

# Coloring With a Limited Paintbox



*Stephanie van Willigenburg*

On October 23, 1852 Francis Guthrie asked his brother Frederick, a student at University College London, England, to ask his professor, Augustus De Morgan, whether four colors sufficed to color the countries of a map such that two countries sharing a border must be colored differently. De Morgan had no answer to what would become known as the four-color problem, but wrote that day to William Rowan Hamilton at Trinity College, Dublin, Ireland, saying:

Query cannot a necessity for five or more be invented.... If you retort with some very simple case which makes me out a stupid animal, I think I must do as the Sphinx did.<sup>1</sup>

Hamilton wrote back on 26 October 1852 (which MacTutor [12] wryly notes was “showing the efficiency of both himself and the postal service”) saying:

---

*Stephanie van Willigenburg is a professor of mathematics at The University of British Columbia. Her email address is [steph@math.ubc.ca](mailto:steph@math.ubc.ca).*

*Communicated by Notices Associate Editor Steven Sam.*

*For permission to reprint this article, please contact:*

*[reprint-permission@ams.org](mailto:reprint-permission@ams.org).*

DOI: <https://doi.org/10.1090/noti2490>

<sup>1</sup>*The Sphinx killed herself after Oedipus solved her riddle.*

I am not likely to attempt your quaternion of colour very soon.

In fact this problem would only be solved in 1976 courtesy of Kenneth Appel, Wolfgang Haken, John Koch, and the first major computer-assisted proof. Along the way many tools were developed to resolve it, including the chromatic polynomial [3], which eventually played no role in the solution but became its own avenue of research. Over the next few pages we will explore the chromatic polynomial, its generalization called the chromatic symmetric function, and two problems that, much like the four-color problem, are straightforward to state but remain open after 25 years. **The chromatic polynomial:** A limited paint palette with an unlimited amount of paint. As we might guess from the name, the chromatic polynomial is a polynomial arising from coloring, and rather than coloring maps we will be coloring graphs. A *graph*  $G$  consists of a set of *vertices*  $V(G)$  and a set of *edges*  $E(G)$  connecting the vertices. The number of edge ends meeting at a vertex  $v$  is its *degree*,  $\deg(v)$ , and two vertices connected by an edge are *adjacent*. The graphs of interest to us are *simple*, which means

- the number of vertices is finite;
- no pair of vertices has more than one edge connecting them (no *multiple edges*);

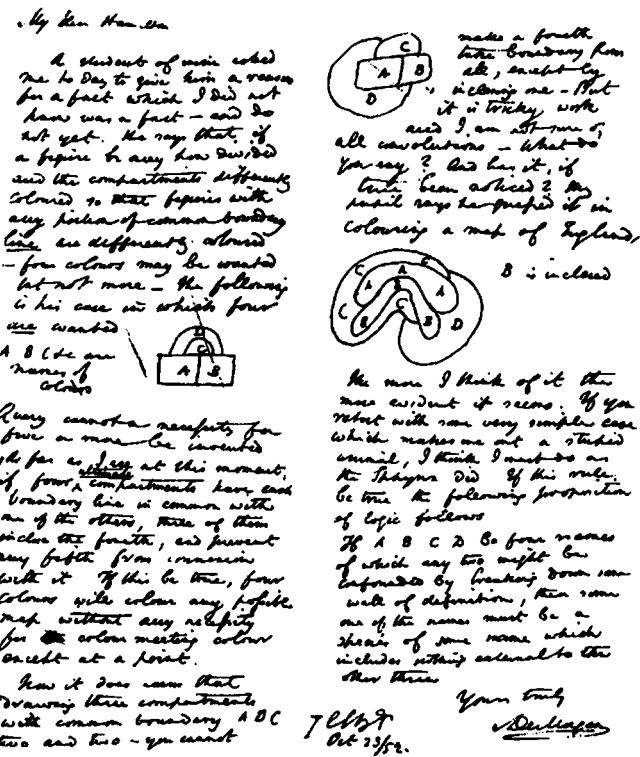


Figure 1. The letter from De Morgan to Hamilton [18].

- no vertex has an edge connecting it to itself (no loops).

We say a graph is *connected* if every pair of vertices has at least one sequence of edges connecting them, and *disconnected* otherwise. Three famous graph families that will be useful to us later are the following, examples of which can be seen in Figure 2.

**Complete graphs:** The *complete graph*  $K_n$ ,  $n \geq 1$ , has  $n$  vertices, each pair of which are adjacent.

**Trees:** A connected graph is a *tree* if every pair of vertices has exactly one sequence of edges connecting them.

**Path graphs:** The *path graph*  $P_n$ ,  $n \geq 1$ , has  $n$  vertices and is the tree with 2 vertices of degree 1, and  $n-2$  vertices of degree 2 for  $n \geq 2$ , and  $P_1 = K_1$ .

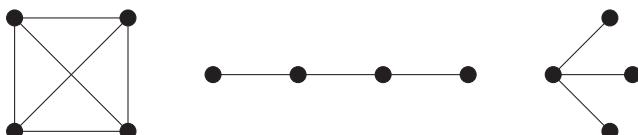


Figure 2. From left to right,  $K_4$ ,  $P_4$ , and the tree  $K_{1,3}$ , also called the claw.

A *proper coloring*  $\kappa$ , with  $k$  colors, of a graph  $G$  is a function

$$\kappa : V(G) \rightarrow \{1, 2, \dots, k\}$$

such that if  $u, v \in V(G)$  are adjacent, then  $\kappa(u) \neq \kappa(v)$ :



With this in mind, the *chromatic polynomial*  $\chi_G(k)$  of a graph  $G$  is the number of proper colorings of the vertices of  $G$  with  $k$  colors.

For example, let us color the vertices of  $K_n$  with  $k$  colors. For our first vertex we have  $k$  colors to choose from. Then since all our vertices are adjacent, by definition, for the second vertex we only have  $k - 1$  colors to choose from, for the third vertex we only have  $k - 2$  colors to choose from, and so on until for the  $n$ -th vertex we only have  $k - n + 1$  colors to choose from. Hence the total number of proper colorings of the  $n$  vertices of  $K_n$  with  $k$  colors is

$$k(k-1)(k-2) \cdots (k-n+1) = \chi_{K_n}(k).$$

As a second example, let us color the vertices of any tree  $T$  that has  $n$  vertices with  $k$  colors. Choose a vertex, any vertex. For it we have  $k$  colors to choose from. For all the vertices adjacent to it we have  $k - 1$  colors to choose from. For all the vertices adjacent to those we have  $k - 1$  colors to choose from, and for all the vertices adjacent to those we have  $k - 1$  colors to choose from, and so on. We note that since every pair of vertices has exactly one sequence of edges connecting them, by definition, we will only have one opportunity to color each vertex using our greedy algorithm, and when we do we will have  $k - 1$  colors to choose from after we color the first vertex. Hence the total number of proper colorings of the  $n$  vertices of  $T$  with  $k$  colors is

$$k(k-1)^{n-1} = \chi_T(k).$$

As we can see from these examples, the chromatic polynomial is in fact a polynomial, even though this may not seem immediate from the definition. Lastly, returning to the motivation for its definition, Birkhoff [3] had been hoping to solve the four-color problem by showing that  $\chi_G(4) > 0$  whenever  $G$  was a *planar* graph, namely a graph that can be drawn in the plane without edges crossing.

**The chromatic symmetric function:** An unlimited paint palette with a limited amount of paint. In order to describe our main protagonist we need to consider commuting variables  $x_1, x_2, \dots$  and extend our proper colorings from  $k$  colors to infinitely many colors, so that now a proper coloring  $\kappa$  of a graph  $G$  is a function

$$\kappa : V(G) \rightarrow \{1, 2, \dots\}$$

such that if  $u, v \in V(G)$  are adjacent, then  $\kappa(u) \neq \kappa(v)$ . However, we must not be misled by the infinite paint palette, because if we have a graph  $G$  with  $n$  vertices, then there is as much information contained in a paintbox with  $n$  colors as there is in a paintbox with infinitely many colors. Instead the crucial enhancement is that while  $\chi_G$  tells us the number of proper colorings with  $k$  colors for all  $k$ ,  $X_G$  defined below will tell us the more refined information of how many times each color is used. Moreover, because we can package all this information into a single symmetric function, as we will see, we can bring to bear the deep,

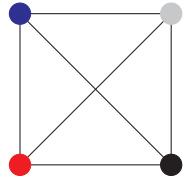
extensive, and well-developed theory of symmetric functions. Thus let us define the chromatic symmetric function, which was introduced in 1995 by Richard Stanley [16] as a generalization of the chromatic polynomial and has been a rich avenue of research ever since. A symmetric function is one where permuting the indices fixes the function, and so our chromatic symmetric function defined below is indeed a symmetric function because permuting the colors fixes the function.

**Definition 1.** For a graph  $G$  with vertex set  $V(G) = \{v_1, v_2, \dots, v_n\}$  the *chromatic symmetric function* is

$$X_G = \sum_{\kappa} x_{\kappa(v_1)} x_{\kappa(v_2)} \cdots x_{\kappa(v_n)}$$

where the sum is over all proper colorings  $\kappa$  of  $G$ .

For example, let us color  $K_4$  in Figure 2. Since every pair of vertices is adjacent, so will need different colors for a proper coloring, this means that all four vertices will need a different color



that can be permuted and hence, omitting details of the calculation,

$$X_{K_4} = 24x_1x_2x_3x_4 + 24x_1x_2x_3x_5 + \cdots + 24x_2x_7x_{439}x_{1001} + \cdots$$

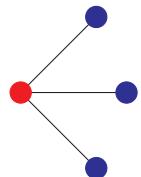
As a second example, let us color the two trees in Figure 2,  $P_4$  and  $K_{1,3}$ . Note that we can always color the vertices four different colors, so again terms such as  $x_1x_2x_3x_4$  will arise in both cases. However, if we try to color the vertices in only two colors, then in  $P_4$  we can color at most two vertices the same color



that can be swapped and hence, omitting details of the calculation,

$$X_{P_4} = 2x_1^2x_2^2 + 2x_1^2x_3^2 + \cdots + 24x_1x_2x_3x_4 + \cdots$$

In contrast, in  $K_{1,3}$  we can color three vertices the same color

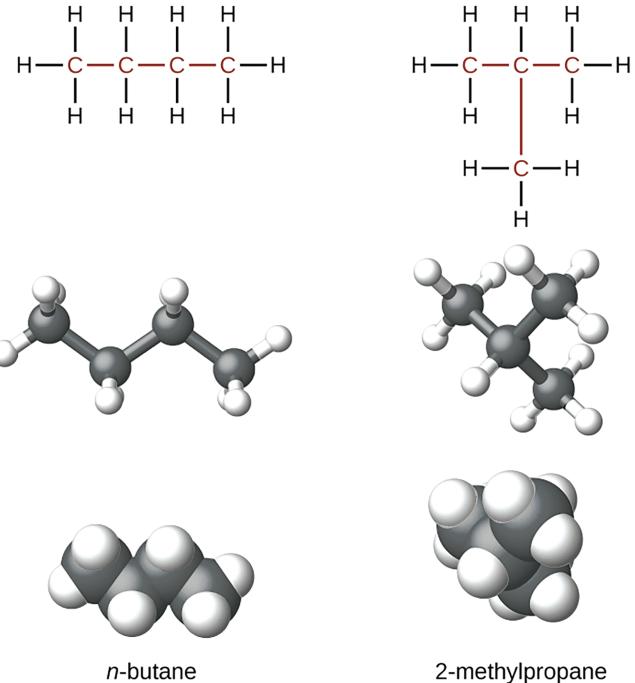


that can be swapped and hence, omitting details of the calculation,

$$X_{K_{1,3}} = x_1^3x_2 + x_1x_2^3 + \cdots + 24x_1x_2x_3x_4 + \cdots$$

In his paper, Stanley noted that if  $x_1 = x_2 = \cdots = x_k = 1$  and  $x_{k+1} = x_{k+2} = \cdots = 0$ , then  $X_G = \chi_G(k)$ , so the chromatic symmetric function naturally generalizes the chromatic polynomial. For example, it is well known that if we evaluate  $\chi_G(-1)$  then we obtain the number of acyclic orientations of  $G$  up to a sign [15], where an acyclic orientation is a placement of directions on the edges of  $G$  in such a way that following the directions never leads to us going round in a cycle. With this in mind,  $X_G$ , when written as a sum of elementary symmetric functions, which we will meet later, refines this result by determining the number of acyclic orientations of  $G$  with  $k$  sinks, where a sink is a vertex  $v$  whose incident edges are all directed towards  $v$ , and we look at the coefficients of those elementary symmetric functions whose indices have  $k$  parts. This result can in turn be further refined, but that is a story for another time [1]. However, there is one particular way in which  $\chi_G(k)$  and  $X_G$  seem to differ, which leads us to our first open problem.

**Distinguishing trees.** Given a hydrocarbon,  $C_nH_{2n+2}$ , how many isomers are there? Isomers have the same chemical formula but different configurations, so equivalently how many different configurations can  $C_nH_{2n+2}$  have? This problem remains open to this day.



**Figure 3.** Butane,  $C_4H_{10}$ , has two isomers [19].

However, we note that if we omit the hydrogens in the butane isomers, then we have two trees. In fact, they are  $P_4$  and  $K_{1,3}$  from earlier. This is true in general: if we omit the hydrogens, then we obtain a tree. So one way to

distinguish hydrocarbons would be to know when two trees are the same, or isomorphic. Informally two graphs are isomorphic if we can redraw one so it is the other. More formally, two graphs  $G_1$  and  $G_2$  are *isomorphic* if and only if there exists a 1-1 correspondence

$$\phi : V(G_1) \rightarrow V(G_2)$$

such that if  $u, v$  are adjacent in  $G_1$ , then  $\phi(u), \phi(v)$  are adjacent in  $G_2$ . In practice it is often easier to prove that two graphs are not isomorphic by finding a trait or *invariant* that differs, such as the number of vertices or edges. For example,  $P_4$  and  $K_{1,3}$  are not isomorphic because  $P_4$  has vertices of degree 2 whereas  $K_{1,3}$  does not. However, they both have the same chromatic polynomial,  $k(k-1)^3$ , so the chromatic polynomial is an example of an invariant that cannot be used to prove that  $P_4$  and  $K_{1,3}$  are not isomorphic. From earlier we know that every tree on  $n$  vertices has chromatic polynomial  $k(k-1)^{n-1}$ , so it can never be used to distinguish non-isomorphic trees with the same number of vertices. However, when we calculated the chromatic symmetric functions of  $P_4$  and  $K_{1,3}$  we saw that

$$X_{P_4} \neq X_{K_{1,3}}$$

so the chromatic symmetric function could be used to distinguish non-isomorphic trees. This was proposed by Stanley [16], and is now known as Stanley's tree problem, which we will formulate below as a conjecture to prove.

**Conjecture 2** (Stanley's tree problem). *If  $T_1$  and  $T_2$  are trees, then*

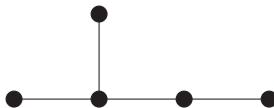
$$X_{T_1} = X_{T_2} \Leftrightarrow T_1 \text{ and } T_2 \text{ are isomorphic.}$$

This has been verified for all trees with up to 29 vertices [10], and has been proved for only a few families of trees:

**Spiders** [13]: A *spider* is a tree consisting of disjoint path graphs and a central vertex joined to a degree 1 vertex in each path graph.

**Caterpillars** [2, 11, 13]: A *caterpillar* is a tree consisting of a path graph with degree 1 vertices joined to (some of) the vertices in the path graph.

For example,



is both a spider and a caterpillar.

A positive outlook. For our second open problem we need to introduce the space that chromatic symmetric functions live in, the algebra of symmetric functions. The algebra of symmetric functions,  $\text{Sym}$ , is a subalgebra of  $\mathbb{Q}[[x_1, x_2, \dots]]$  generated by the  $n$ -th elementary symmetric functions  $e_n$  for  $n \geq 1$

$$e_n = \sum_{i_1 < i_2 < \dots < i_n} x_{i_1} x_{i_2} \cdots x_{i_n},$$

for instance  $e_4 = x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_5 + \dots + x_2 x_{27} x_{439} x_{1001} + \dots$ , and that

$$\text{Sym} = \mathbb{Q}[e_1, e_2, \dots]$$

and is freely generated by the  $e_n$  is known as the *fundamental theorem of symmetric functions*. A basis for  $\text{Sym}$  is given by all *elementary symmetric functions*

$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_k}$$

where  $\lambda = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$  is a list of weakly decreasing positive integers, for instance  $e_{211} = e_2 e_1 e_1$ , and we say that any  $f \in \text{Sym}$  is *e-positive* if it can be written as a positive linear combination of elementary symmetric functions, and is of interest due to its connection to permutation representations. For example,

$$X_{K_{1,3}} = 4e_4 + 5e_{31} - 2e_{22} + e_{211}$$

is not *e-positive*, and  $K_{1,3}$  is the smallest graph whose chromatic symmetric function is not *e-positive*. However,

$$X_{K_4} = 24e_4$$

is *e-positive*. More generally, we can see that

$$X_{K_n} = n! e_n$$

where  $n! = n(n-1)\cdots 321$  because each vertex must be colored a different color, and hence each proper coloring gives a monomial

$$x_{i_1} x_{i_2} \cdots x_{i_n}$$

where all the indices are different. Also the colors of a given proper coloring can be permuted in  $n!$  different ways, and each gives a valid proper coloring. Hence

$$X_{K_n} = n! \sum_{i_1 < i_2 < \dots < i_n} x_{i_1} x_{i_2} \cdots x_{i_n} = n! e_n.$$

The complete graphs are a special case of a family of graphs called unit interval graphs. A *unit interval graph* on  $n$  vertices  $G_{\mathcal{C}}$  is formed as follows. Choose a collection  $\mathcal{C}$  of intervals  $I = [i, j] = \{i, i+1, \dots, j\} \subseteq \{1, 2, \dots, n\} = [n]$  such that no interval contains another. Then  $G_{\mathcal{C}}$  has vertices labeled  $1, 2, \dots, n$  and an edge between vertices  $u, v$  if  $u, v \in I$  for some  $I \in \mathcal{C}$ . For example, let  $G_{\mathcal{C}}$  be a graph on 4 vertices. If  $\mathcal{C} = \{[1, 4]\}$ , then  $G_{\mathcal{C}} = K_4$ , and if  $\mathcal{C} = \{[1, 2], [2, 3], [3, 4]\}$ , then  $G_{\mathcal{C}} = P_4$ . However,  $G_{\mathcal{C}} \neq K_{1,3}$  for any  $\mathcal{C}$ . This leads us to our second open problem that was originally posed as a conjecture by Richard Stanley and John Stembridge in 1993 in terms of partially ordered sets not containing a chain of three elements and one further element incomparable to the chain. While the conjecture's name still bears this terminology, it is easiest for us to state it using unit interval graphs, a reduction that is due to Mathieu Guay-Paquet [9].

**Conjecture 3** (The  $(3+1)$ -free conjecture). *If  $G_{\mathcal{C}}$  is a unit interval graph, then  $X_{G_{\mathcal{C}}}$  is *e-positive*.*

Number of vertices, $n$	1	2	3	4	5	6	7	8	9	10	11	12	13
Number of trees, $T$	1	1	1	2	3	6	11	23	47	106	235	551	1301
$X_T$ is $e$ -positive	1	1	1	1	2	1	3	1	2	2	5	1	4

**Table 1.** Data on  $e$ -positivity and trees.

Like Stanley’s tree problem, this conjecture has been proved for only a few families, for example considering graphs  $G_C$  on  $n$  vertices we have

- Complete graphs, namely  $\mathcal{C} = \{[1, n]\}$ ;
- Path graphs, namely  $\mathcal{C} = \{[1, 2], [2, 3], \dots, [n-1, n]\}$  [16];
- Complements of triangle-free graphs, namely  $\mathcal{C} = \{[1, i], [i-j, n]\}$  [17];
- $K_\alpha$ -chains, namely  $\mathcal{C} = \{[1, \alpha_1], [\alpha_1, \alpha_2], \dots, [\alpha_k, n]\}$  and  $\alpha_1 < \alpha_2 < \dots < \alpha_k < n$  [8];
- Triangular ladders, namely  $\mathcal{C} = \{[1, 3], [2, 4], \dots, [n-2, n]\}$  [5].

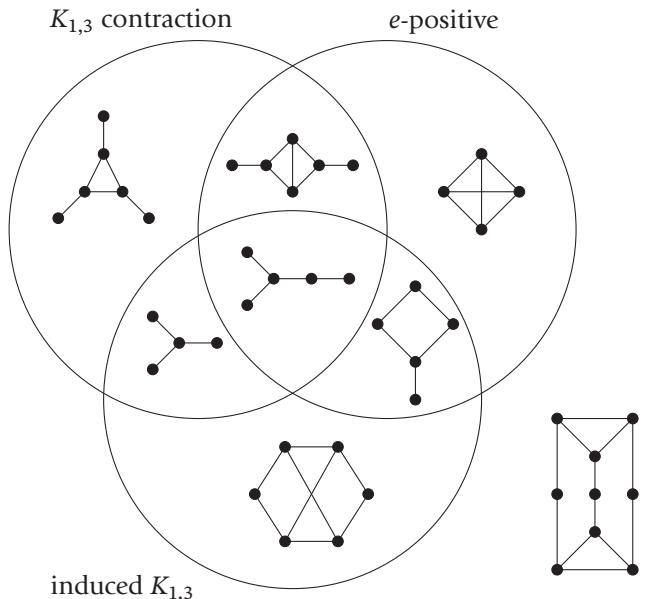
In particular, a number of these and other results have been proved using generalizations of the chromatic symmetric function to quasisymmetric functions [14] or to symmetric functions in noncommuting variables [8]. Furthermore, through the generalization to quasisymmetric functions, chromatic symmetric functions of unit interval graphs have been shown to encode the decomposition into irreducible representations of a symmetric group action on the cohomology of regular semisimple Hessenberg varieties, for example [4].

One thing we can observe from the definition of unit interval graphs is that they can be thought of as a collection of complete graphs that have been lined up in a row and then squashed together. Looking individually at each complete graph we notice that it does not contain  $K_{1,3}$  as an *induced subgraph* (a subset of vertices together with all edges connecting them). This led Stanley to widen his conjecture to propose that the chromatic symmetric function of all graphs that do not contain  $K_{1,3}$  as a *contraction* (shrinking edges and identifying vertices) are  $e$ -positive. This conjecture was recently shown to be false by Samantha Dahlberg, Angèle Foley and the author [6] who proved there exist infinite families of graphs that satisfy every combination of

- they contain  $K_{1,3}$  as an induced subgraph;
- they contain  $K_{1,3}$  as a contraction;
- their chromatic symmetric functions are  $e$ -positive.

Some of these can be seen in Figure 4.

Now that we have met these two famed open problems on chromatic symmetric functions, it is time to tie these two research threads together and end with one final open problem. From our discussion on the  $(3+1)$ -free conjecture we see that the chromatic symmetric function of the tree  $K_{1,3}$  is not  $e$ -positive, although that of the path graph  $P_4$  is. So a natural question to ask is when is the



**Figure 4.** Graphs satisfying every combination.

chromatic symmetric function of a tree  $e$ -positive? The data computed so far by Samantha Dahlberg, Adrian She and the author [7] in Table 1 implies the answer is not often beyond the chromatic symmetric function of the path graphs, and their conjecture is our final open problem, which has been confirmed by Kai Zheng [20] for degree  $\geq 6$ .

**Conjecture 4.** *If a tree contains a vertex of degree  $\geq 4$ , then its chromatic symmetric function is not  $e$ -positive.*

**ACKNOWLEDGMENTS.** The author would like to thank the referees, Farid Aliniaeifard, Niall Christie, Sheila Sundaram, and Victor Wang for helpful suggestions, and especially Richard Stanley whose suggestions, insights, and stories made writing and researching this article all the more chromatic.

## References

- [1] F. Aliniaeifard, V. Wang, and S. van Willigenburg, *The chromatic symmetric function of a graph centred at a vertex*, arXiv:2108.04850v1, 2021.
- [2] José Aliste-Prieto and José Zamora, *Proper caterpillars are distinguished by their chromatic symmetric function*, Discrete Math. 315 (2014), 158–164, DOI 10.1016/j.disc.2013.10.016. MR3130368

- [3] George D. Birkhoff, *A determinant formula for the number of ways of coloring a map*, Ann. of Math. (2) 14 (1912/13), no. 1-4, 42–46, DOI 10.2307/1967597. MR1502436
- [4] Patrick Brosnan and Timothy Y. Chow, *Unit interval orders and the dot action on the cohomology of regular semisimple Hessenberg varieties*, Adv. Math. 329 (2018), 955–1001, DOI 10.1016/j.aim.2018.02.020. MR3783432
- [5] S. Dahlberg, *Triangular ladders  $P_{d,2}$  are e-positive*, arXiv:1811.04885v1, 2018.
- [6] Samantha Dahlberg, Angèle Foley, and Stephanie van Willigenburg, *Resolving Stanley's e-positivity of claw-contractible-free graphs*, J. Eur. Math. Soc. (JEMS) 22 (2020), no. 8, 2673–2696, DOI 10.4171/JEMS/974. MR4118618
- [7] Samantha Dahlberg, Adrian She, and Stephanie van Willigenburg, *Schur and e-positivity of trees and cut vertices*, Electron. J. Combin. 27 (2020), no. 1, Paper No. 1.2, 22. MR4057182
- [8] David D. Gebhard and Bruce E. Sagan, *A chromatic symmetric function in noncommuting variables*, J. Algebraic Combin. 13 (2001), no. 3, 227–255, DOI 10.1023/A:1011258714032. MR1836903
- [9] M. Guay-Paquet, *A modular relation for the chromatic symmetric functions of (3+1)-free posets*, arXiv:1306.2400v1, 2013.
- [10] S. Heil and C. Ji, *On an algorithm for comparing the chromatic symmetric functions of trees*, Australas. J. Combin. 75 (2019), 210–222. MR4008247
- [11] Martin Loebl and Jean-Sébastien Sereni, *Isomorphism of weighted trees and Stanley's isomorphism conjecture for caterpillars*, Ann. Inst. Henri Poincaré D 6 (2019), no. 3, 357–384, DOI 10.4171/AIHPD/74. MR4002670
- [12] MacTutor, *The four colour theorem*, [https://mathshistory.st-andrews.ac.uk/HistTopics/The\\_four\\_colour\\_theorem/](https://mathshistory.st-andrews.ac.uk/HistTopics/The_four_colour_theorem/).
- [13] Jeremy L. Martin, Matthew Morin, and Jennifer D. Wagner, *On distinguishing trees by their chromatic symmetric functions*, J. Combin. Theory Ser. A 115 (2008), no. 2, 237–253, DOI 10.1016/j.jcta.2007.05.008. MR2382514
- [14] John Shareshian and Michelle L. Wachs, *Chromatic quasisymmetric functions*, Adv. Math. 295 (2016), 497–551, DOI 10.1016/j.aim.2015.12.018. MR3488041
- [15] Richard P. Stanley, *Acyclic orientations of graphs*, Discrete Math. 5 (1973), 171–178, DOI 10.1016/0012-365X(73)90108-8. MR317988
- [16] Richard P. Stanley, *A symmetric function generalization of the chromatic polynomial of a graph*, Adv. Math. 111 (1995), no. 1, 166–194, DOI 10.1006/aima.1995.1020. MR1317387
- [17] Richard P. Stanley and John R. Stembridge, *On immanants of Jacobi-Trudi matrices and permutations with restricted position*, J. Combin. Theory Ser. A 62 (1993), no. 2, 261–279, DOI 10.1016/0097-3165(93)90048-D. MR1207737
- [18] Wikimedia Commons, <https://commons.wikimedia.org/wiki/File:DeMorganFourColour.png>.
- [19] Wikimedia Commons, [https://commons.wikimedia.org/wiki/File:CNX\\_Chem\\_20\\_01\\_butaneIsom\\_img.png](https://commons.wikimedia.org/wiki/File:CNX_Chem_20_01_butaneIsom_img.png).
- [20] K. Zheng, *On the e-positivity of trees and spiders*, arXiv:2008.05038v2, 2020.



Stephanie van Willigenburg

#### Credits

Opening image is courtesy of bgblue via Getty.

Figure 1 is courtesy of Wikimedia Commons.

Figure 3 is courtesy of OpenStax. Licensed under Creative Commons Attribution 4.0 International.

All other figures are courtesy of the author.

Photo of Stephanie van Willigenburg is by Niall Christie.



# Call for Proposals RIMS Joint Research Activities 2023-2024

**Application deadline : August 31, 2022, 23:59 (JST)**

**Types of Joint Research Activities** **More Information : RIMS**

\*RIMS Satellite seminars 2023 Int.JU/RC Website

\*RIMS Review seminars 2023 <https://www.kurims.kyoto-u.ac.jp/kyoten/en/>

\*RIMS Workshops Type C 2023

\*RIMS Research Project 2024



京都大學  
KYOTO UNIVERSITY



RIMS Research Institute for  
Mathematical Sciences