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

SIAM Central States Section
University of Missouri, Kansas City
5 October 2024

Chromatic Symmetric Functions of Graphs

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

proper coloring: $f : V \to \mathbb{N}_{>0}$ with $f(i) \neq f(j)$ 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:

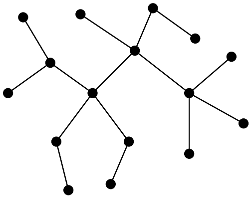
$$\mathbf{X}_G(1^k, 0^\infty) = \text{ number of proper } k\text{-colorings}$$

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)

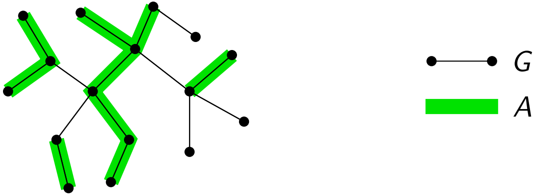
Describing CSFs Finitely

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

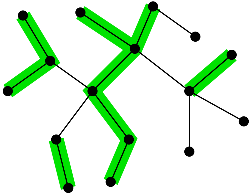


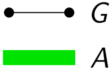
Describing CSFs Finitely

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 $G|_A$

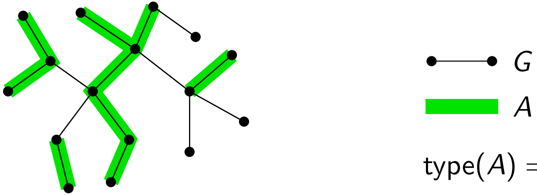




$\text{type}(A) = (6, 3, 2, 2, 1, 1, 1)$

Describing CSFs Finitely

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



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

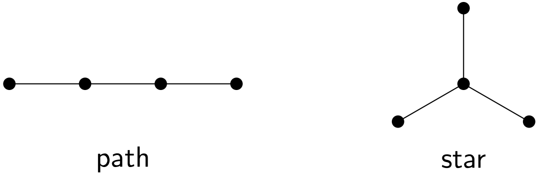
Theorem (Stanley 1995 / JLM–Morin–Wagner 2008)
The CSF of a tree T is determined by the **type numbers**

$$c_\lambda(T) = \#\{A \subseteq E \mid \text{type}(A) = \lambda\}$$

for all partitions λ of n.

Distinguishing Trees with the CSF

Example: The two trees with $n = 4$ are the path and the star.



Type numbers $c_\lambda(T)$:

λ	$(1,1,1,1)$	$(2,1,1)$	$(2,2)$	$(3,1)$	(4)
path	1	3	1	2	1
star	1	3	0	3	1

In fact $\mathbf{X}_T \neq \mathbf{X}_{T'}$ for all non-isomorphic trees T, T' with $n \leq 29$ [Heil–Ji 2019]

Question (Stanley)

Is a tree uniquely determined up to isomorphism by its CSF?

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

Or, stated more broadly:

Can the local structure of a tree be recovered from global data?

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 $X(T) \neq X(T')$?

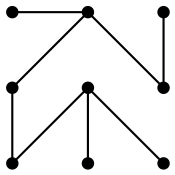
Or, stated more broadly:
Can the local structure of a tree be recovered from global data?

- There are infinitely many pairs of non-acyclic graphs with the same CSFs [Orellana–Scott 2016].
- All trees on n vertices have the same chromatic *polynomial*; the CSF is much stronger.

Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

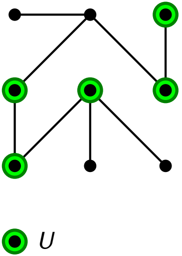
- type numbers (from CSF)



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

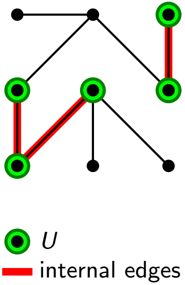
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

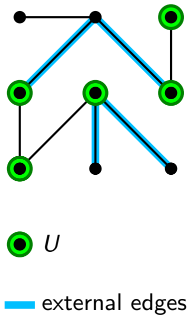
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

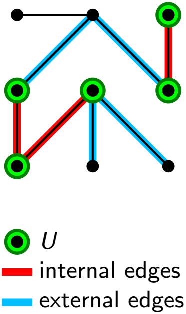
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

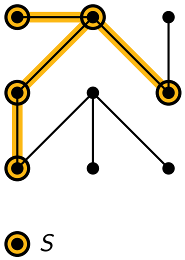
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

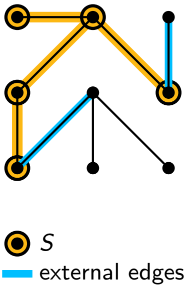
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

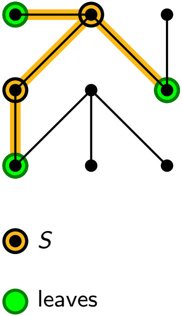
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

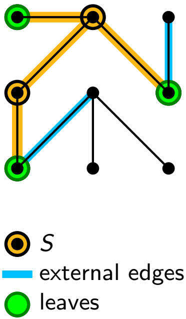
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

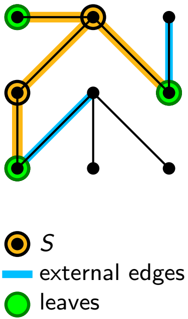
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

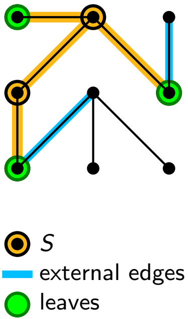
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves
- vertices of degree d



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

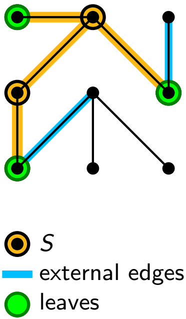
- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves
- vertices of degree d
- paths of length ℓ



Numerical Tree Invariants

Numerical data sets associated with a tree $T = (V, E)$:

- type numbers (from CSF)
- vertex sets $U \subseteq V$ with a vertices, b external edges, and c internal edges
- subtrees S with q edges and r external edges
- subtrees S with q edges and s leaves
- vertices of degree d
- paths of length ℓ



Which of these data sets determine the others?

Relationships Between Tree Invariants

chromatic symmetric function/ type numbers

vertex sets (size, internal edges, external edges)

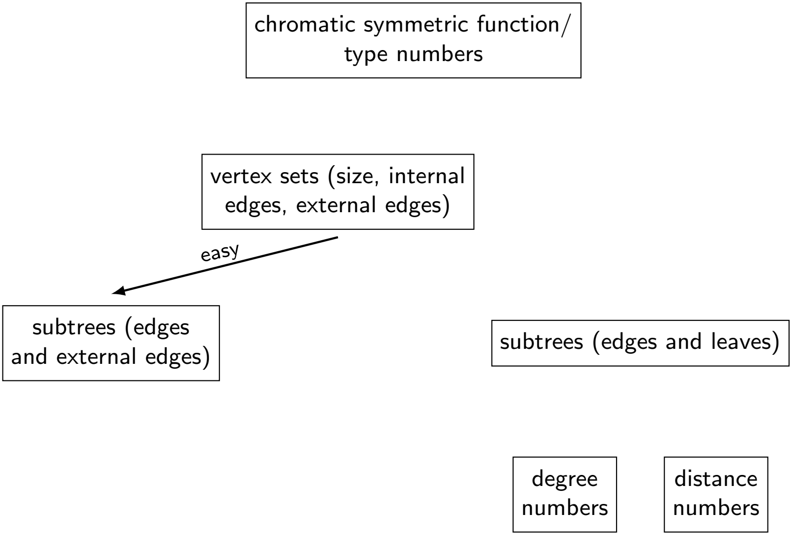
subtrees (edges and external edges)

subtrees (edges and leaves)

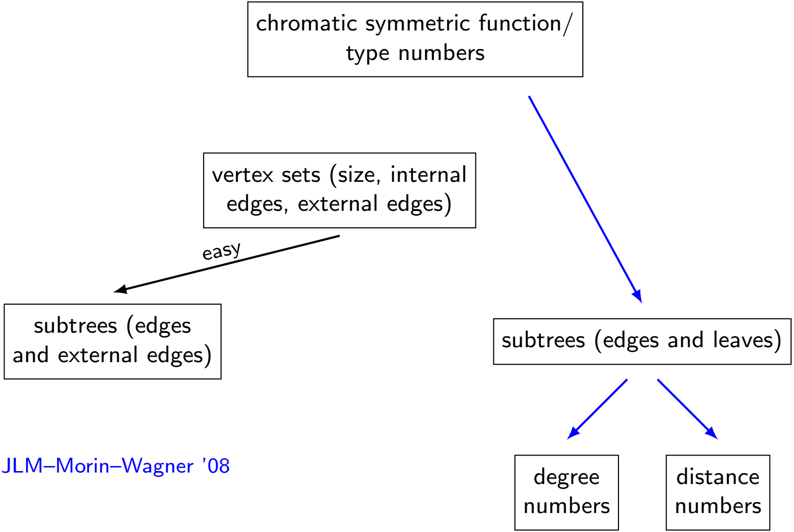
degree numbers

distance numbers

Relationships Between Tree Invariants

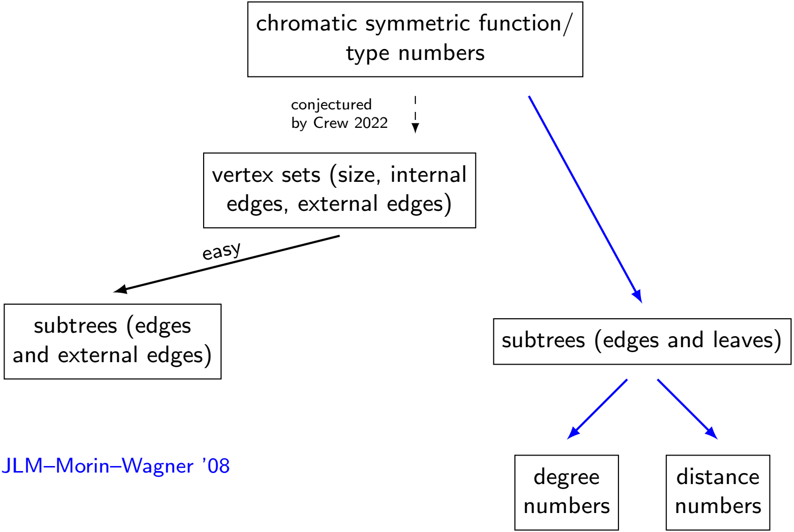


Relationships Between Tree Invariants



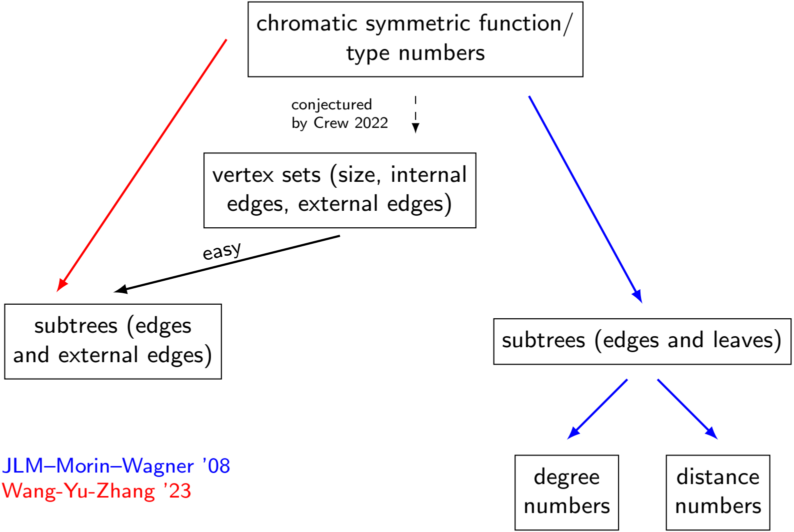
JLM–Morin–Wagner '08

Relationships Between Tree Invariants



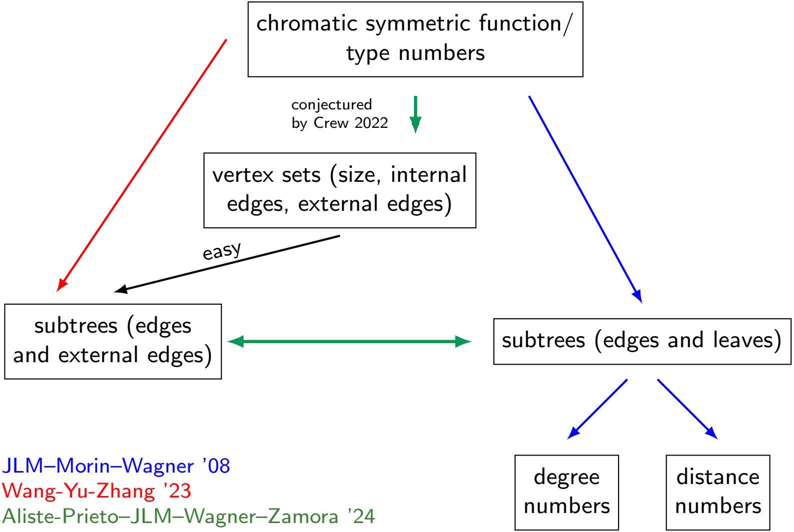
JLM–Morin–Wagner '08

Relationships Between Tree Invariants



JLM–Morin–Wagner '08
Wang-Yu-Zhang '23

Relationships Between Tree Invariants



JLM–Morin–Wagner '08
Wang-Yu-Zhang '23
Aliste-Prieto–JLM–Wagner–Zamora '24

Crew's Conjecture: Obtaining the GDP from the CSF

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

The CSF determines the vertex set data *linearly*.
That is, for every tree T, the numbers

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

(where $d(A)$ and $e(A)$ are the numbers of external and internal edges) are given by the formula

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

where $\omega(\lambda,a,b,c)$ are integers defined independently of T.

Similarly, the two sets of subtree data are linearly equivalent.

Overview of the Proof of Crew's Conjecture

1. Compute the matrices of coefficients

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

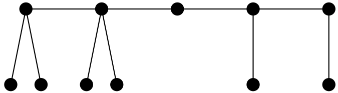
   for $n \leq 7$ or so.

2. Solve the matrix equation $X\Omega = G$ for Ω (there will be a large solution space).
3. Find a ~~needle~~ matrix Ω in the ~~haystack~~ solution space whose entries have a predictable combinatorial form.
4. Finish the proof (which mixes algebra and combinatorics).

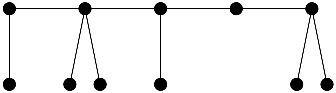
The proof of equivalence of the two subtree data sets is analogous.

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 edge/leaf subtree data.

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

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 [Aliste-Prieto, JLM, Wagner, Zamora 2024]

Reversing any of the irreducible factors in $1 \odot \alpha \odot 1$ produces a caterpillar with the same edge/leaf subtree counts.

For example,

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

In particular, if α has k irreducible factors then $\mathsf{C}(1 \odot \alpha \odot 1)$ 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: Does factorization extend from caterpillars to more general trees?

Further Questions

Question 1: Does factorization extend from caterpillars to more general trees?

Question 2: It is natural to consider the combined subtree data

$$\{\#\text{subtrees } S \text{ with } q \text{ edges, } r \text{ external edges, } s \text{ leaves}\colon q, r, s \in \mathbb{N}\}.$$

- This three-parameter subtree data is **strictly stronger** than either two-parameter data set
- Does it distinguish trees?
- Can we recover it from the CSF? (If so, not linearly.)

Thank you!

References

J. Aliste-Prieto, J.L. Martin, J.D. Wagner, and J. Zamora, *Chromatic symmetric functions and polynomial invariants of trees*, Bull. London Math. Soc., to appear; arXiv:2402.10333, 2024.

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.

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.

Okay, You Asked For It

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

where $\ell(\lambda)$ means the length (= number of parts) of λ, and

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