linear combinations are equal. The following corollary introduces another way to decompose chromatic symmetric functions that will complement Theorem 3.1 to strengthen the method.

Notice that there are six different ways to label the three sides of a triangle $e_1$, $e_2$, and $e_3$. By choosing different permutations of labels, Theorem 3.1 yields three different ways to decompose a graph with a triangle into smaller graphs. We use this fact in the following corollary.

**Corollary 3.2.** Let $G$ be a graph with the adjacent edges $e_1 = vv_1$, $e_2 = vv_2$ and $e_3 = v_1v_2 \notin E(G)$ (that is, $e_1$ and $e_2$ meet at the vertex $v$, but there is no edge connecting $v_1$ to $v_2$). Define...
\[ G_{1,3} = (V(G), (E(G) - \{e_2\}) \cup \{e_3\}) \]
\[ G_{2,3} = (V(G), (E(G) - \{e_1\}) \cup \{e_3\}) \]
\[ G_1 = (V(G), (E(G) - \{e_2\}) \]
\[ G_3 = (V(G), (E(G) - \{e_1, e_2\}) \cup \{e_3\}). \]

Then
\[ X_G = X_{G_{2,3}} + X_{G_1} - X_{G_3}. \]

**Proof.** Let \( G' = (V(G), E(G) \cup \{e_3\}) \). We can apply Theorem 3.1 in two different ways:
\[ X_{G'} = X_{G_{2,3}} + X_{G_{1,3}} - X_{G_3} \]
\[ X_{G'} = X_{G_{1,3}} + X_G - X_{G_1}. \]
This gives the desired equality. \( \square \)

Figure 6 illustrates Corollary 3.2 and Figure 7 shows an application of it.

There is one more corollary to Theorem 3.1 that will be useful.

**Corollary 3.3.** Let \( G \) be a graph containing a triangle formed by the edges \( e_1 = vv_1, e_2 = vv_2 \) and \( e_3 = v_1v_2 \). Define
\[ G_{1,2} = (V(G), (E(G) - \{e_3\}) \]
\[ G_1 = (V(G), (E(G) - \{e_2, e_3\}) \]
\[ G_2 = (V(G), (E(G) - \{e_1, e_3\}) \]
\[ G_3 = (V(G), (E(G) - \{e_1, e_2\}). \]

Then
\[ X_G = 2X_{G_{1,2}} + X_{G_3} - X_{G_1} - X_{G_2}. \]
Proof. If we define
\[ G_{2,3} = (V(G), E(G) - \{e_1\}), \]
then we can apply Theorem 3.1 and Corollary 3.2 to obtain the following equations:
\[ X_G = X_{G_{1,2}} + X_{G_{2,3}} - X_{G_2}, \]
\[ X_{G_{2,3}} = X_{G_{1,2}} + X_{G_3} - X_{G_1}. \]
The desired equality follows from substitution of the second equation into the first. \(\square\)

Corollary 3.3 is illustrated in Figure 8.

\[ \begin{array}{c}
\text{e}_1 \sim \text{e}_2 \quad \text{2} \quad + \quad - \\
\text{e}_3
\end{array} \]

Figure 8: Illustration of Corollary 3.3

4 The chromatic symmetric function of a unicyclic graph

In this section, we study chromatic symmetric functions of unicyclic graphs. In [7], the authors describe two families of unicyclic graphs, squids and crabs. A squid is a connected unicyclic graph having only one vertex of degree greater than 2, while a crab is a connected unicyclic graph in which every vertex not lying on the cycle has degree 1. In [7, Theorem 12], they proved that no two non-isomorphic squids have the same chromatic symmetric function. In [7, Proposition 13] they show a similar result for crabs subject to an additional technical condition. In addition, Martin, Morin and Wagner asked whether two distinct unicyclic graphs could have the same chromatic symmetric function. Figure 9 shows an example of two such graphs.

\[ \begin{array}{c}
\text{\textbullet} \\
\text{\textbullet} \\
\text{\textbullet}
\end{array} \quad \begin{array}{c}
\text{\textbullet} \\
\text{\textbullet} \\
\text{\textbullet}
\end{array} \]

Figure 9: Two unicyclic graphs with the same chromatic symmetric function

These graphs are precisely the same ones we used to illustrate Theorem 3.1 in Section 3. It is also interesting that these graphs do not have the same degree sequence. It was proven in [7] that any two trees with the same chromatic symmetric function share the same degree sequence; here we see that an analogous result does not apply for unicyclic graphs. In fact, not even the number of leaves is determined from the chromatic symmetric function of a unicyclic graph.

In the following proposition we provide a relation between the number of vertices of degree one and two within the cycle in two unicyclic graphs with the same chromatic symmetric function.

Proposition 4.1. If graphs \( G \) and \( H \) are connected unicyclic graphs with a cycle of length \( p \) and \( X_G = X_H \), then
\[(p - 1)L_G + I_G = (p - 1)L_H + I_H\]
where \( L_G \) and \( L_H \) are the number of leaves in \( G \) and \( H \) respectively, and \( I_G \) and \( I_H \) are the number of vertices of degree two on the cycles of \( G \) and \( H \) respectively.
Proof. Let \( n \) be the number of vertices in \( G \) and \( H \). Because \( G \) and \( H \) are unicyclic, the removal of any three edges of \( G \) or \( H \) will be a subgraph with at least three connected components. Therefore, every subgraph of \( G \) or \( H \) of type \((n - 1, 1)\) is a subgraph with either \( n - 1 \) or \( n - 2 \) edges. The only way that the removal of one edge would result in a type \((n - 1, 1)\) subgraph is if the removed edge is adjacent to a leaf. There are two ways that the removal of two edges could result in a type \((n - 1, 1)\) subgraph: either an edge adjacent to a leaf and a cycle edge are both removed, or two edges attached to a vertex of degree two on the cycle are removed. Thus, equating the coefficient of \( p_{(n-1,1)} \) in \( X_G \) and \( X_H \) yields

\[
(-1)^{n-1}L_G + (-1)^{n-2}p_L + (-1)^{n-2}I_G = (-1)^{n-1}L_H + (-1)^{n-2}p_L + (-1)^{n-2}I_H.
\]

Multiplying the above equation by \((-1)^n\) gives the desired result. \(\square\)

So far, we have just one example of a pair of unicyclic graphs (those shown in Figure 9) that have the same chromatic symmetric function. Next, we will prove a theorem that gives a sufficient condition for two graphs to have the same chromatic symmetric function. It will aid us in constructing infinitely many pairs of unicyclic graphs with the same chromatic symmetric function.

**Theorem 4.2.** Let \( G = (V(G), E(G)) \) be a graph that has four vertices \( u, v, w, z \) with the property that \( uz, wz, vw \in E(G) \) and \( uw, vz, uv \notin E(G) \). If there exists a graph automorphism \( \varphi : (V(G), E(G) - wz) \to (V(G), E(G) - wz) \) such that

\[
\varphi(\{u, w\}) = \{v, z\} \quad \text{and} \quad \varphi(\{v, z\}) = \{u, w\}
\]

then

\( H = (V(G), E(G) \cup \{uw\}) \) and \( J = (V(G), E(G) \cup \{vz\}) \) have the same chromatic symmetric function.

**Proof.** Consider the following partition of the spanning subgraphs of \( H \):

- \( H^1 \) is the set of all spanning subgraphs of \( H \) that do not contain the edge \( uw \)
- \( H^2 \) is the set of all spanning subgraphs of \( H \) that do contain the edge \( uw \) but not \( wz \)
- \( H^3 \) is the set of all spanning subgraphs of \( H \) that contain the edges \( uw \) and \( wz \), but not \( uz \)
- \( H^4 \) is the set of all spanning subgraphs of \( H \) that contain the edges \( uw, wz, \) and \( uz \)

We partition the spanning subgraphs of \( J \) similarly:

- \( J^1 \) is the set of all spanning subgraphs of \( J \) that do not contain \( vz \)
- \( J^2 \) is the set of all spanning subgraphs of \( J \) that do contain \( vz \) but not \( wz \)
- \( J^3 \) is the set of all spanning subgraphs of \( J \) that contain \( vz \) and \( wz \), but not \( vw \)
- \( J^4 \) is the set of all spanning subgraphs of \( J \) that contain \( vz, wz, \) and \( vw \)

First, note that \( H^4 \) can be put into a bijection with \( J^4 \) via the map of subgraphs induced by the identification \( V(H) = V(J) \). Because this bijection preserves the vertex partition of a subgraph, the net contribution of \( H^4 \) to \( X_H \) equals the net contribution of \( J^4 \) to \( X_J \).
Next, note that $H^2$ can be put into a bijection with $J^2$ via the map of subgraphs induced by $\varphi$. Because this bijection preserves the vertex partition of a subgraph, the net contribution of $H^2$ to $X_H$ equals the net contribution of $J^2$ to $X_J$.

Note that the contribution of $H^3$ to $X_H$ is the negative of the contribution of $H^4$ to $X_H$. This is because the bijection from $H^3$ to $H^4$ that adds the edge $vw$ to each subgraph in $H^3$ preserves the vertex partition of the subgraph while adding an extra edge. Therefore, the contribution of subgraphs of type $H^3$ to $X_H$ is exactly cancelled by the contribution of the subgraphs of type $H^4$ to $X_H$. Similarly, the contribution of subgraphs of type $J^3$ to $X_J$ is exactly cancelled by the contribution of the subgraphs of type $J^4$.

We see that the contribution of $H^1 \cup H^2 \cup H^3 \cup H^4$ to $X_H$ is equal to the contribution of $J^1 \cup J^2 \cup J^3 \cup J^4$ to $X_J$. Therefore, $X_H = X_J$. \hfill \Box

Notice that Theorem 4.2 can also be used to show that the two unicyclic graphs in Figure 9 do indeed have the same chromatic symmetric function. In fact, Theorem 4.2 actually gives us infinitely many examples of unicyclic graphs having the same chromatic symmetric function. For any two nonisomorphic rooted trees $T_1$ and $T_2$, one can make two copies of each tree and connect the four trees according to the left picture in Figure 10. By Theorem 4.2 this unicyclic graph will have the same chromatic symmetric function as the graph obtained using the connection rule in the right picture in Figure 10.

Figure 10: Creating unicyclic graphs with identical chromatic symmetric function

Although Theorem 4.2 was motivated by the question of whether unicyclic graphs could share a chromatic symmetric function, its usefulness is not limited to unicyclic graphs. On the contrary, one can use Theorem 4.2 to show that graphs such as those in Figure 11 have the same chromatic symmetric function.

Figure 11: Two non-isomorphic graphs with the same chromatic symmetric function

5 Chromatic symmetric functions of trees

In [9], Stanley asks whether a tree $T$ is determined by $X_T$. This question is still open. Throughout this section, we aim to prove results bringing us closer to answering this question. In particular, we will prove a classification theorem of trees related to Stanley’s question.
Let $\mathcal{P}(X)$ denote the power set of a set $X$, and let $\text{Part}(k)$ denote the partitions of an integer $k$. For a tree $T$, we define the function

$$
\theta_T : \mathcal{P}(E(T)) \rightarrow \text{Part}(\#V(T))
$$

by $\theta_T(\{e_1, e_2, \ldots, e_k\}) = \pi(E(T) - \{e_1, e_2, \ldots, e_k\})$, where for $S \subseteq E(G)$, the partition $\pi(S)$ of $\#V(G)$ has as its parts the number of vertices in each connected component of the graph $(V(G), S)$.

For positive integers $\{a_i\}_{1 \leq i \leq r}$ that sum to $n$, we define the rearrange function

$$
\text{re} : (a_1, a_2, \ldots, a_r) \rightarrow \text{Part}(n)
$$

that sends a collection of $r$ integers to the partition having $a_1, a_2, \ldots, a_r$ as its parts (recall that a partition is always written in weakly decreasing order). We define an ordering on the 2-part partitions of $n$ by $(n - i, i) > (n - j, j)$ if $i > j$.

Throughout this section, we will pay particular attention to the image under $\theta_T$ of singleton and 2-element sets of edges. Because $T$ is a tree, for any set $S$ of edges, $\theta_T(S)$ will have $\#S + 1$ parts. In particular, $\theta_T$ sends singleton sets to 2-part partitions of $\#V(T)$, and sends 2-element sets to 3-part partitions of $\#V(T)$. The following lemma can be easily proved by contradiction.

**Lemma 5.1.** Let $T$ be a tree with even order $n$. Then there is at most one edge $e$ in $T$ such that $\theta_T(e) = (n/2, n/2)$. If such an edge exists, it joins two centroids.

Hence, we have that we can tell from $X_T$ if the tree $T$ has one or two centroids.

**Proposition 5.2.** Let $T$ be a tree with order $n$ having the distinct edges $e_a$ and $e_b$. Let $\theta_T(e_a) = (n - i, i)$ and $\theta_T(e_b) = (n - k, k)$ and without loss of generality assume that $i \geq k$. If $k = i$ then $\theta_T(\{e_a, e_b\}) = \text{re}(n - 2i, i, i)$ and if $k < i$ then $\theta_T(\{e_a, e_b\})$ is either $\text{re}(n - i - k, i, k)$ or $\text{re}(n - i, i - k, k)$.

**Proof.** Upon removing $e_a$, the tree $T$ is divided into two connected components containing $n - i$ and $i$ vertices. If $i = k$, then we can only remove $e_b$ from the component with $n - i$ edges, hence the result follows. If $k < i$, then $e_b$ may be in either connected component. Its removal leaves one of these components unchanged. Therefore, either $n - i$ or $i$ must be present in $\theta_T(\{e_a, e_b\})$. Similarly, by removing $e_b$ before $e_a$, we see that either $n - k$ or $k$ must be present in $\theta_T(\{e_a, e_b\})$. However, both $n - i$ and $n - k$ cannot be present in $\theta_T(\{e_a, e_b\})$ since $n - i + n - k \geq n$. Similarly, both $i$ and $n - k$ cannot be present in $\theta_T(\{e_a, e_b\})$, since $i + n - k \geq n$ (recalling the fact that $k < i$). Therefore, $\theta_T(\{e_a, e_b\})$ must either contain both $k$ and $i$ or it must contain $n - i$ and $k$. Since the sum of the numbers in the partition of $n$ must add to $n$, it follows that $\theta_T(\{e_a, e_b\})$ is either $\text{re}(n - i - k, i, k)$ or $\text{re}(n - i, i - k, k)$.

**Definition 5.3.** Let $T$ be a tree. We say that the distinct edges $e_a$ and $e_b$ of $T$ attract if there is a path in $T$ containing both $e_a$ and $e_b$ and having a centroid of $T$ as one endpoint. Otherwise, the edges $e_a$ and $e_b$ repel.

This terminology is motivated by imagining each edge of $T$ as a bar magnet, the positive side of which is pointed towards the centroid(s). In the same way that two magnets attract if the positive side of one is pointed toward the negative side of the other, we say that two edges attract if a path connects the “positive” side of one edge to the “negative” side of the other. It can be checked that this corresponds to the formal definition of *attract* given above. On the other hand, two edges repel if a path connects the “positive” side of one to the “positive” side of the other. Figure 12 shows an example: the dotted edges repel $c$, while the thick edges attract $c$. Vertex $c$ is the centroid of the
tree. From the definition, it can be checked that if $T$ has two centroids, then the edge connecting them attracts every edge in $T$. We proceed by exploring a few properties of edge attraction and repulsion.

**Proposition 5.4.** Let $e_a$ and $e_b$ be distinct edges of a tree $T$ with $\theta_T(e_a) = (n - i, i)$ and $\theta_T(e_b) = (n - k, k)$ with $i \geq k$. Then $\theta_T(e_a, e_b) = re(n - i, i - k, k)$ if and only if $e_a$ and $e_b$ attract.

**Proof.** Suppose we remove $e_a$ from $T$ to form a forest with two connected components: $T_{n-i}$ which contains $n - i$ vertices and $T_i$ which contains $i$ vertices. We split the proof into two cases: either $T_{n-i}$ and $T_i$ contain one centroid of $T$ each, or all the centroids are contained in $T_{n-i}$.

In the first case, the endpoints of $e_a$ must be the centroids of $T$. In this case, by Lemma 5.1 $\theta_T(e_a) = (\frac{n}{2}, \frac{n}{2})$ and $e_a$ attracts every edge of $T$. From Proposition 5.2, $\theta_T(e_a, e_b)$ is either $re(n - \frac{n}{2} - k, \frac{n}{2}, k)$ or $re(n - \frac{n}{2}, \frac{n}{2} - k, k)$. However, these partitions are identical. In this first case, $e_a$ and $e_b$ attract if and only if $\theta_T(e_a, e_b) = (\frac{n}{2}, \frac{n}{2} - k, k)$.

In the second case, only $T_{n-i}$ contains centroids of $T$. If $e_b$ is also contained in $T_{n-i}$, then every path from $e_b$ to a centroid would also be contained in $T_{n-i}$ and would avoid $e_a$. Therefore, $e_a$ attracts $e_b$ if and only if $e_b$ is in $T_i$. By the proof of Proposition 5.2, $e_b$ is in $T_i$ exactly when $\theta_T(e_a, e_b) = (n - i, i - k, k)$.

Notice that from Proposition 5.4 we have that $\theta_T(e_a, e_b)$ is not a three part partition when $n = k$. Thus we have the following corollary.

**Corollary 5.5.** If $\theta_T(e_a) = \theta_T(e_b)$, then $e_a$ and $e_b$ repel.

**Remark 5.6.** We say that an edge $e_a$ separates an edge $e_b$ from a vertex $v$ if $e_a$ and $v$ are in different connected components of the graph after removing the edge $e_a$. If a tree $T$ has a single centroid and $e_a$ separates $e_b$ from the centroid, then $\theta_T(e_a) > \theta_T(e_b)$.

This remark follows from the definition of $\theta_T$. We are now ready to prove the main theorem of this section.

**Theorem 5.7.** Given the set of edges $\{e_1, e_2, \ldots, e_{n-1}\}$ of a tree $T$ with a single centroid and the values $\theta_T(e_i)$ and $\theta_T(e_i, e_k)$ for all distinct edges $e_i, e_k \in T$, then $T$ can be constructed from this data. In other words, the values $\theta_T(e_i)$ and $\theta_T(e_i, e_k)$ determine the tree.

**Proof.** After reindexing the edges, we may assume that $\theta_T(e_i) \geq \theta_T(e_{i+1})$ for all $i < n$. We will construct $T$ by adding edges, one by one, to a forest of $n$ isolated vertices. We will add the edges in the order of their index. During each step, we will use the values of $\theta_T$ given to determine the unique way to add the next edge. This process will create a sequence of forests $T_i$, where $T_i$ is the
resulting forest after adding the first $i$ edges. In the course of the construction, we will see that for $i \geq 1$, $T_i$ consists of $n - 1 - i$ isolated vertices and one tree with $i + 1$ vertices.\footnote{This explains our choice to use the letter $T_i$ to denote the forests, not $F_i$.}

Notice that the first edge placed must have the centroid as an endpoint. Clearly, $T_1$ consists of $e_1$ together with $n - 2$ isolated vertices. Now assume that we have placed the first $i$ edges, and that $T_i$ consists of a tree and $n - 1 - i$ isolated vertices. We wish to place the $(i + 1)^{st}$ edge. We will show two things. First, $e_{i+1}$ must be attached to the tree with $i$ edges in $T_i$ (therefore $T_{i+1}$ will consist of a tree and $n - i$ isolated vertices). Second, we will show that there is one unique permissible way to attach the edge $e_{i+1}$ to the tree in $T_i$.

If we suppose that $e_{i+1}$ is not attached to the tree with $i$ edges in the forest $T_i$, then in $T$ the edge $e_{i+1}$ will be separated from the centroid by an edge $e_j$ with $j > i + 1$. By our indexing of the edges, $\theta_T(e_{i+1}) \geq \theta_T(e_j)$. However, this is forbidden by Remark 5.6. Therefore, $e_{i+1}$ is attached to the tree in $T_i$.

To see where $e_{i+1}$ must be attached, we need only look at which edges it attracts. By Proposition 5.4 we can determine which edges of $T_i$ attract $e_{i+1}$ and which edges of $T_i$ repel $e_{i+1}$ from the values $\theta(e_i, e_j)$ (for $e_j \in T_i$). The set of edges of $T_i$ that attract $e_{i+1}$ form a path from the centroid to the vertex $v$ that is adjacent to $e_{i+1}$. This is evident from the definition of attraction and repulsion. Since there is a unique path from the centroid to the vertex $v$, there is a unique way to attach the edge $e_{i+1}$ to the forest $T_i$. Thus we extend $T_i$ to $T_{i+1}$ by placing $e_{i+1}$ at the end of the path created by edges that attract it.

At first glance, the conditions outlined in Theorem 5.7 seem very strict, and one might wonder whether they may be relaxed. For example, perhaps it is possible to construct a tree in $T_i$ consists of a tree and $n - 1 - i$ isolated vertices. Now assume that we have placed the first $i$ edges, and that $T_i$ consists of a tree and $n - 1 - i$ isolated vertices. We wish to place the $(i + 1)^{st}$ edge. We will show two things. First, $e_{i+1}$ must be attached to the tree with $i$ edges in $T_i$ (therefore $T_{i+1}$ will consist of a tree and $n - i$ isolated vertices). Second, we will show that there is one unique permissible way to attach the edge $e_{i+1}$ to the tree in $T_i$.

If we suppose that $e_{i+1}$ is not attached to the tree with $i$ edges in the forest $T_i$, then in $T$ the edge $e_{i+1}$ will be separated from the centroid by an edge $e_j$ with $j > i + 1$. By our indexing of the edges, $\theta_T(e_{i+1}) \geq \theta_T(e_j)$. However, this is forbidden by Remark 5.6. Therefore, $e_{i+1}$ is attached to the tree in $T_i$.

To see where $e_{i+1}$ must be attached, we need only look at which edges it attracts. By Proposition 5.4 we can determine which edges of $T_i$ attract $e_{i+1}$ and which edges of $T_i$ repel $e_{i+1}$ from the values $\theta(e_i, e_j)$ (for $e_j \in T_i$). The set of edges of $T_i$ that attract $e_{i+1}$ form a path from the centroid to the vertex $v$ that is adjacent to $e_{i+1}$. This is evident from the definition of attraction and repulsion. Since there is a unique path from the centroid to the vertex $v$, there is a unique way to attach the edge $e_{i+1}$ to the forest $T_i$. Thus we extend $T_i$ to $T_{i+1}$ by placing $e_{i+1}$ at the end of the path created by edges that attract it.

At first glance, the conditions outlined in Theorem 5.7 seem very strict, and one might wonder whether they may be relaxed. For example, perhaps it is possible to construct $T$ knowing only $\theta_T(e)$ for all edges $e$ of $T$. This is not the case. Consider the following trees.

![Figure 13: Two trees with the same $\{\theta_T(e)\}_{e \in E(G)}$ data](image)

Notice that both graphs have exactly one edge that divides the vertices into a (4, 3) partition, one that divides the vertices into a (5, 2) partition, and four edges that divide the vertices in a (6, 1) partition. Therefore, knowing $\theta_T(e)$ for all the edges $e$ in a tree $T$ is insufficient information to construct a tree.

However, in the next theorem we show that we can reconstruct $T$ knowing only $\theta_T(e_i, e_j)$ for all pairs of edges $e_i, e_j$ in $T$ in the case that $T$ has only one centroid. In the proof of the theorem we will refer to an edge adjacent to a vertex of degree 1 as a leaf-edge. In the literature it is the vertex that is called a leaf, but to simplify the exposition we will refer to the edge as a leaf-edge.

**Theorem 5.8.** A tree $T$ with a single centroid is uniquely determined by the data $\theta_T(e_i, e_k)$ for all distinct edges $e_i, e_k \in T$.

**Proof.** The number of vertices in the tree is determined by the data $\{\theta_T(e_i, e_j)\}$ by counting the number of edges and adding one. Because there is only one single-centroid tree on $n$ vertices for $n \leq 4$, we may assume that $n > 4$. We will prove the statement by first showing that we can determine precisely which edges of $T$ must be leaf-edges, then using this information to determine $\theta_T(e)$ for all $e \in E(T)$. Invoking Theorem 5.7 will complete the proof.
Let \( L \subseteq E(T) \) consist of all edges \( e_i \in E(T) \) such that

\[
\# \{ e_j : e_j \neq e_i, \theta(\{e_i, e_j\}) = (n-2,1,1) \} \geq 2.
\]

We claim that \( L \) is the set of leaf-edges of \( T \). To show that every leaf-edge is in \( L \), we study two cases: \( T \) has three or more leaves or \( T \) has two leaves. If \( T \) has three or more leaves then for any leaf-edge \( e_i \) there are at least two ways to remove another leaf-edge and get the partition \((n-2,1,1)\). If \( T \) has only two leaves, then \( T \) is a path on \( n > 4 \) vertices. In this case, for any leaf-edge \( e_i \) we get the partition \((n-2,1,1)\) by removing \( e_i \) and the edge incident to it, or by removing \( e_i \) and the other leaf-edge. Hence, any leaf-edge is contained in \( L \). Suppose now that an edge \( e_i \) is not a leaf-edge, we will show that \( e_i \notin L \). Indeed, if \( e_i \) is a leaf-edge, then \( \theta(\{e_i, e_j\}) = (n-2,1,1) \) only if \( e_j \) is a leaf-edge that meets \( e_i \) at a vertex of degree 2. Therefore, unless \( e_i \) is adjacent to two leaf-edges through vertices of degree 2, it is not in \( L \). The only case where this phenomenon could occur is in the path with four vertices, contradicting \( n > 4 \). Therefore, the set \( L \) consists precisely of the leaves of \( T \).

Next, consider any \( e_k \notin L \). We can completely determine \( \theta_T(e_k) \) by studying how the vertices of \( T \) are partitioned when \( e_k \) and a leaf-edge are removed. That is, it is clear that there is at least one leaf-edge in \( T \) that repels \( e_k \), and at least one leaf-edge in \( T \) that attracts \( e_k \). Because every leaf-edge in \( T \) either attracts \( e_k \) or repels \( e_k \), it follows from Proposition 5.4 that \( \{ \theta_T(e_k, e_j) : e_j \in L \} = \{ \text{re}(n-i-1, i, 1), \text{re}(n-i, i-1, 1) \} \) for some \( i \). By letting \( a = \max\{n-i, i\} \), we see that \( \theta_T(e_k) = (a, n-a) \).

We have determined \( \theta_T(e_k) \) for all \( e_k \in E(T) \). Using Theorem 5.7 completes the proof.

To illustrate Theorem 5.8, suppose we are given the following doubleton sets of \( \theta_T \)-images of edges of a tree with a single centroid, and we are asked to construct the tree.

| E_1 | E_2 | E_3 | E_4 | E_5 | E_6 | E_7 | E_8 | E_9 | E_10 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 10  | 2  | 11  | 3  | 4   | 5   | 6   | 7   | 8   | 9    |
| 14  | 11 | 15  | 2   | 17  | 3   | 18  | 4   | 19  | 5    |
| 16  | 13 | 17  | 2   | 20  | 4   | 21  | 6   | 22  | 5    |
| 17  | 12 | 18  | 1   | 20  | 3   | 21  | 1   | 22  | 6    |

Table 5.4: \( \theta_T \)-images of two-element sets

The first step is to identify the leaves. Following the proof of Theorem 5.8, an edge \( e \) is a leaf if and only if there are at least two other edges \( e_i \) and \( e_j \) for which \( \theta_T(e, e_i) = \theta_T(e, e_j) = (11,1,1) \). The leaves of our tree are the edges \( L = \{e_1, e_5, e_6, e_7, e_8, e_{11}\} \). Therefore, \( \theta_T(e) = (12,1) \) for these edges.

The next step is to determine the \( \theta_T \)-images of all other edges in the tree. For an edge \( e \notin L \), we check the set of partitions \( \{ \theta_T(e, e_j) : e_j \in L \} \). For example, this set of partitions for \( e_{10} \) is \( \{ (9,3,1), (10,2,1) \} \). Since 10 is the maximum number in any of these partitions, \( \theta_T(e_{10}) = (10,3) \). We can use similar reasoning to deduce the \( \theta_T \)-image of every edge.

| Edge | E_1 | E_2 | E_3 | E_4 | E_5 | E_6 | E_7 | E_8 | E_9 | E_10 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 11   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1    |
| 12   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2    |
| 13   | 3   | 3   | 3   | 3   | 3   | 3   | 3   | 3   | 3   | 3    |
| 14   | 4   | 4   | 4   | 4   | 4   | 4   | 4   | 4   | 4   | 4    |
| 15   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5    |
| 16   | 6   | 6   | 6   | 6   | 6   | 6   | 6   | 6   | 6   | 6    |

Table 5.4: \( \theta \)-images of the edges in Table 5.4
The next step is to index the edges \( \{e'_i\}_{i=1}^{n-1} \) so that \( \theta_T(e'_i) \geq \theta_T(e'_{i+1}) \). We will use the indexing \( e'_1, e'_2, \ldots, e'_{12} = e_3, e_{12}, e_9, e_{10}, e_2, e_4, e_1, e_6, e_7, e_8, e_{11} \).

This is the order in which the edges will be added. Notice that we have some freedom in our choice of edge ordering. For example, we may swap the position of \( e_9 \) and \( e_{10} \) without changing the result of this example. The forests \( T_1, T_2, \ldots, T_{12} \) are shown below (their isolated vertices are omitted for clarity). To create \( T_{i+1} \) from \( T_i \), we add the edge \( e'_{i+1} \) by checking which edges attract it. The attracting edges are in bold, and both the centroid \( c \) and the newly added edge \( e'_{i+1} \) are labelled.

\[
\begin{align*}
T_1 &= e_3 & T_2 &= e_{12} & T_3 &= e_9 \\
T_4 &= e_{10} & T_5 &= e_2 & T_6 &= e_4 \\
T_7 &= e_1 & T_8 &= e_5 & T_9 &= e_6 \\
T_{10} &= e_7 & T_{11} &= e_8 & T_{12} &= e_{11}
\end{align*}
\]

Figure 5.5: Construction of the tree satisfying the \( \theta_T \)-images in Table 5.4

The forest \( T_{12} \) is the unique tree with a single centroid satisfying the data in Table 5.4. Theorem 5.8 shows that all single-centroid trees are classified by the \( \theta \) images of 2-element sets of edges. One might ask whether we may lift the restriction in Theorem 5.8 that the tree must have a single centroid. Unfortunately, we cannot. Figure 5.6 shows an example of two labelled graphs that have the same \( \theta \) images of 2-element sets of edges. A table of all the 2-element sets is included for reference.

| \( e_1 \) | \( e_2 \) | \( e_3 \) | \( e_4 \) | \( e_5 \) | \( e_6 \) | \( e_7 \) | \( e_8 \) | \( e_9 \) | \( e_{10} \) | \( e_{11} \) | \( e_{12} \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( 7, 6, 1 \)   | \( 10, 3, 1 \)   | \( 12, 1, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   | \( 10, 3, 1 \)   |
The data of the \( \theta_T \)-images of sets of edges of a tree are related to the data of the coefficients of its chromatic symmetric function. To understand this relationship, recall that when we write the chromatic symmetric function of any graph using the basis of power symmetric functions, each coefficient encapsulates information about subgraphs of a certain vertex partition type. For a graph with cycles, it is possible that two subgraphs have the same vertex partition type, but a different parity in their number of edges, so that their contributions to the sum in Equation (2) cancel one another out. This does not occur in forests.

**Proposition 5.9.** For forests \( F \) and \( H \), \( X_F = X_H \) if and only if

\[
\#\{S \subseteq E(F) : \pi(S) = \lambda\} = \#\{S \subseteq E(H) : \pi(S) = \lambda\}
\]

for all partitions \( \lambda \) of \( \#V(F) \).

**Proof.** All induced subgraphs of a forest with a given vertex partition have the same number of connected components, and therefore the same number of edges. Let \( k \) be the number of parts in a partition \( \lambda \) of \( \#V(F) \). Note that \( \#S = \#V(F) - k \). Then the coefficient of \( p_\lambda \) in \( X_F \) is equal to

\[
(-1)^{\#V(F)-k} \#\{S \subseteq E(F) : \pi(S) = \lambda\}.
\]

The proposition follows immediately.

Notice that \( \pi(S) = \theta_F(E(F) - S) \), so a corollary of Proposition 5.9 is the following.

**Corollary 5.10.** For forests \( F \) and \( H \), \( X_F = X_H \) if and only if

\[
\#\{S \subseteq E(F) : \theta_F(S) = \lambda\} = \#\{S \subseteq E(H) : \theta_F(S) = \lambda\}
\]

for all partitions \( \lambda \) of \( \#V(F) \).

It is very tempting to misinterpret Theorem 5.8 and claim that together with Corollary 5.10 it proves that \( X_T \) determines trees with single centroids. After all, it seems that the \( \theta_T \)-images of 2-element sets of \( E(T) \) are determined from \( X_T \) by the above corollary. However, it is important to remember that only the *number* (and not the labels) of pairs of edges giving a certain vertex partition under \( \theta_T \) is determined from \( X_T \). To illustrate this important distinction, consider the trees in Figure 14 and the chart of their \( \theta_T \)-images of 2-element sets. Each 3-part partition appears the same number of times in each chart, but they are arranged differently within the chart. The trees in Figure 14 are not isomorphic and it can be checked that their chromatic symmetric functions are not equal: the entire function is too long to print here, but when written in the power-sum basis the tree on the left has \(-9p_{(8,5,1,1)}\) as a summand while the tree on the right has \(-8p_{(8,5,1,1)}\) as a summand.

What Theorem 5.8 and Corollary 5.10 *do* imply is that if there does exist a pair of distinct trees with the same chromatic symmetric function and a single centroid, then it must also share this special property with the trees in Example 14. That is, the partitions appearing as the \( \theta_T \)-images of 2-element sets must agree, though their relative location within the table might be “scrambled.”
Figure 14: Two trees with nearly identical $\theta_T$ images of 2-element sets

| $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_8$ | $e_9$ | $e_{10}$ | $e_{11}$ | $e_{12}$ | $e_{13}$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|---------|---------|---------|
| 12    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 13    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 1    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 12    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 13    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |
| 1    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1       | 1       | 1       | 1       |

Figure 5.8: $\theta_T$ images of 2-element sets for $T_1$ (above) and $T_2$ (below)

References

[1] T. Chow. *Symmetric function generalizations of graph polynomials*, Ph. D. thesis, MIT (1995).

[2] J. A. Bondy, U. S. R. Murty. *Graph Theory*. Graduate Texts in Mathematics, 244. Springer, New York, 2008.

[3] R. Diestel. *Graph Theory*. Third Edition. Graduate Texts in Mathematics, 173. Springer-Verlag, Berlin, 2005.

[4] D. Gebhard and B. Sagan. *A chromatic symmetric function in noncommuting variables*, J. Alg. Combin. 2 (2001), 227-255.

[5] J. Fouger On symmetric chromatic polynomials of trees, undergraduate thesis, Dartmouth College, 2003.

[6] F. Harary. *Graph Theory*, Addison-Wesley Publishing Co., Reading, Mass.-Menlo Park, Calif.-London 1969.

[7] J. Martin, M. Morin and J. Wagner. *On distinguishing Trees by their Chromatic Symmetric Functions*, Journal of Combinatorial Theory, Series A 115, 237-253 (2008).
[8] M. Morin, *Caterpillars, ribbons, and the chromatic symmetric function*, M.S. thesis, University of British Columbia, 2005.

[9] R. Stanley. *A symmetric function generalization of the chromatic polynomial of a graph*, Adv. Math 111, 166-194 (1995).

[10] R. Stanley, Richard. *Graph colorings and related symmetric functions: ideas and applications. A description of results, interesting applications, and notable open problems*, Discrete Math., 1998

[11] R. Stanley. *Enumerative combinatorics*. Vol. 2. With a foreward by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999.