Chromatic Symmetric Functions and Polynomial Invariants of Trees

José Aliste-Prieto[1] Jeremy L. Martin[2]
Jennifer D. Wagner[3] José Zamora[1]

[1]Universidad Andres Bello, Chile
[2]University of Kansas, USA
[3]Washburn University, USA

CombinaTexas
Texas A&M University
23 March 2024

Chromatic Symmetric Functions of Graphs

Let $G = (V, E)$ be a simple graph with $V = [n] = \{1, \ldots, n\}$.

proper coloring: a function $f : V \to \mathbb{N}_{>0}$ such that $f(i) \neq f(j)$ whenever $ij \in E$.

chromatic symmetric function (CSF): the power series

$$\mathbf{X}_G = \mathbf{X}_G(x_1, x_2, \ldots) = \sum_{\substack{f:V\to\mathbb{N}_{>0} \\ \text{proper}}} x_{f(1)} \cdots x_{f(n)}.$$

- Symmetric and homogeneous of degree n
- Generalizes the chromatic polynomial:

$$\mathbf{X}_G(1^k, 0^\infty) = \text{ number of proper } k\text{-colorings}$$

Chromatic Symmetric Functions of Graphs

- Introduced by Stanley in 1995
- **Related invariants:** Tutte symmetric function / U-polynomial (Noble–Welsh 1999), matroid quasisymmetric function (Billera–Jia–Reiner 2009)
- **Analogues:** noncommutative CSFs (Gebhard–Sagan 2001), quasisymmetric CSFs (Shareshian–Wachs 2016), . . .
- **Applications:** combinatorial Hopf algebras (Aguiar–Bergeron–Sottile 2006), cohomology of Hessenberg subvarieties of flag manifolds (Shareshian–Wachs 2012)

The Merest Glimpse of Symmetric Functions

symmetric function: a power series in $x_1, x_2, \ldots$ that is invariant under permuting the indeterminates.

partition $\lambda \vdash n$: non-increasing sequence of positive integers $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ that add up to n.

monomial SF: $m_\lambda =$ sum of all monomials whose nonzero exponents are the parts of n.

$$m_{322} = x_1^3 x_2^2 x_3^2 + \cdots + x_1^2 x_3^3 x_6^2 + \cdots + x_8^2 x_9^2 x_{13}^3 + \cdots$$

power-sum SF: $p_\lambda = p_{\lambda_1} \cdots p_{\lambda_\ell}$, where $p_k = \sum_{i=1}^{\infty} x_i^k$.

$$p_{322} = (x_1^3 + x_2^3 + \cdots)(x_1^2 + x_2^2 + \cdots)^2$$

- Both $\{m_\lambda \mid \lambda \vdash n\}$ and $\{p_\lambda \mid \lambda \vdash n\}$ are bases for the space of homogeneous degree-n SFs.

Example 1: K_n: complete graph on n vertices

- Proper coloring = injective map $[n] \to \mathbb{N}$

$$\mathbf{X}_{K_n} = \sum_{S \subseteq \mathbb{N},\ |S|=n} n! \prod_{i \in S} x_i = n! e_n = n! m_{1^n}$$

Example 2: $\overline{K_n}$: n vertices, no edges

- Proper coloring = *any* map $[n] \to \mathbb{N}$

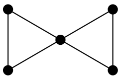
$$\mathbf{X}_{\overline{K_n}} = (x_1 + x_2 + \cdots)^n = p_{1^n}$$

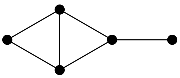
Some Chromatic Symmetric Functions

Example 3: P_3 = path with three vertices

$$\mathbf{X}_{P_3} = \sum_{i<j<k} 6x_ix_jx_k + \sum_{i\neq j} x_i^2x_j = 6m_{111} + m_{21}$$

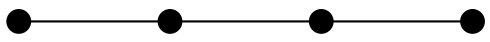
Example 4: [Stanley] These graphs have the same CSF:



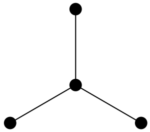


Some Chromatic Symmetric Functions

Example 5: The two trees with four vertices.



path P_4



star St_4

$$\mathbf{X}_{P_4} = 24m_{1111} + 6m_{31} + 2m_{22}$$
$$\mathbf{X}_{St_4} = 24m_{1111} + 6m_{31} + m_{31}$$

Every tree with n vertices has chromatic polynomial $t(t-1)^{n-1}$.
So the CSF is a strictly stronger invariant. **How much stronger?**

Stanley's Uniqueness Problem

Question (Stanley)
Is a tree uniquely determined up to isomorphism by its CSF?

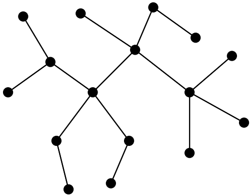
I.e., if T, T' are non-isomorphic trees, must $X(T) \neq X(T')$?

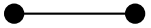
No one really has any idea (although some experts have opinions).

- The answer is yes for $n \leq 29$ [Heil–Ji 2019].
- Also yes for various very special classes of trees

Let $G = (V, E)$ be a graph, $n = |V|$, $A \subseteq E$

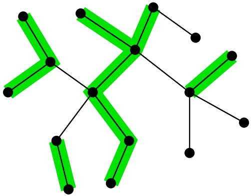
type of A = partition of n whose parts are component sizes of $G|_A$

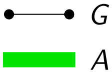




Let $G = (V, E)$ be a graph, $n = |V|$, $A \subseteq E$

type of A = partition of n whose parts are component sizes of $G|_A$

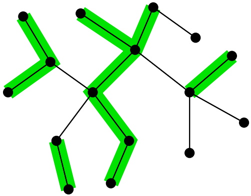


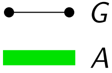


Expanding CSFs in the Power Sum Basis

Let $G = (V, E)$ be a graph, $n = |V|$, $A \subseteq E$

type of A = partition of n whose parts are component sizes of $G|_A$

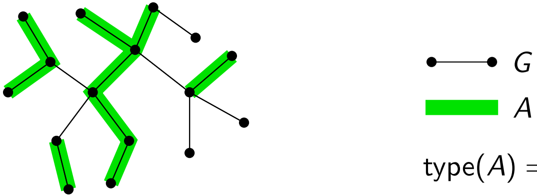




$\text{type}(A) = (6, 3, 2, 2, 1, 1, 1)$

Expanding CSFs in the Power Sum Basis

Let $G = (V, E)$ be a graph, $n = |V|$, $A \subseteq E$
type of A = partition of n whose parts are component sizes of $G|_A$



$\text{type}(A) = (6, 3, 2, 2, 1, 1, 1)$

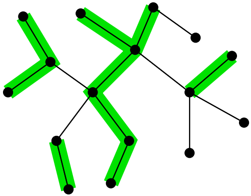
Theorem (Stanley 1995)

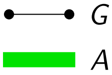
$$\mathbf{X}_G = \sum_{A \subseteq E} (-1)^{n-|A|} p_{\text{type}(A)}.$$

Expanding CSFs in the Power Sum Basis

Let $G = (V, E)$ be a graph, $n = |V|$, $A \subseteq E$

type of A = partition of n whose parts are component sizes of $G|_A$





$\text{type}(A) = (6, 3, 2, 2, 1, 1, 1)$

Theorem (Stanley 1995)

$$\mathbf{X}_G = \sum_{A \subseteq E} (-1)^{n-|A|} p_{\text{type}(A)}.$$

Corollary If $T = (V, E)$ is a tree, then $\ell(\text{type}(A)) = n - |A|$, so there is no cancellation:

$$[p_\lambda]\mathbf{X}_T = c_\lambda(T) = (-1)^{\ell(\lambda)} \#\{A \subseteq E \mid \text{type}(A) = \lambda\}.$$

The Subtree Polynomial

Let $T = (V, E)$ be a tree. For a subtree $S \subseteq T$, define

$$\begin{aligned} e(S) &= \text{ number of edges of } S \\ \ell(S) &= \text{ number of leaf edges of } S \end{aligned}$$

(Henceforth "subtree" means "subtree with at least one edge.)

The **subtree polynomial** (STP) of T is

$$\mathbf{S}_T = \sum_{\text{subtrees } S \subseteq T} q^{e(S)} r^{\ell(S)}$$

For instance:

$$\begin{aligned} \mathbf{S}_{P_n} &= (q^{n-1} + 2q^{n-2} + \cdots + (n-2)q^2)r^2 + (n-1)qr \\ \mathbf{S}_{St_n} &= (qr+1)^{n-1} - 1 \end{aligned}$$

The Subtree Polynomial

Theorem [JLM–Morin–Wagner 2008]
The STP can be obtained linearly from the CSF:

$$[q^i r^j]\mathbf{S}_T = \sum_{\lambda \vdash n} \phi(\lambda, i, j) c_\lambda(T).$$

where $\phi(\lambda, i, j)$ is independent of T.

Corollary
The CSF of a tree T determines its degree and distance sequences, i.e., the numbers

$$\#\{v \in V \colon \operatorname{deg}(v) = k\}, \qquad \#\{(v, w) \colon \operatorname{dist}(v, w) = k\}.$$

Distinguishing Power of the STP

The STP is a strictly weaker isomorphism invariant than the CSF.
These two trees are an **S-pair** (same STP, different CSF):

n	11	12	13	14	15	16	17	18
trees	235	551	1301	3159	7741	19320	48629	123867
S-pairs	1	1	1	5	1	7	17	15

The Generalized Degree Polynomial of a Tree

The **generalized degree polynomial (GDP)** of T is

$$\mathbf{G}_T = \mathbf{G}_T(x, y, z) = \sum_{A \subseteq V} x^{|A|} y^{d(A)} z^{e(A)}.$$

where

$$e(A) = \text{ number of internal edges (both endpoints in } A)$$
$$d(A) = \text{ number of external edges (one endpoint in } A)$$

- In particular, $d(\{v\}) = \deg(v)$.

Conjecture [Crew 2022]
The CSF of a tree determines its GDP.

The Half-Generalized Degree Polynomial

The **half-generalized degree polynomial (HDP)** of T is

$$\mathbf{H}_T = \mathbf{H}_T(x,y,z) = \sum_{\substack{A \subseteq V \\ T[A] \text{ connected}}} y^{d(A)} z^{e(A)}$$

$$= \sum_{\text{subtrees } S \subseteq T} y^{d(S)} z^{e(S)}$$

$$= \left(\text{sum of terms of } \mathbf{G}_T \text{ of the form } x^{c+1} y^b z^c\right)\Big|_{x=1}$$

Theorem [Wang-Yu-Zhang 2023]
The CSF of a tree determines its HDP.

(Key tool: use Stanley's formula for $\dfrac{\partial \mathbf{X}_T}{\partial p_k}$.)

Our Results

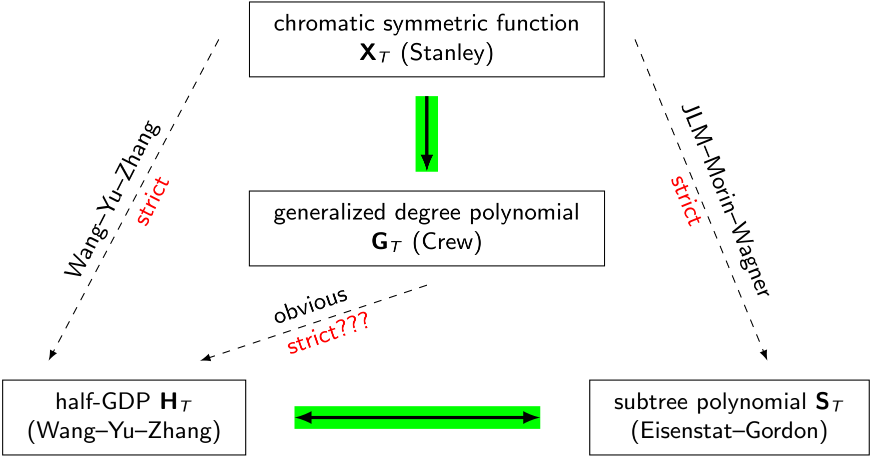
Theorem [Aliste-Prieto, JLM, Wagner, Zamora 2024^+]

1. *The CSF of a tree determines its GDP linearly*.
(This is Crew's conjecture.)

2. *The HDP and the STP of a tree determine each other.*
(Together, these two results imply the 2008 theorem of JLM–Morin–Wagner.)

3. *There exist arbitrarily large sets of trees with the same STP.*
(This implies a 2006 conjecture of Eisenstat and Gordon.)

Our Results


chromatic symmetric function
$\mathbf{X}_T$ (Stanley)
generalized degree polynomial
$\mathbf{G}_T$ (Crew)
half-GDP $\mathbf{H}_T$
(Wang–Yu–Zhang)
subtree polynomial $\mathbf{S}_T$
(Eisenstat–Gordon)
Wang–Yu–Zhang
strict
JLM–Morin–Wagner
strict
obvious
strict???

Theorem [APMWZ 2024^+]
The coefficients

$$g_T(a,b,c) = \#\{A \subseteq V(T) \colon \ |A| = a,\ d(A) = b,\ e(A) = c\}$$

of $\mathbf{G}_T$ are given by

$$g_T(a,b,c) = \sum_{\lambda \vdash n} c_\lambda(T)\omega(\lambda,a,b,c)$$

where $c_\lambda(T) = [p_\lambda]\mathbf{X}_T$ and

$$\omega(\lambda,a,b,c) = (-1)^{n-b-1} \sum_{\mu \vdash a} \binom{a-\ell(\mu)}{c} \binom{\lambda}{\mu} \binom{n-\ell(\lambda)+\ell(\mu)-a}{n-b-c-1}$$

and

$$\binom{\lambda}{\mu} := \prod_{i=1}^{n} \binom{\#\text{ of parts of } \lambda \text{ equal to } i}{\#\text{ of parts of } \mu \text{ equal to } i}$$

How to Prove Crew's Conjecture

1. Hope that the conjecture was true.
2. Compute the matrices of coefficients

   $$X = [c_\lambda(T)]_{T \in \mathcal{T}_n,\ \lambda \vdash n} \qquad G = [g_T(a,b,c)]_{T \in \mathcal{T}_n,\ \lambda \vdash n}$$

   Do this until the computer gets tired.
3. Solve the matrix equation $X\Omega = G$ for Ω (there will be a large solution space).
4. Find a ~~needle~~ matrix Ω in the ~~haystack~~ solution space whose entries have a nice combinatorial form.
5. Prove a theorem.

Equivalence of the HDP and STP

The HDP and the STP can be written as

$$\mathbf{H}_T = \sum_{b,c} h_T(b,c)\, y^b\, z^c, \qquad \mathbf{S}_T = \sum_{i,j} s_T(i,j)\, q^i\, r^j$$

where

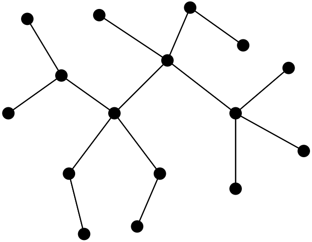
$$h_T(b,c) = |\{U \subseteq T\colon\ d(U) = b,\ e(U) = c\}|,$$
$$s_T(i,j) = |\{S \subseteq T\colon\ e(S) = i,\ \ell(S) = j\}|.$$

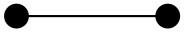
Computational evidence ($n \leq 18$) suggested to us that

$$\mathbf{H}_T = \mathbf{H}_{T'} \iff \mathbf{S}_T = \mathbf{S}_{T'}.$$

However, the needle-in-a-haystack method leads to very ugly matrices (with non-integer entries).

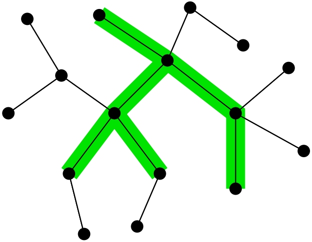
Equivalence of the HDP and STP

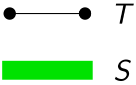




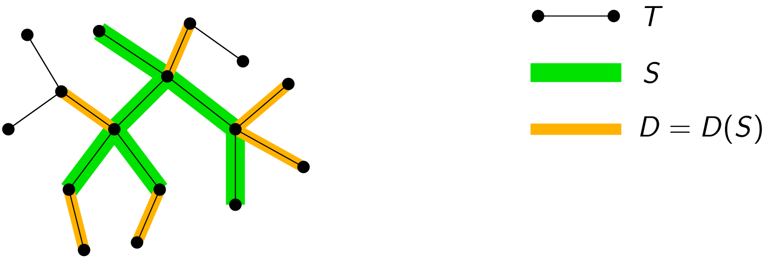
T

Equivalence of the HDP and STP

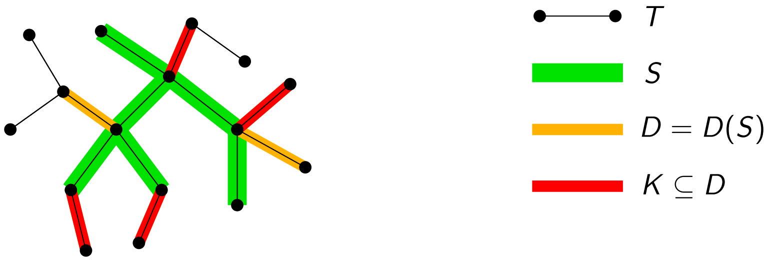




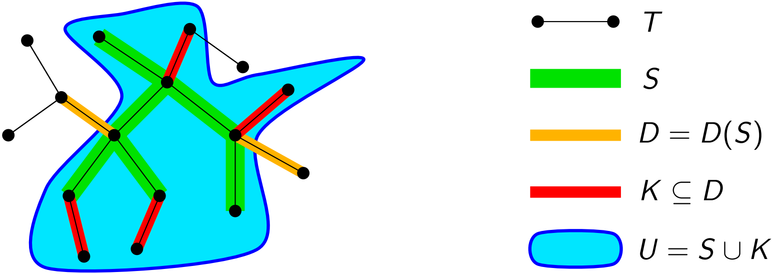
Equivalence of the HDP and STP



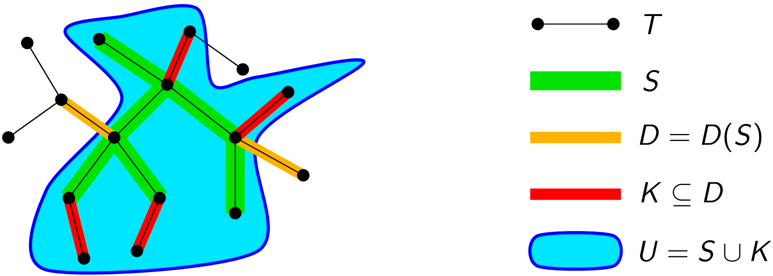
Equivalence of the HDP and STP



Equivalence of the HDP and STP



Equivalence of the HDP and STP



We have a bijection

$$\{(S, K) \colon S \subseteq T,\ K \subseteq D(S)\} \xrightarrow{\xi} \{(U, K) \colon U \subseteq T,\ K \subseteq L(U)\}$$

$$(S, K) \longmapsto (S \cup K, K)$$

$$(U \setminus K, K) \longleftarrow (U, K)$$

Equivalence of the HDP and STP

The bijection implies the equalities

$$\sum_{b=k}^{n-1-a} \binom{b}{k} h(a,b) = \sum_{j=k}^{n-1} \binom{j}{k} s(a+k,j)$$

for all a, k. In matrix form, $MH = NS$, where

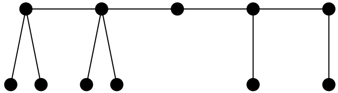
$$H = [h(a,b)]_{a,b=1}^{n}, \qquad S = [s(i,j)]_{i,j=1}^{n}.$$

- Entries of M and N are binomial coefficients
- M is unitriangular, hence invertible over $\mathbb{Z}$
- $\det N = n!$ by Gessel–Viennot lattice path theory

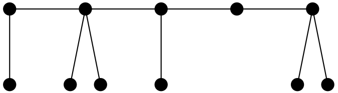
In particular, H and S, hence $\mathbf{H}_T$ and $\mathbf{S}_T$, determine each other.
(But $M^{-1}N$ and $N^{-1}M$ are not combinatorially nice!)

Caterpillars

A tree is a **caterpillar** if deleting all its leaves produces a path.



C(3, 3, 1, 2, 2)



C(2, 3, 2, 1, 3)

Caterpillars are indexed by compositions with both first and last parts > 1, up to reversal.

Eisenstat and Gordon conjectured that for gap-free polynomials $p(x)$, the caterpillars arising from $(a + bx)p(x)$ and $(b + ax)p(x)$ have the same STP:

$$(2 + 1x)(1 + x + x^3) = \mathbf{2} + 3x + x^2 + 2x^3 + \mathbf{1}x^4 \rightsquigarrow (\mathbf{3}, 3, 1, 2, \mathbf{2})$$
$$(1 + 2x)(1 + x + x^3) = \mathbf{1} + 3x + 2x^2 + x^3 + \mathbf{2}x^4 \rightsquigarrow (\mathbf{2}, 3, 2, 1, \mathbf{3})$$

Caterpillars and Unique Factorization

For compositions $\alpha = (a_1, \ldots, a_k)$ and $\beta = (b_1, \ldots, b_m)$, define

$$\alpha \cdot \beta = (a_1, \ldots, a_k, b_1, \ldots, b_m)$$
$$\alpha \odot \beta = (a_1, \ldots, a_{k-1}, a_k + b_1, b_2, \ldots, b_m)$$
$$\alpha \circ \beta = \beta^{\odot a_1} \cdot \beta^{\odot a_2} \cdots \beta^{\odot a_k}$$

Example

$(2,1) \circ (2,1) = (2,1)^{\odot 2} \cdot (2,1)^{\odot 1} = (2,3,1) \cdot (2,1) = (2,3,1,2,1)$
$(2,1) \circ (1,2) = (1,2)^{\odot 2} \cdot (1,2)^{\odot 1} = (1,3,2) \cdot (1,2) = (1,3,2,1,2)$

Fact [Billera–Thomas–van Willigenburg 2006]

Every composition α admits a unique irreducible factorization $\alpha = \alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_k$.

The Eisenstat-Gordon Conjecture

Theorem [APMWZ]

Reversing any of the irreducible factors preserves the subtree polynomial of $\mathrm{C}(1 \odot \alpha \odot 1)$.

For example,

$$\begin{cases} (2,1) \circ (2,1) = (\mathbf{2},3,1,2,\mathbf{1}) \\ (2,1) \circ (1,2) = (\mathbf{1},3,2,1,\mathbf{2}) \end{cases} \implies \mathrm{C}(\mathbf{3},3,1,2,\mathbf{2}) = \mathrm{C}(\mathbf{2},3,2,1,\mathbf{3}).$$

In particular, if α has k irreducible factors then $\mathrm{C}(1 \odot \alpha \odot 1)$ is one of at least 2^{k-1} non-isomorphic caterpillars with the same subtree polynomial.

The case $k = 2$ implies the Eisenstat-Gordon conjecture.

Further Questions

Question 1: Is the GDP really a stronger invariant than the HDP?

To our surprise, the answer is "no" for $n \leq 18$. But,

- we don't know how one might construct two trees with the same HDP but different GDPs, and
- we don't see how to recover the rest of the GDP from the terms that appear in the HDP.

Further Questions

Question 2: Does factorization extend from caterpillars to more general trees?

Further Questions

Question 3: What can be said about the invariant

$$\mathbf{F}_T(y,z,r) = \sum_{\text{subtrees } S \subseteq T} y^{d(S)} z^{e(S)} r^{\ell(S)} \ ?$$

- $\mathbf{F}_T(y,z,1) = \mathbf{H}_T$ and $\mathbf{F}_T(1,q,r) = \mathbf{S}_T$.
- However, **F** is **strictly stronger** than either **H** or **S**.
- In fact, we have found no pair $T \ncong T'$ with $\mathbf{F}_T = \mathbf{F}_{T'}$.

Thank you!

Oh, and please read our preprint!

J. Aliste-Prieto, J.L. Martin, J.D. Wagner, and J. Zamora,
Chromatic symmetric functions and polynomial invariants of trees,
arXiv:2402.10333, 2024.

References

M. Aguiar, N. Bergeron, and F. Sottile, *Combinatorial Hopf algebras and generalized Dehn-Sommerville relations*, Compos. Math. **142** (2006), 1–30.

L.J. Billera, N. Jia, and V. Reiner, *A quasisymmetric function for matroids*, Eur. J. Combin. **30** (2009), 1727–1757.

L.J. Billera, H. Thomas, and S. van Willigenburg, *Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon Schur functions*, Adv. Math. **204** (2006), 204–240.

L. Crew, *A note on distinguishing trees with the chromatic symmetric function*, Discrete Math. **345** (2022), Paper No. 112682

D. Eisenstat and G. Gordon, *Non-isomorphic caterpillars with identical subtree data*, Discrete Math. **306** (2006), 827–830.

D.D. Gebhard and B.E. Sagan, *A chromatic symmetric function in noncommuting variables*, J. Algebraic Combin. **13** (2001), no. 3, 227–255.

I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. Math. **58** (1985), 300–321.

S. Heil and C. Ji, *On an algorithm for comparing the chromatic symmetric functions of trees*, Australas. J. Combin. **75** (2019), 210–222.

J.L. Martin, M. Morin, and J.D. Wagner, *On distinguishing trees by their chromatic symmetric functions*, J. Combin. Theory Ser. A **115** (2008), 237–253.

S.D. Noble and D.J.A. Welsh, *A weighted graph polynomial from chromatic invariants of knots*, Ann. Inst. Fourier (Grenoble) **49** (1999), 1057–1087.

R. Orellana and G. Scott, *Graphs with equal chromatic symmetric functions*, Discrete Math. **320** (2014), 1–14.

R.P. Stanley, *A symmetric function generalization of the chromatic polynomial of a graph*, Adv. Math. **111** (1995), 166–194.

J. Shareshian and M.L. Wachs, *Chromatic quasisymmetric functions*, Adv. Math. **295** (2016), 497–551.

Y. Wang, X. Yu, and X.-D. Zhang, *A class of trees determined by their chromatic symmetric functions*, arXiv:2308.03980, 2023.